Available online at www.sciencedirect.com

ScienceDirect

EJSO 42 (2016) 123–131

EJSO
the Journal of Cancer Surgery

www.ejso.com

Predicting recurrence after curative resection for gastric cancer: External validation of the Italian Research Group for Gastric Cancer (GIRCG) prognostic scoring system

L.C. Barchi [a,*], O.K. Yagi [a,1], C.E. Jacob [a,1], D.R. Mucerino [a,1], U. Ribeiro Jr. [a,1], D. Marrelli [b,2], F. Roviello [b,2], I. Cecconello [a,1], B. Zilberstein [a,1]

[a] *Digestive Surgery Division, Department of Gastroenterology, University of Sao Paulo School of Medicine – USP Brazil, Av. Dr. Enéas de Aguiar 255 – 9° andar 05403-000, São Paulo, SP, Brazil*

[b] *Department of General Surgery and Surgical Oncology, University of Siena, Via Banchi di Sotto, 55, 53100 Siena, SI, Italy*

Accepted 13 August 2015
Available online 3 September 2015

Abstract

Background: Most nomograms for Gastric Cancer (GC) were developed to predict overall survival (OS) after curative resection. The Italian Research Group for Gastric Cancer (GIRCG) prognostic scoring system (PSS) was designed to predict the recurrence risk after curative treatment based on pathologic tumor stage and treatment performed (D1–D2/D3 lymphadenectomy). This study was carried out to externally validate the GIRCG's PSS.

Patients and methods: Adopting the same criteria used by GIRCG to build the PSS, 185 patients with GC operated with curative intention were selected. The median follow-up period was 77.8 months (1.93–150.8) for all patients and 102.5 months (60.9–150.8) for patients free of disease. The NRI (net reclassification improvement) was calculated to estimate the overall improvement in the reclassification of patients using the PSS in place of the TNM stage system.

Results: GC recurrence occurred in 70 (37.8%) patients. The mean time to recurrence was 22.2 (range 1.9–98.1) months. For patients with recurrence, the gain in the proportion of reclassification was 0.257 ($p < 0.001$), indicating an improvement of 26%. For patients without recurrence, the gain in the proportion of reclassification was −0.122 ($p < 0.001$), indicating a worsening of 12%. The NRI calculated was 0.135 ($p = 0.0527$).

Conclusion: The GIRCG's PSS, which predicts the likelihood of recurrence after radical surgical treatment for GC, is more accurate than TNM system to predict recurrence mainly for high-risk patients. Yet, the PSS does not have the same effectiveness for low-risk patients, overestimating the chance of recurrence occurs even for disease-free patients.

Keywords: Gastric cancer; Recurrence; Lymphadenectomy; Stomach neoplasm; Gastrectomy; Adenocarcinoma; Nomogram

Introduction

Gastric cancer (GC) remains a very poor prognosis disease. Although its incidence has decreased,[1] it still ranks 2nd place in cancer deaths throughout the world,[2] with Japan, South Korea, China, Latin America (Chile) and Eastern Europe locations with the highest incidence. It is estimated that approximately one million people die

* Corresponding author. Tel.: +55 11 3061 7560.
E-mail addresses: leandrobarchi@hotmail.com (L.C. Barchi), okyagi@uol.com.br (O.K. Yagi), cejacob@uol.com.br (C.E. Jacob), drmucerino@uol.com.br (D.R. Mucerino), ulyssesribeiro@terra.com.br (U. Ribeiro), daniele.marrelli@unisi.it (D. Marrelli), franco.roviello@unisi.it (F. Roviello), icecconello@hotmail.com (I. Cecconello), brunozilb@uol.com.br (B. Zilberstein).
[1] Tel.: +55 11 3061 7560.
[2] Tel.: +39 0577 232111.

http://dx.doi.org/10.1016/j.ejso.2015.08.164
0748-7983/

worldwide every year[3] and in Brazil in 2010 there were 22,035 deaths from stomach cancer.[4] Fortunately, thanks to excellent screening programs and early detection, there has been an increase in patients' survival with GC, mainly in Japan and South Korea.[5]

Although surgery is still considered the only curative treatment, considering the extent of resection and lymphadenectomy performed,[6] some prognostic benefits depend not only on multimodal treatments but also patient characteristics and the tumor itself. Recent studies on molecular biology and target therapies are in addition to other factors known to determine the prognosis of patients.[7,8] Yet, these studies are expensive and restricted to major technological centers. The ability to predict prognosis is essential in order to establish the best treatment plan and organize follow-up strategies.

Several nomograms have been created to predict overall survival (OS) in patients undergoing potentially curative surgery for GC.[9–12] They are statistical forecasting tools for patient individual analysis based on known prognostics variables, including the extent of the treatment, which provide the overall probability of a specific outcome. To date, few nomograms for GC were created and validated: Han et al. analyzed 7954 patients who underwent D2 gastrectomy for gastric cancer at Seoul National University Hospital (SNUH) and constructed a nomogram that predicts 5 and 10 year overall survival. Additional external validation was performed using the data set (n 2500) from Cancer Institute Ariake Hospital (CIAH) in Tokyo. Validation using the SNUH and CIAH data sets revealed good discrimination and calibration, suggesting good clinical utility. This nomogram improved individualized predictions of survival.[9] Kattan et al. using 1039 patients who underwent R0 resection for GC at the Memorial Sloan-Kettering Cancer Center (MSKCC) created a nomogram that showed superiority in predicting disease-specific survival (DSS) at 5 and 9 years to survival prediction of the TNM stage system (TNM) 10. This nomogram was subsequently validated as a precise tool for prediction of DSS in the United States, Holland, and Germany[13–15] and more recently in China.[16] Marrelli et al. using a logistic regression model with 536 patients developed a prognostic scoring system (PSS) of the Italian Research Group for Gastric Cancer (GIRCG) that predicts recurrence after R0 resection for GC. This PSS was more effective in predicting recurrence when compared to TNM.[11] These results were validated by the same group analyzing 635 patients, but not externally by another center.[17] It is unknown if the effectiveness of the GIRCG scoring system to predict recurrence is reproducible in a different population of patients. The centers in question (GIRCG and Hospital das Clínicas, School of Medicine, University of São Paulo – HCFMUSP) are comparable in number of cases and OS and use similar strategies in the surgical treatment for GC. Both perform D2 lymphadenectomy (D2Ly) as recommended by the Japanese School with curative intent.[18] The aim of this study is to validate the PSS advocated by GIRCG in patients operated for gastric cancer with R0 resection in the HCFMUSP.

Patients and methods

In order to validate the GIRCG's PSS the 185 patients selected met the same criteria used by GIRCG to build the score. Inclusion criteria were: patients who underwent potentially curative surgery with R0 resections. Patients who had disseminated disease or distant metastases, even being submitted to palliative resections were not included in the study. As well as patients with less than 5 years of follow up, patients who died due to perioperative complications (within 30 days) or who eventually died of other causes not related to GC. Tumors of the esophagogastric junction (EGJ) and *Linitis plastica* were also excluded. The first GIRCG prognostic score considered pT and pN status according to the 6th edition of the TNM classification (2005). With the introduction of the new TNM classification (7th edition),[19] the scoring system was recalculated according to the new criteria, excluding EGJ neoplasms.[20] In the current AJCC staging, EGJ tumors are staged as esophageal and not as gastric. Regarding *Linitis plastica*, it has been associated with invariable poor prognosis in a recent report and therefore could not be included in the present study.[21]

Of the 185 patients with GC operated with curative intention in University of São Paulo, School of Medicine (HCFMUSP) from January 2001 to December 2007, 106 patients (57.3%) were male and 79 were female (42.7%) with a mean age of 59.6 years (±12.8 SE 0.94 range 22–88 years). The median follow-up period was 77.8 months (range 1.93–150.8) for all patients and 102.5 months (range 60.9–150.8) for patients free of disease. Tumor's location was divided into proximal (PGT), middle (MGT) and distal thirds (DGT). For most DGT subtotal gastrectomy (STG) was performed. For all PGT, most MGT and some DGT total gastrectomy (TG) was preferred. The decision was made according to tumor's size and histopathological type, in order to obtain a R0 resection. Regarding the safety margin is recommended at least 2–3 cm of proximal margin in patients with early gastric cancer, 5 cm in advanced lesions of the differential type, and 7–10 cm in the undifferentiated advanced lesions. Whenever necessary, frozen section examination of the proximal and distal margins was performed, and the totalization of the gastrectomy was accomplished.[22] All patients were reconstructed using the Roux-en-Y technique. Tumors were classified according to TNM 7th ed.[19]

In all patients was held complete preoperative clinical staging through upper endoscopy with biopsies, computed tomography of the chest and abdomen, oncomarkers (CEA, CA19.9). Diagnostic laparoscopy was performed in selected cases, when there was doubt about the tumor resection and the presence of peritoneal carcinomatosis.

Peritoneal cytology was held in all cases, and if positive, patients were excluded from the study. None of the patients included in this study received neoadjuvant therapy. The same tests were used in postoperative period, every 6 months in the first 2 years and every year until the fifth year after surgery. Patients were considered disease free if all imaging studies and laboratory were normal.

After being classified by pTNM system (7th ed.) all patients were entered into the Excel program demonstrated in Fig. 1, also considering the extent of Ly performed, as recommended by the GIRCG nomogram.[11] Patients who received D1 Ly were allocated in the *LIMITED* cell, while patients treated with D2/D3 Ly were allocated in the *EXTENDED* cell. D2 Ly has been routinely used with curative intent in our institution for several years. As most of our patients tumors are indeed advanced cases, that was the type of the lymphadenectomy performed in this series. In order to enable the comparison between the scoring system and the TNM system, patients were divided according to the value of the score obtained, considering the risk of recurrence, and the TNM stage, considering the survival probability. The categorical variable "TNM stage" was regrouped into three categories of risk for recurrence using the curves estimated by the Kaplan–Meier method. The categories that showed near curves were grouped, thus forming three types of risk of recurrence: Mild (stages Ia, Ib, IIa and IIb), moderate (IIIa and IIIb) and intensive (IIIc). For the continuous variable "score", three categories were created and applied the same method used in the regrouping stage, also coming to three risk categories: low (Group I: 0–30), moderate (Group II: 31–69) and high (Group III: 70–100).

The distribution of patients in groups according to the score and the TNM stage obtained are shown in Table 1.

Table 1
Frequency distribution of variables in prediction and study outcome.

Variables – n (%)	(n = 185)
Stage: Risk	
Mild (Ia ou Ib ou IIa ou IIb)	96 (51.9%)
Moderate (IIIa ou IIIb)	60 (32.4%)
Intensive (IIIc)	29 (15.7%)
Score: Risk	
Group I (0–30)	85 (45.9%)
Group II (31–69)	46 (24.9%)
Group III (70–100)	54 (29.2%)

Statistical analysis

Quantitative variables were expressed as median, mean, standard deviation (SD) and standard error (SE) and the qualitative as absolute and relative frequencies (percentages). Differences between the recurrent patients and non-recurrent patients were verified using Mann–Whitney and Fisher tests. Survival curves and the cumulative proportion of patients with recurrence were estimated by the method of Kaplan–Meier considering the period from and after tumor resection, to recurrence or the last follow-up for patients with at least five years of follow-up time.

The NRI (net reclassification improvement) was calculated to estimate the overall improvement in the reclassification of patients using the score system in place of the TNM stage system. It is a statistic model that measures the improvement in prediction performance gained by adding a marker to a set of baseline predictors for predicting a binary outcome.[23]

All statistical analyzes were performed using SPSS version 19 and adopted a significance level of 0.05.

Results

Only 46 (24.9%) patients had early gastric cancer, 23 (12.4%) restricted to the mucosa (T1a) and 23 (12.4%) with submucosa invasion (T1b). Twenty-five (13.5%) patients had muscularis propria invasion (T2), 32 (17.3%) subserosa (T3), 77 (41.6%) serosa (T4a) and five (2.7%) affected adjacent structures (T4b). When we analyzed lymph node status, 79 (42.7%) patients were considered N0, 27 (14.6%) N1, 34 (18.4%) N2 and 45 (24.3%) N3. The average number of lymph nodes resected was 42.3 (±17.7 SE 1.3 range 15–114). The average number of positive lymph nodes was 4.8 (±7.67 SE 0.56 range 0–44). Lymph Node Ratio (LNR), which is the reason of positive lymph nodes and examined nodes, was 0.12 (±0.18 SE 0.01 range 0–0.96). The survival curve for each stage is demonstrated in Fig. 2A. The mean score obtained was 44.7 (±34.6 SE 2.55 range 2.5–99.7). Association between

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	NAME	LOCATION			DEPTH OF INVASION				NODAL STATUS					LYMPHADENECTOMY		EQ	SCORE
2		LOWER	MIDDLE	UPPER	T1	T2	T3	T4	N0	N1	N2	N3a	N3b	LIMITED	EXTENDED		
3	EXAMPLE 1	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
4																	
5																	
6																	

Figure 1. GIRCG prognostic score using Excel. See instructions and download the model at the website: www.gircg.it
In the cell EQ (P3) insert the formula: = – 1.978 – 0.366(C3) + 2.123(D3) + 0.912(F3) + 1.363(G3) + 2.307 (H3) + 1.333(J3) + 1.634(K3) + 3.616(L3) + 4.728 (M3) – 1.290(O3)
In the cell **SCORE (Q3)**, insert the formula: = 1/(1 + 1/EXP (P3))*100
The numbers that appears in the cell SCORE corresponds to the estimated risk of recurrence for each patient.

clinicopathological variables, score and recurrence are shown in Table 2.

The curve corresponding to each survival score group was calculated by Kaplan–Meier (KM) method and is shown in Fig. 2B. We notice a clear difference between the three groups, where, after 5 years, in group I about 85% of patients were without evidence of recurrence while in groups II and III approximately 59% and only 27% of patients were free of disease, respectively ($p < 0.001$). When we divide patients into 10 subgroups according to score, as well as in GIRCG patients, most HCFMUSP patients stood at the extremes of the range. The minority was located in the middle of the range, with an intermediate risk of recurrence (Fig. 2C).

GC recurrence occurred in 70 (37.8%) of 185 patients. The mean age of patients with recurrence was 60 years (±12.5; SE 1.49, range 33–88). The mean time to recurrence was 22.2 (range 1.9–98.1) months. The average values of the scores of patients with recurrence and disease-free were, respectively, 68.2 (±29.7 range 3.7–97.7) and 30.5 (±29.4 range 2.57–97.7). We noted a strong relationship between the incidence of recurrence and score level value obtained, ie, the risk of recurrence increased remarkably with the score values (Fig. 2D). The

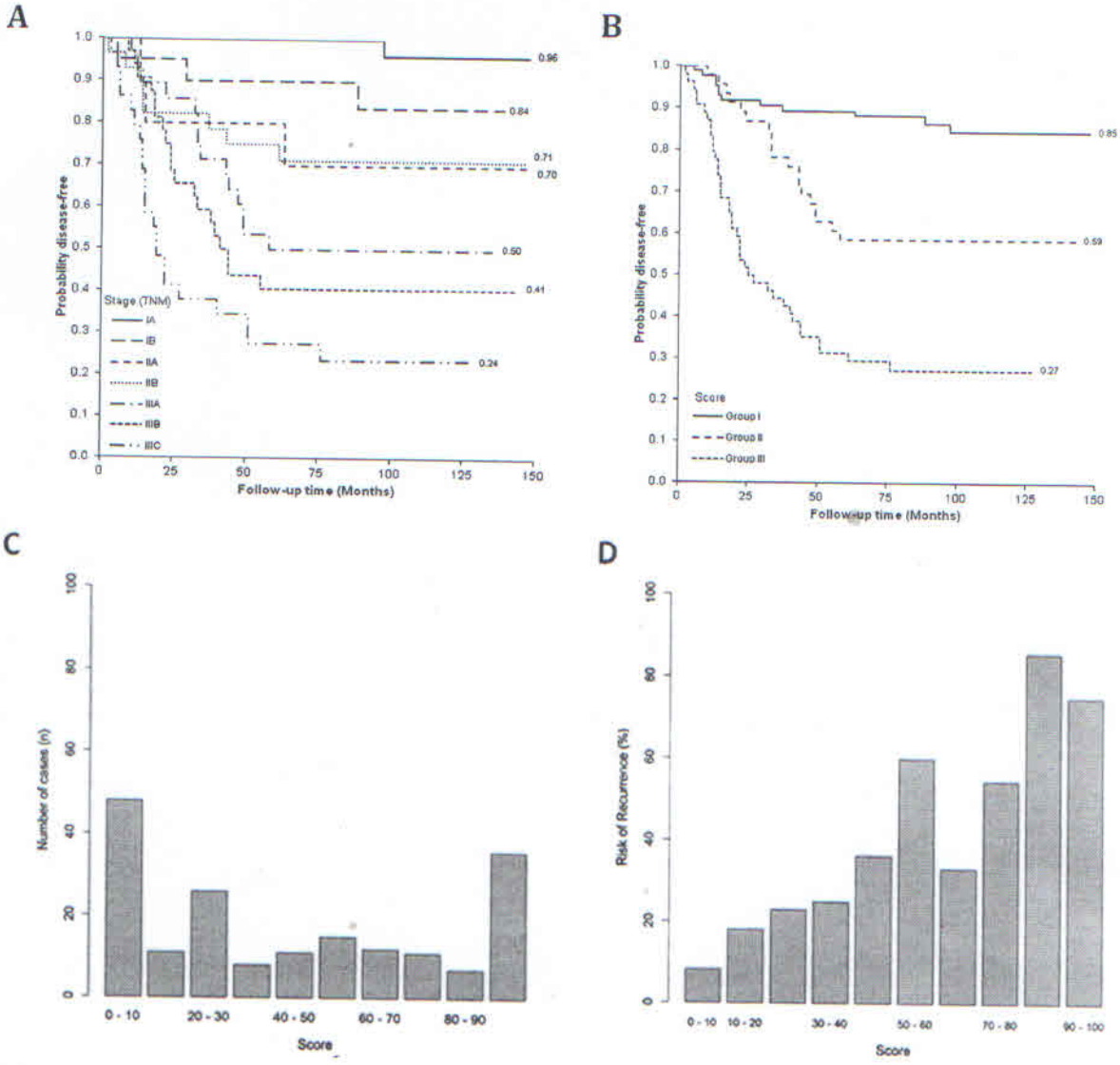

Figure 2. All patients (185) received R0 GC resection at HCFMUSP from 2001 to 2007. **A.** Kaplan–Meier method (KM) of DSS (TNM 7° edition). Each line represents the survival of patients within a single TNM stage ($p < 0.001$) **B.** KM of GIRCG prognostic score. Each line represents the survival of patients within a subgroup of score stage ($p < 0.001$) **C.** Distribution of score levels. Most patients are placed in the extreme of the range, whereas few patients were classified as intermediate risk group. **D.** Incidence of recurrence is score subgroups. A strong correlation between incidence of recurrence and score levels was observed.

Table 2
Correlation between clinicopathological variables, score and recurrence.

Variable	Cases (n)	With recurrence (n = 70)	Without recurrence (n = 115)	p
Gender				
Male	106 (57.3)	41 (58.6)	65 (56.5)	*0.785*
Female	79 (42.7)	29 (41.4)	50 (43.5)	
Age (years)		60 ± 12.5	59.5 ± 13	*0.513*
Tumor location				
Proximal	18 (9.7)	7 (10)	11 (9.6)	*0.521*
Middle	47 (25.4)	21 (30)	26 (22.6)	
Distal	120 (64.9)	42 (60)	78 (67.8)	
Gastrectomy				
Subtotal	88 (47.6)	27 (38.6)	61 (53)	*0.056*
Total	97 (52.4)	43 (61.4)	54 (47)	
Nodal status				
Positive	106 (57.3)	8.9 ± 9.6 (SE 1.14)	1.9 ± 4.3 (SE 0.40)	***<0.001***
Examined		45 ± 18.8 (SE 2.24)	40.4 ± 16.8 (SE 1.57)	*0.079*
LNR		0.2 ± 0.22 (SE 0.03)	0.05 ± 0.09 (SE 0.01)	***<0.001***
Stage (TNM)				
IA	38 (20.5)	1 (1.4)	37 (32.2)	***<0.001***
IB	20 (10.8)	3 (4.3)	17 (14.8)	
IIA	10 (5.4)	3 (4.3)	7 (6.1)	
IIB	28 (15.1)	8 (11.4)	20 (17.4)	
IIIA	28 (15.1)	14 (20)	14 (12.2)	
IIIB	32 (17.3)	19 (27.1)	13 (11.3)	
IIIC	29 (15.7)	22 (31.4)	7 (6.1)	
Score				
Group I [0–30]	85 (45.9)	12 (17.1)	73 (63.5)	***<0.001***
Group II [31–69]	46 (24.9)	19 (27.1)	27 (23.5)	
Group III [≥70]	54 (29.2)	39 (55.7)	15 (13)	

Data are number (%) of patients; Mean ± Stander Deviation; SE (Stander Error).

cumulative risk of recurrence is shown in Fig. 3. Patients with high scores had greater chance of recurrence in a short time, while patients with low score remained free of disease for a long period. Eight (4.3%) patients with score higher than 90 showed no recurrence and only four (2.1%) patients with score lower than 10 recurred (Fig. 4A). The comparison between the predicted risk of recurrence according to the values of the score obtained in GIRCG and HCFMUSP patients are shown in Fig. 4B. Curves are related and overlapping in most parts, with the exception of the score values between 60 and 70, where an inversion occurs.

The association between TNM stage groups and recurrence are represented in Table 3A. We found significant association between recurrence and stage ($p < 0.0001$), where the relapse percentage in the Mild group (15.6%) was significantly lower than in the Moderate group (55%) that was significantly lower than in the Intensive group (75.9%). The association between score groups and recurrence are represented in Table 3B. We found significant association between score and recurrence ($p < 0.0001$), where the relapse percentage in group I (14.1%) was significantly lower than in group II (41.3%) that was significantly lower than in group III (72.2%).

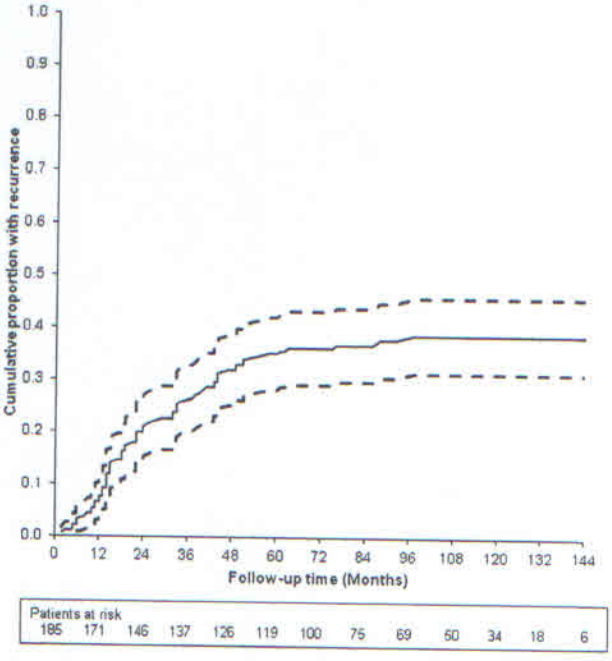

Figure 3. Cumulative risk of recurrence (KM) in 5 and 10 years of follow-up is, respectively, 35.1% (±7%) e 38.7% (±7%). The numbers of patient at risk in each subperiod is indicated below column time.

To evaluate the effectiveness of the scoring system for predicting recurrence in comparison to the TNM, the NRI was calculated to estimate the overall improvement in the reclassification of patients using the score system in place of the TNM stage system. In Table 4 is the analysis of reclassification of risk of recurrence using score groups compared to TNM stage groups. In the sample of 70 patients who had recurrence, using score groups, the risk was increased for 18 patients (moving up) and in no case was observed risk reduction (moving down). Thus, the gain in the proportion of reclassification was 0.257, significantly greater than zero ($p < 0.001$), indicating an improvement around 26% in the classification of patients with recurrence. In the sample of 115 patients without recurrence, using the score groups, the risk was increased for 15 patients (moving up) and in one case was observed risk reduction (moving down). Thus the gain in the proportion of reclassification was −0.122, significantly less than zero ($p < 0.001$), indicating a worsening of around 12% in the classification of patients without recurrence. The NRI calculated is 0.135 ($p = 0.0527$).

Discussion

Surgery remains the primary treatment with curative intent for GC. Currently, the prognosis is estimated based on the TNM system.[19] Other factors such as age, gender, comorbidities, tumor diameter,[24] Bormann classification, LNR,[25] degree of cell differentiation, vascular invasion, DNA ploidy[25,26] and type of treatment performed,[24,26]

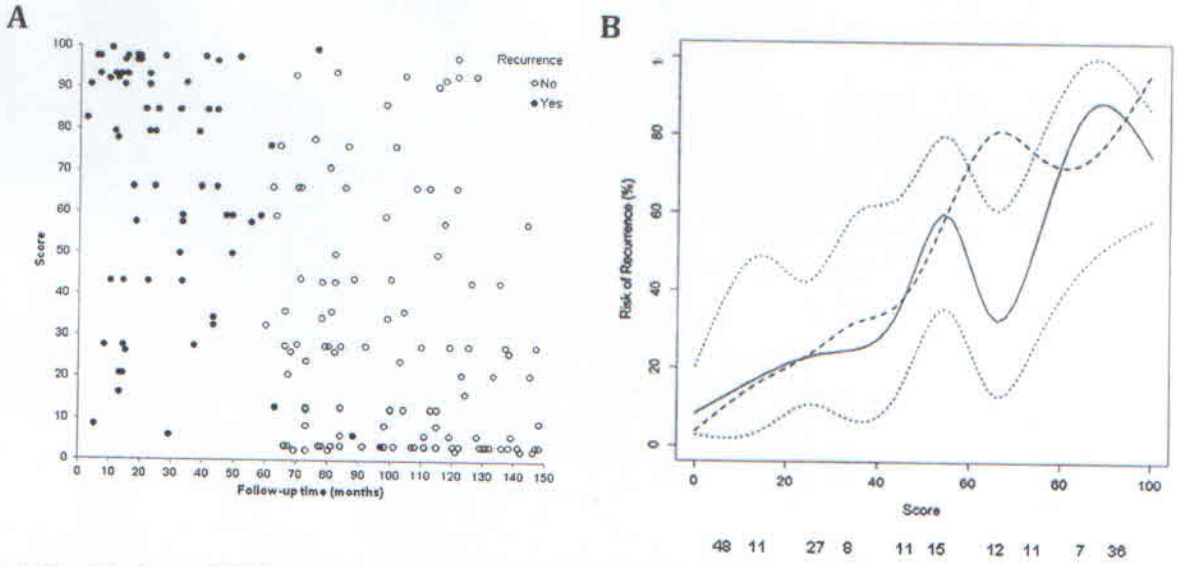

Figure 4. **A**. Correlation between GIRCG prognostic score and follow-up time (months) in 185 patients treated in HCFMUSP. Most patients with high score level relapsed within 2 years, whilst patients with low score level remained disease-free for a long follow-up time. **B**. Linear definition of he risk of recurrence according to score level. Predicted GIRCG patients (broken line), observed HCFMUSP patients (solid line) and confidence interval (dotted line). The distribution of HCFMUSP patients is shown below score column.

which may have an impact on prognosis, are not routinely used. A few years ago, some other factors were reported as relevant in predicting survival of GC, such as free tumor cells in the peritoneum,[27] microsatellite instability,[28–30] loss of tumor suppressor genes[31] and multiplicity of molecular markers.[32] The integration of these variables in nomograms produced more accurate models for predicting recurrence and DSS than the TNM system. That is because they reflect not only tumor characteristics, but also the status of the host and are capable to incorporate more clinicopathological parameters than the TNM.[33] In order to considerer a nomogram effective and routine clinically applicable, it needs to be validated by different large studies, in different places and populations, analyzing the pathological behavior and prognosis of a given disease.[14–16,34–37] Unlike most nomograms created for GC, which basically predict the probability of survival, the GIRCG's nomogram simply seeks to predict the risk of disease relapse after curative treatment based on pathologic tumor stage and treatment performed (D1–D2/D3 Ly).

The results obtained in this study were similar to the Italian group study: There were more males (57.3% HCFMUSP patients vs. 63.2% GIRCG patients.). The mean age was 59.7 vs. 63 years. There was a higher prevalence of tumors located in the distal third (65% vs. 51%). Recurrence occurred in about 38% of HCFMUSP patients vs. 50% of GIRCG patients, with a median time of 22 vs. 19 months. In both groups, the majority of patients relapsed within 2 years after surgery. Only 2.1% of patients in this study and 2.4% of GIRCG patients recurred after 5 years of follow up. There was a slight preference performing TG in HCFMUSP patients in relation to STG (52.4% vs. 47.6%), what did not occur in the Italian group (46.6% vs. 53.4%). The median overall and disease free patients follow-up time were higher in HCFMUSP patients than the GIRCG (77.8 vs. 56 months and 102.5 vs. 94 months), respectively. In both centers there was a higher incidence of advanced tumors – T3/T4 (58.9% vs. 46.3%) and without lymph node involvement – N0 (42.7% vs. 40.6%). When we compare scores values, the results resembled a lot. The vast majority of patients had high or low scores values, i.e., with high or low risk of recurrence. The minority received intermediate scores values, with moderate risk of recurrence. When we overlap the predicted and observed recurrence curves according to the score level between the two studies, we can observe significant similarity in practically all score subgroups, except scores values between 60 and 70, where an inversion in the

Table 3A
Correlation between TNM stage groups and recurrence in 185 patients.

Recurrence	Stage: Risk		
	Mild (n = 96)	Moderate (n = 60)	Intensive (n = 29)
No	81 (84.4%)	27 (45%)	7 (24.1%)
Yes	15 (15.6%)	33 (55%)	22 (75.9%)
p-Value	*<0.0001*		

Table 3B
Correlation between score groups and recurrence in 185 patients.

Recurrence	Score: Risk		
	Group I (n = 85)	Group II (n = 46)	Group III (n = 54)
No	73 (85.9%)	27 (58.7%)	15 (27.8%)
Yes	12 (14.1%)	19 (41.3%)	39 (72.2%)
p-Value	*<0.0001*		

Table 4
Concordance between TNM stage groups and score groups according to recurrence.

Recurrence (n = 70)			
Stage: Risk	Score: Risk		
	Group I (0–30)	Group II (31–69)	Group III (70–100)
Mild (Ia ou Ib ou IIa ou IIb)	12	1	2
Moderate (IIIa ou IIIb)	0	18	15
Intensive (IIIc)	0	0	22
18 events moving up	0 event moving down		52 events not moving
Without recurrence (n = 115)			
Stage: Risk	Score: Risk		
	Group I (0–30)	Group II (31–69)	Group III (70–100)
Mild (Ia ou Ib ou IIa ou IIb)	72	7	2
Moderate (IIIa ou IIIb)	1	20	6
Intensive (IIIc)	0	0	7
15 events moving up	1 event moving down		99 events not moving

NRI = [(18/70) − (0/70)] + [(1/115) − (15/115)] = **0.135** (*p = 0.0527*).

curve of HCFMUSP patients occurs (Fig. 4B). This might be explained because only eight (4.3%) HCFMUSP patients met this subgroup score level, and only 25% showed recurrence, whereas 73% of GIRCG patients allocated in this subgroup relapsed. We speculate that with larger number of patients in this range of scores (60–70) the curves of the two studies would approach. In fact, a special feature of the score is to strongly reduce the number of cases in the intermediate group (between 30 and 70), and concentrate them in the extreme of the range. In this group with few patients, only one event (recurrence) changes the percentage of survival significantly.

For patients with recurrence, using Marrelli's score increased significantly the prediction of recurrence when compared to the TNM system (26%). However, it overestimated the chance of recurrence occurs even for disease-free patients (12%). This demonstrates that the score is useful mainly for high-risk patients, but does not have the same effectiveness for low-risk patients.

The MSKCC nomogram has been validated by numerous studies in the West, and more recently in the east. Petters et al. (2009) validated this nomogram in Dutch patients, regardless the extent of lymphadenectomy performed.[14] In a study conducted by Novotny et al. (2006) in Germany, the calibration of the nomogram comparing predicted survival with actual survival was excellent.[15] Chen et al. (2013) conducted a cohort study analyzing 979 Chinese patients who underwent R0 resection for GC to validate the MSKCC nomogram. They concluded that even in Eastern populations, the nomogram was able to predict DSS with greater efficacy when compared to the TNM.[16] Moreover, the same results were not obtained by Koc et al. (2009) that analyzing 65 Turkish patients operated for GC with R0 resection did not achieve the same accuracy to validate the MSKCC nomogram.[36]

Hirabayashi et al. (2014) analyzed 5.196 patients operated for locally advanced GC without serosa invasion with R0 resections in multiple centers in Japan and created a nomogram including lymphovascular invasion as an independent prognostic factor. This cohort study demonstrated superiority in predicting OS when compared to TNM.[37]

The main peculiarity of the GIRCG score is that it is specifically designed to predict the risk of recurrence. This might fit better with the purpose of the score (patient counseling, planning additional therapies and tailored follow-up), with respect to the estimation of survival probability. Another potential utility is that it can overcome some limitations of the TNM system, such as the heterogeneity within the stages. For instance, stage IIB includes T4aN0, T3N1, and T2N2 patients, who have different prognosis, while the scoring system assigns risk according to pT and pN separately.[38] Furthermore, the type of lymphadenectomy performed directly impacts the risk of recurrence. According to the GIRCG scoring system, a patient with a distal tumor, T2N1 undergone D1 Ly would have 57% risk of recurrence, while when performing D2 Ly in this same patient, this chance would drop to 26%. Patients with higher risks of relapse should be followed closer, with more frequent endoscopies, CTs, serum oncomarkers, etc. The purpose of follow-up programs is to assess the long-term complications, collect data on survival and clinical course of the disease and to diagnose as soon as possible, any potential recurrence. Ideally, recurrence should be diagnosed in an asymptomatic and still treatable stage of disease. However, when it occurs, is a condition with an unfavorable prognosis in most cases.

Yet, Marrelli's score has some limitations. First, he does not consider patient comorbidities as variables. Although it is difficult to categorize and quantify the risk of comorbidities due to immense diversity of them, it is expected that patients with severe chronic diseases, with weakened immune system have more chance to relapse and therefore worse prognosis, independent of tumor stage and treatment received. Second, adjuvant treatments were not considered to build the scoring system in Italian patients, because they were not standardized at the time of score calculation. As such, their potential impact of risk of recurrence cannot be assessed. At the time of publication (2005), there were no protocols that revealed real benefit chemo/radiotherapy (CRT) in patients undergoing R0 resection. To date, numerous studies have proven increased OS and DSS with adjuvant CRT in patients undergoing curative surgery for GC.[39–42] It seems reasonable to include high-risk patients into aggressive multimodal treatments protocols in order to try diminishing relapse rate. Nowadays, the results of patients in advanced stages (III and IV) are still unsatisfactory even with adequate surgery and extended lymphadenectomy, especially in western countries. This tailored treatment has been approached differently around the

world. For example, in Asia, the most commonly used treatment is adjuvant chemotherapy; in the United States, the favored treatment is CRT; and in Europe, neoadjuvant therapy is mostly used.[43] In a study conducted by Dikken et al. (2014) using MSKCC nomogram in patients who have undergone R0 resection for GC plus received adjuvant CRT, there was a 20% increase in observed survival when compared to expected survival for patients who received post-operative CRT. These data support the need to update nomograms with the incorporation of multimodal strategies.[44]

There are also some limitations of this study. The GIRCG scoring system was validated using retrospective data from a single Brazilian high-volume institution. To nationwide use is necessary that other Brazilian centers perform this validation. Moreover, many patients that were used to validate the Italian nomogram received adjuvant CRT (indicated in T3/T4 and/or N+). It is unknown if this could have some impact in our results.

In conclusion, the GIRCG's PSS, which predicts the likelihood of recurrence after radical surgical treatment for GC, is more accurate than TNM system to predict recurrence mainly for high-risk patients. Yet, the PSS does not have the same effectiveness for low-risk patients, overestimating the chance of recurrence occurs even for disease-free patients. The TNM stage system is simple, easily reproducible and has high prognostic accuracy and therefore should not be replaced, but used in combination with the score, identifying high-risk groups and then individualize patients when necessary. Marrelli's score could be useful in daily clinical practice, providing better post-operative treatment planning and follow-up.

Acknowledgments

The authors thank Márcio Diniz (Laboratory of Epidemiology and Statistics from Department of Gastroenterology, School of Medicine – University of São Paulo) and Sandra Malagutti for statistical support.

Conflict of interest statement

The authors declare no conflict of interest.

References

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, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005;**55**:74–108.
3. Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol* 2006;**24**:2137–50.
4. Instituto Nacional de Câncer José Alencar Gomes da Silva. Magnitude do câncer no Brasil: incidência, mortalidade e tendência. *Informativo Vigilância do Câncer* 2012;1–26. n. 3, mai./nov.
5. Park CH, Song KY, Kim SN. Treatment results for gastric cancer surgery: 12 years' experience at a single institute in Korea. *Eur J Surg Oncol* 2008;**34**:36–41.
6. Songun I, Putter H, Kranenbarg EM, et al. Surgical treatment of gastric cancer: 15-year follow-up results of the randomised nationwide Dutch D1–D2 trial. *Lancet Oncol* 2010;**11**:439–49.
7. Bang YJ, Van Cutsem E, Feyereislova A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastroesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* 2010;**376**:687–97.
8. Wu J, Liu X, Cai H, Wang Y. Prediction of tumor recurrence after curative resection in gastric carcinoma based on bcl-2 expression. *World J Surg Oncol* 2014;**12**:40.
9. Han DS, Suh YS, Kong SH, et al. Nomogram predicting long-term survival after d2 gastrectomy for gastric cancer. *J Clin Oncol* 2012; **30**:3834–40.
10. Kattan MW, Karpeh MS, Mazumdar M, et al. Post-operative nomogram for disease-specific survival after an R0 resection for gastric carcinoma. *J Clin Oncol* 2003;**21**:3647–50.
11. Marrelli D, De Stefano A, de Manzoni G, et al. Prediction of recurrence after radical surgery for gastric cancer: a scoring system obtained from a prospective multicenter study. *Ann Surg* 2005;**241**: 247–55.
12. Costa MLV, Ribeiro KCB, Machado MAC, Costa ACLV, Montagnini AL. Prognostic score in gastric cancer: the importance of a conjoint analysis of clinical, pathologic, and therapeutic factors. *Ann Surg Oncol* 2005;**13**(6):843–50.
13. Gold JS, Tang LH, Gönen M, et al. Utility of a prognostic nomogram designed for gastric cancer in predicting outcome of patients with R0 resected duodenal adenocarcinoma. *Ann Surg Oncol* 2007;**14**:3159–67.
14. Peeters KC, Kattan MW, Hartgrink HH, et al. Validation of a nomogram for predicting disease-specific survival after an R0 resection for gastric carcinoma. *Cancer* 2005;**103**:702–7.
15. Novotny AR, Schuhmacher C, Busch R, et al. Predicting individual survival after gastric cancer resection: validation of a U.S.–derived nomogram at a single high-volume center in Europe. *Ann Surg* 2006;**243**:74–81.
16. Chen D, Jiang B, Xing J, et al. Validation of the memorial sloan-kettering cancer center nomogram to predict disease-specific survival after r0 resection in a chinese gastric cancer population. *Plos One* 2013;**8**(10):e76041. http://dx.doi.org/10.1371/journal.pone.0076041.
17. Marrelli D, Morgagni P, de Manzoni G, et al. Italian Research Group for Gastric Cancer (IRGGC). External validation of a score predictive of recurrence after radical surgery for non-cardia gastric cancer: results of a follow-up study. *J Am Coll Surg* 2015 Aug;**221**(2):280–90.
18. Maruyama K, Sasako M, Kinoshita T, et al. Should systematic lymph node dissection be recommended for gastric cancer? *Eur J Cancer* 1998;**34**:1480–9.
19. Edge SB, Compton CC, Fritz AG, Greene FL, Trotti A. *AJCC cancer staging manual*. 7th ed. New York: Springer; 2010.
20. Marrelli D, Caruso S, Roviello F. Prognostic factors and score systems in gastric cancer. In: de Manzoni G, Roviello F, Siquini W, editors. *Surgery in the multimodal management of gastric cancer*. Milan: Springer; 2012;, p. 35–42.
21. Pedrazzani C, Marrelli D, Roviello F, et al. Gastric linitis plastica: which role for surgical resection? *Gastric Cancer* 2012 Jan;**15**(1): 56–60. http://dx.doi.org/10.1007/s10120-011-0063-z [Epub 2011 Jun 30].
22. Zilberstein B, Malheiros C, Jacob CE, et al. Brazilian consensus in gastric cancer: guidelines for gastric cancer in Brazil. *ABCD Arq Bras Cir Dig* 2013;**26**(1):2–6.

23. Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics Med* 2008;**27**: 157–72.
24. Maruyama K. The most important prognostic factors for gastric cancer patients: a study using univariate and multivariate analyses. *Scand J Gastroenterol* 1987;**2**(Suppl. 133):63.
25. Siewert JR , Bottcher K, Stein HJ, et al. Relevant prognostic factors in gastric cancer: ten-year results of the German gastric cancer study. *Ann Surg* 1998;**228**:449–61.
26. Korenaga D, Haraguchi M, Okamura T, et al. DNA ploidy and tumor invasion in human gastric cancer: histopathologic differentiation. *Arch Surg* 1989;**124**:314–8.
27. Ribeiro Jr U, Safatle-Ribeiro AV, Zilberstein B, et al. Does the intraoperative peritoneal lavage cytology add prognostic information in patients with potentially curative gastric resection? *J Gastrointest Surg* 2006 Feb;**10**(2):170–6. discussion 176–7.
28. Keller G, Rotter M, Vogelsang H, et al. Microsatellite instability in adenocarcinomas of the upper gastrointestinal tract: relation to clinicopathological data and family history. *Am J Pathol* 1995;**147**:593–600.
29. Nakashima H, Inoue H, Mori M, et al. Microsatellite instability in Japanese gastric cancer. *Cancer* 1995;**75**:1503–7.
30. Seruca R, Santos NR, David L, et al. Sporadic gastric carcinomas with microsatellite instability display a particular clinicopathologic profile. *Int J Cancer* 1995;**64**:32–6.
31. Lee HS, Lee HK, Kim HS, et al. Tumour suppressor gene expression correlates with gastric cancer prognosis. *J Pathol* 2003;**200**:39–46.
32. Akama Y, Yasui W, Yokozaki H, et al. Frequent amplification of the cyclin E gene in human gastric carcinomas. *Jpn J Cancer Res* 1995; **86**:617–21.
33. Vickers AJ, Kattan MW, Daniel S. Method for evaluating prediction models that apply the results of randomized trials to individual patients. *Trials* 2007;**5**(8):14.
34. Song KY, Park YG, Jeon HM, Park CH. A nomogram for predicting individual survival of patients with gastric cancer who underwent radical surgery with extended lymph node dissection. *Gastric Cancer* 2014 Apr;**17**(2):287–93.
35. Strong VE, Song KY, Park CH, et al. Comparison of gastric cancer survival following R0 resection in the United States and Korea using an internationally validated nomogram. *Ann Surg* 2010;**251**:640–6.
36. Koc M, Dizen H, Ozalp N, Keskek M, Karakose N, Tez M. External validation of a US-derived nomogram that predicts individual survival after gastric cancer resection. *Langenbecks Arch Surg* 2009; **394**:755–6.
37. Hirabayashi S, Kosugi S, Isobe Y, et al. Development and external validation of a nomogram for overall survival aftercurative resection in serosa- negative, locally advanced gastric cancer. *Ann Oncol* 2014 Jun;**25**(6):1179–84.
38. Marrelli D, Morgagni P, de Manzoni G, et al. Italian Research Group for Gastric Cancer (IRGGC). Prognostic value of the 7th AJCC/UICC TNM classification of noncardia gastric cancer: analysis of a large series from specialized western centers. *Ann Surg* 2012 Mar;**255**(3): 486–91.
39. Miceli R, Tomasello G, Bregni G, Bartolomeo MD, Pietrantonio F. Adjuvant chemotherapy for gastric cancer: current evidence and future challenges. *World J Gastroenterol* 2014 April 28;**20**(16):4516–25.
40. Paoletti X, Oba K, Burzykowski T, et al. Benefit of adjuvant chemotherapy for resectable gastric cancer: a meta-analysis. *JAMA* 2010; **303**:1729–37. http://dx.doi.org/10.1001/jama.2010.534 PMID: 20442389.
41. Sasako M, Sakuramoto S, Katai H, et al. Five year outcomes of a randomized phase III trial comparing adjuvant chemotherapy with S-1versus surgery alone in stage II or III gastric cancer. *J Clin Oncol* 2011;**29**:4387–93.
42. Noh SH, Park SR, Yang HK, et al. Adjuvant capecitabine and oxaliplatin (xelox) for gastric cancer after D2 gastrectomy: final results from the CLASSIC trial. *Ann Oncol* 2013;**24**:iv14.
43. Marrelli D, Polom K, de Manzoni G, et al. Multimodal treatment of gastric cancer in the west: where are we going? *World J Gastroenterol* 2015 July 14;**21**(26):7954–69.
44. Dikken JL, Coit DG, Baser RE, et al. Performance of a nomogram predicting disease-specific survival after an R0 resection for gastric cancer in patients receiving postoperative chemoradiation therapy. *Int J Radiat Oncol Biol Phys* 2014 Mar 1;**88**(3):624–9.