PREDICTION OF SURVIVAL IN PATIENTS WITH HEAD AND NECK CANCER

Robert Jan Baatenburg de Jong, PhD[1], Jo Hermans, PhD[2], Jan Molenaar[3], Jeroen Johannes Briaire[1], Saska le Cessie, PhD[2]

[1] Department of Otolaryngology and Head and Neck Surgery, Leiden University Medical Center, P. O. Box 9600, 2300RC Leiden, Holland. E-mail: r.j.baatenburg_de_jong@lumc.nl
[2] Department of Medical Statistics, Leiden University Medical Center, P. O. Box 9600, 2300RC Leiden, Holland
[3] Datamanager ONCDOC (Oncology Documentation), Leiden University Medical Center, P. O. Box 9600, 2300RC Leiden, Holland

Accepted 22 January 2001

Abstract: *Background.* In patients with head and neck squamous cell carcinoma (HNSCC) the estimated prognosis is usually based on the TNM classification. The relative weight of the three contributing parameters is often not completely clear. Moreover, the impact of other important clinical variables such as age, gender, prior malignancies, etc is very difficult to substantiate in daily clinical practice. The Cox-regression model allows us to estimate the effect of different variables simultaneously. The purpose of this study was to design a model for application in new HNSCC patients. In our historical data-base of patients with HNSCC, patient, treatment, and follow-up data are stored by trained oncological data managers. With these hospital-based data, we developed a statistical model for risk assessment and prediction of overall survival. This model serves in clinical decision making and appropriate counseling of patients with HNSCC.

Patients and Methods. All patients with HNSCC of the oral cavity, the pharynx, and the larynx diagnosed in our hospital between 1981 and 1998 were included. In these 1396 patients, the prognostic value of site of the primary tumor, age at diagnosis, gender, T-, N-, and M-stage, and prior malignancies were studied univariately by Kaplan-Meier curves and the log-rank test. The Cox-regression model was used to investigate the effect of these variables simultaneously on overall survival and to develop a prediction model for individual patients.

Results. In the univariate analyses, all variables except gender contributed significantly to overall survival. Their contribution remained significant in the multivariate Cox model. Based on the relative risks and the baseline survival curve, the expected survival for a new HNSCC patient can be calculated.

Conclusions. It is possible to predict survival probabilities in a new patient with HNSCC based on historical results from a dataset analyzed with the Cox-regression model. The model is supplied with hospital-based data. Our model can be extended by other prognostic factors such as co-morbidity, histological data, molecular biology markers, etc. The results of the Cox-regression may be used in patient counseling, clinical decision making, and quality maintenance. *Head Neck* **23:** 718–724, 2001.

Keywords: prognostic model; survival; head and neck squamous cell carcinoma; univariate analysis; multivariate analysis; Cox-regression; TNM-classification

Correspondence to: R. J. Baatenburg de Jong

The TNM grading system has been the mainstay of cancer outcome prediction in patients with head and neck squamous cell carcinoma (HNSCC) for many years. The TNM consists of (1) the size of the primary tumor (Tis, T1, T2, T3, T4),

Table 1. Baseline demographic and tumor data of the study population.

(Sub)sites	Number of Patients
Lip	139
Oral cavity	292
Oropharynx	153
Nasopharynx	41
Hypopharynx	141
Esophagus	218
Glottic carcinoma	480
Supraglottic carcinoma	198
Total number of patients 1981–1998	1662
Excluding Tis and esophagus	
Lip	138
Oral cavity	286
Oropharynx	152
Nasopharynx	41
Hypopharynx	141
Glottic carcinoma	442
Supraglottic carcinoma	196
Total number of patients in this study	1396
Gender	
Male	1105
Female	291
Total	1396
Age categories	
<50 yr.	188
50–59	371
60–69	436
≥70	401
Total	1396
T-stage	
T1	516
T2	369
T3	208
T4	279
Tx	24
Total	1396
N-stage	
N0	985
N1	148
N2	180
N3	82
Nx	1
Total	1396
Distant metastasis	
M0	1378
M1	17
Mx	1
Total	1396
Treatment	
Irradiation only	802
Surgery only	223
Surgery and postop irradiation	251
Otherwise	120
Total	1396
Year of diagnosis	
1981–85	308
1986–90	365
1991–95	418
1996–98	305
Total	1396
Prior malignancies	
None	1257
Single prior malignancy	111
Multiple prior malignancies	28
Total	1396

(2) description of regional (N0, N1, N2a, N2b, N2c, N3), and (3) distant metastasis (M0, M1). Each combination of these three variables can be seen as a bin into which patients with these characteristics are placed. This is called the TNM-bin model[1] and consists of 60 bins (5X6X2). One of the characteristics of a bin model is that the number of bins increases rapidly with the number of variables. For example, when we add the variable "primary tumor site" (lip, oral cavity, nasopharynx, oropharynx, hypopharynx, glottic larynx, supraglottic larynx, and esophagus), the result is 480 bins. In fact we should also add relevant parameters like age, gender, race, and histological grade. This would result in thousands of bins with very few patients in most of the bins even in large cohorts.

The clustering of various subsets of TNM categories into four stages was undertaken in an attempt to stratify according to prognosis and to simplify communication. This clustering will of course result in a loss of accuracy because, for example, a patient with a T4N0M0 carcinoma may biologically be very different from a patient with a T1N2M0 carcinoma, whereas both tumors are stage IV diseases. Furthermore, the stage groupings were created based on presumed prognosis: no prospective, multivariate analysis was performed to create the four stage groupings from the various combinations of T, N, and M.[2] In addition, the stage grouping does not take other prognostic parameters, such as age, gender, and prior malignancies, into account.

It is clear that, although the TNM classification harbors very important clinical information, it is not as useful in daily clinical practice as we want it to be. Most importantly, it is not very useful for prediction of outcome in an individual cancer patient at initial presentation.

Multivariable techniques such as Cox regression have provided methods to predict survival time from diagnosis until an endpoint (usually death) for decades. However, complexity and unfamiliarity with these methods at the level of physicians and lack of sufficient clinically relevant data prevented use in daily otolaryngeal practice until now. The SOKAL score,[3] which is used in

Table 2. Multivariate analysis (FU limited to 120 months).

Parameters	Regression coefficient B	*p* value	Relative risk (exp (B))	95% Confidence interval for relative risk	
(Sub)site		.000			
Nasopharynx	.000		1.000		
Lip	.008		1.008	.545	1.865
Oral cavity	.695		2.003	1.169	3.432
Oropharynx	.429		1.535	.884	2.667
Hypopharynx	.625		1.868	1.079	3.236
Glottic carcinoma	.119		1.127	.647	1.961
Supraglottic ca	.250		1.284	.744	2.217
Gender		.67			
Female	.000		1.000		
Male	.040		1.041	.865	1.252
Age		.000			
<50 yr.	.000		1.000		
50–59	.312		1.366	1.024	1.824
60–69	.594		1.812	1.370	2.396
≥70 yr.	.997		2.710	2.057	3.568
T-stage		.000			
T1	.000		1.000		
T2	.391		1.478	1.189	1.839
T3	.657		1.930	1.505	2.475
T4	.860		2.363	1.853	3.014
Tx	.804		2.236	1.316	3.797
N-Stage		.000			
N0	.000		1.000		
N1	.401		1.493	1.178	1.894
N2	.725		2.065	1.639	2.601
N3	1.044		2.841	2.143	3.766
Distant metastasis		.000			
M0	.000		1.000		
M1	1.821		6.178	3.579	10.662
Prior malignancy[7]		.000			
No	.000		1.000		
Yes	.640		1.897	1.515	2.375

haemato-oncology, is an example of an existing prognostic score. However, this system uses risk categories as do most other prognostic systems. Although superior to, eg, the TNM staging (because they are usually based on survival analysis), such systems lose information due to simplification as well.

The purpose of this study was to design a model in which the prognostic value of age, gender, site of the primary tumor, and T-, N-, and M-stage and prior malignancies is integrated to predict survival probabilities. This would allow for a more precise and individual prediction of outcome in HNSCC patients. Therefore, this model may help in clinical decision making and appropriate counseling of patients with HNSCC.

PATIENTS AND METHODS

Between January 1981 and December 1998, 1662 patients with primary HNSCC were diagnosed in our hospital (Table 1). Patients with esophageal cancer ($n = 218$) were excluded because (1) the number of patients with incomplete TNM staging was relatively large and (2) because prognosis is extremely poor. Patients with carcinoma in situ ($n = 51$; including 3 patients with carcinoma in situ of the esophagus) were excluded as well because the prognosis of these patients is very good and because these patients did not contribute to subclasses other than N0 and M0. The study sample contained 1396 patients with SCC of the lip, oral cavity, pharynx, and larynx.

Study Design. Data were collected from our hospital-based cancer registry system (ONCDOC), which was established in 1981. In this registry system, patient, treatment, and follow-up data of each cancer patient in our hospital are stored by trained oncological data-managers. These data are retrieved from the patients file and the hos-

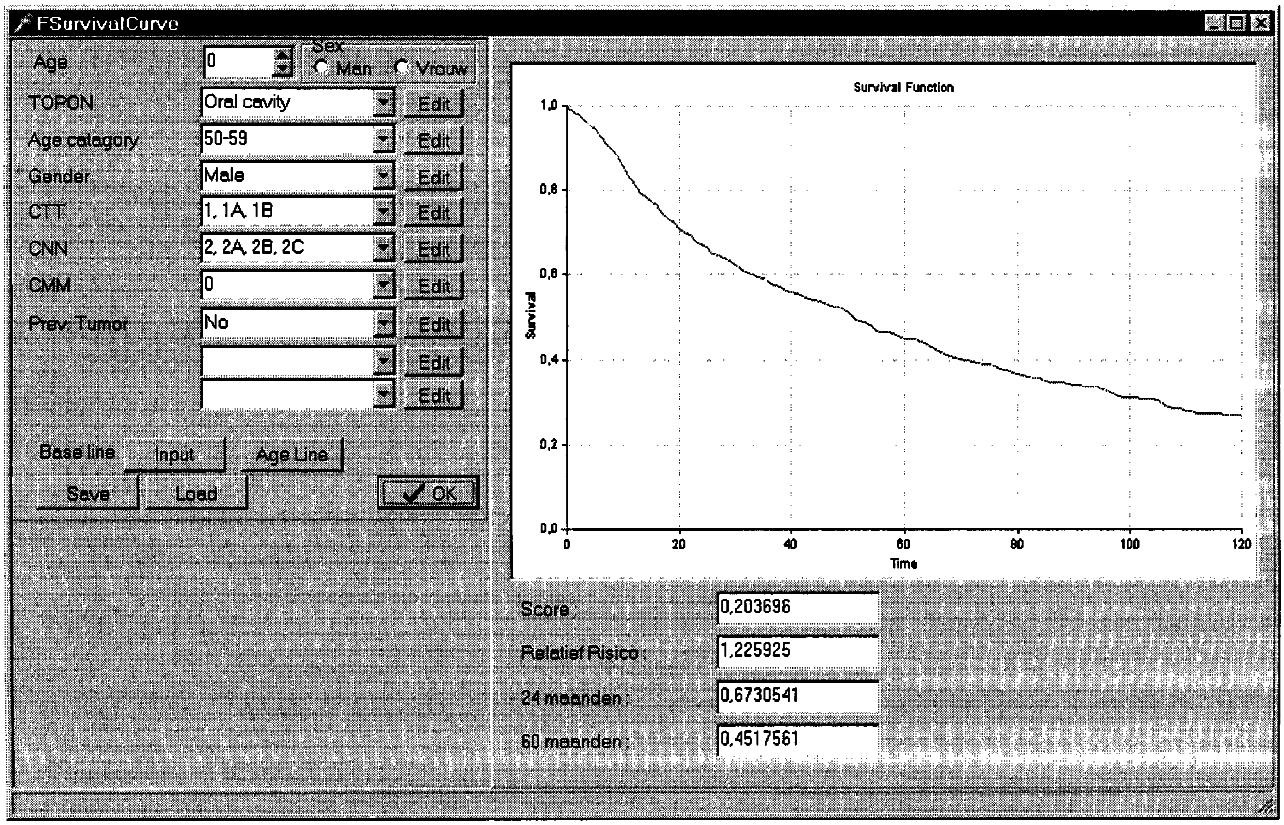

FIGURE 1. Expected survival for a 55-year-old male patient with a T1N2M0 carcinoma of the oral cavity (stage IV) and without a history of prior malignancies. The expected 2- and 5-year overall survival rates are 67% and 45%, respectively.

pital-based data-system[1]. ONCDOC also performs an independent and active follow-up. Primarily, the patients file is used for this purpose. When patients are lost to follow up, ONCDOC will contact the family's doctor and/or the Dutch Registry of Births, Deaths, and Marriages. In this way follow-up is as complete as possible.

One of the other tasks of ONCDOC is quality control: the TNM stage, which is applied to a particular patient, is checked retrospectively by ONCDOC. When discrepancies exist, ONCDOC and the physician will discuss the TNM stage until agreement is reached. Patients were staged according the UICC manual. In 1981 the second edition was used; from 1982–1988 the third edition was used; from 1989–1992 the fourth edition was used; from 1993–1998 the revised fourth edition was used.

Choice of Factors. We included only simple and basic variables (age, gender, site of the primary tumor, T-, N-, and M-stage, and prior malignancies) available for all patients and available before treatment. Prior malignancies were defined as all preceding malignant tumors except for basal cell and squamous cell carcinoma of the skin.

Based on the therapeutic nil hypothesis,[4–8] treatment was not considered as a prognostic factor in this study (see Discussion).

Statistical Analysis. The only endpoint in this study was overall survival. Follow-up of patients was limited to 10 years after diagnosis because the number of patients with longer follow-up was small and these low numbers may theoretically cause unjust differences in survival. The prognostic value of site of the primary tumor, age at diagnosis, gender, prior malignancies, and T-, N-, and M-stage on survival were studied univariately by Kaplan-Meier curves and the log-rank test. The Cox-regression model was used to investigate the effect of these variables simultaneously on overall survival and to develop a prediction model for individual patients.

[1] ZIS by HISCOM, The Netherlands

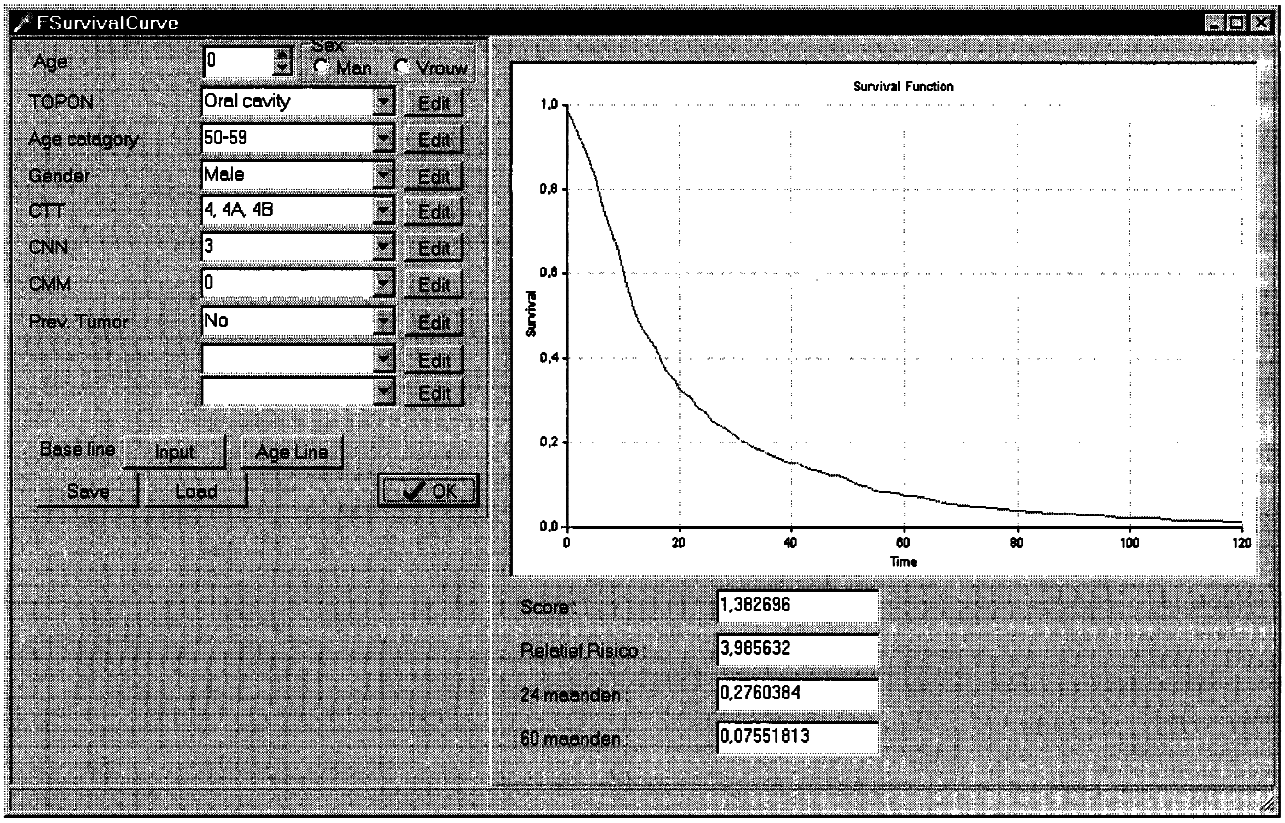

FIGURE 2. Expected survival for a 55-year-old male patient with a T4N3M0 carcinoma of the oral cavity (stage IV) without a history of prior malignancies. The expected 2- and 5-year survival rates are 27% and 7%, respectively.

Validation. The Cox model was validated in two ways: (1) for each patient a prognostic index is calculated by adding the regression coefficients (Table 2) of the risk factors. A high prognostic index indicates a poor prognosis. Based on this prognostic index the patients were divided into five groups (varying from patients with a favorable prognosis to patients with a poor prognosis), and in each of the five groups the average expected survival curve was compared with the observed Kaplan-Meier curve. (2) To judge the stability of the prognostic index derived from the Cox regression, we applied a split-sample technique.[9] For this purpose the total cohort was randomly split into two groups: two-thirds of the patients was used to develop the prognostic index; the other one-third was used to check the predictive value of the index. The results in the latter group confirmed the predictions based on the former group (data not shown). Hence, only data for the entire data set are presented.

Software. Data were analyzed in SPSS® (version 10.0). Dedicated software was designed in Borland Delphi® to allow for easy input of a new patient for whom a prediction is needed. The calculated survival curve for that individual is then displayed onscreen (Figures 1 and 2).

RESULTS

Follow-Up. The median follow-up time was 5.2 years (range 0–120 months). Only 6% of the patients were lost to follow-up.

Univariate Analysis. In the univariate analyses, all variables except gender contributed significantly to overall survival (Table 3).

Multivariate Analysis. Having established the univariate relationship between age, gender, site of the primary tumor, T-, N-, and M-stage, prior malignancies, and survival, the next step was to examine how these variables perform in a multiple Cox regression analysis. Patients with Nx ($n = 1$) or Mx ($n = 1$) were excluded from this analysis. All variables (except gender) remained significant in the multivariate Cox model. Whereas

Table 3. Univariate analysis for all variables.

Variable	Subcategory	5-Year survival probability	*p* Value log rank test
(Sub)site	Lip	.71	.0000
	Oral cavity	.35	
	Oropharynx	.36	
	Nasopharynx	.53	
	Hypopharynx	.31	
	Larynx-glottic	.69	
	Larynx-supraglottic	.51	
Gender	Male	.54	.06
	Female	.45	
Age	<50 years	.65	.0000
	50–59	.58	
	60–69	.52	
	≥70	.39	
T-stage	1	.72	.0000
	2	.52	
	3	.36	
	4	.26	
	X	.46	
N-stage	0	.63	.0000
	1	.32	
	2	.26	
	3	.11	
M-stage	0	.53	.0000
	1	.00	
Prior malignancy	Yes	.33	.0000
	No	.54	

the subsite "lip" was characterized by the best prognosis in the univariate analysis, in the multivariate analysis the "nasopharynx" appeared to be slightly more favorable. The regression coefficients (B), and estimated relative risks (exp (B)) are given in Table 2. RR is the relative risk of dying compared with the reference category in which the risk is set to 1.0. Values of RR above 1.0 indicate higher risks than for the reference category. Confidence intervals (95%) for the relative risks of dying are also reported. For each variable the category with the best prognosis is chosen as reference.

Based on these relative risks and the baseline survival function, which is calculated in the Cox model, the expected survival for a new HNSCC patient can be calculated. For example, a patient with a very good prognosis (T1N0M0 nasopharynx, female less than 50 years old) the Cox model yields an estimated 2-year survival probability of .94 and a 5-year survival probability of .88, whereas for a male patient, for example, older than 70 years with T4N3M1 glottic cancer the estimated 2-year survival is virtually 0%.

DISCUSSION

Accurate and individualized estimation of survival in patients with HNSCC would undoubtedly lead to an improvement in therapeutic and care strategies, minimizing risks of under-treatment and over-treatment. We therefore developed a model for predicting outcome on the basis of the patients' clinical characteristics before treatment. This model was meant to assist doctors to make decisions for the newly diagnosed patient with HNSCC.

The current TNM system is not very useful in this respect. The limitations of the TNM system were well described by Byron J Bailey, Chairman of the Committee to study the TNM classification of the Laryngeal Cancer Association[10]: "Physicians are focused on optimal treatment while patients are interested in their prognosis, and the TNM is not designed to provide answers to either sets of questions. At the present time, the TNM system is neither a roadmap for patient management nor is it a crystal ball with the answers sought by patients."

The TNM system combines information concerning the extent of anatomic tumor spread into stages that have different estimates of survival. Although these stages stratify patients, their predictive power in any given patient is variable.[2,11,12] We feel that, for the sake of simplification, stage grouping leads to loss of important information, and we agree with Piccirillo[2] that stage groupings should be based on multivariate studies of the relative impact of the T, N, and M categories. However, with modern statistical techniques and software, simplification is no longer necessary and stage grouping may be abandoned: Cox regression analysis may be used for calculation of relative risks of dying for each variable in the TNM classification adjusting for other variables. In our model, predictions are made for each individual variable and not for each bin separately. The RR of each variable is then multiplied and combined with the estimated baseline survival curve to produce an estimate of survival for a new cancer patient.

Because this analysis would be rather cumbersome when performed in every individual new cancer patient, we designed dedicated software for a PC to allow for easy input of data of new cancer patients. With pull-down menus for each variable, an estimate of survival is obtained within 15 seconds. Internal validation by comparing estimates with observed Kaplan-Meiers and

by the split-sample technique was very satisfactory. In the near future our predictions will be validated in a different population.

Our data as collected by ONCDOC were specifically intended for research according to a protocol and therefore characterized by consistency, accuracy, availability, and completeness. Hoever, this is an observational study with a potential for bias and systematic errors in treatment assignment. The goal of this study was however not to compare treatment options but to integrate diagnostic data to improve our prognostic accuracy. In any index for prognostic stratification, choices between treatment options, which are not under control of the investigator, will influence outcome. This systematic error cannot be eliminated and is the reason why this study is based on the therapeutic nil hypothesis.

The data presented here show that the outcome for a given patient can be presented by a simple score that makes use of known risk factors. The risk of death increased cumulatively in proportion to the number and weight of risk factors present.

In 1993 Burke and Henson[1] formulated 12 criteria for selecting a prognostic system. Our model fulfills most of these criteria. However, we did not study relapse predictions even though our model is suitable for that purpose also. In addition, this study was based on the nil hypothesis, and therefore our model does not allow for grouping by treatment. Evaluation of different treatment strategies can only be done by randomized trials. Results of such trials may only be included in a prognostic index under strict conditions. The only assumption we made is that the RR of the variables studied are multiplicative. Interactions between variables may play a role and will be subject of further research. With these possible shortcomings in mind, we feel that our model may serve as the enhanced prognostic system Burke and Henson[1] proposed.

Finally, it is important to realize that the clinical variables incorporated in this model are simple variables that do not reflect the biological heterogeneity of HNSCC. Our goal is to identify other prognostic markers and to study their independence in our model. When such a variable remains significant in the multivariate Cox analysis it may be added to the model.

CONCLUSIONS

It is possible to predict survival probabilities in a new patient with HNSCC based on historical results from a data-set analyzed with the Cox-regression model. The results of the Cox-regression may be used in patient counseling, clinical decision making, and quality maintenance.

REFERENCES

1. Burke HB, Henson DE. Criteria for prognostic factors and for an enhanced prognostic system. Cancer 1993;72(10): 3131–3135.
2. Piccarillo JF. Purposes, problems and proposals for progress in cancer staging. Arch Otolaryngol Head Neck Surg 1995;121:145–149.
3. The Italian Co-Operative CML Study Group. Prognostic discrimination in "good-risk" chronic granulocytic leukemia. Blood 1984;63:789–799.
4. Feinstein AR. Clinical biostatistics XIV. Clin Pharmacol Ther 1972;13:285–297.
5. Feinstein AR. Clinical biostatistics XV. Clin Pharmacol Ther 1972;13:442–457.
6. Feinstein AR. Clinical biostatistics XVI. Clin Pharmacol Ther 1972;13:609–624.
7. Feinstein AR. Clinical biostatistics XVII. Clin Pharmacol Ther 1972;13:755–768.
8. Picarillo JF, Wells CK, Sasaki CT, Feinstein AR. New clinical severity staging system for cancer of the larynx. Ann Otol Rhinol Laryngol 1994;103:83–92.
9. The Chronic Leukemia Working Party of the European Group for Blood and Marrow Transplantation. Lancet 1998;352:1087–1092.
10. Bailey BJ. Beyond the 'new' TNM classification. Arch Otolaryngol Head Neck Surg 1991;117:369–370.
11. Ahmad K, Kim YH, Fayos JV. Reliability of the AJCC's staging system as a prognostic indicator. Acta Oncol 1987;26(3):173–174.
12. Fielding LP, Henson DE. Multiple prognostic factors and outcome analysis in patients with cancer. Cancer 1993;71(7):2426–2429.