

Comparison of Clinical and Pathological Staging in Head and Neck Squamous Cell Carcinoma

Results From Intergroup Study ECOG 4393/RTOG 9614

Wayne M. Koch, MD; John A. Ridge, MD, PhD; Arlene Forastiere, MD; Judith Manola, MS

Objectives: To compare the results of clinical and pathological staging for a large cohort of patients with head and neck squamous cell carcinoma (HNSCC) and to examine patterns and ramifications of the disparity between staging methods.

Design: Prospective inception cohort (median follow-up, 7 years).

Setting: Multi-institutional cooperative group study (Eastern Cooperative Oncology Group 4393/Radiation Therapy Oncology Group 9614) involving 17 academic medical centers.

Patients: A total of 560 patients with new-onset or recurrent HNSCC enrolled during a 7-year period.

Interventions: Surgical resection with curative intent with or without adjuvant or previous radiotherapy or chemotherapy.

Main Outcome Measures: Clinical staging and pathological staging and the component TN tumor categories

were compared with overall and disease-specific survival. Association of survival with staging was derived by means of the proportional hazards model.

Results: Of the 501 cases in which both clinical and pathological staging was available, a disparity was found between at least 1 component tumor category assigned by the 2 methods in almost 50% of cases. Both methods showed a strong association of stage with overall survival for the cohort at large. However, pathological nodal category was a superior predictor ($P < .001$ vs $P = .005$), whereas there was an advantage to pathological tumor category in predicting disease-specific survival ($P = .01$).

Conclusions: Both staging methods are useful in predicting survival, whereas information gained at neck dissection regarding nodal metastases provides some refinement in prognostic results. These findings demonstrate the need for enhanced methods of tumor assessment and apparent benefit of data gathered at neck dissection for accurate disease assessment and stratification.

Arch Otolaryngol Head Neck Surg. 2009;135(9):851-858

Author Affiliations:

Department of Otolaryngology–Head and Neck Surgery, The Johns Hopkins University, Baltimore, Maryland (Drs Koch and Forastiere); Head and Neck Surgery, Fox Chase Cancer Center, Philadelphia, Pennsylvania (Dr Ridge); and Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, Massachusetts (Ms Manola).

CLINICAL DATA REGARDING a large cohort of patients with head and neck squamous cell carcinoma (HNSCC) have been gathered through the auspices of 2 cooperative groups, the Eastern Cooperative Oncology Group (EGOG) and the Radiation Therapy Oncology Group (RTOG). Study ECOG 4393/RTOG 9614 has as its first objective to determine the clinical utility of molecular detection of cancer cells in tumor margins, which continues under analysis. An independent second objective of the protocol is to determine the incidence of *TP53* (OMIM *191170) mutation in HNSCC and seek associations between *TP53* status and clinical outcome. This analysis forms the basis of a recent publication.¹ The current analysis was motivated by the observation that pathological N category, but not clinical T or N category, was associated with *TP53*

mutation status. The reasons for this variance were not immediately apparent, and so we set out to explore differences between the 2 staging paradigms. Our hypothesis was that pathological stage should correspond more faithfully to biological aggressiveness of disease, resulting ultimately in greater precision in predicting outcome. This study was undertaken to further elucidate the significance of clinical and pathological stage and the component T and N categories as prognostic factors for overall and disease-specific survival. This is, to our knowledge, the first such report in the literature to compare clinical and pathological stage in a large cohort of patients with HNSCC.

METHODS

From January 17, 1996, to October 18, 2002, 560 patients with HNSCC were enrolled in a prospective multicenter study involving 17

member institutions of ECOG and RTOG. The protocol was approved by the cooperative groups and by each participating institutional investigational review board. Subjects provided written informed consent. Individuals with newly diagnosed or recurrent HNSCC were eligible if the treatment plan included primary surgical extirpation with curative intent. The cohort included 118 patients in whom cancer at presentation had persisted or recurred after previous therapy (typically radiotherapy or wide local excision).

SAMPLE COLLECTION

Tumor and margin samples were collected during the operation. Demographic and clinical data were collected from participating institutions by the ECOG and RTOG data managers perioperatively and at scheduled intervals during the follow-up period. Follow-up data from patients who were alive without recurrence at the time of the analysis were censored at the last follow-up contact with those patients. Clinical stage was evaluated by each local treating surgeon and reported at study entry by means of American Joint Committee on Cancer criteria.² Clinical stage of patients with recurrent or persistent cancer was to be that of the index presentation, following clinical staging convention. At the time of subject enrollment, computerized imaging (computed tomography and magnetic resonance imaging) was typically performed and results were included in the staging, but positron emission tomographic scanning was not used routinely. Pathology reports were submitted to ECOG/RTOG and the findings were reviewed, pathological stage was determined, and findings were tabulated. At 6-month intervals for the first 3 years and annually thereafter, the status of each subject, which included information about recurrent or second primary cancer and subsequent treatment, was reported to the ECOG central office.

STATISTICAL ANALYSIS

Analysis was performed at the ECOG Statistical Center. Descriptive statistics were used to characterize patients at baseline. Survival curves were derived by the Kaplan-Meier method,³ and differences were examined by means of the log-rank test.⁴ Survival was defined as the time from study entry to death or to the last follow-up visit. Progression-free survival was defined as the time from study entry to death or recurrence. Data from patients who were alive without recurrence at the time of the analysis were censored at the last follow-up contact with those patients. The Fisher exact test,⁵ unpaired, 2-sided *t* test, and Mehta exact test for ordered categorical data⁶ were used to compare patient categories.

Proportional hazards models⁷ were used to assess the univariate prognostic significance of tumor variables on overall and disease-specific survival. *P* values shown with hazard ratios are from the likelihood ratio test. Hazard ratios were calculated relative to a reference group as shown for univariate analyses. To better elucidate cause-specific mortality, the method developed by Gray⁸ was used to examine competing causes of death, distinguishing between death from head and neck cancer and death from other or unknown causes.

To explore further the association between stage and survival, multivariable models were developed. Factors of potential interest included sex, age, site of the primary tumor, ECOG performance status, smoking history, alcohol use at study entry, cell differentiation, status of disease at study entry (eradicated, recurrent, residual disease after previous therapy, or untreated), and whether the tumor extended into adjacent structures. To reduce the impact of missing data, factors were introduced into the model as level-specific indicator variables, with a level for missing values. Models were evaluated by means of the Akaike information

criterion,⁹ which moderates the effect of the number of covariates on changes in the log likelihood by taking the number of covariates into account. First, the full model was evaluated; factors were then removed if their elimination resulted in a model with improved Akaike information criterion.

RESULTS

A total of 560 patients were registered to the study. Fifteen patients were ineligible because of concurrent cancer within 5 years (1 patient), consent given after surgery (1), metastases at study entry (2), distant metastases not assessed (2), no surgery performed (3), or no HNSCC identified at resection (6), and comparable stage information (both T and N category) was missing for 43 patients. One patient was missing a valid survival time interval. The remaining 501 patients are the focus of this analysis. This group is more extensive than the group analyzed for *TP53* mutation status² because inclusion was not dependent on the availability or integrity of tissues for molecular analysis.

Survival was evaluated as of June 16, 2008, and was defined as the time from registration to death or the last follow-up visit. Among the 501 patients, 289 had died. Median follow-up among patients still alive was 7 years. Characteristics of patients included in this analysis are provided in **Table 1**, along with numbers of deaths, median survival by category, 95% confidence intervals, and *P* values for the log-rank test.

Tables 2, 3, and 4 show the relationship between clinical and pathological T category, N category, and overall stage, respectively. Perfect concordance between clinical and pathological T category was 52.2%; N category, 53.5%; and overall stage, 54.9%. In particular, nearly 40% (38 of 97) of the cases judged to be T4 clinically were found to be in a lower T category at pathological evaluation, whereas more than 40% of cases (43 of 102) found to be T4 on pathological evaluation had been classified at less than T4 clinically (Table 2). Only 69.7% of cases judged to be N0 clinically (145 of 208) were found to be N0 pathologically (true-negative clinical staging), which resulted in 30.3% false-negative clinical staging (82 cases could not be scored for pN category because no neck dissection had been performed). Twenty-six of 172 cases (15.1%) found not to have tumor in dissected lymph nodes (pN0) had been clinically judged to have nodal metastasis (cN+) (false-positive clinical staging) (Table 3).

Table 5 provides overall survival by each level of clinical and pathological T and N category and by combined stage for all patients with known stage. The numbers of patients and deaths may therefore vary among categories. There was a strong association between both clinical and pathological stage and overall survival. The median survival for those judged N0 by pathological evaluation was 7.8 years, compared with 5.9 years for those judged N0 on the basis of clinical criteria.

Table 6 provides survival by each level of clinical and pathological T and N category and by overall stage for the 380 patients who were registered with a disease status of "untreated" at study entry. These were assessed alone because it was assumed that the TNM

Table 1. Patient Characteristics and Univariate Survival Analyses

Characteristic	No. (%) ^a	Deaths, No.	Years After Surviving Surgery, Median (95% CI)	Log-Rank P Value
Age, y				
<55	165 (32.9)	73	8.4 (5.3-∞)	<.001
55-64	130 (25.9)	72	4.7 (2.8-7.4)	
≥65	206 (41.1)	144	3.3 (2.9-4.2)	
Sex				
Female	140 (27.9)	79	5.6 (3.5-6.8)	.45
Male	361 (72.1)	210	4.3 (3.2-5.3)	
Weight loss previous 6 mo, %				
<5.0	340 (74.2)	180	5.3 (4.2-7.3)	<.001
5.0-9.9	64 (14.0)	46	2.6 (1.5-5.1)	
10.0-19.9	39 (8.5)	26	1.7 (0.6-5.7)	
≥20	15 (3.3)	14	1.2 (0.6-1.5)	
Unknown ^b	43			
Smoking history				
Never smoked	92 (18.8)	45	6.6 (3.8-∞)	.15
Pipe/cigar	18 (3.7)	11	2.4 (2.0-∞)	
Cigarettes, pack-y				
<20	60 (12.3)	31	5.1 (3.2-∞)	
20-40	131 (26.8)	78	4.0 (2.8-5.6)	
>40	188 (38.4)	120	3.7 (2.9-5.7)	
Unknown ^b	12			
Average alcohol consumption, mL/wk				
<300	262 (56.5)	142	5.6 (4.0-6.8)	.006
300-960	95 (20.5)	65	2.8 (1.6-4.3)	
>960	107 (23.1)	63	3.7 (3.0-7.3)	
Unknown ^b	37			
Performance status				
0	365 (75.3)	191	5.3 (4.0-7.5)	<.001
1	102 (21.0)	76	2.4 (1.8-4.0)	
2	14 (2.9)	11	1.2 (0.3-5.9)	
3	3 (0.6)	3	4.2 (0.6-4.7)	
4	1 (0.2)	1	2.1 (NA)	
Unknown ^b	16			
Cell differentiation				
Well differentiated	105 (22.2)	58	5.3 (3.2-7.9)	.74
Moderately differentiated	273 (57.7)	158	4.8 (3.4-5.7)	
Poorly differentiated	93 (19.7)	56	3.8 (2.6-6.0)	
Undifferentiated	2 (0.4)	1	NYR (5.6-∞)	
Unknown ^b	28			
Site				
Oral cavity	217 (43.5)	121	4.3 (3.4-7.3)	<.001
Oropharynx	108 (21.6)	56	6.0 (4.7-∞)	
Hypopharynx	36 (7.2)	27	2.1 (1.3-3.4)	
Larynx	112 (22.4)	69	4.0 (3.0-5.7)	
Other	23 (4.6)	12	6.8 (2.6-∞)	
>1 Primary tumor	3 (0.6)	3	0.4 (0.1-1.3)	
Unknown ^b	2			
Extension of primary tumor to adjacent structures				
No	271 (55.6)	138	5.9 (5.0-8.1)	<.001
Yes	216 (44.4)	144	3.2 (2.2-3.8)	
Unknown ^b	14			

Abbreviations: CI, confidence interval; NA, not applicable; NYR, not yet reached.

^aBecause of rounding, percentages may not total 100.

^bPatients in the "Unknown" category were excluded from calculation of the percentages.

Table 2. Concordance Between Clinical and Pathological T Category^a

Clinical T Stage	Pathological T Stage					Total
	1	2	3	4	In Situ	
1	69	19	3	8	1	104
2	51	85	13	19	0	170
3	11	48	47	15	0	123
4	2	16	18	59	1	97
X	0	0	2	1	0	3
In situ	0	0	0	0	1	1
Total	133	168	83	102	3	498

^aFrequency was missing for 3 patients. Concordance rate was 52.2%.

Table 3. Concordance Between Clinical and Pathological N Category^a

Clinical N Stage	Pathological N Stage					Total
	0	1	2	3	X	
0	145	29	33	1	82	290
1	15	27	38	0	2	82
2	9	13	92	0	1	115
3	2	0	5	2	0	9
X	1	0	1	0	1	3
Total	172	69	169	3	86	499

^aFrequency was missing for 2 patients. Concordance rate was 53.5%.

Table 4. Concordance Between Clinical and Pathological Overall Stage^a

Clinical Stage	Pathological Stage				Total
	I	II	III	IV	
I	20	5	8	9	42
II	19	24	9	21	73
III	8	15	30	49	102
IV	4	11	23	146	184
Total	51	55	70	225	401

^aFrequency was missing for 100 patients. Concordance rate was 54.9%.

category provided at study entry would be most closely associated with prognosis for these index-case patients. Again, among this subset of patients, there was a strong association between both clinical and pathological stage and overall survival.

Figure 1 and **Figure 2** show Kaplan-Meier survival curves for overall survival for all patients and for untreated patients only by clinical and pathological T and N category. Note the wider separation of curves for pathological categorization of nodal status than for clinical category figures. **Figure 3** and **Figure 4** show the cumulative incidence of death due to disease (competing cause of death: death from unrelated or unknown cause) for the entire population and also for patients entering the study with newly diagnosed cancers, demonstrating the association between rate of death and overall stage, both clinical and pathological. In each case, the survival curves

Table 5. Impact of Stage on Overall Survival, All Patients

Stage	No.	Deaths, No. ^a	Median (95% CI), y	Log-Rank P Value
Clinical				
T category				
1	104	46	NYR (5.3-∞)	<.001
2 ^b	171	93	5.1 (3.5-9.7)	
3	124	79	3.7 (2.4-5.3)	
4	98	98	2.1 (1.3-3.6)	
Pathological				
T category				
1	133	65	6.5 (4.1-∞)	.003
2	168	94	5.6 (3.5-7.3)	
3	83	54	3.1 (2.1-5.3)	
4	102	67	2.9 (1.7-4.0)	
Clinical				
N category				
0	290	153	5.9 (4.7-7.8)	.002
1	82	47	4.5 (2.1-7.7)	
2	115	78	2.4 (1.6-3.8)	
3	9	7	2.0 (0.4-5.1)	
Pathological				
N category				
0	173	79	7.8 (5.7-∞)	<.001
1	69	42	4.0 (1.9-7.4)	
2	170 ^c	118	2.1 (1.4-2.9)	
3	3	1	NYR (2.0-∞)	
Clinical stage				
I	42	15	NYR (5.9-∞)	<.001
II	73	37	6.0 (3.1-∞)	
III	102	51	7.3 (3.9-9.4)	
IV	184	125	2.6 (1.8-4.0)	
Pathological				
stage				
I	51	21	NYR (5.7-∞)	.004
II	55	30	5.7 (3.2-9.4)	
III	70	36	6.5 (3.9-8.4)	
IV	225	141	2.8 (1.9-4.0)	

Abbreviations: CI, confidence interval; NYR, not yet reached.

^aDeaths among patients in categories not shown are excluded.

^bOne patient in each category had an unknown pathological stage; these patients were excluded from Table 2 but are included here.

^cOne patient with the pN2 stage was missing the clinical stage. That patient is included here but was excluded from Table 3.

for death owing to disease were separated more widely and associated with stage compared with curves for death due to other or unknown cause. For death owing to disease, the association of survival with pathological stage was stronger than the association with clinical stage.

Table 7 provides hazard ratios and 95% confidence intervals from the optimal multivariable model developed. We conclude that age, ECOG performance status, clinical tumor category, and pathological nodal category are significant prognostic factors for overall survival.

We explored whether subgroups that demonstrated disparity between clinical and pathological stage had corresponding differences in outcome. If patients with perfect T or N category concordance were compared with those without concordance (in which pathological staging differed from clinical), no statistically significant differences were identified (data not shown). Disparate groupings included the 8 cases evaluated as cT1 and pT4, 18 cases evaluated as cT4 and pT1 or

Table 6. Impact of Stage in Previously Untreated Patients

Stage	No.	Deaths, No. ^a	Median (95% CI), y	Log-Rank P Value
Clinical T category				
1	68	29	NYR (5.3-∞)	<.001
2	131	61	9.7 (5.2-∞)	
3	97	60	4.1 (2.6-6.6)	
4	83	54	2.4 (1.5-5.1)	
Pathological T category				
1	103	47	7.7 (5.1-∞)	<.001
2	127	63	6.6 (4.3-∞)	
3	63	42	2.6 (1.8-6.5)	
4	79	50	3.7 (1.9-4.5)	
Clinical N category				
0	204	97	7.6 (5.7-∞)	.005
1	68	38	5.1 (2.1-∞)	
2	98	63	2.8 (1.7-5.3)	
3	6	4	2.3 (0.4-∞)	
Pathological N category				
0	128	52	9.4 (6.5-∞)	<.001
1	56	33	4.5 (2.3-7.7)	
2	140	92	2.6 (1.8-3.8)	
3	3	1	NYR (2.0-∞)	
Clinical stage				
I	27	10	NYR (5.9-∞)	<.001
II	55	22	NYR (4.2-∞)	
III	81	40	7.7 (4.0-∞)	
IV	158	102	2.9 (2.0-5.1)	
Pathological stage				
I	42	16	NYR (5.9-∞)	.01
II	39	17	7.8 (4.1-∞)	
III	54	29	5.2 (2.9-7.7)	
IV	186	112	3.7 (2.4-5.3)	

Abbreviations: CI, confidence interval; NYR, not yet reached.

^aDeaths among patients in categories not shown are excluded.

pT2, 33 cases evaluated as cN0 but pN2, and 24 cases evaluated as cN1 or cN2 and pN0. Median overall survival and P values for the log-rank test for these groups of patients are given in **Table 8**. The discordant groups are small, and the statistical power to detect differences may be limited. However, cases for which nodal status changed from N0 to N+ and those changing from N+ to N0 had significantly different length of survival compared with cases correctly classified by clinical method, indicating the greater prognostic accuracy of pathological nodal assessment.

COMMENT

The results of our analysis reconfirm that both clinical and pathological staging are strongly predictive of clinical outcome, despite the fact that components of the 2 forms of staging are discordant in nearly one-half of cases. Among all patients, pathological T and N categories were slightly more closely associated with deaths from head and neck cancer (disease-specific survival) than was clinical stage. Pathological nodal category was a somewhat better predictor of overall survival. This was also true when previously untreated patients were considered alone. There were 274 patients who were discordant in either T or N category, and 181 (66.1%) of these had discordant overall stage. The other 93 patients had a T or N

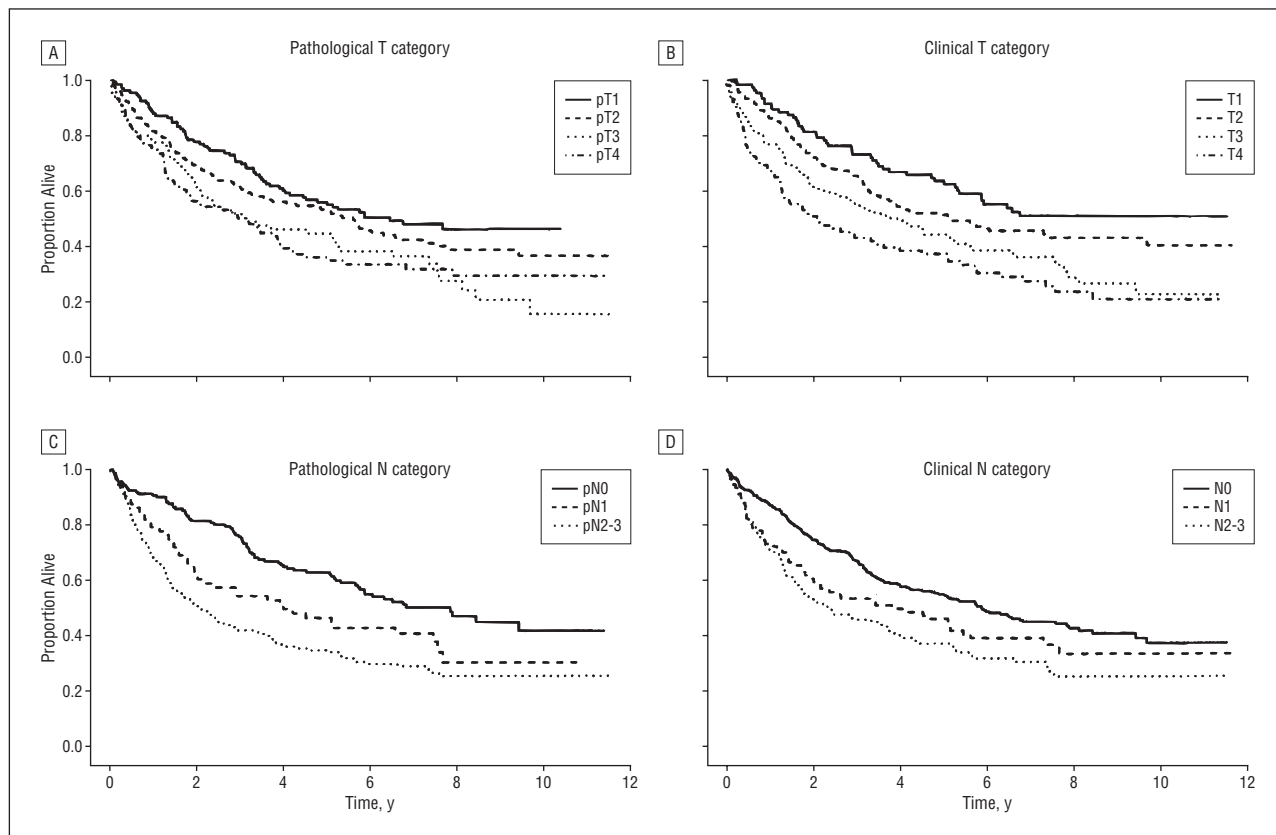


Figure 1. Overall survival for all patients by pathological (A) and clinical (B) T category and by pathological (C) and clinical (D) N category. Details about the groups are provided in Table 5. The clinical T category performed somewhat better than the pathological T category in distinguishing groups with different survival, whereas the pathological N category outperformed the clinical N category, although both clinical and pathological categorizations were effective. Note the wider separation of curves for nodal category assessed by the pathological vs the clinical method.

shift that did not result in a change to overall stage (change either did not affect overall stage or canceled reciprocal change in the other category). This may explain, in part, why discordances in T and N components were associated with detectable differences in outcome whereas comparison of the method of overall staging did not result in significant differences in survival. Alternatively, the results might be interpreted as demonstrating deficiencies in the overall staging approach (Table 7).

There are a number of reasons for differences between clinical and pathological staging. Judgment of size and depth of cancer penetration form the bases of most components of the primary T category. Clinical T category is based on size as judged by physical and radiographic examination. These methods have limited accuracy, making overestimation and underestimation possible. On the other hand, pathological evaluation of T category is done by the pathologist either at the time of gross inspection of the resected specimen or after fixation, sectioning, and microscopic analysis. Pathologic assessment, then, may be flawed by separation or filleting of tumor fragments at the time of resection and by shrinkage in the fixation process. For our population as a whole, the number of cases that had lower pathological than clinical T category was balanced by those that had greater pathological than clinical T status. In general, minor adjustments in T category comparing cT with pT did not result in meaningful differences in the association with

survival. It would seem that, when a tumor thought on clinical evaluation to be T4 (usually because of suspected bone or cartilage involvement) was found to be of a lesser pathological class, the adjustment should result in a more accurate and better prognosis for that group. Similarly, cases judged not to be T4 clinically yet found to be T4 pathologically would be expected to have substantially worse prognosis. However, as seen in Table 8, differences in survival associated with these adjustments did not achieve statistical significance, although the mean survivals of these groups appear to match the hypothesis. It is noteworthy that the pT4 group had a better survival outcome than the pT3 group. The reason for this unexpected result is not clear. The number of cases within subgroupings limits the power of the comparisons. Although cT category had a lower *P* value for its association with overall survival, when the association between T category and disease-specific survival was evaluated, pT had a slight advantage (shown by more evenly distributed and wider separation between pT and cT category curves in Figure 3A and Figure 4A vs Figure 3B and Figure 4B).

Clinical nodal categorization is a simple process based primarily on the size of suspicious lymph nodes, combined with radiographic features such as loss of fatty hilum, central necrosis, or increased vascularity. (Note, again, that all cases were accrued and staged before the availability of positron emission tomographic and com-

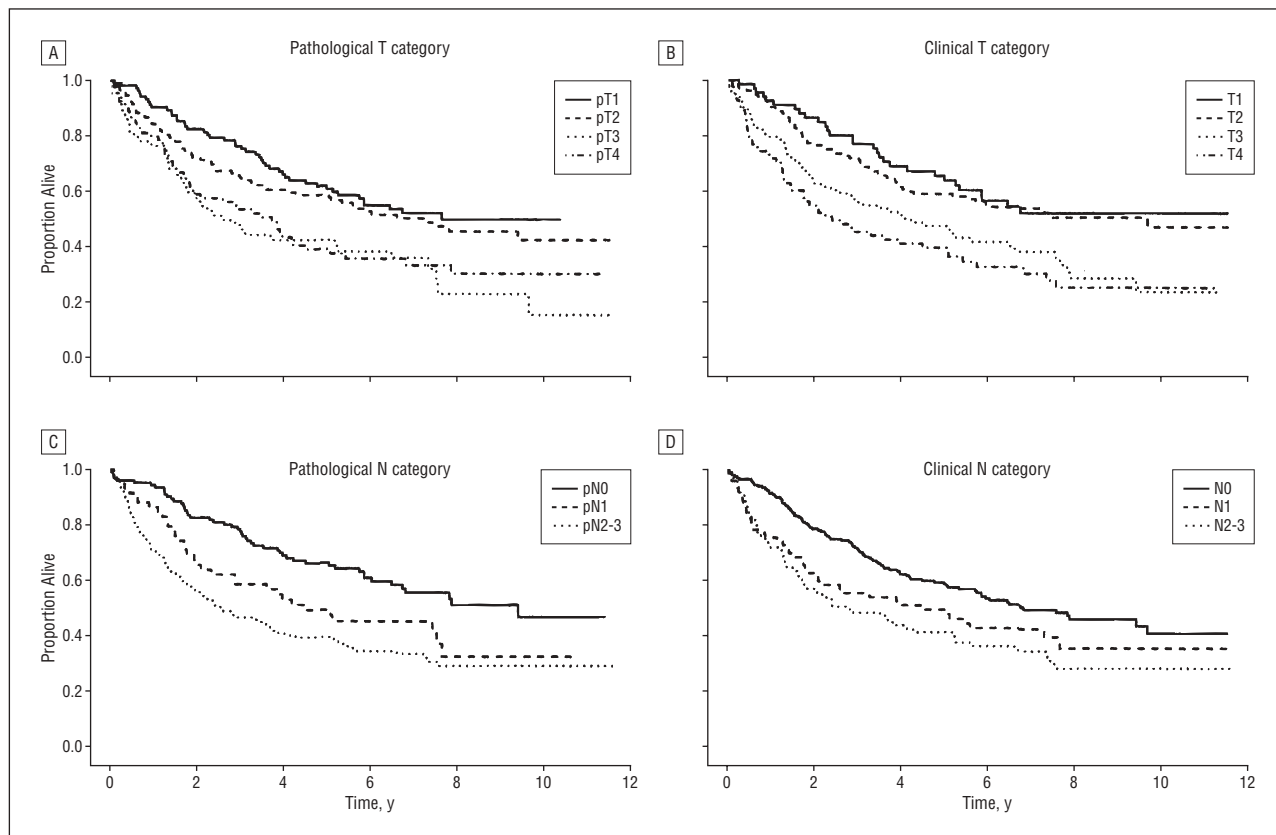


Figure 2. Overall survival for newly diagnosed (previously untreated) patients by pathological (A) and clinical (B) T category and by pathological (C) and clinical (D) N category.

puted tomographic scanning.) None of the radiographic features is necessarily indicative of actual tumor involvement, and early nodal involvement may not be identified by these means. The decision of when to do a neck dissection was based on the clinical judgment of the participating surgeons. All were faculty members of major cancer centers, and it is expected that a similar rationale was used when a decision was made to dissect a cN0 neck (need for exposure of vessels for free tissue transfer; greater than 20% likelihood of occult nodal involvement on the basis of primary tumor site, stage, or thickness; history of previous treatment [neck dissection or radiotherapy]; availability of posttreatment adjuvant treatment options, etc). No neck dissection was performed in 82 cN0 cases. It is possible that bias was introduced in our results in that the dissected nodes would be selected for cases expected to have a greater likelihood of tumor involvement, thereby inflating the apparent value of neck dissection to upgrade nodal category by identifying occult disease.

Nevertheless, on the basis of the availability of more detailed information derived from surgical nodal sampling, a stronger association of pN than cN category with clinical outcome was demonstrated by the results. Comparison of cases in which nodal status changed from cN0 to pN+ (presence of occult nodal disease) with cases judged cN0 that were truly pN0 showed a significantly poorer outcome associated with pathological upgrade ($P=.02$; Table 8). Similarly, cases judged suggestive of metastatic disease on the basis of nodal size (cN+), yet

found not to harbor tumor (pN0), had better survival than those judged rightly to have metastases by both methods ($P=.01$; Table 8). As a result, Kaplan-Meier survival curves with log-rank statistic show the pN category for the entire group to be more accurately associated with overall survival (Tables 5, 6, and 7; Figure 1 and Figure 2C vs D). Patel and Lydiatt,¹⁰ in their review of staging in head and neck cancer, stated that “the adverse effect of subclinical nodal disease has not been settled.”¹⁰ Our results indicate that accurate assessment of cN0 disease does contribute substantially to accurate prognostication.

The results of analysis of stage related to cause of death demonstrate that tumor stage affects death owing to disease, but not death from other causes or unknown cause (Figures 3 and 4). It may be implied that successful treatment for more extensive disease does not lead to earlier death owing to morbidity of treatment, nor is more extensive disease associated with higher comorbidity that would lead to earlier death. The independence of stage and comorbidity is also indicated in our multivariable model in which performance status at study entry (a rough measurement of comorbidity) and tumor stage are independent prognostic factors.

We hypothesized that inclusion of cases of recurrent cancer for which clinical stage remains that of index disease at initial presentation would introduce substantial inaccuracy. As such, it was expected that only newly diagnosed cases would have accurate clinical stage. However, our results indicate no substantial decrement in association of clinical stage with outcome when all cases—

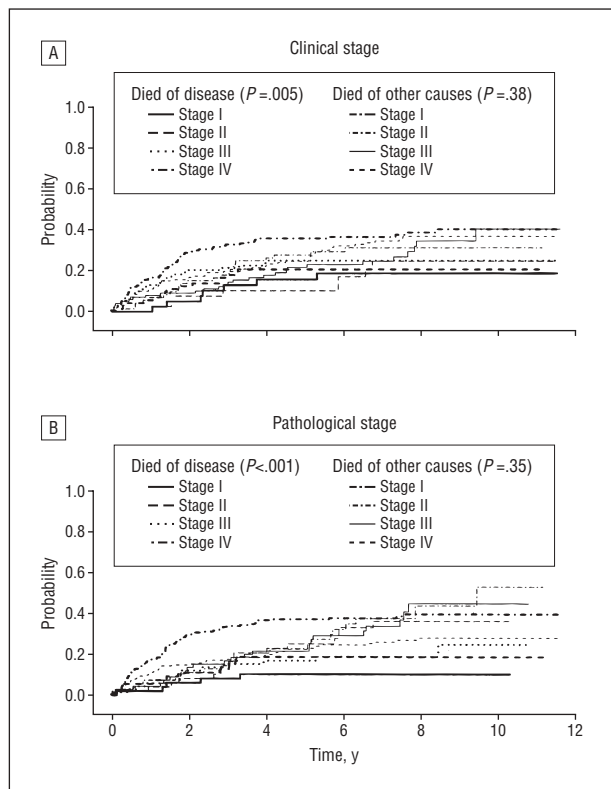


Figure 3. Competing cause of death for all patients. The cumulative incidence of death from disease and other causes over time is given by overall clinical (A) and pathological (B) stage. When deaths are classified as owing to disease or owing to other causes (unrelated, n=73; unknown, n=70), advancing clinical and pathological stages demonstrate an association with increased probability of death owing to disease, with pathological stage having a stronger association than clinical stage. As expected, neither clinical nor pathological stage was associated with different rates of death owing to other or unknown causes. (Note that the lines for death owing to other causes fall between the early and late categories for death owing to disease.)

recurrent and newly diagnosed—were considered together. This could be because some recurrent cases were initially small, localized lesions amenable to simple salvage excision while others were extensive advancements from a small index lesion.

The results shown in Table 1 demonstrate the association with survival of several key clinical factors not included in the staging system. Not surprisingly, patients who experience significant weight loss before diagnosis and treatment and those with poorer performance status have a shorter median survival. It is interesting that smoking status did not correlate significantly with survival, whereas alcohol consumption had a strong prognostic value. The primary site of cases in our cohort was strongly associated with survival in ways that we have noted previously; particularly, cases of oropharyngeal cancer have a good prognosis, whereas those of the hypopharynx, which are most often associated with heavy smoking and drinking, have the poorest survival.

There are few reports of large cohorts of patients with HNSCC for which both clinical and pathological stage are available for comparison. De Waal and colleagues¹¹ published retrospective data of neck nodal stage for 186 patients. Similar to our findings, the overall sensitivity of clinical staging was 80.1% compared with the stan-

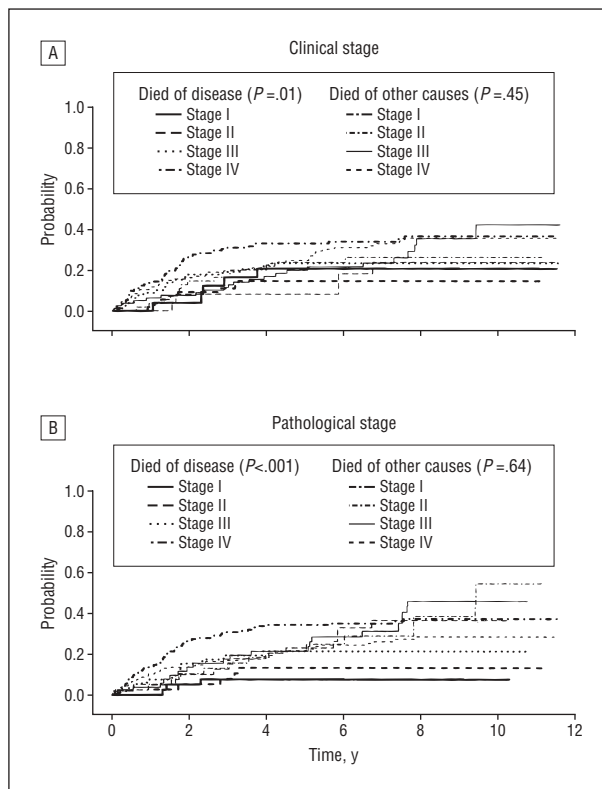


Figure 4. Competing cause of death for previously untreated patients. The cumulative incidence of death from disease and other causes over time is given by overall clinical (A) and pathological (B) stage.

Table 7. Multivariable Model of Factors Prognostic for Overall Survival

Factor	Hazard Ratio (95% CI)	P Value
Age, y		
55-64	1.57 (1.12-2.26)	.008
≥65	2.04 (1.52-2.67)	<.001
Performance status		
1	1.57 (1.18-2.07)	.002
2-4	2.33 (1.37-3.99)	.002
Missing	0.87 (0.40-1.87)	.72
Clinical T category		
2	1.32 (0.92-1.91)	.14
3	1.77 (1.20-2.60)	.004
4	2.10 (1.40-3.17)	<.001
Missing	2.40 (0.74-7.76)	.15
Pathological N category		
1	1.79 (1.22-2.63)	.003
2-3	2.39 (1.78-3.20)	<.001
Missing	1.58 (1.08-2.31)	.02

Abbreviation: CI, confidence interval.

dard criterion of pathological stage, whereas specificity was only 52.2%. There were occult nodal metastases in 32% of elective neck dissection specimens, making the specificity of intraoperative staging on the N0 neck 33.3% with sensitivity of 72.4%.

Our study is limited by the retrospective nature of determination of pathological staging. Cases were accrued from 18 institutions, each with its own approach to reporting surgical specimen results. Uniform processing and

Table 8. Strongly Discordant Cases

	No.	Deaths, No. ^a	Median (95% CI)	Log-Rank P Value
cT1 to pT4				.08
Nondiscordant	490	284	4.5 (3.5-5.4)	
Discordant	8	2	NYR (3.7-∞)	
cT4 to pT1-2				.44
Nondiscordant	480	274	4.5 (3.5-5.6)	
Discordant	18	12	5.1 (0.5-∞)	
cN0 to pN+				.02
Nondiscordant	145	67	7.8 (5.7-∞)	
Discordant	63	39	3.7 (1.9-6.6)	
cN0 to pN2				.65
Nondiscordant	466	268	4.5 (3.6-5.6)	
Discordant	33	20	3.7 (1.4-∞)	
cN+ to pN0				.01
Nondiscordant	177	120	2.1 (1.6-3.7)	
Discordant	26	11	5.6 (4.5-∞)	
cN1-2 to pN0				.16
Nondiscordant	475	278	4.2 (3.5-5.4)	
Discordant	24	10	5.6 (4.1-∞)	
cT4, cT any				.60
pT<4	38	27	2.4 (1.0-6.3)	
pT4	102	67	2.9 (1.7-4.0)	
cT<4				.34
pT<4	360	195	5.3 (4.1-6.6)	
pT4	43	27	3.8 (2.6-7.9)	

Abbreviations: CI, confidence interval; NYR, not yet reached.

^aThe number of patients who were evaluable for various stage categories differed, and the number of deaths among patients excluded varied among categories.

reporting protocols would have provided more reliable results and might have produced a clearer advantage for pathological T category. Another limitation is that our cases accrued before the regular availability and use of positron emission tomographic scanning. The improved sensitivity of positron emission tomography for occult nodal disease might improve the accuracy of clinical staging; however, specificity may be reduced.

CONCLUSIONS

There is often a substantial difference between pathological and clinical categorization, resulting in changes to overall staging 45% of the time (concordance is only 55%). However, the median survival for clinical and pathological staging is very similar, with nearly completely overlapping confidence intervals. Thus, one can reliably use either pathological or clinical stage to predict cancer-specific survival. There is not a major difference in the 2 staging methods as tools for prognostication. This is likely because of flaws in both systems, the plurality and complexity of factors involved in overall survival of a patient, and mutual cancellations of small changes in stage within a large cohort.

The pN category is a better individual predictor of overall survival than is pT, cT, or cN. The histologic assessment of lymph nodes provides accuracy in disease assessment that enhances outcome prediction. Nodal status as judged through the findings at neck dissection yields

demonstrably more accurate predictive information for HNSCC, particularly when major corrections are made from N0 to N+ and N+ to N0.

Submitted for Publication: December 1, 2008; final revision received April 22, 2009; accepted April 26, 2009.

Correspondence: Wayne M. Koch, MD, Department of Otolaryngology–Head and Neck Surgery, The Johns Hopkins University School of Medicine, 601 N Caroline St, Baltimore, MD 21287 (wkoch@jhmi.edu).

Author Contributions: Drs Koch, Ridge, and Forastiere had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Koch and Forastiere. *Acquisition of data:* Koch, Ridge, and Manola. *Analysis and interpretation of data:* Koch, Ridge, and Manola. *Drafting of the manuscript:* Koch, Ridge, and Manola. *Critical revision of the manuscript for important intellectual content:* Ridge, Forastiere, and Manola. *Statistical analysis:* Manola. *Obtained funding:* Koch. *Administrative, technical, and material support:* Ridge and Forastiere. *Study supervision:* Koch.

Financial Disclosure: None reported.

Funding/Support: This study was coordinated by the Eastern Cooperative Oncology Group (Robert L. Comis, MD, Chair) and supported in part by Public Health Service grants CA23318, CA66636, CA21115, CA16116, and CA27525; grant R01 DE013152 from the National Institute of Dental and Craniofacial Research; and grants from the National Cancer Institute and the Department of Health and Human Services.

Disclaimer: The contents of this report are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute.

Previous Presentation: This study was presented at the Seventh International Conference on Head and Neck Cancer; July 22, 2008; San Francisco, California.

REFERENCES

- Poeta ML, Manola J, Goldwasser MA, et al. TP53 mutations and survival in squamous-cell carcinoma of the head and neck. *N Engl J Med.* 2007;357(25):2552-2561.
- American Joint Committee on Cancer. *AJCC Cancer Staging Manual.* 6th ed. New York, NY: Springer-Verlag; 2002.
- Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc.* 1958;53:457-481.
- Peto R, Peto J. Asymptotically efficient rank invariant test procedures. *J R Stat Soc Ser A.* 1972;135(2):185-206.
- Cox DR. *Analysis of Binary Data.* London, England: Methuen & Co, Ltd; 1970.
- Mehta CR, Patel NR, Tsiatis AA. Exact significance testing to establish treatment equivalence with ordered categorical data. *Biometrics.* 1984;40(3):819-825.
- Cox DR. Regression models and life tables (with discussion). *J R Stat Soc Ser B.* 1972;34(2):187-220.
- Gray RJ. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat.* 1988;16(3):1141-1154.
- Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr.* 1974;19(6):716-723.
- Patel SG, Lydiatt WM. Staging of head and neck cancers: is it time to change the balance between the ideal and the practical? *J Surg Oncol.* 2008;97(8):653-657.
- De Waal PJ, Fagan JJ, Isaacs S. Pre- and intra-operative staging of the neck in a developing world practice. *J Laryngol Otol.* 2003;117(12):976-978.