

Preliminary study of surgical audit for overall survival following gastric cancer resection

Yoshio Haga · Koji Ikejiri · Yasuo Wada ·
Masakazu Ikenaga · Hitoshi Takeuchi

Received: 23 July 2013 / Accepted: 13 January 2014 / Published online: 6 February 2014
© The International Gastric Cancer Association and The Japanese Gastric Cancer Association 2014

Abstract

Background Previous studies for surgical audit have focused on short-term outcomes, such as perioperative mortality. There has been no gold standard how to evaluate quality of care for long-term outcomes in surgical oncology. This preliminary study aims to propose a method for surgical audit targeting long-term outcome following gastrectomy for gastric cancer.

Methods We prospectively investigated a set of variables relating to physiologic conditions, tumor characteristics and operations in patients who underwent gastrectomy for gastric cancer between June 2005 and July 2008 in 18

referral hospitals in Japan. Overall survival (OS) is the endpoint. Cox hazard regression analysis was used to generate a model to predict OS. The calibration and discrimination power of the model were assessed using the Hosmer-Lemeshow (H-L) test and area under the receiver-operating characteristic curve (AUC), respectively. The ratio of observed-to-estimated 5-year OS rates (OE ratio) was defined as a measure of quality.

Results Among 762 patients analyzed, 697 (91 %) completed the 5-year follow-up. The constructed model for OS exhibited a good discrimination power (AUC, 95 % confidence interval 0.89, 0.86–0.91), which was significantly better than that for the UICC stage (0.81, 0.77–0.84). This model also demonstrated a good calibration power (H-L: $\chi^2 = 27.2$, $df = 8$, $P = 0.77$). The OE ratios among the participating hospitals revealed no significant variation between 0.74 and 1.1.

Conclusions The current study suggests the possibility of surgical audit for postoperative OS in gastric cancer. Further studies including high-volume centers will be necessary to validate this idea.

Keywords Stomach · Carcinoma · Surgery · Survival rate · Prediction

Y. Haga (✉)

Institute for Clinical Research, National Hospital
Organization Kumamoto Medical Center, 1-5 Ninomaru,
Kumamoto 8600008, Japan
e-mail: epass2006@ybb.ne.jp

Y. Haga

Department of International Medical Cooperation, Graduate
School of Medical Sciences, Kumamoto University,
Kumamoto, Japan

K. Ikejiri

Department of Surgery, NHO Kyushu Medical Center,
Fukuoka, Japan

Y. Wada

Department of Surgery, NHO Himeji Medical Center,
Himeji, Japan

M. Ikenaga

Department of Surgery, NHO Osaka National Hospital,
Osaka, Japan

H. Takeuchi

Department of Surgery, NHO Iwakuni Clinical Center,
Iwakuni, Japan

Introduction

Surgical audit has become increasingly important in the modern healthcare system. Previous risk models for surgical audit have focused on short-term outcomes, such as postoperative mortality. Copeland et al. [1] from the UK devised a scoring system for a comparative audit of postoperative mortality and morbidity, designated as the Physiological and Operative Severity Score for the

enumeration of Mortality and morbidity (POSSUM). In the USA, the Department of Veterans Affairs established the National Surgical Quality Improvement Program in 1994 [2]. This program was expanded to private sector hospitals in collaboration with the American College of Surgeons [3, 4]. In Japan, we developed a prediction model for the postoperative mortality of general gastrointestinal surgery, designated as Estimation of Physiologic Ability and Surgical Stress (E-PASS) [5–7]. We recently modified this model, termed mE-PASS, to maintain the discrimination power with a reduced number of variables [8]. We also demonstrated its efficacy in a specialty field of surgery for colorectal carcinoma [9], gastric carcinoma [10], liver carcinoma [11], and choledochocystolithiasis [12].

Although these efforts have made significant progress in outcome research, surgical audit for long-term outcome has rarely been investigated in any type of cancer. In this study, we aimed to generate a prediction model for overall survival (OS) following gastrectomy for gastric cancer and propose a method to gauge long-term outcome.

Patients and methods

Study design

This study was conducted as part of the multicenter prospective cohort study ‘Establishment of quality of care among hospitals in digestive surgery’ (UMIN000001410). This cohort study was conducted to investigate the effectiveness of mE-PASS for surgical audit regarding postoperative mortality and to generate a new model to predict long-term survival following surgery for digestive cancers. The results of mE-PASS have been previously reported [8]. This article addresses only the long-term outcome of gastric cancer resection. The protocol required institutional investigators to obtain written informed consent from all participating patients and was approved by the Central Ethics Committee of the National Hospital Organization (NHO), Japan, on 1 November 2004. Eighteen NHO hospitals registered patients for this analysis.

Patients

Eligibility criteria were patients undergoing any type of elective resection for gastric cancer in the operating room. Exclusion criteria were as follows: (1) patients who did not sign the consent forms, (2) those who had concomitant cancer of different organs, (3) those who had a history of cancer in the previous 5 years, (4) those who received concomitant surgery in different surgical fields, and (5) those who received endoscopic resection [8]. Patient entries were made between 1 April 2005 and 8 April 2007.

Surgical procedures

Surgical procedures for gastric cancer have been standardized for a long time in Japan. In guidelines by the Japanese Gastric Cancer Association (JGCA) [13], systematic lymph node dissection, termed D2, has been recommended for most tumors, except for stage IA. Endoscopic resection has been recommended for lesions of differentiated-type adenocarcinoma without ulcerative findings, of which the depth of invasion is clinically diagnosed to be confined to the mucosal layer (cT1a) and a diameter of less than 2 cm. Wedge resection of the stomach is occasionally applied when endoscopic resection fails to obtain a sufficient resection margin against the indicated lesions. D1 gastrectomy has been recommended for other lesions of cT1a and no lymph node metastasis. D1+ gastrectomy has been recommended for a lesion of differentiated adenocarcinoma where the depth of invasion is confined to the submucosal layer and the lesion is less than 1.5 cm in diameter without regional lymph node metastasis. These policies have been widely distributed throughout the country.

Data collection

Data of 10 variables for E-PASS and 20 variables for POSSUM were prospectively collected along with 18 variables of pathological findings and 6 variables regarding surgical procedures. Pathological and surgical findings were recorded according to the JGCA classification [14]. Data for UICC’s M category and N category were also collected, and we analyzed data using the UICC stage [15]. Data were inputted in a database at the website of the Japan Clinical Research Assistant Center (JCRAC), Tokyo, Japan. Postoperative complications were classified according to Clavien’s classification [16]. The primary end point was postoperative OS. Patients were followed every 3–6 months by the attending doctor. To ensure the reliability of these measurements, institutional investigators were required to periodically input survival status as well as confirmed dates for survival or death in the database. Follow-up data entries were discontinued on 16 April 2012 so that the last patient registered could be followed up to 5 years after surgery.

Statistical analysis

All statistical analyses were performed using the SPSS 17.0 (SPSS Inc., Chicago, IL, USA) software program. OS rates were calculated by the Kaplan-Meier method. Univariate analysis for OS was performed using the log-rank test. Two-tailed *P* values less than 0.05 were considered significant. Cox’s proportional hazard analysis was used to

generate a prediction model for OS. To assess proportionality of the model, we plotted $-\log[-\log(-S(t))]$ curves for the independent variables of the newly devised Cox model [17]. As previously reported, model discrimination and calibration were evaluated by the area under the receiver-operating characteristic curve (AUC) and the Hosmer-Lemeshow goodness-of-fit test, respectively [18]. Categorical variables between groups were compared using the chi-square test with Yates' correction for continuity where appropriate. Correlation between interval and interval variables was analyzed by Spearman's rank correlation (ρ). Correlation between different continuous variables was analyzed by Pearson's correlation coefficient (r).

The ratio of observed-to-estimated 5-year OS rates (OE ratio) was used as a metric of the quality of care between hospitals. When an OE ratio of a hospital is greater than 1, the observed survival rate is higher than the expected rate,

Table 1 Demographic data of group A ($n = 762$)

| | |
|---|------------------|
| Age, median (range), years | 69 (33–93) |
| Gender, %male | 70 |
| Surgical procedure, no. (%) | |
| Distal gastrectomy | 469 (62) |
| Total gastrectomy | 222 (29) |
| Total gastrectomy with splenectomy | 33 (4) |
| Cardiofundectomy | 20 (3) |
| Total gastrectomy with distal pancreatectomy | 8 (1) |
| Wedge resection of stomach | 7 (1) |
| Others | 3 (0.3) |
| Lymph node dissection ^a , D0:D1:D2:D3, no. | 19:303:423:17 |
| Histology ^b , no. (%) | |
| Papillary adenocarcinoma | 29 (4) |
| Tubular adenocarcinoma, well-differentiated type | 159 (21) |
| Tubular adenocarcinoma, moderately differentiated type | 211 (28) |
| Poorly differentiated adenocarcinoma, solid type | 98 (13) |
| Poorly differentiated adenocarcinoma, non-solid type | 148 (19) |
| Signet-ring cell carcinoma | 84 (11) |
| Mucinous adenocarcinoma | 21 (3) |
| Others | 12 (2) |
| UICC TNM stage, I:II:III:IV, no. | 375:157:165:65 |
| Percentage of elderly patients | |
| Postoperative complications of Clavien's grading ≥ 3 , no. (%) | 53 (7) |
| Postoperative 30-day mortality, no. (%) | 1 (0.13) |
| Postoperative in-hospital mortality rate, no. (%) | 8 (1) |
| Five-year overall survival rates for stage, I:II:III:IV, % | 87, 70, 27, 18 % |

^{a,b} Lymph node dissection and histology were classified according to the Japanese classification of gastric carcinoma (Ref. [13])

Table 2 Univariate analysis of prognostic factors for overall survival in gastric carcinoma resection

| Factor | No. | 5-Year OS rate (%) | χ^2 value | <i>P</i> value |
|--|-----|--------------------|----------------|----------------|
| Age, years | | | | |
| <70 | 399 | 74 | 47 | <0.001* |
| 70–79 | 265 | 60 | | |
| ≥ 80 | 98 | 40 | | |
| Gender | | | | |
| Female | 228 | 69 | 2 | 0.13 |
| Male | 534 | 63 | | |
| Diabetes mellitus | | | | |
| Absence | 681 | 66 | 4 | 0.058 |
| Presence | 81 | 53 | | |
| Severe heart disease ^a | | | | |
| Absence | 749 | 65 | 6 | 0.012* |
| Presence | 13 | 32 | | |
| Severe pulmonary disease ^a | | | | |
| Absence | 738 | 66 | 15 | <0.001* |
| Presence | 24 | 33 | | |
| Glasgow Coma Scale | | | | |
| 15 | 733 | 66 | 7.3 | 0.0070* |
| <15 | 29 | 44 | | |
| Respiratory history ^b | | | | |
| No dyspnea | 669 | 68 | 50 | <0.001* |
| Dyspnea on exertion | 84 | 42 | | |
| Limiting dyspnea | 6 | 33 | | |
| Dyspnea at rest | 3 | 0 | | |
| Cardiac signs ^b | | | | |
| No failure | 624 | 66 | 3 | 0.29 |
| Diuretic, digoxin, antianginal or hypertensive therapy | 128 | 58 | | |
| Peripheral edema; warfarin therapy | 10 | 60 | | |
| Performance status | | | | |
| 0 | 464 | 76 | 140 | <0.001* |
| 1 | 231 | 53 | | |
| 2 | 53 | 29 | | |
| 3 | 11 | 0 | | |
| 4 | 3 | 0 | | |
| ASA class | | | | |
| 1 | 227 | 79 | 70 | <0.001* |
| 2 | 427 | 64 | | |
| 3 | 107 | 39 | | |
| 4 | 1 | 0 | | |
| Chest radiograph, lung ^b | | | | |
| Normal | 696 | 67 | 40 | <0.001* |
| Mild COAD | 52 | 46 | | |
| Moderate COAD | 10 | 20 | | |
| Fibrosis or consolidation | 4 | 0 | | |
| Chest radiograph, cardiomegaly ^b | | | | |

Table 2 continued

| Factor | No. | 5-Year OS rate (%) | χ^2 value | <i>P</i> value |
|--|-----|--------------------|----------------|----------------|
| Normal | 689 | 67 | 37 | <0.001* |
| Borderline cardiomegaly | 71 | 49 | | |
| Cardiomegaly | 2 | 0 | | |
| Systolic blood pressure, mmHg ^b | | | | |
| 110–130 | 359 | 66 | 6 | 0.094 |
| 131–170 or 100–109 | 380 | 65 | | |
| >170 or 90–99 | 17 | 41 | | |
| <90 | 6 | 67 | | |
| Pulse (beats/min) ^b | | | 5 | 0.19 |
| 50–80 | 655 | 66 | | |
| 81–100 or 40–49 | 104 | 56 | | |
| 101–120 | 2 | 50 | | |
| <40 or >120 | 1 | 0 | | |
| White cell counts, $\times 1.0^{12}/l^b$ | | | 1.5 | 0.47 |
| 4–10 | 684 | 65 | | |
| 10.1–20 or 3.1–4 | 67 | 60 | | |
| >20 or ≤ 3.0 | 10 | 70 | | |
| Hemoglobin, g/dl ^b | | | 115 | <0.001* |
| 13–16 | 352 | 79 | | |
| 11.5–12.9 or 16.1–17 | 190 | 64 | | |
| 10–11.4 or 17.1–18 | 122 | 51 | | |
| >18 | 97 | 31 | | |
| Urea, mg/dl ^b | | | 5 | 0.177 |
| ≤ 20 | 672 | 66 | | |
| 20.1–27 | 73 | 52 | | |
| 27.1–40 | 9 | 56 | | |
| >40 | 7 | 43 | | |
| Albumin, g/dl | | | 43 | <0.001* |
| ≥ 3.0 | 706 | 67 | | |
| <3.0 | 51 | 30 | | |
| Serum sodium level, mmol/l ^b | | | 127 | <0.001* |
| ≥ 136 | 731 | 67 | | |
| 131–135 | 23 | 4 | | |
| 126–130 | 7 | 0 | | |
| Serum potassium level, mmol/l ^b | | | 9 | 0.033* |
| 3.5–5.0 | 729 | 66 | | |
| 3.2–3.4 or 5.1–5.3 | 15 | 68 | | |
| 2.9–3.1 or 5.4–5.9 | 15 | 18 | | |
| ≤ 2.8 or ≥ 6.0 | 2 | 50 | | |
| Depth of invasion ^c | | | 280 | <0.001* |
| T1 (M or SM) | 358 | 89 | | |
| T2 (MP) | 234 | 57 | | |
| T3 (SS) | 142 | 27 | | |
| T4 (SE or SI) | 28 | 12 | | |
| Distant metastasis ^c | | | 131 | <0.001* |
| Absence | 697 | 69 | | |
| Presence | 95 | 18 | | |

Table 2 continued

| Factor | No. | 5-Year OS rate (%) | χ^2 value | <i>P</i> value |
|--|-----|--------------------|----------------|----------------|
| No. of metastatic lymph nodes ^c | | | 278 | <0.001* |
| 0 | 425 | 85 | | |
| 1–2 | 107 | 62 | | |
| 3–6 | 94 | 37 | | |
| >6 | 136 | 25 | | |
| Lymphatic invasion ^d | | | 171 | <0.001* |
| No invasion | 299 | 88 | | |
| Minimal invasion | 184 | 61 | | |
| Moderate invasion | 169 | 51 | | |
| Marked invasion | 107 | 28 | | |
| Venous invasion ^d | | | 114 | <0.001* |
| No invasion | 458 | 78 | | |
| Minimal invasion | 153 | 50 | | |
| Moderate invasion | 106 | 41 | | |
| Marked invasion | 41 | 27 | | |
| Macroscopic type ^d | | | 176 | <0.001* |
| Non-invasive | 519 | 79 | | |
| Invasive ^e | 243 | 35 | | |
| Location ^d | | | 54 | <0.001* |
| Lower | 270 | 66 | | |
| Middle | 214 | 77 | | |
| Upper | 157 | 63 | | |
| ≥ 2 regions | 121 | 43 | | |
| Circumferential involvement ^d | | | 200 | <0.001* |
| Absence | 651 | 73 | | |
| Presence | 111 | 16 | | |
| Histology ^d | | | 4 | 0.042* |
| Non-diffuse type | 530 | 67 | | |
| Diffuse type ^f | 232 | 60 | | |

OS overall survival, ASA American Society of Anesthesiologists, COAD chronic obstructive airways disease

^a The definitions are based on the E-PASS scoring system as previously described (see Ref. [5])

^b The definitions are based on the POSSUM scoring system as previously described (see Ref. [1])

^c The definitions are based on the UICC TMN staging system as previously described (see Ref. [15])

^d The definitions are based on Japanese classification of gastric carcinoma (JCGA) as previously described (see Ref. [13])

^e Invasive types include type 3 (infiltrative ulcerative), type 4 (diffuse infiltrative), and some forms of type 5 (unclassifiable), such as the so-called type IIc-like advanced cancer

^f Diffuse type includes the non-solid type of poorly differentiated adenocarcinoma and signet-ring cell carcinoma in the JCGA

P values were calculated using log-rank test. * *P* < 0.05

indicating that the quality of care is better than expected. In contrast, an OE ratio smaller than 1 indicates a poorer performance. For the comparison of OE ratios between

Table 3 Equation to predict postoperative overall survival in gastric carcinoma

$$S(t, x) = \{e^{-0.0654t}\}^{\exp(R - 3.6)}$$

t : postoperative observation period (years)

$$R = 0.59X_1 + 0.44X_2 + 0.61X_3 + 0.43X_4 + 0.50X_5 + 0.41X_6$$

X_1 : Depth of invasion

- 1 for T1 (mucosa or submucosa); 2 for T2 (muscularis propria); 3 for T3 (subserosa);
4 for T4 (exposed beyond the serosa or invasion to adjacent organs)

X_2 : Nodal status

- 0 for no metastasis; 1 for N1 (1–2 regional nodes metastasis);
2 for N2 (3–6 regional nodes metastasis); 3 for N3 (≥ 7 regional nodes metastasis)

X_3 : Circumferential involvement^a (0 for absence; 1 for presence)

X_4 : Age score

- 1 for <70 years; 2 for 70–79 years; 3 for ≥ 80 years

X_5 : American Society of Anesthesiologists physical status classification (1–4)

X_6 : Serum sodium level

- 1 for >135 mmol/l; 2 for 131–135 mmol/l; 4 for 126–130 mmol/l
8 for < 126 mmol/l

^a Circumferential involvement indicates carcinoma involvement of all of four parts of the lesser and greater curvatures and the anterior and posterior walls (see Ref. [15])

hospitals, a sample size determination was made based on the hypothesis that a difference between the 5-year OS rate of 72 %, the baseline rate of our subjects, and that of 42 % was significant. If we determined the α error to be 0.05 and the β error to be 0.2, the sample size needed was 42. Therefore, we compared OE ratios between hospitals that registered more than 42 patients.

Results

Of the 796 patients enrolled, three were found not to have undergone resection and were therefore excluded from the analysis. Of the remaining 793 patients, 31 lacked enough data for the analysis. Accordingly, we analyzed 762 patients (group A) to develop a model to predict overall survival. Among them, 697 patients (91 %) completed the 5-year follow-up (group B). The reason for failure to follow up was disconnection due to a patient moving, transfer to another hospital, or disappearance without notice. We analyzed the accuracy of the model using patients with full follow-up data. Table 1 shows the demographic data of group A. Among the patients, 98 patients (13 %) were 80 years or older. Median values for numbers of dissected lymph nodes were 4 for D0, 19 for D1, 30 for D2, and 44 for D3.

Univariate analysis for postoperative OS identified 13 physiologic- and 9 tumor-related variables as significant from the 29 variables (Table 2). Using these significant variables, we performed a stepwise increase method of the Cox regression analysis to obtain a prediction model for OS (Table 3). We designated this model as the Estimation of Postoperative Overall Survival for Gastric Cancer (EPOS-GC). To assess whether the adequate variables were incorporated into the model, we checked the proportionality of the independent variables (Fig. 1). The curves did not cross each other and remained parallel in all the variables.

The EPOS-GC yielded the high χ^2 value of 519 in the Cox hazard analysis, which was comparable to that of the model where all 22 significant variables were incorporated (χ^2 value 563). Furthermore, the predicted 5-year survival rates of the EPOS-GC were significantly correlated with those of the model utilizing all 22 significant variables ($\rho = 0.93$, $n = 697$, $P < 0.0001$). As shown in Fig. 2, the EPOS-GC demonstrated good discrimination power as the AUC (95 % CI) was 0.89 (0.86–0.91), which was significantly higher than that for the UICC TNM stage of 0.81 (0.77–0.84) and that for the residual disease status of 0.60 (0.56–0.65). Furthermore, the AUC (95 % CI) for the model utilizing all 22 significant variables was 0.89 (0.86–0.91), being the same as that of the EPOS-GC. The calibration power of the EPOS-GC was also good as judged by the Hosmer-Lemeshow test ($\chi^2 = 23.5$, $df = 8$, $P = 0.77$).

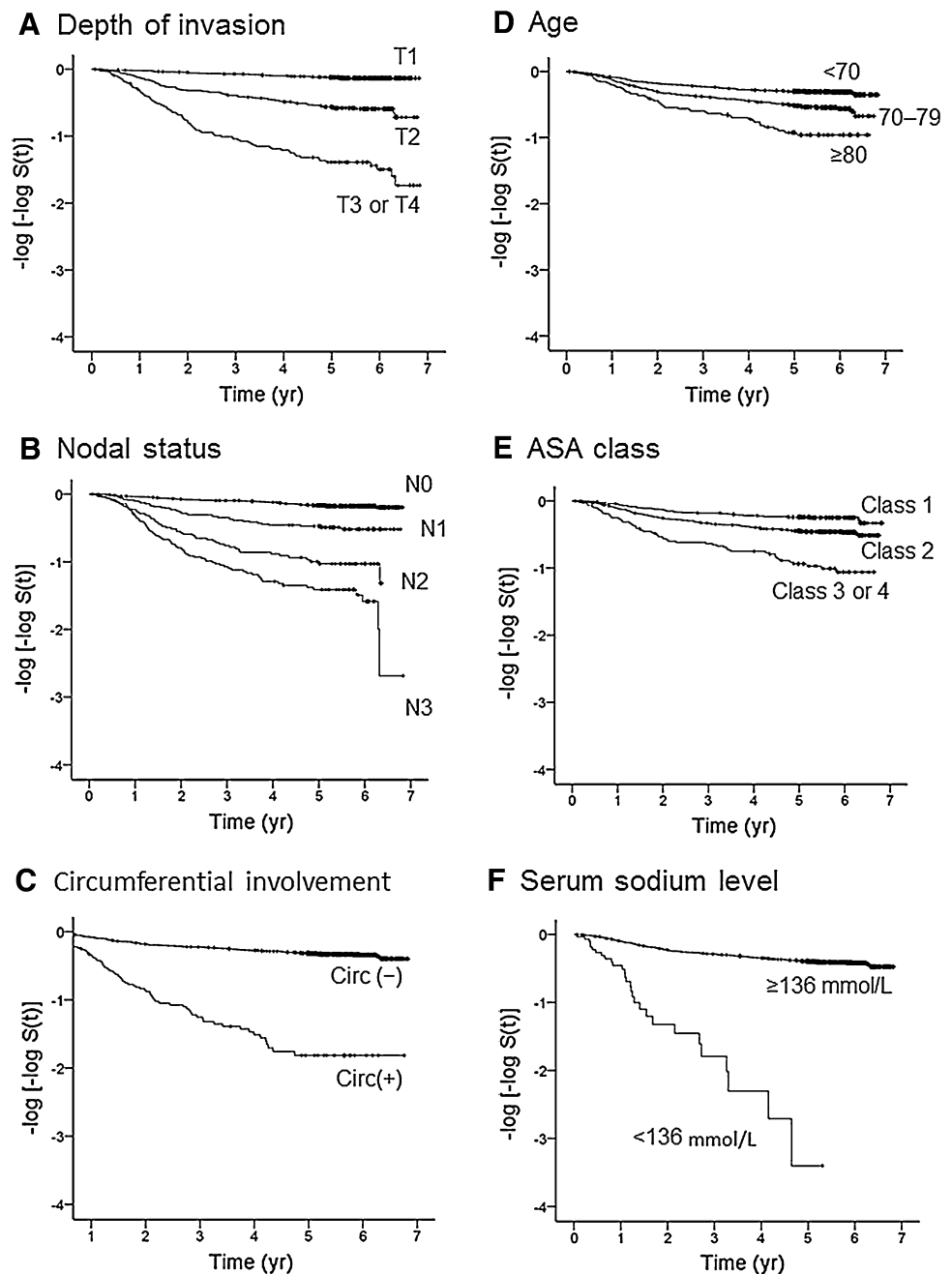
Subsequently, we investigated the effects of predicted OS rates on survival in each TNM stage (Fig. 3). In stage I patients, the observed 5-year OS rates were 0 % at the predicted OS rates (Y) < 0.5 ($n = 3$). The observed rates increased to 38 % at $0.5 \leq Y < 0.7$ ($n = 16$) and 89 % at $Y \geq 0.7$ ($n = 311$). All patients with $Y < 0.7$ ($n = 19$) were 70 years or older, 13 of whom (68 %) were 80 years or older. In stage IV patients, the observed 5-year OS rates were only 7.7 % at $Y < 0.5$ ($n = 52$), but it increased markedly when the predicted rates (Y) increased. The patients at $0.5 \leq Y < 0.7$ ($n = 8$) included five patients of 70 years or older, but all the patients at $Y \geq 0.7$ ($n = 3$) were under 70 years old.

Subsequently, we determined the OE ratios of the participating hospitals and found that they did not vary significantly between hospitals (Table 4). When we checked the OE ratios of the model utilizing all the significant variables, they were significantly correlated with those of the EPOS-GC ($r = 0.79$, $n = 10$, $P = 0.0061$).

Discussion

In this study, we generated a model to predict postoperative OS using Cox hazard analysis in light of physiologic status

Fig. 1 Assessing proportionality of EPOS-GC variables. In the case that the number of cases in a category was ≤ 50 , the category was merged into the next category. For example, in the depth of invasion, the T4 category ($n = 28$) was merged into the T3 category ($n = 142$). $S(t)$ survival function, *Circ* circumferential involvement



and tumor characteristics, and compared the quality of care between hospitals. Although the findings were not validated in an external data set, this report may lead to staging a risk-adjusted surgical audit for long-term outcome in gastric carcinoma surgery.

Previous studies for surgical audit have been focused on short-term outcome, such as perioperative mortality. Nevertheless, some concerns would surface among surgical oncologists if solely a comparative audit regarding perioperative care was performed. Hospitals that tend to perform operations with fewer nodal dissections would get higher rankings. Therefore, surgical audit regarding long-term

survival should be performed simultaneously in order to confer decent evaluation to the referral hospitals. When evaluating long-term survival, we must keep in mind the patients' physiological conditions. As an unprecedented aging society has come to developed countries, how to deal with the growing number of individuals with multiple chronic conditions (MCC) is a major task for the modern healthcare system [19]. These people sometimes go back and forth between long-term care facilities and hospitals. It is often difficult to predict what disease will terminate the patient's life. Therefore, it is inappropriate to compare stage-stratified OS rates between hospitals as a metric of

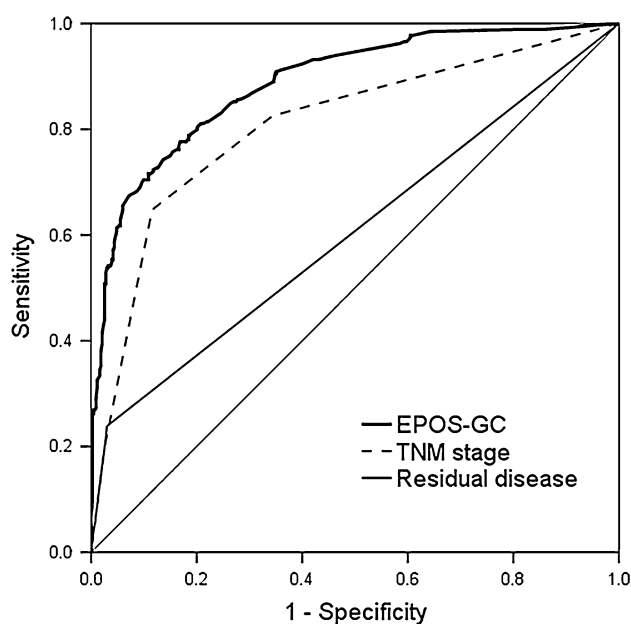


Fig. 2 Receiver-operating characteristic curve analysis to predict 5-year overall survival. The 5-year death rate of the current model, the EPOS-GC, was obtained by (1—the predicted 5-year survival rate) and was used for this analysis. UICC definitions were used for the TNM stage and residual disease status (Ref. [14]). *EPOS-GC* Estimation of Postoperative Overall Survival for Gastric Cancer

the quality of care. Hospitals treating higher rates of MCC patients will get lower rankings. In contrast, hospitals selecting non-MCC patients will get high rankings. Therefore, risk adjustment regarding the physiological status and tumor characteristics is essential for surgical audit. Costa et al. [20] constructed a prognostic score for gastric cancer using clinical, pathological, and therapeutic variables in order to obtain a comprehensive prognostic parameter. This score, which consists of six variables, revealed a better predicting power for OS than the TNM stage. In this study, we proceeded with this idea and obtained the equation for predicted OS rates using a Cox analysis. The EPOS-GC allows us to calculate the OE ratios between centers, leading to the comparative audit. Of course, it is conceivable that all potential confounding factors should be incorporated for the surgical audit. Nevertheless, the EPOS-GC gained the same AUC value as the model in which all the significant variables were incorporated. Furthermore, the predicted 5-year survival rates of EPOS-GC highly correlated with those of the model utilizing all the significant variables. Accordingly, the OE ratios of the model utilizing all the significant variables significantly correlated with those of the EPOS-GC. Therefore, we considered that the EPOS-GC can be used for surgical audit instead of the model utilizing all the significant variables. The advantage to using the stepwise model is that a reduced number of variables would be

required in a future audit survey if the model is universally validated.

Stage-stratified 5-year OS rates in this study were lower than those in the previous nationwide survey. The 5-year OS in stage III patients was 27 % in this study, while the national registry in 2002 reported it as 41 % [21]. These results may be, at least in part, attributable to the higher rates of elderly people in this study; this study included 13 % elderly patients aged 80 years or older, but the previous survey included only 8 % of this group. Actually, stage III patients in the current study included 15 % (25/165) of patients aged 80 years or older.

In the current model, the depth of invasion and nodal status of the TNM stage were identified as expected as an independent variable, but distant metastasis was not chosen by the Cox regression analysis. This may be because distant metastasis is strongly associated with the depth of invasion and nodal status; therefore, no significant contribution augmenting the accuracy of the model would be added by this factor. Instead, circumferential involvement was identified as an independent variable, which may be unassociated with the depth of invasion or nodal status. Most cases of linitis plastica were included in this category. For physiological status, age and ASA class were identified as independent variables as anticipated. Surprisingly, serum sodium levels were also selected by this analysis. Since many critical conditions are associated with hyponatremia, including heart failure, liver cirrhosis, and renal failure, this variable may augment the predictive power in another direction.

In the process of constructing the EPOS-GC, we did not include variables relating to surgery, such as blood loss during the operation or the level of nodal dissection. This is because the purpose of the model was surgical audit. We consider that inclusion of surgery-related variables might interfere with the results of OE ratios. On the other hand, we wish to know whether the inclusion of surgery-related variables would raise the predictive power for OS. Special attention was paid to the ratio of metastatic lymph nodes (RML), since previous studies demonstrated that RML was superior in predicting OS than the number of metastatic nodes [22, 23]. Consequently, inclusion of surgery-related variables including RML did not augment the AUC values as compared with the current model. Therefore, we understand that there is no merit to adding surgery-related variables.

The current study did not demonstrate a significant difference of the OE ratios between hospitals. The reason may be explained by the permeation of standardized surgical procedures for gastric cancer. Since gastric cancer has long been one of the most frequent subjects for surgical oncology in Japan, most surgeons have obtained enough of a caseload in their carriers. Therefore, there may be no significant difference between the hospitals.

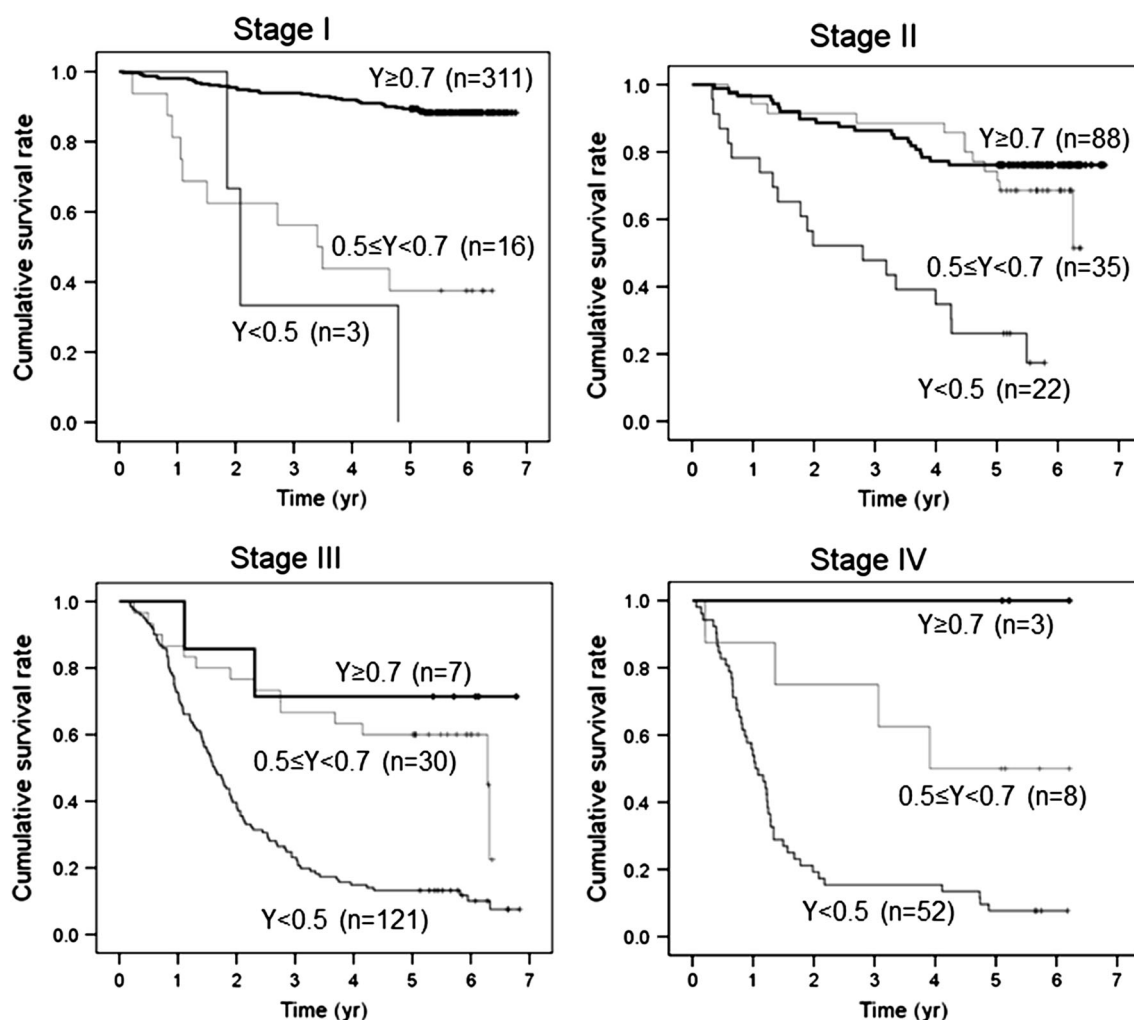


Fig. 3 Effects of predicted overall survival rates on survival in each stage. Overall survival (OS) curve was plotted according to the predicted OS rates (Y) determined by the EPOS-GC in each stage

Table 4 Comparative audit of postoperative overall survival among centers

| Hospital | No. of patients | Observed-5-year survivors | Estimated-5-year survivors | OE ratio (95 %CI) | <i>P</i> value |
|----------|-----------------|---------------------------|----------------------------|-------------------|----------------|
| A | 50 | 29 | 32 | 0.91 (0.66–1.2) | 0.54 |
| B | 42 | 17 | 23 | 0.74 (0.47–1.1) | 0.19 |
| C | 48 | 29 | 28 | 1.0 (0.74–1.4) | 0.98 |
| D | 72 | 49 | 49 | 1.0 (0.8–1.3) | 1.0 |
| E | 76 | 46 | 48 | 0.96 (0.75–1.2) | 0.74 |
| G | 53 | 34 | 34 | 1.0 (0.75–1.3) | 1.0 |
| I | 165 | 99 | 111 | 0.89 (0.76–1.1) | 0.17 |
| J | 61 | 48 | 44 | 1.1 (0.89–1.3) | 0.40 |

OE ratio observed-to-estimated survival rates ratio

P values were calculated by the chi-square test with Yates' correction for continuity where appropriate

The main limitation of this study is that participating hospitals did not include high-volume hospitals for gastric cancer surgery, since this study was conducted within the NHO hospitals. Furthermore, the study

protocol required informed consent from the patients; therefore, all the patients who qualified were not incorporated into this study. Validation studies will be necessary including high-volume centers with a complete

enrollment of the qualified patients. Another limitation was that this study was done only in Japan. Since D2 gastrectomy has now prevailed throughout the world [24–28], validation studies are presumable in other countries.

In conclusion, this preliminary study suggests the possibility of surgical audit for long-term outcome. If validated in the future, the current methodology would open the door to a new field of outcome research.

Acknowledgments This work was supported by a grant for NHO Multi-Center Clinical Research for Evidence-Based Medicine.

The authors declared no conflicts of interests.

The authors wish to thank all of the institutional investigators listed below for their unstinting exertions in collecting data.

H. Naito, MD, NHO Hokkaido Cancer Center; M. Oohara, MD, NHO Hokkaido Cardiovascular Center; A. Nagase, MD, NHO Dokusoku Hospital; T. Kageyama, MD, NHO Kasumigaura Medical Center; M. Nakamura, MD, NHO Takasaki General Medical Center; H. Tsunoda, MD, NHO Kofu National Hospital; S. Koike, MD, NHO Matsumoto Medical Center; K. Kondo, MD, NHO Nagoya Medical Center; O. Kimura, MD, NHO Yonago Medical Center; M. Koseki, MD, NHO Kure Medical Center; T. Furuya, MD, NHO Kanmon Medical Center; T. Nakata, MD, NHO Nagasaki Medical Center; H. Matsuzaki, MD, NHO Kumamoto Medical Center; Y. Ogura, MD, NHO Minami Kyushu National Hospital.

References

- Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. *Br J Surg*. 1991;78:355–60.
- Khuri SF, Daley J, Henderson W, Hur K, Demakis J, Aust JB, et al. The Department of Veterans Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. *Ann Surg*. 1998;228:491–507.
- Fink AS, Campbell DA, Mentzer RM, Henderson WG, Daley J, Bannister J, et al. The National Surgical Quality Improvement Program in non-VA hospitals: initial demonstration of feasibility. *Ann Surg*. 2002;236:344–53.
- Khuri SF, Henderson WG, Daley J, Jonasson O, Jones RS, Campbell DA Jr, et al. Successful Implementation of the Department of Veterans Affairs' National Surgical Quality Improvement Program in the Private Sector: the Patient Safety in Surgery Study. *Ann Surg*. 2008;248:329–36.
- Haga Y, Ikei S, Ogawa M. Estimation of Physiologic Ability and Surgical Stress (E-PASS) as a new prediction scoring system for postoperative morbidity and mortality following GI surgery. *Surg Today*. 1999;29:219–25.
- Haga Y, Wada Y, Takeuchi H, Sameshima H, Kimura O, Furuya T. Estimation of surgical costs using a prediction scoring system Estimation of Physiologic Ability and Surgical Stress. *Arch Surg*. 2002;137:481–5.
- Haga Y, Wada Y, Takeuchi H, et al. Estimation of physiologic ability and surgical stress (E-PASS) for a surgical audit in elective digestive surgery. *Surgery*. 2004;135:586–94.
- Haga Y, Ikejiri K, Wada Y, Takahashi T, Ikenaga M, Akiyama N, et al. A multicenter prospective study of surgical audit systems. *Ann Surg*. 2011;253:194–201.
- Haga Y, Wada Y, Ikenaga M, Takeuchi H, Ikejiri K. Evaluation of modified Estimation of Physiologic Ability and Surgical Stress (mE-PSS) in colorectal carcinoma surgery. *Dis Colon Rectum*. 2011;54:1293–300.
- Haga Y, Wada Y, Takeuchi H, Ikejiri K, Ikenaga M, Kimura O. Evaluation of modified Estimation of Physiologic Ability and Surgical Stress in gastric carcinoma surgery. *Gastric Cancer*. 2012;15:7–14.
- Haga Y, Ikejiri K, Takeuchi H, Ikenaga M, Wada Y. Value of general surgical risk models for predicting postoperative liver failure and mortality following liver surgery. *J Surg Oncol*. 2012;106:898–904.
- Haga Y, Wada Y, Takeuchi H, Furuya T. Evaluation of modified Estimation of Physiologic Ability and Surgical Stress in patients undergoing surgery for choledochocystolithiasis. *World J Surg*. 2013;. doi:10.1007/s00268-013-2383-0.
- Association Japanese Gastric Cancer. Japanese gastric cancer treatment guidelines 2010 (ver. 3). *Gastric Cancer*. 2011;14:113–23.
- Association Japanese Gastric Cancer. Japanese classification of gastric carcinoma—2nd English edition. *Gastric Cancer*. 1998;1:10–24.
- Sobin LH, Gospodarowicz MK, Wittekind C. TNM classification of malignant tumours. 7th ed. Hoboken: Wiley-Blackwell; 2009.
- Clavien PA, Barkun J, de Oliveira ML, Vauthey JN, Dindo D, Schulick RD, et al. The Clavien–Dindo classification of surgical complications: five-year experience. *Ann Surg*. 2009;250:187–96.
- Rees M, Tekkis PP, Welsh FK, O'Rourke T, John TG. Evaluation of long-term survival after hepatic resection for metastatic colorectal cancer: a multifactorial model of 929 patients. *Ann Surg*. 2008;247:125–35.
- Nakamura T. Cox proportional hazard model. In: Tango T, editor. 6th ed. Tokyo: Asakura Publishing Co., Ltd. 2010 (in Japanese).
- Parekh AK, Goodman RA, Gordon C, Koh HK. Managing multiple chronic conditions: a strategic framework for improving health outcomes and quality of life. *Public Health Rep*. 2011;126:460–71.
- Costa ML, de Cássia Braga Ribeiro K, Machado MA, Costa AC, Montagnini AL. Prognostic score in gastric cancer: the importance of a conjoint analysis of clinical, pathologic, and therapeutic factors. *Ann Surg Oncol*. 2006;13:843–50.
- Nashimoto A, Akazawa K, Isobe Y, Miyashiro I, Katai H, Kodera Y, et al. Gastric cancer treated in 2002 in Japan: 2009 annual report of the JGCA nationwide registry. *Gastric Cancer*. 2013;16:1–27.
- Bando E, Yonemura Y, Taniguchi K, Fushida S, Fujimura T, Miwa K. Outcome of ratio of lymph node metastasis in gastric carcinoma. *Ann Surg Oncol*. 2002;9:775–84.
- Deng J, Liang H, Wang D, Sun D, Ding X, Pan Y, et al. Enhancement the prediction of postoperative survival in gastric cancer by combining the negative lymph node count with ratio between positive and examined lymph nodes. *Ann Surg Oncol*. 2010;17:1043–51.
- Sierra A, Regueira FM, Hernández-Lizoáin JL, Pardo F, Martínez-González MA, A-Cienfuegos J. Role of the extended lymphadenectomy in gastric cancer surgery: experience in a single institution. *Ann Surg Oncol*. 2003;10:219–26.
- Lamb P, Sivashanmugam T, White M, Irving M, Wayman J, Raimes S. Gastric cancer surgery—a balance of risk and radicality. *Ann R Coll Surg Engl*. 2008;90:235–42.
- Deguli M, Sasako M, Ponti A, Italian Gastric Cancer Study Group. Morbidity and mortality in the Italian Gastric Cancer Study Group randomized clinical trial of D1 versus D2 resection for gastric cancer. *Br J Surg*. 2010;97:643–9.
- Ott K, Lordick F, Blank S, Büchler M. Gastric cancer: surgery in 2011. *Langenbecks Arch Surg*. 2011;396:743–58.
- Hanna GB, Boshier PR, Knaggs A, Goldin R, Sasako M. Improving outcomes after gastroesophageal cancer resection: can Japanese results be reproduced in Western centers? *Arch Surg*. 2012;147:738–45.