Quality assurance in surgical oncology

Peeters, K.C.M.J.

5

Validation of a nomogram for predicting disease-specific survival following a R0 resection for gastric cancer

Koen C.M.J. Peeters, Michael W. Kattan, Henk H. Hartgrink, Elma Klein Kranenbarg, Martin S. Karpeh, Murray F. Brennan and Cornelis J.H. van de Velde

Cancer. 2005 Feb 15;103(4):702-7

ABSTRACT

A statistical model for predicting disease-specific survival in gastric cancer patients, based on a single US institution experience, was tested for validity in a different set of patients treated at different institutions. Four hundred and fifty-nine patients from the Dutch Gastric Cancer trial comparing D1 to D2 lymph node dissection, were analysed. Discrimination ability of nomogram with respect to 5 and 9 year disease-specific survival probabilities was superior to that of the AJCC stage. There was considerable heterogeneity of risk within many of the AJCC stages. Calibration plots suggested that predicted probabilities from the nomogram corresponded closely to actual disease-specific survival. The gastric cancer nomogram performed well when applied to patients treated in a large number of institutions. The nomogram provided predictions that discriminated better than AJCC stage, regardless the extent of lymph node dissection. Patient counselling and adjuvant therapy decision making should benefit from use of the nomogram.

INTRODUCTION

Although the incidence is declining in Western Europe(1), gastric cancer remains the second most common cause of cancer death worldwide.(2) Surgery is the only curative treatment. The influence of extent of gastric and lymph node resection is debated.(3-5) Adjuvant chemo-radiation has been proposed as well and tested in an attempt to improve local control and survival. The US Intergroup study by the Southwest Oncology Group showed a significant overall survival benefit after postoperative chemoradiation (36 versus 27 months median overall survival in the surgery alone-group), which lead to standardisation of this regimen in the United States.(6) The trial was criticized however for the suboptimal surgery employed and the level of unresected nodal disease. Surgical undertreatment, as observed in this trial, clearly undermined survival.(7)

Although treatment delivered determines patient's prognosis to a large extent, other factors such as patient characteristics, age and sex, the stage of disease at presentation, and tumour location and morphology play a substantial role. Current staging modalities, that solely focus on depth of tumour invasion and the presence of nodal disease, do not take these factors into account. Nomograms have been developed to address this problem: they are predictive

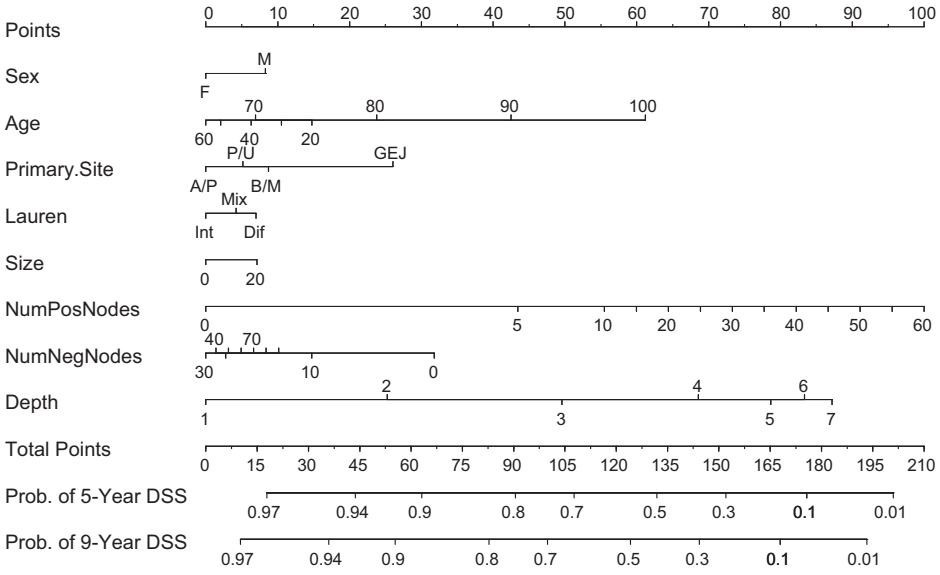

Figure 1. Nomogram for disease-specific survival

Instructions for Physician:
Locate the patient's sex on the **Sex** axis. Draw a line straight upwards to the **Points** axis to determine how many points towards gastric cancer-specific death the patient receives for his or her sex. Repeat this process for the other axes, each time drawing straight upward to the **Points** axis. Sum the points achieved for each predictor and locate this sum on the **Total Points** axis. Draw a line straight down to the disease-specific survival axes to find the patient's probability of surviving gastric cancer assuming he or she does not die of another cause first.

tools for the individual patient based on known prognostic variables including the extent of surgical treatment. Nomograms aid in patient counselling, follow-up scheduling and clinical trial determination and have been developed in soft tissue sarcoma(8), prostate(9-12), renal cell(13), pancreatic(14), and breast cancer.(15) The statistical model developed for gastric cancer (see figure 1) was able to predict the individual patient's probability for disease-specific 5 and 9 year survival after an R0 resection for gastric cancer in a single institution US patient population involving 1039 patients treated from 1985 to 2002.(16)

The purpose of this study was to assess the validity of this prediction tool when applied to patients with a different stage of disease at presentation, differing (surgical) treatment at different institutions. We also compared the discriminating value of the nomogram to the AJCC staging system.

PATIENTS AND METHODS

Patients were enrolled in the Dutch Gastric Cancer trial. This trial was undertaken between August 1989 and July 1993 and randomized gastric cancer patients, coming from 80 Dutch hospitals, between a limited (D1) and an extended (D2) lymph node dissection as recommended by the Japanese Research Society for the Study of Gastric Cancer.(17;18) The results of this trial have been published.(19-21) For the present analysis, patients were considered eligible if they had underwent an R0 resection, i.e. a resection with negative margins without any evidence of tumour spillage (n = 633). In agreement with our previous report, the following prognostic variables were assembled for use in validating the nomogram: age, sex, primary site (distal one third, middle one third, proximal one third, and gastroesophageal junction), Lauren histotype (diffuse, intestinal, mixed), number of positive lymph nodes resected, number of negative lymph nodes resected, and depth of invasion as defined by the standard nomenclature.(22) Patients with suspected vs. definite serosal invasion are distinguished in the nomogram. However, pathologic analysis from the Dutch trial did not distinguish between these depths. For purposes of nomogram validation, we calculated the nomogram prediction assuming a point half way between these two points on the nomogram. Patients with one or more missing values were excluded (Lauren histotype, n = 126; size, n = 19; primary site, n = 41), leaving 459 patients that had values for all nomogram predictor variables, AJCC stage, and follow-up. For each of these patients, the nomogram 5 and 9 year disease-specific survival probabilities were computed and compared with the AJCC stage on the basis of discrimination ability, as measured by the concordance index. Disease-specific survival was estimated using the Kaplan-Meier method.

Nomogram validation comprised two activities. First, discrimination was quantified with the concordance index.(23) Similar to the area under the receiver operating characteristic curve, but appropriate for censored data, the concordance index provides the probability

that, in a randomly selected pair of patients in which one patient dies before the other, the patient who died first had the worse predicted outcome from the nomogram.

Second, calibration was assessed. This was done by grouping patients with respect to their nomogram-predicted probabilities and then comparing the mean of the group with the observed Kaplan-Meier estimate of disease-specific survival. All analyses were performed using S-plus 2000 Professional software (Statistical Sciences, Seattle, WA) with the Design and Hmisc libraries added.(24)

RESULTS

Table 1 depicts the patient and tumor characteristics of the 459 eligible patients with all the information available for the nomogram calculation. With a median follow-up of 10 years, 194 of the 459 patients had died of disease. Disease specific survival by AJCC stage grouping is shown in figure 2, suggesting a reasonable number of patients alive at both 5 and 9 years for nomogram validation. The concordance index for the nomogram was 0.77. Calibration of the nomogram, as shown in figure 3, appeared to be accurate for both the 5- and 9-year predictions.

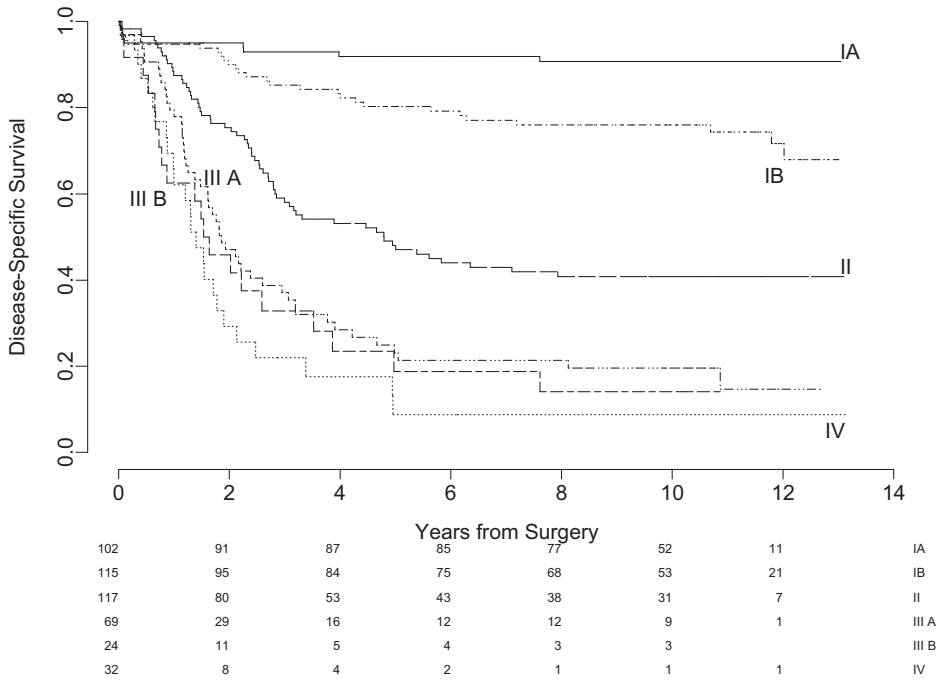

Figure 2. Disease specific survival by AJCC stage grouping

Table 1. Patient and tumor characteristics of all patients with available information on nomogram predictor variables

	N	%
Sex		
Male	270	59
Female	189	41
Primary Site		
A/P	199	43
B/M	191	42
GEJ	69	15
Lauren		
Mixed	17	4
Intestinal	337	73
Diffuse	105	23
Stage		
IA	102	22
IB	115	25
II	117	26
III A	69	15
III B	24	5
IV	32	7
Depth		
Mucosa	81	13
Submucosa	100	16
Propria musclaris	93	15
Subserosa	215	34
Suspected/definite serosal invasion	132	21
Adjacent organ involvement	12	2
Number of Negative Nodes		
Minimum	0	
1st Quartile	13	
Median	21	
Mean	24	
3rd Quartile	32	
Maximum	105	
Number of Positive Nodes		
Minimum	0	
1st Quartile	0	
Median	1	
Mean:	3.5	
3rd Quartile	5	
Maximum	28	
Size (cm)		
Minimum	0	
1st Quartile	3	
Median	4	
Mean	5	
3rd Quartile	6	
Maximum	24	
Age (years)		
Minimum:	31	
1st Quartile:	57	
Median:	66	
Mean:	64	
3rd Quartile:	73	
Maximum:	84	

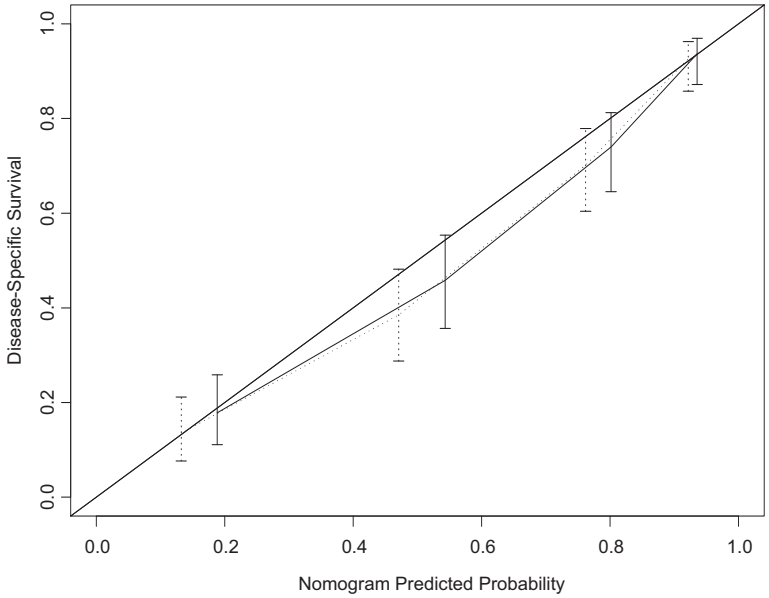

Figure 3. Calibration curves for the nomogram. X-axis is nomogram predicted probability. Patients were grouped by quartiles of predicted risk. Y-axis is actual disease-specific survival as estimated by the Kaplan-Meier method. Solid line is performance of the 5-year prediction; dotted line represents 9-year prediction. Vertical bars represent 95% confidence intervals. For each quartile of both nomogram predictions, the 95% confidence intervals overlap the diagonal "ideal" line, where predicted would exactly match actual disease-specific survival

We compared predictions from the nomogram with those obtained by using the AJCC stage groupings. Individual AJCC stage groups and nomogram predictions were compared for their ability to rank the patients (e.g. concordance index). Nomogram discrimination was superior to that of AJCC stage grouping (concordance index 0.77 vs. 0.75 $P < .001$, Z-test). This difference is difficult to appreciate clinically, and therefore, figure 4 illustrates the discrepancies between the two prediction methods. Within each AJCC stage grouping is a histogram of nomogram-predicted probabilities, illustrating heterogeneity within many of the stages.

DISCUSSION

Patient prognosis is currently estimated on the basis of AJCC staging, and not on other factors like age, sex or morphology that may have an impact on disease-specific survival. Integrating these variables in a nomogram has yielded a model that is a more accurate predictor for disease specific survival than is AJCC stage. This study validates the predictive value of the nomogram, previously tested in a single US institution.(16) The difference in concordance index between the nomogram and the AJCC staging is not great, and may therefore appear

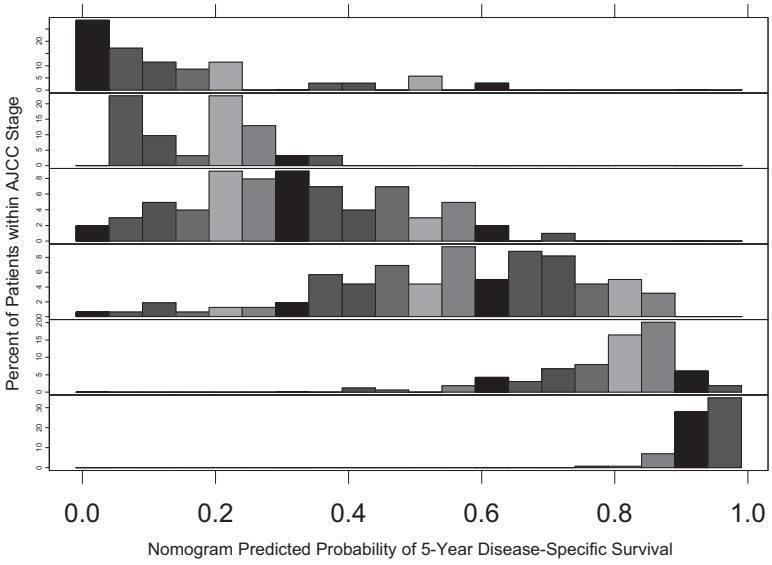

Figure 4. Nomogram predicted probabilities within each of the AJCC stages. Numbers in parentheses for each stage indicate number of patients within that stage. Note the large variation in nomogram predicted probability present within many of the stages

clinically irrelevant. However, figure 4 shows the clinical meaningfullness and benefits of nomogram predictions: patients within different AJCC stages with heterogeneous prognosis are successfully discerned, using the nomogram. Apparently, the present AJCC staging system is unable to identify subsets of patients with homogeneous prognoses. Accurate prediction can aid in individual patient counselling and in follow-up scheduling. It may also play a role in designing future trials, identifying subsets of patients within known AJCC stages that have a different prognosis, and likewise a potential for different response to novel adjuvant treatment regimens. It is important that this model, shown to be valuable in a single institution US patient population, is valid in a multicenter European gastric cancer patient population. The type of gastric cancer management depends largely on where the patient is being treated: many US gastric cancer patients receive postoperative chemoradiation(6), whereas adjuvant treatment is not the norm in Europe. In the current patient population as well as the original group of patients used to develop the nomogram, no adjuvant treatment was given, and the surgical treatment consisted of D1 and D2 dissection in all validation patients. This is more extensive surgery than undertaken in the general US patient population. The American College of Surgeons evaluated surgical treatment of over 18,000 gastric cancer patients between 1982 and 1987 and concluded that dissection of the celiac nodes occurred in only 14% of the cases.(25) Among the 3,804 patients having a curative resection, only 695 (18%) had dissection of the nodes along the celiac axis, hepatic artery, or splenic artery (N2 nodes).(26) Stage of disease differs between the current patient population and the US patients that were analysed in our previous report with less cases of advanced disease in the present

patient population because we included R0 patients only. Despite these major discrepancies between the series, the nomogram predicted accurately, superior to AJCC stage, for disease specific survival in a patient population treated in as many as 80 hospitals, consistent with common surgery in the Netherlands.

Patients in the present analysis were derived from the Dutch Gastric Cancer trial, comparing D1 to D2 dissection. The nomogram predicted well in this series despite the fact that type of dissection was not a variable, per se, in the nomogram. The likely reason for this favourable outcome is that the numbers of positive and negative nodes are predictor variables in the model. Thus far, there is still no overall difference in survival rates between the arms of the Dutch trial.(21) Consequently, considering the type of resection as an input variable for nomogram construction does not seem to have additional value. Defining the extent of lymph node dissection (i.e. D1 or D2) requires intra-operative identification of all 16 lymph node stations as defined by the Japanese Research Society for the Study of Gastric Cancer (JRSGC).(17;18) Identification and subsequent resection of all these separate stations may contribute to improving clinical outcome, even in Western patients considering recent publications that focus on adequate lymph node removal with critical organ resection, thus minimising postoperative morbidity and mortality.(27-29) Notwithstanding the efforts of improving locoregional control through extended nodal dissection, the surgical effort of meticulous dissection is not routinely performed in Western gastric cancer patients, especially not outside the framework of clinical trials. Including the type of resection as a mandatory input variable in the predictive nomogram would therefore make the nomogram less applicable in daily practise. However, the basis of the initial nomogram was an institution where extended lymph node dissection is performed in the majority, but not all, of patients. By requiring only the *numbers* of negative and positive lymph nodes resected for the nomogram computation without specifying their location, we believe that the extent of lymph node dissection is sufficiently addressed.

In conclusion, the gastric cancer nomogram performed well when applied to a validation dataset of patients, coming from a large number of institutions with different stage of disease, treated with a focus on lymph node clearance. The nomogram provided predictions that discriminated better than AJCC stages, regardless of the extent of lymph node dissection, and illustrated the heterogeneity of risk within many stages. With the availability of this external validation, individual patient counselling and tailored adjuvant therapy decision making should be encouraged using the nomogram, freely available in software from www.nomograms.org.

REFERENCE LIST

1. Ekstrom AM, Hansson LE, Signorello LB, Lindgren A, Bergstrom R, Nyren O. Decreasing incidence of both major histologic subtypes of gastric adenocarcinoma--a population-based study in Sweden. *Br.J.Cancer* 2000;83:391-6.
2. Parkin DM, Pisani P, Ferlay J. Estimates of the worldwide incidence of 25 major cancers in 1990. *Int. J.Cancer* 1999;80:827-41.
3. van de Velde CJ, Peeters KC. The gastric cancer treatment controversy. *J.Clin.Oncol.* 2003;21:2234-6.
4. Maruyama K, Okabayashi K, Kinoshita T. Progress in gastric cancer surgery in Japan and its limits of radicality. *World J.Surg.* 1987;11:418-25.
5. Maruyama K, Sasako M, Kinoshita T, Sano T, Katai H, Hada M et al. Should systematic lymph node dissection be recommended for gastric cancer? *Eur.J.Cancer* 1998;34:1480-9.
6. Macdonald JS, Smalley SR, Benedetti J, Hundahl SA, Estes NC, Stemmermann GN et al. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. *N.Engl.J.Med.* 2001;345:725-30.
7. Hundahl SA, Macdonald JS, Benedetti J, Fitzsimmons T. Surgical treatment variation in a prospective, randomized trial of chemoradiotherapy in gastric cancer: the effect of undertreatment. *Ann. Surg.Oncol.* 2002;9:278-86.
8. Kattan MW, Heller G, Brennan MF. A competing-risks nomogram for sarcoma-specific death following local recurrence. *Stat.Med* 2003;22:3515-25.
9. Kattan MW, Zelefsky MJ, Kupelian PA, Cho D, Scardino PT, Fuks Z et al. Pretreatment nomogram that predicts 5-year probability of metastasis following three-dimensional conformal radiation therapy for localized prostate cancer. *J Clin Oncol* 2003;21:4568-71.
10. Kattan MW. Nomograms are superior to staging and risk grouping systems for identifying high-risk patients: preoperative application in prostate cancer. *Curr.Opin.Urol.* 2003;13:111-6.
11. Cagiannos I, Karakiewicz P, Eastham JA, Ohori M, Rabbani F, Gerigk C et al. A preoperative nomogram identifying decreased risk of positive pelvic lymph nodes in patients with prostate cancer. *J Urol.* 2003;170:1798-803.
12. Diblasio CJ, Kattan MW. Use of nomograms to predict the risk of disease recurrence after definitive local therapy for prostate cancer. *Urology* 2003;62 Suppl 1:9-18.
13. Kattan MW, Reuter V, Motzer RJ, Katz J, Russo P. A postoperative prognostic nomogram for renal cell carcinoma. *Journal of Urology* 2001;166:63-7.
14. Brennan MF, Kattan MW, Klimstra D, Conlon K. Prognostic nomogram for patients undergoing resection for adenocarcinoma of the pancreas. *Ann Surg* 2004;240:293-8.
15. Van Zee KJ, Manasseh DM, Bevilacqua JL, Boolbol SK, Fey JV, Tan LK et al. A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. *Ann.Surg.Oncol* 2003;10:1140-51.
16. Kattan MW, Karpeh MS, Mazumdar M, Brennan MF. Postoperative nomogram for disease-specific survival after an R0 resection for gastric carcinoma. *J Clin Oncol* 2003;21:3647-50.
17. Kajitani T. The general rules for the gastric cancer study in surgery and pathology. Part I. Clinical classification. *Jpn.J.Surg.* 1981;11:127-39.
18. The general rules for the gastric cancer study in surgery ad pathology. Part II. Histological classification of gastric cancer. *Jpn.J.Surg.* 1981;11:140-5.
19. Bonenkamp JJ, Songun I, Hermans J, Sasako M, Welvaart K, Plukker JT et al. Randomised comparison of morbidity after D1 and D2 dissection for gastric cancer in 996 Dutch patients. *Lancet* 1995;345:745-8.
20. Bonenkamp JJ, Hermans J, Sasako M, van de Velde CJ. Extended lymph-node dissection for gastric cancer. Dutch Gastric Cancer Group. *N.Engl.J.Med.* 1999;340:908-14.
21. Hartgrink HH, van de Velde CJ, Putter H, Bonenkamp JJ, Klein KE, Songun I et al. Extended Lymph Node Dissection for Gastric Cancer: Who May Benefit? Final Results of the Randomized Dutch Gastric Cancer Group Trial. *J Clin Oncol* 2004.
22. Karpeh MS, Leung D.H.Y., Brennan MF. Cancer of the Stomach. In: DeVita V.T.Jr., Hellman S, Rosenberg S.A. Cancer: Principles and Practise of Oncology. Lippincott Williams & Wilkins, 2001:1092-125.

23. Harrell FE, Jr., Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA* 1982;247:2543-6.
24. Harrell FE, Jr. Regression Modeling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis. New York, NY: Springer-Verlag, 2001.
25. Wanebo HJ, Kennedy BJ, Chmiel J, Steele G, Jr., Winchester D, Osteen R. Cancer of the stomach. A patient care study by the American College of Surgeons. *Ann.Surg.* 1993;218:583-92.
26. Wanebo HJ, Kennedy BJ, Winchester DP, Fremgen A, Stewart AK. Gastric carcinoma: does lymph node dissection alter survival? *J Am Coll.Surg.* 1996;183:616-24.
27. Marubini E, Bozzetti F, Miceli R, Bonfanti G, Gennari L. Lymphadenectomy in gastric cancer: prognostic role and therapeutic implications. *Eur.J.Surg.Oncol.* 2002;28:406-12.
28. Siewert JR, Bottcher K, Stein HJ, Roder JD. Relevant prognostic factors in gastric cancer: ten-year results of the German Gastric Cancer Study. *Ann.Surg.* 1998;228:449-61.
29. Sue-Ling HM, Johnston D, Martin IG, Dixon MF, Lansdown MR, McMahon MJ et al. Gastric cancer: a curable disease in Britain. *BMJ* 1993;307:591-6.