World J. Surg. 24, 1550–1558, 2000
DOI: 10.1007/s002680010276

WORLD
Journal of
SURGERY
© 2000 by the Société
Internationale de Chirurgie

Evaluation of the Maruyama Computer Program Accuracy for Preoperative Estimation of Lymph Node Metastases from Gastric Cancer

Stefano Guadagni, M.D.,[1] Giovanni de Manzoni, M.D.,[2] Marco Catarci, M.D.,[3] Marco Valenti, M.D.,[4] Gianfranco Amicucci, M.D.,[1] Giancarlo De Bernardinis, M.D.,[1] Claudio Cordiano, M.D.,[2] Manlio Carboni, M.D.,[3] Keiichi Maruyama, M.D.[5]

[1]Department of Surgery, University of L'Aquila, via Vetoio (Blocco XI), Coppito, 67010 L'Aquila, Italy
[2]First Department of General Surgery, University of Verona, Piazza Stefani 1, 37126 Verona, Italy
[3]Second Surgical Clinic, University of Rome "La Sapienza," viale del Policlinico 155, 00161 Rome, Italy
[4]Department of Medical Statistics, University of L'Aquila, via Vetoio (Blocco XI), Coppito, 67010 L'Aquila, Italy
[5]Department of Surgical Oncology, National Cancer Center Central Hospital, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104, Japan

Abstract. Controversy still exists about the optimal lymph node (LN) dissection for potentially curable gastric cancer. For rational LN dissection it is important to know the incidence of metastasis at each LN station. For this purpose a computer program was developed using data from 4302 primary gastric cancers treated at the National Cancer Center Hospital in Tokyo between 1969 and 1989. To evaluate the accuracy of the computer program, the differences between the individual reports generated by the computer and the stored data were investigated in 282 Italian patients submitted to curative gastrectomy and D_2 or more extended LN dissections for gastric cancer. Receiver operating characteristic (ROC) analysis was used to assess the sensitivity and specificity of the program for predicting LN metastases in each of the 16 regional LN stations. The computer program showed good predictive ability for LN metastases in most of the 16 LN stations, as the areas under the curve ranged from 0.741 (station 15) to 0.944 (station 8), with a mean of 0.856. A critical cutoff point of 18% of the program's expected percentage was the value maximizing the validity of the prediction. Using an "absolute" cutoff point of 0%, the overall rate of false-negative (FN) predictions in 176 N+ patients was 11.9%; of these, 11 (6.2%) were absolute FNs, in which the program totally failed to estimate LN metastases; the remaining 10 cases (5.7%) were relative FNs because the specific prediction was positive for a different depth of stomach invasion. The low number of D_3/D_4 lymphadenectomies in the historical database may affect the low estimate of metastases to N_3/N_4 nodes generated by the program. Based on these data, the program predicts with good accuracy the extent of LN metastases from gastric cancer, but it is not recommended for directing the surgeon to perform more extensive lymphadenectomy.

In the Western world the 5-year survival rate of patients with resected gastric cancer is around 30% [1, 2]. On the other hand, during 1979–1982 an overall 5-year survival rate of 63.5% has been reported in the Japanese nationwide registry [3] reaching 71% for the latest period [4]. Japanese surgeons claim that, apart from early detection, gastric resection with extended (D_2) lymph node dissection markedly contributes to these results. Nevertheless, the therapeutic value of extended lymphadenectomy during gastric cancer surgery is still controversial [5–7], although Western reports also support its effectiveness for advanced stages (II and IIIa) [8–10].

Lymph node involvement is an important prognostic variable in gastric cancer [11], and most local failures of surgical treatment are believed to be sustained by insufficient nodal clearance. Unfortunately, current staging methods, such as percutaneous ultrasonography, endoluminal ultrasonography (EUS), computed tomography (CT), and dynamic CT, do not offer an accurate preoperative estimation of lymph node involvement [12]. Preoperative prediction of lymph node metastases in an individual patient could be gathered from browsing through historical cases.

A computer program designated to estimate the incidence of lymph node metastases, the expected prognosis, and other information based on the most significant preoperative prognostic factors was developed by Maruyama's group [13]. The current version of the computer program is based on a database of 4302 cases of primary gastric cancer treated at the National Cancer Center Hospital, Tokyo, between 1969 and 1989. In 1992 the accuracy of the Maruyama computer program for preoperative assessment of lymph node metastases was evaluated in a group of 222 German patients with gastric cancer [14]. The study concluded that computer predictions provided perioperative information of therapeutic value.

To confirm the German analysis and to establish if this system is a valuable tool for directing the surgeon to perform more extensive lymphadenectomy, the differences between the individual report and the Maruyama program stored data were investigated in 282 Italian patients, checking the computer estimation against postoperative histologic findings.

Materials and Methods

The basic principle and mathematic concept used in the Maruyama program were described in 1989 [13]. The Italian Institu-

Correspondence to: S. Guadagni, M.D., e-mail: guadagni@rdn.it

Table 1. Regional lymph nodes of the stomach classified by the JRSGC into numbered stations.

Station no.	Description
1	Right paracardial LN
2	Left paracardial LN
3	LN along the lesser curvature
4	LN along the greater curvature
5	Suprapyloric LN
6	Infrapyloric LN
7	LN along the left gastric artery
8	LN along the common hepatic artery
9	LN around the celiac artery
10	LN at the splenic hilum
11	LN along the splenic artery
12	LN in the hepatoduodenal ligament
13	LN on the posterior surface of the pancreatic head
14	LN along the superior mesenteric vein and artery
15	LN along the middle colic vessels
16	LN in the aortic hiatus and around the abdominal aorta

JRSGC: Japanese Research Society for Gastric Cancer [15]; LN: lymph nodes.

tions (L'Aquila, Rome, and Verona surgical departments) received the current version of the program in June 1989 and started a study to evaluate its performances.

Patients included in the study were those who underwent total or subtotal gastrectomy with en bloc D_2 lymphadenectomy, resecting nodes in stations 1 to 12, or D_3/D_4 lymphadenectomy, extended to nodes in stations 13 to 16. Patients who underwent limited lymphadenectomy (D_0/D_1) were excluded because of the incompleteness of the excised lymph node groups. Gastric stump cancer and cardia cancer were also excluded. At September 1994 a total of 282 patients (80.8% of 349 total patients treated during the same period) entered the study.

Immediately after the surgical procedure, the resected specimens were examined according to the rules of the Japanese Research Society for Gastric Cancer (JRSGC) [15] and the TNM-UICC staging system [16]. All lymph nodes were labeled according to the numbering system of the JRSGC, as presented in Table 1. Mean number of nodes examined by the pathologist per patient was 41 (range 19–78). Concerning Japanese cases collected in the historical database, the mean number was of 30 nodes. Tumor location, depth of invasion and nodal status of Italian patients and Japanese patients are listed in Table 2. There were significantly fewer patients without lymph node metastases in the Italian group than in the Japanese group ($p < 0.0001$). The two series were different also with respect to tumor location and depth of invasion.

To evaluate the computer program performances, eight prognostic variables—gender, age, macroscopic type, location, position, diameter, World Health Organization (WHO) histologic type, depth of invasion—were obtained from the clinicopathologic records of each patient. In our institutions, depth of invasion was classified by histopathologic examination (pT1, pT2, pT3, pT4) and not by the classification (mm, sm, pm, ss, s1, s2, s3) used in the program. Therefore pT1 was considered as m or sm, pT2 always as pm, pT3 as s2, and pT4 as s3. Undifferentiated cancer was recorded as poorly differentiated, and tumors involving the entire stomach were assigned to their prevalent location.

Table 2. Comparison of clinicopathologic details of 4302 gastric cancer patients operated on at the National Cancer Center Hospital, Tokyo, and 282 Italian patients.

Parameter	4302 Japanese patients	282 Italian patients	Significance
Location			
Upper third	830 (19.3%)	68 (24.1%)	$\chi^2 = 38.81$;
Middle third	1845 (42.9%)	68 (24.1%)	$p = 0.000001$
Lower third	1627 (37.8%)	146 (51.8%)	
Depth of invasion			
pT1	1746 (40.6%)	67 (23.8%)	$\chi^2 = 50.4$;
pT2	729 (16.9%)	83 (29.4%)	$p = 0.000001$
pT3	1231 (28.6%)	103 (36.5%)	
pT4	596 (13.9%)	29 (10.3%)	
Nodal status			
pN0	2250 (52.3%)	106 (37.6%)	$\chi^2 = 22.3$;
pN1, pN2	2052 (47.7%)	176 (62.4%)	$p = 0.000002$

Program Output

The program gives information on the expected frequency of metastasis in the 16 lymph node stations in three columns (Table 3): one in the center for the degree of parietal involvement, one for a more superficial layer, and one for a deeper layer (e.g., "PM" in the center as the selected depth of invasion, "SM" on the left, and "S1" on the right). The need for other categories is related to a possible uncertainty in preoperative or intraoperative assessment of the depth of stomach involvement.

Statistical Analysis

The chi-square test was used for testing differences between the two case series. For each patient in the Italian series, we attributed the status of lymph node metastasis for each of 16 lymph node stations on the basis of the postoperative histologic findings. The absence/presence of metastasis was entered in a logistic regression model as a dependent variable. Then we assumed, on the basis of the Maruyama program fitting the patient's characteristics, the expected percentage of lymph node metastasis for that patient for each of 16 lymph node stations. The Maruyama program's expected percentage was entered in the logistic regression model as an independent (predictive) variable.

Logistic regression estimates the probability of lymph node metastasis for each patient at the value of the independent variable. Higher values of this estimated probability are assumed to be associated with the event (lymph node metastasis). A receiver operating characteristic (ROC) curve was constructed for each lymph node station by varying the cutoff point that determinates which estimated event probabilities are considered to predict the event. The area under the ROC curve, which has been described as one of the best indexes of detectability [17], represents the probability of concordance between the predicted probability and the "true" (according to the gold standard, i.e., postoperative histology) diagnosis of lymph node metastasis. The area under the ROC curve, as determined by the trapezoidal rule, is calculated by the statistic C:

$$C = [nc + 0.5(t - nc - nd)]/t$$

where t is the total of pairs with different responses, nc is the number of concordant pairs, nd is the number of discordant pairs

Table 3. Analysis printout after processing of the clinicopathologic data in a specific patient and prognosis and LN metastases of patients with same background treated by resection in National Cancer Center Hospital (Tokyo), 1969–1983.

A. Patient L.A.			
Gender and age	F 60 ± 5 years		
Type/depth	B2/PM		
Location	A A (position)		
Maximum diameter	2.0 ± 2.5 cm		
Histologic type	SIG		
B. Other patients			
Depth of invasion	SM	PM	S1
Type of cancer	2A	B2	B2
5-Year survival rate (%)	86.6	60.0	37.0
Greenwood 5% error (%)	18.0	29.0	29.2
No. of cases	18	15	12
Lymph node metastasis[a]			
LN-1[b]	0/18 (0.0%)	0/14 (0.0%)	0/11 (0.0%)
LN-2	0/1 (0.0%)	0/1 (0.0%)	0/0 (0.0%)
LN-3	0/18 (0.0%)	3/15 (20.0%)	3/12 (25.0%)
LN-4	2/18 (11.0%)	0/15 (0.0%)	4/12 (33.0%)
LN-5	0/18 (0.0%)	1/15 (7.0%)	0/12 (0.0%)
LN-6	5/18 (28.0%)	4/15 (27.0%)	1/12 (8.0%)
LN-7	1/18 (6.0%)	1/14 (7.0%)	3/12 (25.0%)
LN-8	1/18 (6.0%)	1/12 (7.0%)	3/12 (25.0%)
LN-9	0/17 (0.0%)	0/12 (0.0%)	1/11 (8.0%)
LN-10	0/1 (0.0%)	0/0 (0.0%)	0/1 (0.0%)
LN-11	0/9 (0.0%)	0/6 (0.0%)	0/3 (0.0%)
LN-12	0/8 (0.0%)	0/8 (0.0%)	2/8 (17.0%)
LN-13	0/5 (0.0%)	0/3 (0.0%)	0/3 (0.0%)
LN-14	0/2 (0.0%)	0/2 (0.0%)	0/0 (0.0%)
LN-15	0/3 (0.0%)	0/0 (0.0%)	0/0 (0.0%)
LN-16	0/4 (0.0%)	0/4 (0.0%)	0/2 (0.0%)
Cause of death			
Living now	16 (89.0%)	10 (67.0%)	5 (2.0%)
Unknown	0 (0.0%)	0 (0.0%)	1 (8.0%)
Peritoneal dissemination	1 (6.0%)	0 (0.0%)	2 (17.0%)
Hepatic metastases	0 (0.0%)	2 (13.0%)	1 (8.0%)
Local recurrence	0 (0.0%)	1 (7.0%)	1 (8.0%)
Distant metastases	0 (0.0%)	0 (0.0%)	0 (0.0%)
Recurrence	0 (0.0%)	0 (0.0%)	0 (0.0%)
Direct death	0 (0.0%)	0 (0.0%)	0 (0.0%)
Other cancer	1 (6.0%)	1 (7.0%)	1 (8.0%)
Other disease	0 (0.0%)	1 (7.0%)	1 (8.0%)
Curability at surgery			
Abs curative	16 (89.0%)	11 (73.0%)	6 (50.0%)
Rel curative	2 (11.0%)	1 (7.0%)	3 (25.0%)
Rel noncurative	0 (0.0%)	0 (0.0%)	2 (17.0%)
Abs noncurative	0 (0.0%)	3 (20.0%)	1 (8.0%)

F: female; B2: type 2 (ulcerated carcinomas with sharply demarcated and raised margins) according to Borrmann's classification; PM: muscularis propria; A: antrum; A (position): anterior wall; SIG: signet-ring cell carcinoma; SM: submucosa; S1: suspected serosal invasion; 2A: superficial elevated type (early gastric cancer); Abs curative: absolute curative resection; Rel: relative.

[a] Metastatic/dissected. Percents in parentheses are the percent metastases among all cases.

[b] JRSGC stations [15].

[18]. All calculations were performed using the SAS/STAT statistical software [19].

Results

Table 4 shows the values for the area under the ROC curves for each of the 16 lymph node stations. It indicates an overall good predictive ability of the Maruyama test for predicting lymph node metastases in specific lymph node stations.

Table 5 reports pairs of sensitivity and specificity values for several critical (cutoff point) Maruyama program expected percentage values obtained by the ROC analysis for each of 16 lymph node stations. The highest validity performances (i.e., the "better" sensitivity–specificity value pair) are highlighted in Table 5. A critical cutoff point of 18% of the Maruyama program expected percentage maximizes the test validity. For example, considering station 1, a Maruyama score of more than 18% enables, when targeting as "positive" (correct preoperative diagnosis), almost 86% of patients with metastatic disease, whereas a Maruyama score lower than 18% enables targeting as "negative" almost 80% of patients whose disease is nonmetastatic.

If we want to diminish the false-negative proportion (i.e., increase sensitivity), we must assume as critical a lower cutoff point of the Maruyama program expected percentage. For example, for station 1 using a cutoff point of 14% a Maruyama score of more than 14% targets as "positive" almost 93% of patients with metastatic disease;

Table 4. Areas under the receiver operating characteristic curve for the Maruyama program's estimated positivity percentage at each of 16 LN stations in 176 patients.

LN station[a]	Frequency of patients with dissected LN station	Area under the ROC curve	SE	Chi-square statistic for model fitting and p
1	167 (94.8%)	0.858	0.040	44.1; $p < 0.001$
2	79 (44.9%)	0.811	0.086	39.6; $p < 0.001$
3	171 (97.1%)	0.872	0.037	40.2; $p < 0.001$
4	168 (95.4%)	0.890	0.033	39.9; $p < 0.001$
5	155 (88.1%)	0.868	0.039	41.4; $p < 0.001$
6	160 (90.9%)	0.849	0.051	38.6; $p < 0.001$
7	164 (93.2%)	0.933	0.030	46.1; $p < 0.001$
8	141 (80.1%)	0.944	0.039	47.7; $p < 0.001$
9	141 (80.1%)	0.879	0.028	45.2; $p < 0.001$
10	123 (69.9%)	0.857	0.055	40.8; $p < 0.001$
11	139 (78.9%)	0.880	0.046	41.9; $p < 0.001$
12	102 (57.9%)	0.864	0.087	43.3; $p < 0.001$
13	27 (15.3%)	0.813	0.104	40.7; $p < 0.001$
14	12 (6.8%)	0.794	0.221	38.5; $p < 0.001$
15	6 (3.4%)	0.741	0.143	14.3; $p < 0.05$
16	28 (15.9%)	0.840	0.118	37.4; $p < 0.001$

SE: standard error; ROC: receiver operating characteristic.
[a]JRSGC [15] numbering system.

on the other hand, such a cutoff point is clearly nonspecific, giving a false-positive proportion of almost 50%.

Table 6 reports the false-positive rate (FPR), false-negative rate (FNR), positive predictive value (PPV), and negative predictive value (NPV) for each lymph node station, given the sensitivity/specificity values of the 18% cutoff reported in Table 5. The FNR averages around 15% per node. The NPVs are in the range of 80%, indicating that prediction of a negative node status has almost an 80% chance of being correct.

Table 7 shows the clinicopathologic features of 21/176 cases (11.9%) in which the computer program gave false-negative (FN) predictions using an "absolute" cutoff point of 0%. Among a total of 282 patients, only the 176 patients with positive nodes were taken into account for statistical calculation of the computer program FN prediction rate.

Table 8 reports a comparative evaluation between postoperative histologic lymph node findings and the preoperative computer predictions for the above-mentioned 21 cases. As reported in the column headed Program validity (Table 8), in 10/176 patients (5.7%) the FN prediction was considered to be relative because the program showed a positive prediction in the other two columns. In 11 cases (6.2%) the program totally failed to predict lymph node metastases (absolute FN). Among these 11 patients, 6 (6/176, 3.4 per cent) would have undergone inadequate treatment if the extent of lymphadenectomy had been based on the computer prediction alone. Particularly in these six cases the program failed to predict lymph node metastases to nodes at the splenic hilus (station 10, N_3) in one case, along the splenic artery (station 11, N_3) in one case, at the hepatoduodenal ligament (station 12, N_3) in three cases, and in the paraaortic nodes (station 16, N_4) in one case.

Discussion

The controversy over the value of extended lymph node dissection for treatment of gastric cancer patients is fiercely debated. The available (Cape Town [20], Hong Kong [21], Dutch [5], MRC [22, 23]) prospective randomized trials and the German observational study [8] do not convince surgeons at work in centers where extended lymph node dissections can be performed with low morbidity and mortality. There are several inconsistencies, such as the tendency to perform intermediate-type dissections and to retrieve an insufficient number of lymph nodes, contamination and noncompliance during lymph node dissection, the specific experience (surgical and postoperative) of each team involved, and the morbidity and mortality related to splenectomy and distal pancreatectomy [6, 7, 24, 25]. It has been suggested [24] that the real value of extended lymph node dissection seems to be more correctly expressed by less recently published studies assessing prognostic factors in patients with gastric cancer [1, 8, 11, 26].

The role of extended lymphadenectomy can be also better understood when considering if survival can be expected after dissection of involved lymph nodes. A study [27] on this subject was recently done at the National Cancer Center Hospital, Tokyo. Among 1281 potentially curative resections for advanced gastric cancer, the histologic results of each nodal station were evaluated with special reference to survival, independent of other prognostic factors. The 5-year survival rates of patients with positive perigastric nodes (first tier, stations 1–6 according to JRSGC [15]) ranged from 18% (station 2) to 43% (station 3); and the survival of those with positive nodes in the second tier (stations 7–11) ranged from 11% (station 11) to 31% (station 7). On the assumption that the patient would not have survived if these metastatic nodes had been left in situ, D_2 dissection was beneficial for many patients with advanced gastric cancer. Among the nodes in the third tier, those in the hepatoduodenal ligament (station 12) showed relatively high dissection efficiency (5-year survival 24%), whereas all the patients with positive nodes behind the pancreas head (station 13) died of disease even after extended dissection.

The question of whether, and to what extent, D_2 dissection is capable of inducing a survival benefit for patients with gastric cancer seems to be less important for general surgical practice than to determine which nodes should be effectively dissected in each individual patient, as the incidence and site of positive nodes and the efficacy of dissection vary considerably depending on the size and location of the primary tumor. The diagnostic accuracy of

Table 5. Estimated relation between given critical values (cutoff points) of the Maruyama program's expected percentage of LN positivity, and the corresponding operating points (sensitivity, specificity) of the fitted ROC curve for each of 16 LN stations.

LN station[a]	Maruyama program's percent cutoff point	Sensitivity (%)	Specificity (%)
1	22	77.4	90.2
	18	**88.3**	**75.5**
	14	93.4	49.3
2	22	79.1	93.7
	18	**89.2**	**78.4**
	14	95.0	47.2
3	22	59.0	95.3
	18	**74.2**	**74.4**
	14	86.4	39.5
	10	94.1	26.6
4	22	79.0	95.2
	18	**89.4**	**73.6**
	14	96.6	34.1
5	22	75.1	94.8
	18	**84.4**	**73.9**
	14	95.2	36.2
6	22	79.6	91.1
	18	**88.1**	**75.2**
	14	94.6	38.6
7	22	77.4	93.3
	18	**86.0**	**76.9**
	14	97.1	34.5
8	**22**	**83.8**	**96.5**
	18	**90.5**	**72.2**
	14	98.9	41.1
9	22	77.7	94.5
	18	**87.6**	**76.2**
	14	94.9	40.0
10	22	77.2	90.0
	18	**81.9**	**75.9**
	14	93.4	63.0
11	22	79.6	91.2
	18	**80.4**	**76.9**
	14	94.0	61.2
12	22	74.6	93.3
	18	**83.8**	**75.0**
	14	95.6	59.9
13	22	71.0	91.6
	18	**79.6**	**69.9**
	14	80.5	31.4
14	22	66.4	88.4
	18	**73.5**	**65.2**
	14	80.3	26.0
15	**22**	**68.8**	**83.4**
	18	70.3	60.6
	14	81.3	36.2
16	22	79.1	96.1
	18	**88.9**	**78.6**
	14	91.0	40.1

The highest performance is highlighted in bold.
[a]JRSGC [15] numbering system.

preoperative assessment of the N category for gastric cancer is not excellent, even using endoscopic ultrasonography (EUS) [12]. As with other imaging methods, EUS can detect only enlarged lymph nodes in close vicinity to the gastric wall. Nodules that are invaded but not enlarged cannot be differentiated. With modern videolaparoscopy and laparoscopic ultrasonography it is possible to inspect suspicious lymph nodes and to obtain biopsy specimens from various sites, as well as from the region around the celiac axis [12]. Feasibility in each individual patient and potential hazards, such as free tumor cell implantation into the abdominal wall, limit the effectiveness of this method for determining which nodes should be dissected.

For this purpose, it has been reported [14] that the Maruyama computer program is a valid tool in the general surgeon's hands. Our experience with the computer program suggested several considerations. For cautious use of the program it is of paramount importance to look at the three columns as a whole (Table 3: Depth of invasion) to obtain the highest degree of accuracy. This need arises from the uncertainty of preoperative assessment of the depth of cancer invasion. Lightdale [28] found a concordance of 42% when comparing CT scanning to surgical pathology. The concordance grew to 92% with EUS, but up to now the use of EUS is restricted to centers that already have sufficient experience with this technique. In the current version of the program, the number of D_3/D_4 lymphadenectomies seems to be insufficient to avoid false-negative predictions in N_3/N_4 nodes. This lack could reduce the accuracy and the usefulness of the program in patients with advanced gastric cancer. On the other hand, it should be stressed that in our series, using an "absolute" cutoff point of 0%, there were only six cases (3.4%) in which the extent of lymph node dissection suggested by the computer analysis would have been inadequate. Moreover, one of these cases was lower-third advanced cancer with metastases at the splenic hilus, which is a truly uncommon finding in both Japanese and European series. The other absolute false-negative predictions with inadequate surgical treatment can be avoided by increasing the number of D_3/D_4 lymphadenectomies in the Japanese database. The marked difference between the number of patients considered in this study or the German analysis [14] and the number of patients stored in the historical Japanese database may represent only a theoretic bias when evaluating the computer program's accuracy regarding positive predictions. In contrast, it should be highlighted that the higher prevalence of metastatic cancers and penetrating tumors in Western countries' gastric cancer populations, with respect to the Japanese gastric cancer population, is likely to increase the positive predictive value of the Maruyama program in Italian and other Western countries' surgical settings, as the positive predictive value increases for higher prevalence values. The difference in staging the depth of the lesion between the computer system and the classification system adopted in Western institutions does not influence the performance of the program, as we considered the computer estimates for three levels of depth of invasion. The division of pT2 cases into subgroups used by the computer system, was considered of prognostic interest by the 1993 *TNM Supplement* [29] but was not adopted in the latest TNM version (1997) [30]. On the other hand, the latest version of the Japanese classification of gastric carcinoma [31] takes into account the subclassification of pT1 and pT2 cases, which adds controversy to the differences between Japanese and Western staging systems [32, 33].

All the above-mentioned considerations were not pointed out in the previous German analysis [14] of the program's accuracy, which also clearly accounted for the lower number of observations compared with our study. It should be taken into consideration that the use of postoperative findings to obtain the program's predictions in our study is in major contrast to the previously mentioned studies [13, 14] in which the preoperative findings of the prognostic variables were used. In particular, the present study showed more exactly in which cases the program could be inaccurate. The percentage of false-negative predictions detected in the German analysis [14] (4/154 patients with lymph node metas-

Table 6. Predictive values of the Maruyama program using a cutoff of >18%, by LN station.

LN station[a]	Total no. of patients	Prevalence of metastatic nodes	FPR (%)	FNR (%)	PPV (%)	NPV (%)	Accuracy (%)
1	167	104 (62.3%)	23.8	11.7	86.0	80.0	83.8
2	79	50 (63.3%)	20.6	10.0	88.2	82.1	86.1
3	171	107 (62.6%)	25.6	13.0	85.3	77.4	82.4
4	168	106 (63.0%)	25.8	10.4	85.6	80.7	83.9
5	155	99 (63.9%)	26.8	15.1	84.8	73.2	80.6
6	160	102 (63.7%)	24.1	11.8	86.5	78.6	83.7
7	164	104 (63.8%)	23.3	13.5	86.5	76.6	82.9
8	141	89 (63.1%)	26.9	10.1	85.1	80.8	83.7
9	141	86 (61.0%)	23.6	12.8	85.3	79.2	83.0
10	123	78 (63.4%)	24.4	17.9	85.3	70.8	79.7
11	139	83 (59.7%)	21.4	19.2	84.8	73.3	79.8
12	102	63 (61.8%)	25.6	15.8	84.1	74.3	80.4
13	27	16 (59.2%)	27.2	25.0	80.0	66.7	74.1
14	12	7 (58.3%)	40.0	28.5	71.4	60.0	66.7
15	6	4 (66.6%)	50.0	25.0	75.0	50.0	66.7
16	28	17 (60.7%)	27.2	11.8	83.3	80.0	82.1

FPR: false-positive rate; FNR: false-negative rate; PPV: positive predictive value; NPV: negative predictive value.
[a]JRSGC [15] numbering system.

Table 7. Clinicopathologic features of 21 cases in which the computer program gave false-negative predictions using an "absolute" cutoff point of 0%.

Patient	Gender	Age (years)	M.T.	D.I.	M.D. (mm)	Loc.	Pos.	Hist.	Lymph
SC	M	46	B3	T2	20	A	G	POR	D2
TG	M	65	B2	T3	40	M	L	POR	D3
BF	F	65	B3	T2	40	C	L	POR	D2
ME	F	70	B3	T2	46	A	L	POR	D2
VE	F	78	B3	T2	17	A	L	POR	D2
GS	F	62	B3	T3	80	C	G	SIG	D2
GS	F	57	2C	T1sm	50	M	L	MOD	D2
SC	M	46	2C	T1m	25	A	A	SIG	D2
FF	M	42	2B	T1m	30	A	A	SIG	D3
RI	M	75	2C	T1sm	15	A	A	POR	D2
CM	M	61	2B	T1sm	20	C	P	MOD	D2
LA	F	60	B2	T2	20	A	A	SIG	D2
BC	F	62	B3	T2	30	A	A	MOD	D3
MG	M	71	B2	T2	60	A	A	WEL	D2
VA	M	69	B3	T3	35	C	G	MOD	D3
CA	F	50	B1	T3	70	A	L	MOD	D2
FE	M	48	B2	T3	40	C	L	SIG	D3
CI	F	70	B3	T3	55	A	C	PAP	D3
BM	F	55	B3	T2	40	C	L	SIG	D3
CA	M	79	B2	T3	60	C	P	POR	D3
SS	F	48	B4	T2	140	C	G	SIG	D3

M.T.: macroscopic type according to the Borrmann's classification and to the JRSGC [15]; D.I.: depth of invasion according to the JRSGC; (m: mucosa; sm: submucosa); M.D.: maximum diameter; Loc.: location; (A: lower third; M: middle third; C: upper third); Pos.: position (L: lesser curvature; G: greater curvature; A: anterior wall; P: posterior wall; C: all the circumference); Hist.: World Health Organization (WHO) histologic type (POR: poorly differentiated adenocarcinoma; SIG: signet-ring cell carcinoma; MOD: moderately differentiated adenocarcinoma; WEL: well differentiated adenocarcinoma; PAP: papillary adenocarcinoma); Lymph: lymph node dissection (D2: complete removal of N1 and N2 lymph nodes; D3: complete removal of N1, N2, and N3 lymph nodes).

tases, 2.6%) was lower than that in our experience (21/176, 11.9%). In the German study the accuracy of lymph node metastases prediction was 96% for stations 13 to 16, 89% for stations 7 to 12, and 82% for stations 1 to 6. In our study, the accuracy of lymph node metastases prediction was 72.4% for stations 13 to 16, 81.6% for stations 7 to 12, and 83.4% for stations 1 to 6. Thus, we can only partially confirm the conclusions reported in the German paper.

In conclusion, the Maruyama computer program predicts with good accuracy the extent of lymph node metastases from gastric cancer, but it fails to direct exactly the surgeon to perform a more extensive lymph node dissection. Rather than being a system for computed-aided surgery, the present version of the Maruyama computer program, if properly used, could potentially allow accurate prediction of patients with limited nodal disease.

Résumé

L'étendue idéale de la lymphadénectomie (LA) en cas de cancer gastrique potentiellement curable reste controversée. Pour

Table 8. Comparative evaluation between the postoperative histologic lymph nodal findings and computer predictions for 21 patients for whom the program gave false-negative estimations using an "absolute" cutoff point of 0%.

	Actual nodal metastases			Computer estimation					
Patient	N1	N2	N3–4	No.	N1	N2	N3–4	Program validity	Surgical implication
S.C.	**5**	—	—	10	**5 (0/10)**	—	—	Absolute FN	—
	6				6 (2/10)				
T.G.	3	—	—	15	3 (10/15)	—	—	Relative FN	—
	4				4 (3/15)		**5**: s1 (2/13), s3 (1/9)		
	5				**5 (0/14)**				
B.F.	3	—	**16**	17	3 (1/17)	—	**16 (0/1)**	Absolute FN	Inadequate treatment
	4				**4 (0/17)**				
M.E.	3	—	**11**	27	3 (8/27)	—	**11 (0/12)**	Absolute FN	Inadequate treatment
V.E.	**4**	—	**12**	11	**4 (0/11)**	—	**12 (0/8)**	Absolute FN	Inadequate treatment
	6				6 (2/11)				
G.S.	1	**5**	16	73	1 (38/70)	**5 (0/65)**	16 (6/24)	Relative FN	—
	2				2 (21/70)			**5**: s3 (4/47)	
	4				4 (19/71)				
G.S.	3	—	—	38	3 (4/38)	—	—	Absolute FN	—
	5				**5 (0/37)**				
S.C.	3	**1**	—	25	3 (1/25)	**1 (0/25)**	—	Absolute FN	—
	6	**7**			**6 (0/25)**	**7 (0/25)**			
F.F.	**6**	**7**	—	23	**6 (0/23)**	**7 (0/23)**	—	Absolute FN	—
R.I.	**6**	—	—	13	**6 (0/13)**	—	—	Relative FN	—
								6: pm (4/12)	
C.M.	**1**	—	—	14	**1 (0/14)**	—	—	Absolute FN	—
L.A.	3	8	—	15	3 (3/15)	8 (1/12)	—	Relative FN	—
	4				**4 (0/15)**			**4**: sm (2/18), s1 (4/12)	
	6				6 (4/15)				
B.C.	4	**7**	15	17	4 (8/17)	**7 (0/17)**	15 (1/2)	Relative FN	—
	6				6 (5/17)			**7**: sm (1/24), s1 (11/38)	
M.G.	3	7	—	11	3 (3/11)	7 (1/11)	—	Relative FN	—
	4				**4 (0/11)**			**4**: sm (6/39), s1 (14/33)	
	6				6 (5/11)				
V.A.	1	6	**12**	55	1 (26/54)	6 (3/47)	**12 (0/14)**	Absolute FN	Inadequate treatment
	2	7	16		2 (11/51)	7 (19/51)	16 (1/14)		
	3	8			3 (29/54)	8 (7/47)			
	4	9			4 (9/54)	9 (11/49)			
		11				11 (9/45)			
C.A.	3	—	**10**	27	3 (13/27)	—	**10 (0/3)**	Absolute FN	Inadequate treatment
F.E.	3	7	**16**	19	3 (9/19)	7 (4/18)	**16 (0/2)**	Relative FN	—
		9				9 (2/15)		**16**: s1 (1/3), s3 (2/3)	
		11				11 (2/13)			
C.I.	—	8	**13**	24	—	8 (12/20)	**13 (0/5)**	Relative FN	—
								13: s1 (1/7)	
B.M.	3	**7**	—	17	3 (1/17)	**7 (0/16)**	—	Relative FN	—
								7: sm (3/27), s1 (3/10)	
C.A.	1	7	**12**	44	1 (20/44)	7 (11/41)	**12 (0/11)**	Absolute FN	Inadequate treatment
	3	10	**13**		3 (23/44)	10 (2/34)	**13 (0/5)**		
	4	11	15		4 (6/43)	11 (4/33)	15 (1/2)		
S.S.	3	—	—	3	3 (1/2)	—	—	Relative FN	—
	4				**4 (0/2)**			**4**: s1 (3/4)	

In the "Actual nodal metastases" section, the involved stations for each N group, identified by numbers (1–16), are reported. In the "Computer estimation" section, the analysis output for the same stations is reported together with the number of similar cases stored in the database program. The false-negative predicted stations are highlighted in boldface. The ratio between detected metastatic cases and dissected cases is reported in parentheses.

FN: false negative; s1: suspected serosal invasion; s3: invasion to contiguous structures; pm: muscularis propria.

réaliser une LA "rationnelle", il importe de connaître l'incidence des métastases à chaque station. A cette fin, on a développé un registre informatique comprenant l'information concernant 4302 cas de cancer de l'estomac primitif traités à l'Hôpital National Cancer Center à Tokyo entre 1969 et 1989. Pour évaluer la précision d'un tel programme informatique, on a examiné la différence entre le rapport individuel généré par ordinateur et les données stockées chez 282 patients italiens ayant eu une gastrectomie pour cancer gastrique à visée curative avec une lymphadénectomie D2 ou plus. Par l'étude des caractéristiques receveur/opérateur (ROC), on a analysé la sensibilité et la spécificité du programme pour détecter les métastases lymphatiques dans chacun des 16 sites lymphatiques régionaux. Le programme informatique a montré une bonne valeur prédictive pour les métastases dans la plupart des 16 sites ganglionnaires, car l'aire sous la courbe allait de 0.741 (site 15) à 0.944 (site 8), avec une moyenne de 0.856. Le point critique du

programme pour optimiser la validité de la prédiction était de 18%. En utilisant une valeur limite "absolue" de 0%, le taux global de faux négatives (FN) chez 176 patients N+ a été de 11.9%; parmi ceux-là, onze (6,2%) étaient des FN absolus, dans lequel le programme a complètement échoué dans l'estimation des métastases ganglionnaires; les dix cas restants (5.7%) étaient des FN rélatifs, car la prédiction spécifique était positive pour chaque profondeur différente d'invasion gastrique. Un nombre restreint de lymphadénectomies D_3-D_4 dans la banque de données historiques pourrait être à la base de l'estimation basse de métastases aux ganglions N_3-N_4 générée par ce programme. Basée sur ces données, on voit que ce programme prédit avec une bonne précision l'étendue des métastases lymphatiques du cancer gastrique mais qu'il ne peut être utilisé pour diriger le chirurgien vers une meilleure lymphadénectomie.

Resumen

Todavía existe controversia sobre cuál es la disección ganglionar (DG) óptima en el tratamiento del cáncer gástrico potencialmente curable. Para realizar una DG razonable es muy importante conocer la incidencia de metástasis en cada grupo o estación ganglionar. Con tal propósito, se desarrolló un programa de computador con base en 4.302 casos de cáncer gástrico primario tratados en el National Cancer Center Hospital de Tokio entre los años 1969 y 1989. Para determinar su grado de precisión, se investigó la diferencia entre el informe individual generado por el computador y los datos almacenados en 282 pacientes italianos sometidos a gastrectomía curativa y DG D_2 o más amplia. Se utilizó el método de Características de Operación del Receptor en el análisis estadístico para determinar la sensibilidad y especificidad del programa en la predicción de metástasis en la DG en cada uno de los 16 grupos ganglionares. El programa demostró buena capacidad de predicción de metástasis para cada uno de los 16 grupos ganglionares, puesto que las áreas por debajo de la curva oscilaron entre 0.741 (grupo 15) y 0.944 (Grupo 8), con promedio de 0.856. Un punto de corte crítico de 18% del porcentaje esperado del programa resultó ser el valor de predicción de validez máxima. Utilizando un punto de corte "absoluto" de 0%, la tasa global de negativos falsos (NF) en 176 pacientes N+ fue 11.9%; de éstos, 11 (6.2%) fueron NF absolutos, en los cuales el programa falló totalmente en la estimación de metástasis en la DG; los restantes 10 casos (5.7%) fueron NF relativos, porque la predicción específica fue positiva para una diferente profundidad de invasión del estómago. El reducido número de linfadenectomías D_3-D_4 en la base de datos puede condicionar la baja predicción de metástasis en los ganglios N_3-N_4 generada por el programa. Con base en estos datos, el programa predice con buen grado de certeza las metástasis ganglionares en el cáncer gástrico, pero no se lo recomienda para dirigir al cirujano en cuanto a la realización de la linfadenectomía más adecuada.

References

1. Akoh, J.A., MacIntyre, I.M.C.: Improving survival in gastric cancer: review of 5-year survival rates in English language publications from 1970. Br. J. Surg. *79*:293, 1992
2. Wanebo, H.J., Kennedy, B.J., Chmiel, J., Steele, G., Winchester, D., Osteen, R.: Cancer of the stomach: a patient care study by the American College of Surgeons. Ann. Surg. *218*:583, 1993
3. Japanese Research Society for Gastric Cancer and Miwa Cancer Registry: Treatment Results of Gastric Carcinoma in Japan, 1979–1982, Miwa, K., editor, no. 39. Tokyo, Mitamura Press, 1995, pp. 1–33
4. Maruyama, K., Sasako, M., Kinoshita, T., Okajima, K.: Effectiveness of systematic lymph node dissection in gastric cancer surgery. In Gastric Cancer, Nishi, M., Ichikawa, H., Nakajima, T., Maruyama, K., Tahara, E., editors, Tokyo, Springer-Verlag, 1993, pp. 293–305
5. Bonenkamp, J.J., Hermans, J., Sasako, M., van de Velde, C.J.H., Dutch Gastric Cancer Group: Extended lymph-node dissection for gastric cancer. N. Engl. J. Med. *340*:908, 1999
6. Guadagni, S., Catarci, M., De Manzoni, G., Kinoshita, T.: D1 versus D2 dissection for gastric cancer. Lancet *345*:1517, 1995
7. Roder, J.D., Bonenkamp, J.J., Craven, J., van de Velde, C.J.H., Sasako, M., Böttcher, K., Stein, H.J.: Lymphadenectomy for gastric cancer in clinical trials: update. World J. Surg. *19*:546, 1995
8. Siewert, J.R., Böttcher, K., Roder, J.D., Busch, R., Hermanek, P., Meyer, H.J.: Prognostic relevance of systematic lymph node dissection in gastric carcinoma: German Gastric Carcinoma Study Group. Br. J. Surg. *80*:1015, 1993
9. Pacelli, F., Doglietto, G.B., Bellantone, R., Alfieri, S., Sgadari, A., Crucitti, F.: Extensive versus limited lymph node dissection for gastric cancer: a comparative study of 320 patients. Br. J. Surg. *80*:1153, 1993
10. De Manzoni, G., Verlato, G., Guglielmi, A., Laterza, E., Genna, M., Cordiano, C.: Prognostic significance of lymph node dissection in gastric cancer. Br. J. Surg. *83*:1604, 1996
11. Maruyama, K.: The most important prognostic factors for gastric cancer patients: a study using univariate and multivariate analyses. Scand. J. Gastroenterol. *22*(Suppl. 133):63, 1987
12. Sendler, A., Dittler, H.J., Feussner, H., Nekarda, H., Bollschweiler, E., Fink, U., Helmbergrger, H., Hofler, H., Siewert, J.R.: Preoperative staging of gastric cancer as precondition for multimodal treatment. World J. Surg. *19*:501, 1995
13. Kampschöer, G.H.M., Maruyama, K., van de Velde, C.J.H., Sasako, M., Kinoshita, T., Okabayashi, K.: Computer analysis in making preoperative decisions: a rationale approach to lymph node dissection in gastric cancer patients. Br. J. Surg. *76*:905, 1989
14. Bollschwëiler, E., Böttcher, K., Hoëlscher, A.H., Sasako, M., Kinoshita, T., Maruyama, K., Seiwert, J.R.: Preoperative assessment of lymph node metastases in patients with gastric cancer: evaluation of the Maruyama computer program. Br. J. Surg. *79*:156, 1992
15. Japanese Research Society for Gastric Cancer: The general rules for the gastric cancer study in surgery and pathology. Jpn. J. Surg. *11*:127, 1981
16. Hermanek, P., Sobin, L.H.: UICC: TNM Classification of Malignant Tumours. New York, Springer-Verlag, 1987, p. 43
17. Swets, J.A.: Measuring the accuracy of diagnosis systems. Science *240*:1285, 1988
18. Hanley, J.A., McNeil, B.J.: The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology *143*:29, 1982
19. SAS/STAT Software: Changes and Enhancements through Release 6.11. Cary, NC, SAS Institute, 1996, p. 1104
20. Dent, D.M., Madden, M.V., Price, S.K.: Randomized comparison of R1 and R2 gastrectomy for gastric carcinoma. Br. J. Surg. *75*:110, 1988
21. Robertson, C.S., Chung, S.C.S., Woods, S.D.S., Griffin, S.M., Raimes, S.A., Lau, J.T.F., Li, A.K.C.: A prospective randomized trial comparing R1 subtotal gastrectomy with R3 total gastrectomy for antral cancer. Ann. Surg. *220*:176, 1994
22. Cuschieri, A., Fayers, P.M., Fielding, J., Craven, J., Bancewicz, J., Joypaul, V., Cook, P.: Postoperative morbidity and mortality after D1 and D2 resections for gastric cancer: preliminary results of the MRC randomised controlled surgical trial: the Surgical Cooperative Group. Lancet *347*:995, 1996
23. Cuschieri, A., Weeden, S., Fielding, J., Bancewicz, J., Craven, J., Joypaul, V., Sydes, M., Fayers, P.M., Surgical Co-operative Group: Patient survival after D_1 and D_2 resections for gastric cancer: long-term results of the MRC randomized surgical trial. Br. J. Cancer *79*:1522, 1999
24. Sano, T., Martin, I.G.: Lymphadenectomy and pancreatico-splenectomy in gastric cancer surgery. Lancet *348*:195, 1996
25. Sasako, M.: Risk factors for surgical treatment in the Dutch gastric cancer trial. Br. J. Surg. *84*:1567, 1997
26. Shiu, M.H., Moore, E., Sanders, M., Huvos, A., Freedman, B., Goodbold, J., Chaiyaphruk, S., Wesdorp, R., Brennan, M.: Influence of the

extent of resection on survival after curative treatment of gastric carcinoma: a retrospective multivariate analysis. Arch. Surg. *122*:1347, 1987
27. Sasako, M., McCulloch, P., Kinoshita, T., Maruyama, K.: New method to evaluate the therapeutic role of lymph node dissection for gastric cancer. Br. J. Surg. *82*:346, 1995
28. Lightdale, C.J.: Endoscopic ultrasonography in the diagnosis, staging and follow-up of esophageal and gastric cancer. Endoscopy *25*(Suppl. 1):297, 1992
29. Hermanek, P., Henson, D.E., Hutter, R.V.P., Sobin, L.H.: UICC. TNM Supplement 1993. A Commentary on Uniform Use. Tokyo, Springer-Verlag, 1993, pp. 120–121
30. Sobin, L.H., Wittekand, C.H.: UICC. TNM Classification of Malignant Tumors, 5th edition, Heidelberg, Springer-Verlag, 1997, pp. 60–61
31. Japanese Gastric Cancer Association: Japanese classification of gastric carcinoma: 2nd English edition. Gastric Cancer *1*:10, 1998
32. Hermanek, P.: The second English edition of the Japanese classification of gastric cancer: a Western commentary. Gastric Cancer *2*:79, 1999
33. Sasako, M., Aiko, T.: Replay to Professor Hermanek's comments on the new Japanese classification of gastric carcinoma. Gastric Cancer *2*:83, 1999