Retrospective Cohort Study

INTERNATIONAL JOURNAL OF SURGERY

OPEN

An international multi-institution real-world study of the optimal surveillance frequency for stage II/III gastric cancer: the more, the better?

Dong Wu, MD[a,b], Jun Lu, MD, PhD[a,b], Jia Lin, MD[a,b], Bin-bin Xu, MD[a,b], Zhen Xue, MD[a,b], Hua-Long Zheng, MD[a,b], Guo-sheng Lin, MD[a,b], Jiao-bao Huang, MD[a,b], Li-li Shen, MD[a,b], Chao-Hui Zheng, MD, PhD[a,b], Ping Li, MD, PhD[a,b], Jian-Wei Xie, MD, PhD[a,b], Jia-Bin Wang, MD, PhD[a,b], Jian-Xian Lin, MD, PhD[a,b], Qi-Yue Chen, MD, PhD[a,b], Long-Long Cao, MD, PhD[a,b], Yu-Bin Ma, MD[c,*], Mark J. Truty, MD, MS[d,*], Chang-Ming Huang, MD[a,b,*]

Background: Due to lacking evidence on surveillance for gastric cancer (GC), this study aimed to determine the optimal postsurgical surveillance strategy for pathological stage (pStage) II/III GC patients and compare its cost-effectiveness with traditional surveillance strategies.

Methods: Prospectively collected data from stage II/III GC patients ($n = 1661$) who underwent upfront surgery at a large-volume tertiary cancer center in China (FJMUUH cohort) between January 2010 and October 2015. For external validation, two independent cohorts were included, which were composed of 380 stage II/III GC patients at an tertiary cancer center in U.S.A (Mayo cohort) between July 1991 and July 2012 and 270 stage II/III GC patients at another tertiary cancer center in China (QUAH cohort) between May 2010 and October 2014. Random forest models were used to predict dynamic recurrence hazards and to construct individual surveillance strategies for stage II/III GC. Cost-effectiveness was assessed by the Markov model.

Results: The median follow-up period of the FJMUUH, the Mayo, and QUAH cohorts were 55, 158, and 70 months, respectively. In the FJMUUH cohort, the 5-year recurrence risk was higher in pStage III compared with pStage II GC patients ($P < 0.001$). Our novel individual surveillance strategy achieved optimal cost-effectiveness for pStage II GC patients (ICER = \$490/QALY). The most intensive NCCN surveillance guideline was more cost-effective (ICER = \$983/QALY) for pStage III GC patients. The external validations confirmed our results.

Conclusion: For patients with pStage II GC, individualized risk-based surveillance outperformed the JGCTG and NCCN surveillance guidelines. However, the NCCN surveillance guideline may be more suitable for patients with pStage III GC. Even though our results are limited by the retrospective study design, the authors believe that our findings should be considered when recommending postoperative surveillance for stage II/III GC with upfront surgery in the absence of a randomized clinical trial.

Keywords: cost-effectiveness, gastric cancer, real-world study, surveillance

Introduction

Gastric cancer (GC) ranks fifth for incidence and fourth for mortality globally. GC has a high recurrence rate despite the recent advances in chemotherapy and molecular targeted therapy[1–3].

The current National Comprehensive Cancer Network (NCCN) guidelines and Japanese gastric cancer treatment guidelines (JGCTG) recommend routine follow-up to detect recurrence on time[4,5]. However, the development of existing follow-up strategies was mainly based on experts' consensus and researches on recurrence patterns due to the lack of direct evidence and unified standards. On the one hand, it is unreasonable to carry out the same postsurgical surveillance for all pathological stage (pStage) II/III GC patients. On the other hand, unnecessary

[a]Department of Gastric Surgery, [b]Department of General Surgery, Fujian Medical University Union Hospital, Fuzhou, [c]Department of Gastrointestinal Surgery, The Affiliated Hospital of Qinghai University, Xining, People's Republic of China and [d]Department of Surgery, Mayo Clinic, Rochester, USA

Wu D., Lu J., and Lin J. contributed equally to this work and should be considered co-first author.

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

*Corresponding authors. Address: Department of Gastric Surgery, Fujian Medical University Union Hospital, No. 29 Xinquan Road, Fuzhou, 350001, Fujian Province, People's Republic of China. Tel.: +865 918 336 3366, fax: +865 918 332 0319. E-mail: hcmlr2002@163.com (C.-M. Huang); Department of Surgery, Mayo Clinic, 200 First St. Southwest, Rochester, MN 55905, USA. Tel.: +507 718 4954. E-mail: truty.mark@mayo.edu (M.J. Truty); Department of Gastrointestinal Surgery, The Affiliated Hospital of Qinghai University, No. 29 Tongren Road, Xining 810000, Qinghai Province, People's Republic of China. Tel.: +138 972 83522. E-mail: 1026412012@qq.com (Y.-B. Ma).

International Journal of Surgery (2023) 109:4101–4112

Received 3 September 2022; Accepted 21 August 2023

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.lww.com/international-journal-of-surgery.

Published online 5 October 2023

http://dx.doi.org/10.1097/JS9.0000000000000731

surveillance also increases the economic burden on low-risk patients. Therefore, the development of individualized postoperative follow-up strategies for patients with pStage II/III GC is of great clinical significance and economic value.

In addition to the early detection of recurrence, the optimal follow-up strategy should also balance cost and effectiveness. However, there are no criteria for the arrangement of follow-up in different guidelines, and the optimal postoperative follow-up strategy for patients with pStage II/III GC remains to be determined. Therefore, our international multi-institution study first adopted a random survival forest (RSF) and Markov model to evaluate the dynamic risk of recurrence and the cost-effectiveness of postoperative follow-up. We aimed to establish optimal individualized surveillance strategies for patients with pStage II and pStage III GC, providing a theoretical basis for follow-up.

HIGHLIGHTS

- Our international multi-institution study first adopted a random survival forest model to predict the dynamic risk of recurrence and evaluated cost-effectiveness among different surveillance strategies through the Markov model.
- For patients with pStage II GC, individualized risk-based surveillance (12 visits) outperformed the JGCTG and NCCN surveillance guidelines.
- For patients with pStage III GC, the NCCN surveillance guideline maybe more suitable.

Methods

Based on a prospective GC database, 2417 patients with primary gastric adenocarcinoma who underwent curative gastrectomy at FJMUUH between January 2010 and October 2015 were identified.

The inclusion criteria of the FJMUUH cohort were as follows: (1) Eastern Cooperative Oncology Group (ECOG) scores of 0 (asymptomatic) or 1 (symptomatic but completely ambulatory); (2) pathologically confirmed AGC (pStage II and III, except pT4b); (3) no distant metastasis or invasion of adjacent organs (e.g. pancreas, spleen, liver, and colon) detected intraoperatively or postoperatively; and (4) D2 lymph node dissection of GC.

The exclusion criteria of the FJMUUH cohort were as follows: (1) American Society of Anesthesiologists (ASA) grade >2; (2) remnant gastric or neuroendocrine cancer; (3) history of neoadjuvant chemotherapy; (4) palliative surgery; and (5) death within 30 days after surgery.

Ultimately, 1661 patients were included in the FJMUUH cohort. Moreover, we included a western cohort (Mayo cohort) of 380 patients diagnosed between July 1991 and July 2012 and a Chinese cohort (QUAH cohort) of 270 patients who underwent gastrectomy between May 2010 and October 2014 (eFig 1, Supplemental Digital Content 3, http://links.lww.com/JS9/B131) as external validation cohorts. All surgeons had rich experience in performing curative gastrectomy (more than 100 operations annually). This study passed the Institutional Review Board (IRB number: 2022KY106) and followed the STROCSS criteria and guidelines[6]. Due to the retrospective and observational design, the IRB waived the need for informed consent for this study. The study was registered at Clinical Trials.gov.

Cancer staging

GC was defined according to the American Joint Committee on Cancer (AJCC) pathological stage (pStage) II to III, excluding T4b tumors (stage II: T2N1-2M0, T3N0-1M0, and T4aN0M0; stage III: T2N3M0, T3N2-3M0, and T4aN1-3M0). Consistent with previous studies[7,8], to avoid the misjudgment of recurrence, patients with distant metastasis (M1) were excluded in this study. The Japanese Classification was used to categorize tumors as localized, locally advanced, or metastatic[9]. Further categorization into staging subgroups was done based on the AJCC Cancer Staging Manual, 8th edition[10].

Definition

Recurrence-free survival (RFS) was defined as the period from surgery to the first detection of recurrence. Recurrences were categorized according to the involved site into locoregional, peritoneal, distant, or multiple[7] (Specific definitions were seen in Supplemental File 1 named 'Supplementary of Definition and Statistical method', Supplemental Digital Content 1, http://links.lww.com/JS9/B129). All recurrences were documented by pathological diagnosis, radiological imaging, or both.

Follow-up schedule

The median follow-up times of the training cohort and the two independent external validation cohorts were 55, 158, and 70 months, respectively. Follow-up was conducted according to the 2018 Japanese Gastric Cancer Treatment Guidelines (ver. 5)[5]. In general, follow-up was performed every 3 months for the first 2 years and every 6 months thereafter, within 5 years. Follow-up care consisted of the physical examination, laboratory tests, chest radiographs, abdominal ultrasound, abdominopelvic CT images, and annual gastroscopy. Further, MRI and PET-CT were performed whenever deemed necessary. A recurrence diagnosis was confirmed when radiological evidence was positive. Patients were followed up until death or until the cut-off dates of April 2020 in the FJMUUH cohort, July 2017 in the Mayo cohort, and January 2019 in the QUAH cohort. Patients who were lost to follow-up and died of surgery were censored.

Traditional follow-up strategies

In the present study, we searched two widely used guidelines, the NCCN guideline and JGCTG, for postoperative management of patients with GC (details were seen in Supplemental File 1 named 'Supplementary of Definition and Statistical method', Supplemental Digital Content 1, http://links.lww.com/JS9/B129).

Statistical analysis

The data are presented as mean ± SD for continuous variables and as numbers for categorical variables. The distributions of each continuous and categorical variable were compared using the Student's t-test, χ^2 test, or categorical Fisher's exact test, as appropriate for each variable. Cox proportional hazards models were used to estimate hazard ratios (HRs) for disease recurrence.

Random survival forest model

We used the RSF model to simulate the monthly time-specific recurrence probabilities for patients with pStage II and pStage III GC. The RSF methodology produced a risk-adjusted survival curve for each patient group. The RSF methodology was processed using R in the Forest SRC package[11,12].

Establishment of a risk-based follow-up strategy

We assumed the total number of visits to range from 5 to 20 times. Follow-up visits were assigned each month based on the monthly probability of recurrence[13] (Supplementary Table 1, Supplemental Digital Content 2, http://links.lww.com/JS9/B130). The second column of Supplementary Table 1 shows the monthly probability of recurrence. Given the specific number of follow-ups (12 visits for pStage II GC), risk-based follow-ups were allocated to each month in the third column (details of principle were seen in Supplemental File 1 named 'Supplementary of Definition and Statistical method', Supplemental Digital Content 1, http://links.lww.com/JS9/B129).

Evaluation of different strategies

Delayed detection months analysis

A hypothetical cohort of 1000 GC patients was generated to compare the effectiveness of surveillance among the above-mentioned strategies by calculating the sum of the delayed detection months. Delayed detection months were defined as the duration from the occurrence of failure to the next-nearest follow-up (details was seen in Supplemental File 1 named 'Supplementary of Definition and Statistical method', Supplemental Digital Content 1, http://links.lww.com/JS9/B129).

Cost-effectiveness analysis

Markov decision analytic models were generated to analyze the cost-effectiveness of various follow-up strategies at each stage (eFigure 2, Supplemental Digital Content 4, http://links.lww.com/JS9/B132)[14,15]. Parameters, including baseline clinical estimates and utility values required to build the model, were either derived from previous studies[16–20] or expert opinions if published data were unavailable. We used the surveillance and treatment costs reported in 2019 by the Medical Insurance Administration Bureau in China and the USA (Supplementary Table 2, Supplemental Digital Content 2, http://links.lww.com/JS9/B130).

All *P*-values were 2-sided, with values <0.05 considered statistically significant. Statistical analyses were performed using R version 4.0.1 and TreeAge Pro 2011.

Results

Comparison of baseline of clinicopathological characteristics between pStage II and pStage III in the FJMUUH and external validation cohorts

The clinicopathological characteristics of the FJMUUH, QUAH, and Western validation cohorts are seen in Table 1 and Supplementary Table 3, Supplemental Digital Content 2, http://links.lww.com/JS9/B130. There were 1661, 270, and 380 patients in the FJMUUH, QUAH, and Western validation cohorts, respectively. The FJMUUH cohort included 532 (32.0%) pStage II and 1,129 (68.0%) pStage III GC patients (Table 1). There were 164 (43.2%) pStage II and 216 (56.8%) pStage III GC patients in the Mayo cohort, and 105 (38.9%) pStage II and 165 (61.1%) pStage III GC patients in the QUAH cohort (Supplementary Tables 3 and 4, Supplemental Digital Content 2, http://links.lww.com/JS9/B130).

In the FJMUUH cohort, there were significant differences between pStage II and pStage III in terms of sex, tumor differentiation, pathological T stage and N stage, tumor location, tumor size, peripheral nerve invasion, lymphatic vascular invasion, and the number of examined lymph nodes (Table 1). Similarly, in the Western validation cohort, there were significant differences between pStage II and pStage III in terms of age, tumor differentiation, pathological T and N stages, tumor location, tumor size, lymphatic vascular infiltration, and the number of examined lymph nodes (Supplementary Table 3, Supplemental Digital Content 2, http://links.lww.com/JS9/B130). The differences between pStage II and pStage III GC of the QUAH cohort were detailed in Supplementary Table 4 (Supplemental Digital Content 2, http://links.lww.com/JS9/B130).

Univariable and multivariable analyses associated with disease recurrence in the FJMUUH cohort

Surgical approach, age, tumor differentiation, pathological T stage, pathological N stage, tumor location, tumor size, peripheral nerve infiltration, lymphatic vascular infiltration, adjuvant chemotherapy (AC), and Clavien–Dindo grade were associated with RFS in the training cohort (Table 2). Multivariate analysis showed that only pathological T stage, pathological N stage, tumor size, and AC were independent prognostic factors for RFS (Table 2).

Risk-adjusted Recurrence Probability in pStage III and pStage II patients of the FJMUUH cohort

In the FJMUUH cohort, 36.4% of patients experienced postoperative recurrence. Those independent prognostic factors (T stage, pathological N stage, tumor size, AC) were included in the RSF model for risk adjustment. The variables affecting disease recurrence in the RSF model are ranked according to their relative importance in eFig 3, Supplemental Digital Content 5, http://links.lww.com/JS9/B133. Figure 1 showed that the adjusted 5-year recurrence risk of pStage III GC was significantly higher than that of pStage II GC (59.6 vs. 21.3%, respectively, $P < 0.001$). The same trend was observed for distant, peritoneal, and local recurrence (pStage III vs. pStage II: distant recurrence, 34.9 vs. 10.0%; peritoneal recurrence, 18.7 vs. 3.5%; local recurrence, 14.2 vs. 3.7%, all *P*-values <0.001).

Furthermore, Figure 2 and eFig 4 (Supplemental Digital Content 6, http://links.lww.com/JS9/B134) illustrate the dynamic recurrence hazard between pStage II and pStage III GC. The overall recurrence rate of pStage II GC is relatively low, whereas most recurrences occur within the first year without an apparent peak. However, the recurrence of pStage III GC peaked at 8.9 months, and most recurrences occurred within 2 years, with a long hem to the right. Stratified analysis showed no obvious peak time in peritoneal and local recurrences for pStage II, while the peak time was 12 months for pStage III. Meanwhile, the recurrence hazard of patients with pStage III GC was significantly higher than that of patients with pStage II GC in peritoneal and local recurrence. As for distant recurrence, the recurrence hazard

Table 1

Patient characteristics between pStage II and pStage III in FJMUUH cohort.

Clinical characteristics	FJMUUH cohort ($n=1661$)	pStage II ($n=532$)	pStage III ($n=1129$)	*P*
Age *n* (%)				> 0.999
< 65 years	1029 (62.0)	330 (62.0)	699 (61.9)	
≥ 65 years	632 (38.0)	202 (38.0)	430 (38.1)	
BMI *n* (%)				0.147
< 25 kg/m^2	1228 (73.9)	377 (70.9)	851 (75.4)	
≥ 25 kg/m^2	390 (23.5)	140 (26.3)	250 (22.1)	
Unknown	43 (2.6)	15 (2.82)	28 (2.48)	
Sex *n* (%)				**0.007**
Male	1230 (74.1)	417 (78.4)	813 (72.0)	
Female	431 (25.9)	115 (21.6)	316 (28.0)	
Comorbidity *n* (%)				0.651
No	1200 (72.2)	380 (71.4)	820 (72.6)	
Yes	461 (27.8)	152 (28.6)	309 (27.4)	
ASA *n* (%)				0.468
I	945 (56.9)	310 (58.3)	635 (56.2)	
II	716 (43.1)	222 (41.7)	494 (43.8)	
ECOG scores *n* (%)				0.859
0	208 (12.5)	65 (12.2)	143 (12.7)	
1	1453 (87.5)	467 (87.8)	986 (87.3)	
Surgical approach *n* (%)				0.297
OG	284 (17.1)	83 (15.6)	201 (17.8)	
LG	1377 (82.9)	449 (84.4)	928 (82.2)	
Postoperative complication *n* (%)				0.493
<III	1594 (96.0)	515 (96.8)	1079 (95.6)	
≥ III	64 (3.9)	16 (3.0)	48 (4.3)	
Unknown	3 (0.2)	1 (0.2)	2 (0.2)	
Adjuvant *n* (%)				0.684
No	425 (25.6)	140 (26.3)	285 (25.2)	
Yes	1236 (74.4)	392 (73.7)	844 (74.8)	
Oncological characteristics				
Histology *n* (%)				< **0.001**
High/ moderate differentiated	584 (35.2)	258 (48.5)	326 (28.9)	
low differentiated/ undifferentiated	1056 (63.6)	269 (50.6)	787 (69.7)	
G_x	21 (1.3)	5 (0.94)	16 (1.42)	
Pathologic T Stage *n* (%)				< **0.001**
T1	37 (2.2)	36 (6.77)	1 (0.09)	
T2	133 (8.0)	106 (19.9)	27 (2.39)	
T3	730 (43.9)	312 (58.6)	418 (37.0)	
T4a	761 (45.8)	78 (14.7)	683 (60.5)	
Pathologic N Stage *n* (%)				< **0.001**
N0	255 (15.4)	249 (46.8)	6 (0.53)	
N1	283 (17.0)	206 (38.7)	77 (6.82)	
N2	391 (23.5)	69 (13.0)	322 (28.5)	
N3a	478 (28.8)	8 (1.50)	470 (41.6)	
N3b	254 (15.3)	0 (0.00)	254 (22.5)	
Pathological TNM Stage of 8th AJCC *n* (%)				< **0.001**
IIa	264 (15.9)	264 (49.6)	0 (0.0)	
IIb	268 (16.1)	268 (50.4)	0 (0.0)	
IIIa	416 (25.0)	0 (0.0)	416 (36.8)	
IIIb	450 (27.1)	0 (0.0)	450 (39.9)	
IIIc	263 (15.8)	0 (0.0)	263 (23.3)	
Tumor Location *n* (%)				**0.002**
Lower	639 (38.5)	224 (42.1)	415 (36.8)	
Middle	351 (21.1)	106 (19.9)	245 (21.7)	
Upper	430 (25.9)	148 (27.8)	282 (25.0)	
Mixed	241 (14.5)	54 (10.2)	187 (16.6)	
Types of gastrectomy *n* (%)				**0.001**
Total gastrectomy	1023 (61.6)	298 (56.0)	725 (64.2)	
Partial gastrectomy	638 (38.4)	234 (44.0)	404 (35.8)	
Tumor size *n* (%)				< **0.001**
< 50 mm	751 (45.2)	338 (63.5)	413 (36.6)	
≥ 50 mm	910 (54.8)	194 (36.5)	716 (63.4)	

Table 1

(*Continued*)

Clinical characteristics	FJMUUH cohort ($n=1661$)	pStage II ($n=532$)	pStage III ($n=1129$)	*P*
PNI *n* (%)				**<0.001**
No	1157 (69.7)	427 (80.3)	730 (64.7)	
Yes	504 (30.3)	105 (19.7)	399 (35.3)	
LVI *n* (%)				**<0.001**
No	1074 (64.7)	426 (80.1)	648 (57.4)	
Yes	587 (35.3)	106 (19.9)	481 (42.6)	
The number of examined lymph nodes (*n*)	36.9 ± 14.2	35.5 ± 13.4	37.6 ± 14.5	**0.021**

($P < 0.05$) values are in bold.

of pStage III GC was still significantly higher than that of pStage II GC with a peak time of 10.2 months for pStage III and 6.1 months for pStage II.

Establishment of risk-based surveillance strategies of the FJMUUH cohort

Based on the monthly recurrence hazard, we conducted individualized risk-based surveillance according to predetermined rules. Figure 3 shows the risk-based arrangements for patients with pStage II and pStage III GC at a given total number of follow-up visits (range, 5–20) within 5 years after surgery. Taking the follow-up schedule of 12 visits for patients with pStage II as an example, the number of yearly visits was 4, 3, 2, 1, and 2 from year 1 to year 5, respectively. The follow-up assignments for this example are presented in Supplementary Table 1 (Supplemental Digital Content 2, http://links.lww.com/JS9/B130). Generally, the 5-year distribution of follow-up assignments was even for patients with pStage II, whereas follow-up assignments were centered in the first 2 years for patients with pStage III.

Compare the effectiveness of different surveillance strategies of the FJMUUH cohort

Figure 4 illustrates the comparison of recurrence detection efficacy between individualized risk-based surveillance (dark blue curve) and traditional surveillance guidelines (red and green curves). As shown in Figure 4, the number of delayed detection months of recurrence gradually decreased as the total number of follow-up visits increased from 5 to 20. For pStage II GC, individualized risk-based surveillance shortened the delayed detection months to a large extent with the same total number of follow-up visits. For example, the detection efficacy of individualized risk-based surveillance (12 visits) outperformed the most intensive surveillance recommended by the JGCTG guidelines (15 visits; Fig. 4A). For pStage III GC, no obvious advantages can be seen in individualized risk-based surveillance. Thus, only when more follow-up visits (17 visits) are arranged can a detection efficacy comparable to the most intensive surveillance recommended by the JGCTG and NCCN guidelines be achieved (Fig. 4B).

Cost-effectiveness analysis of the FJMUUH cohort

Using the least-intensity follow-up strategy recommended by the NCCN guidelines as a reference, an individualized risk-based surveillance strategy achieved optimal cost-effectiveness with an ICER of $490 per QALY for pStage II GC compared with the follow-up strategies recommended by the JGCTG and NCCN guidelines. For patients with pStage III GC, the most intensive surveillance recommended by the NCCN guidelines (14 visits) outperformed individualized risk-based surveillance (17 visits) ($983 per QALY vs. $1266 per QALY) (Supplementary Table 5, Supplemental Digital Content 2, http://links.lww.com/JS9/B130).

External validation

Similarly, in the Mayo cohort and QUAH cohort (Supplementary Table 6, Supplemental Digital Content 2, http://links.lww.com/JS9/B130), individualized risk-based surveillance for pStage II GC established in our study outperformed other surveillance strategies recommended by the JGCTG and NCCN guidelines in terms of cost-effectiveness (ICER was $5509 per QALY and $274 per QALY, respectively). The most intensive surveillance recommended by the NCCN guidelines (14 visits) showed better cost-effectiveness than the individualized risk-based surveillance (Mayo cohort: $6793 per QALY vs. $9352 per QALY; QUAH cohort: $481 per QALY vs. $765 per QALY) for pStage III GC.

Recommended strategies for pStage II and pStage III patients

Based on the findings mentioned above, individualized risk-based surveillance of 12 visits within 5 years was recommended for patients with pStage II GC. Due to the high recurrence rate of patients with pStage III GC and considering the cost-effectiveness and regularity of surveillance, the most intensive surveillance strategy recommended by the NCCN guidelines (14 visits) was proven to be the best choice. The specific time points for each visit of pStage II and pStage III GC patients were shown in Figure 5.

Discussion

Although GC treatment has dramatically improved, the prognosis of stage II/III GC patients remains poor[21]. Thus, it is critical to develop optimum surveillance strategies for the early detection of recurrence. In this international multicenter study, we simulated the postoperative monthly recurrence hazards using the RSF model. Based on the dynamic risk of recurrence, we constructed an individualized risk-based surveillance schedule that could optimize the efficacy of recurrence detection. In a specific subgroup, individualized risk-based surveillance can detect recurrence more efficiently with more cost-effectiveness compared with the traditional surveillance strategies recommended by the JGCTG and NCCN guidelines.

Previous studies have shown significant differences between pStage II and pStage III GC in terms of gene expression,

Table 2

Univariable and multivariable analysis of factors associated with disease recurrence.

	FJMUUH cohort					
	Univariate analysis			Multivariate analysis		
	HR	95% CI	*P*	HR	95% CI	*P*
Surgical approach *n* (%)						
OG	1			1		
LG	0.8	0.7–1	**0.040**	0.83	0.68–1.02	0.080
Age *n* (%)						
< 65 years	1			1		
≥ 65 years	1.2	1–1.4	**0.025**	1.1	0.93–1.31	0.260
BMI *n* (%)						
< 25 kg/m^2	1			—		
≥ 25 kg/m^2	0.9	0.8–1.1	0.360	—	—	—
Sex *n* (%)						
Male	1			—		
Female	1	0.8–1.2	0.834	—	—	—
Comorbidity *n* (%)						
No	1			—		
yes	1	0.9–1.2	0.717	—	—	—
ASA *n* (%)						
I	1			—		
II	1.1	0.9–1.3	0.317	—	—	—
ECOG scores *n* (%)						
0	1			—		
1	1	0.8–1.2	0.674	—	—	—
The number of examined lymph nodes *n* (%)	1	1–1	0.054	—	—	—
Histology *n* (%)						
High/moderate differentiated	1			1		
Low differentiated/ undifferentiated	1.3	1.1–1.6	**0.002**	0.86	0.72–1.04	0.110
Pathologic T stage *n* (%)						
T1	1			1		
T2	0.7	0.3–1.6	0.405	0.82	0.36–1.86	0.640
T3	1.3	0.7–2.7	0.409	1.33	0.65–2.73	0.440
T4a	3	1.5–6	**0.002**	2.18	1.07–4.47	**0.030**
Pathologic N stage *n* (%)						
N0	1			1		
N1	1.7	1.1–2.6	**0.020**	1.92	1.24–2.99	< **0.001**
N2	2.7	1.8–4	< **0.001**	2.76	1.85–4.11	< **0.001**
N3a	5.5	3.8–7.9	< **0.001**	4.99	3.41–7.3	< **0.001**
N3b	10.8	7.4–15.8	< **0.001**	8.8	5.87–13.18	< **0.001**
Tumor location *n* (%)						
Lower	1			1		
Middle	1.2	0.9–1.4	0.177	0.96	0.77–1.19	0.700
Upper	1	0.8–1.2	0.898	0.91	0.73–1.12	0.370
Mixed	1.6	1.3–2	< **0.001**	1.02	0.8–1.29	0.900
Tumor size, mm						
< 50	1			1		
≥ 50	2	1.7–2.3	< **0.001**	1.25	1.04–1.5	**0.020**
PNI *n* (%)						
No	1			1		
Yes	1.5	1.2–1.7	< **0.001**	1.1	0.92–1.31	0.310
LVI *n* (%)						
No	1			1		
Yes	1.6	1.4–1.9	< **0.001**	1.07	0.9–1.28	0.430
Adjuvant chemotherapy *n* (%)						
No	1			1		
Yes	0.8	0.6–0.9	**0.002**	0.67	0.55–0.81	< **0.001**
Clavien–Dindo grade *n* (%)						
< III	1			1		
≥ III	1.7	1.1–2.4	**0.008**	1.21	0.82–1.77	0.330

($P < 0.05$) values are in bold.

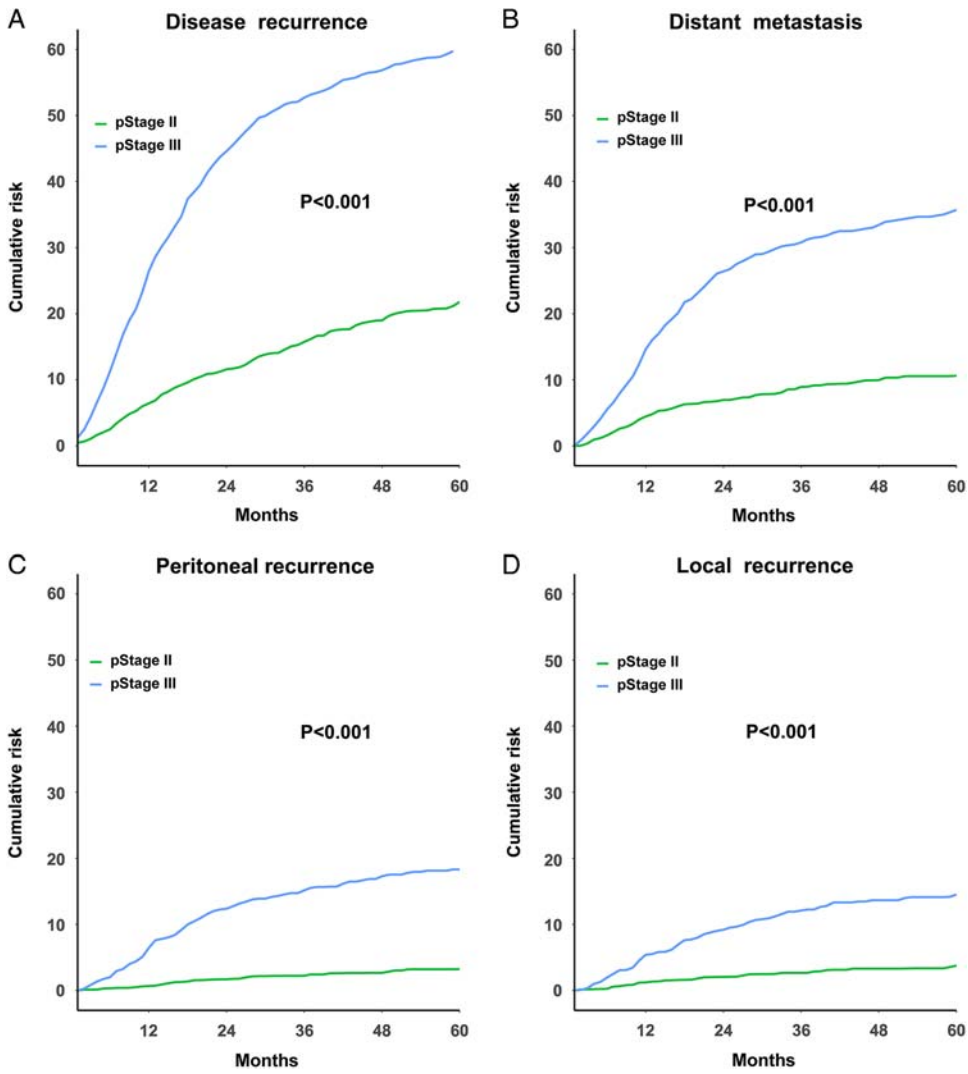

Figure 1. Risk-adjusted survival curves of different staging groups in gastric cancer patients. Disease recurrence probabilities (A); distant metastasis probabilities (B); peritoneal recurrence probabilities (C) and local recurrence probabilities (D). Number of patients: pStage II, $n = 532$; pStage III, $n = 1129$.

chemosensitivity, and prognosis. Cui *et al.*[22] reported that pStage III GC had different gene signatures and gene expression characteristics compared with pStage II GC. From the perspective of chemosensitivity, Noh *et al.* showed that only pStage II patients could achieve a survival benefit from adjuvant chemotherapy based on the XELOX regimen. In terms of prognosis, Sano *et al.*[23] confirmed that the 5-year overall survival (OS) of patients with pStage III GC was significantly lower than that of pStage II GC. Our study conforms with previous studies that showed a significantly lower overall recurrence hazard (local, distant, and peritoneal recurrence) of pStage II GC within 2 years than that of pStage III GC. Given the remarkable differences between pStage II and pStage III GC, the unified surveillance strategy recommended by current guidelines may not be applicable to both stages. Therefore, optimal individualized surveillance strategies for patients with different pathological stages are needed for the optimum utilization of medical resources. Previous studies have explored postoperative follow-up strategies for GC through dynamic hazard curves of different recurrence types[24–26]. However, conventional methods were merely based on the risk of recurrence, without considering the cost-effectiveness. Moreover, the essential indicator (quality-of-life adjusted life years, QALY) is a comprehensive measure of patients' quality and duration of life, which cannot be further stratified by recurrence types. Thus, in this study, we explored the optimal total follow-up assignments for patients with pathologic stage II and III based on the overall risk of recurrence and evaluated them with economic models. In the future, it is necessary to improve the economic models and design more precise follow-up strategies (i.e. adjusting follow-up items based on dynamic changes of different recurrence types).

Early detection of recurrence is conducive to treating recurrent lesions (intervention including radiotherapy, chemotherapy, and salvage surgery) as early as possible. Timely detection, early elimination, and prevention of the dissemination of recurrent lesions improve the quality-of-life and prolong survival. Previous studies[27,28] indicated that radical surgery or chemotherapy may improve the prognosis of recurrent GC compared with palliative

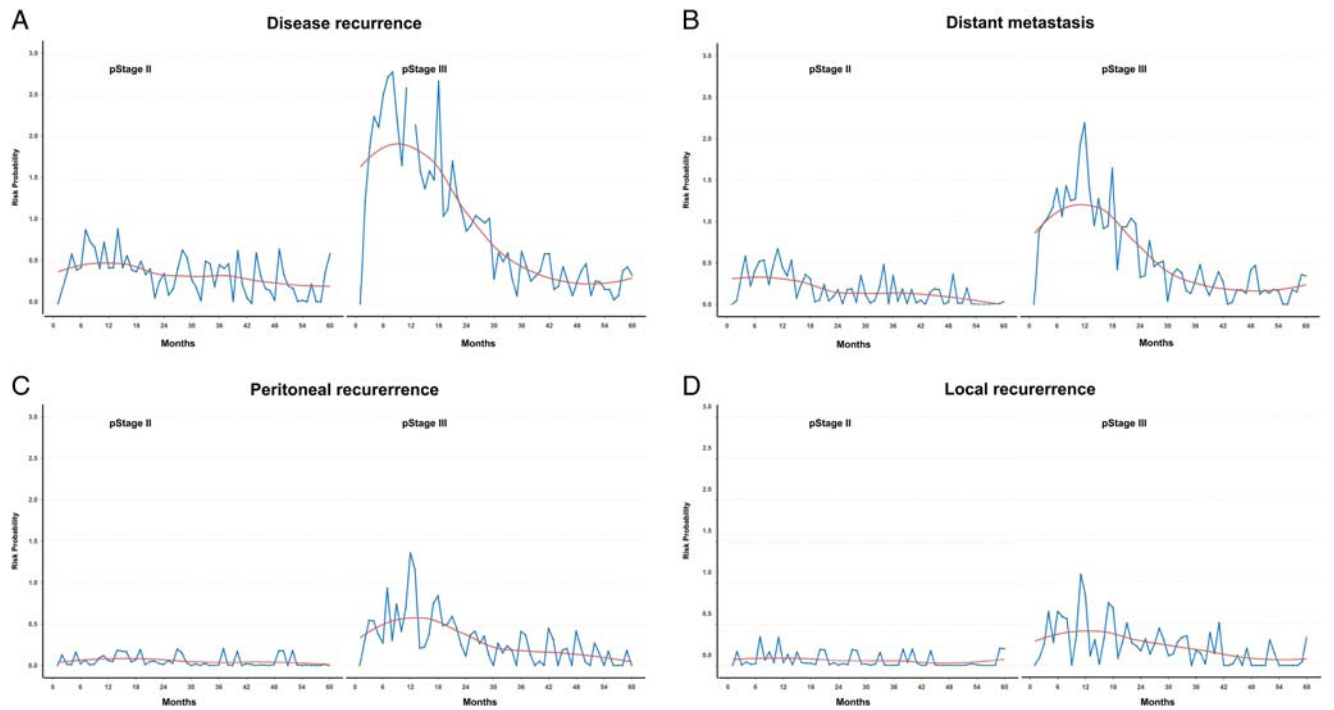

Figure 2. Time-specific occurrence probabilities of different staging groups in gastric cancer patients. Disease recurrence (A); distant metastasis (B); peritoneal recurrence (C) and local recurrence (D). Number of patients: pStage II, $n = 532$; pStage III, $n = 1129$.

care. Ideally, if we only consider the efficacy of recurrence detection, the closer the follow-up is, the higher the detection efficacy of recurrence is. However, it is unrealistic in clinical practice and is not the best option for patients, either. Therefore, a rational arrangement of postoperative surveillance for pStage II/III GC is crucial for the efficient detection of recurrence. Our

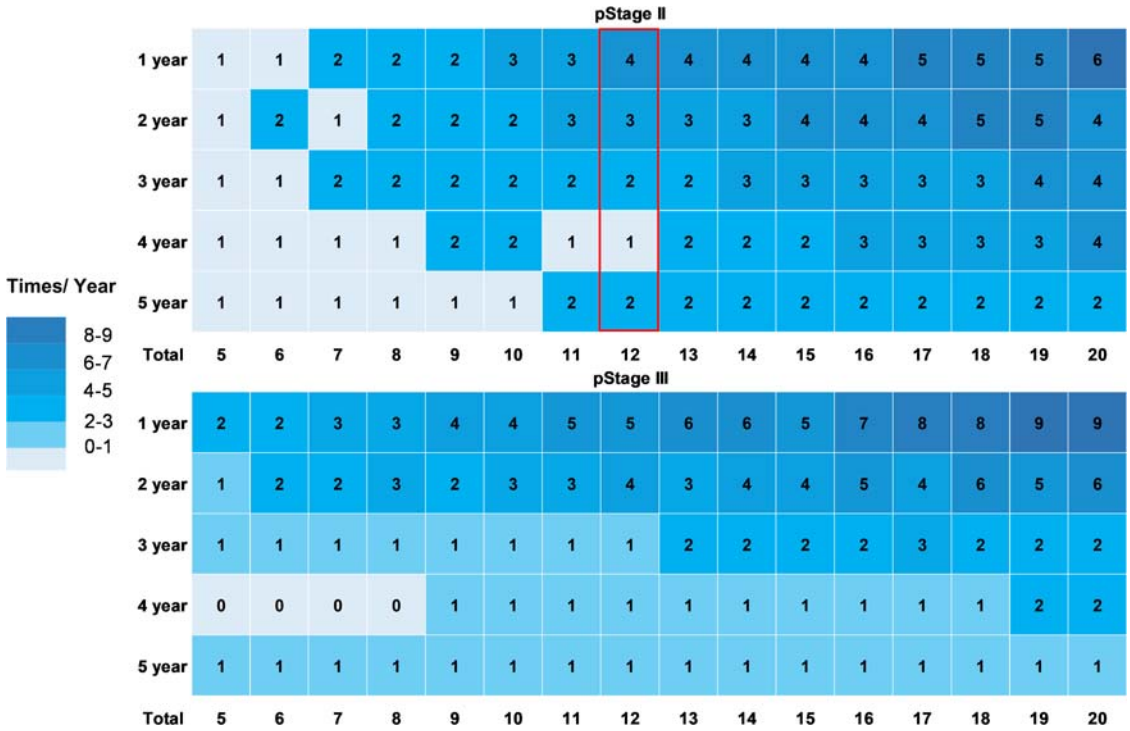

Figure 3. The risk-based surveillance arrangements varied 5–20 visits for early detection of disease recurrence. The follow-up arrangements for a total of 12 visits for pStage II are highlighted in the red box.

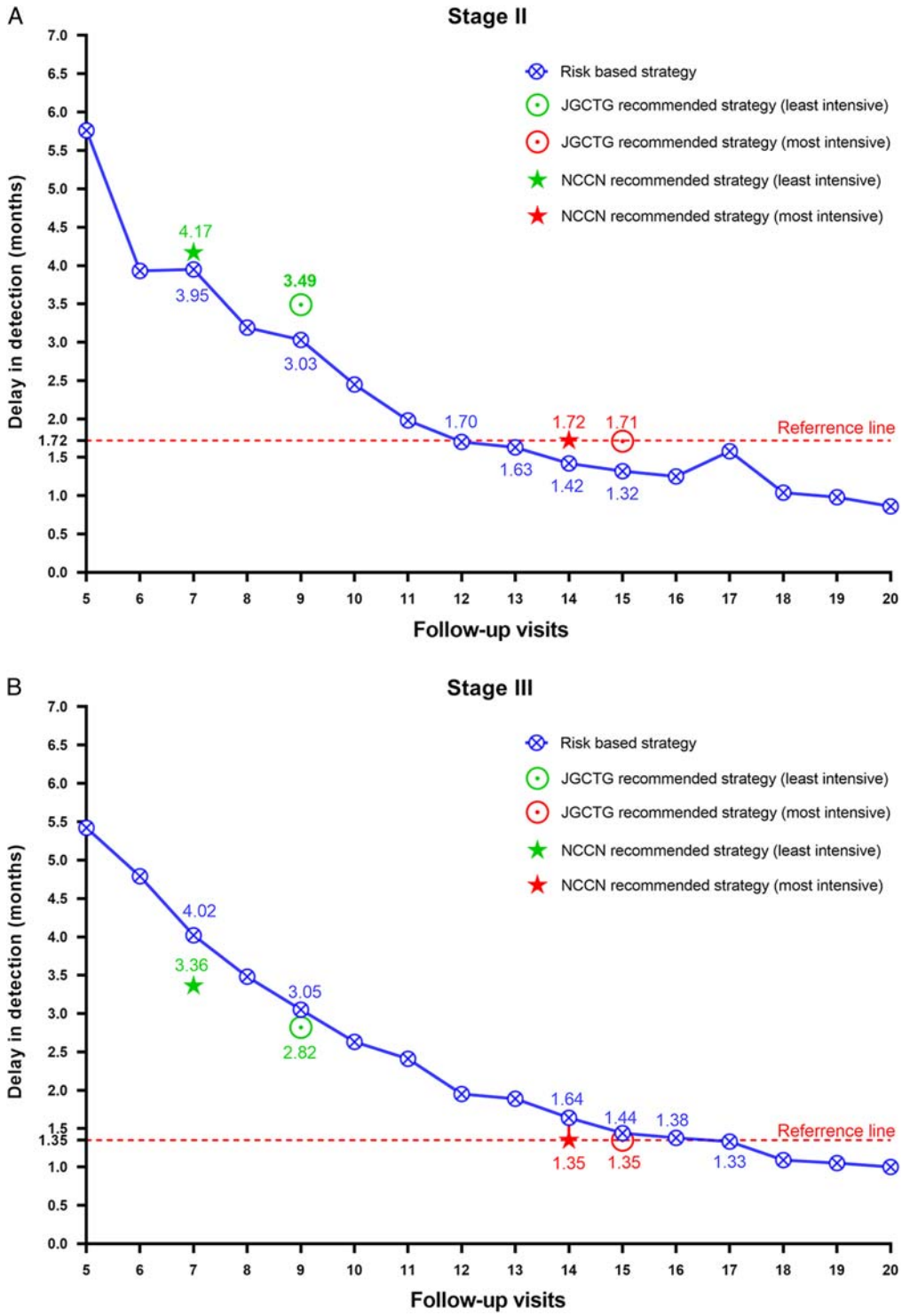

Figure 4. Comparisons of delays in the detection of disease recurrence in different surveillance strategies for each staging group. Delays in the detection of disease recurrence in risk-based surveillance arrangements (blue curve) compared with the control follow-up strategies (the red horizontal lines, represent the most intensive surveillance strategies according to NCCN and JGCTG) for patients in pStage II (A), pStage III (B). Note that the delays in detection was calculated using a hypothetical cohort of 1000 gastric cancer patients with the same features with our study population.

study simulated the monthly recurrence risk of patients with pStage II/III GC through the RSF model and constructed an individualized risk-based follow-up schedule with different follow-up frequencies (from 5 to 20 visits) according to the prespecified rules. This study also compared the detection efficacy of each surveillance strategy. Taking the most intensive follow-up strategy recommended by the JGCTG and NCCN guidelines as a reference, we found that the minimum frequency of

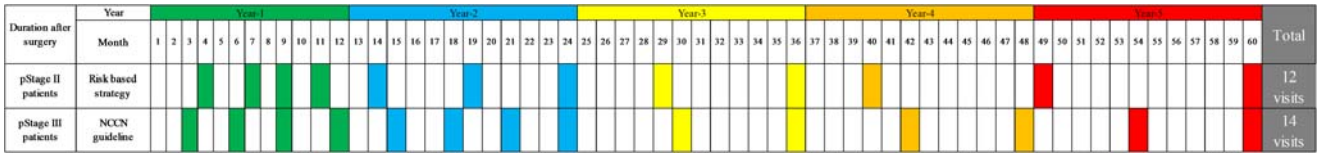

Figure 5. Postoperative recommended follow-up for pStage II and pStage III patients.

individualized risk-based follow-up strategies with better detection efficacy for pStage II and pStage III patients were 12 and 17 follow-up visits, respectively (Fig. 4).

Currently, GC poses a global medical and economic burden with a gradually increasing cost of postoperative follow-up and subsequent treatment[29,30]. Previous studies reported that about 1 in 10 U.S. cancer survivors delay follow-up care because they cannot afford associated medical bills, even if they have medical insurance coverage[31]. Individualized risk-based follow-up strategies for patients with different risks can significantly reduce the medical and economic burden on both patients and society[32,33]. Therefore, a reasonable follow-up strategy should not only emphasize detection efficacy, but also consider cost-effectiveness, that is, to maximize the early detection efficacy through limited and reasonably planned follow-up visits. Due to the scarce health economics literature on the cost-effectiveness of postoperative follow-up of GC, we constructed a Markov model to simulate the evolution of the individual status of patients with different surveillance strategies after the initial treatment. Considering the costs and quality-of-life adjusted life years (QALY), we found that individualized risk-based surveillance strategies for pStage II GC (12 visits) had better QALY and ICER, with fewer follow-ups. Additionally, patient compliance influences the actual clinical outcome. Although greater benefits may be achieved, intensive follow-up is difficult to implement in clinical practice, leading to worse outcomes. Therefore, after trading off cost-effectiveness and clinical feasibility, we concluded that the most intensive surveillance strategy recommended by the NCCN guidelines (14 visits) may be more suitable for pStage III GC than the individualized risk-based surveillance strategy (17 visits). However, individualized risk-based surveillance (12 visits) may be the best alternative for patients with pStage II GC. The above recommendations were also confirmed in the Mayo cohort and the QUAH cohort.

There are several limitations in this study. Above all, because this is an international multicenter real-world study, differences in demographic characteristics and postoperative management are inevitable between patients according to the regions. Even so, our results are robust and reliable in either Eastern or Western cohorts. Of course, it still needs further validation by large-scale prospective studies. Second, we only investigating GC disease progression without considering other types of gastric malignancies, such as remnant GC and neuroendocrine tumors. Therefore, the results cannot be projected to other cancer types. Third, although all patients completed at least 5 years of follow-up, the conclusions could not be extrapolated to GC patients who have longer postoperative survival periods. Likewise, existing guidelines do not recommend routine surveillance due to the rare disease recurrence after 5 years[34]. Fourth, the results of this study was mainly based on planned follow-up according to current guidelines. How to combine symptomatic follow-up with planned follow-up still needs to be explored in the future. Finally, considering the different prognosis of patients who receive neoadjuvant chemotherapy (NACT) or not, this study excluded patients with NACT. Further studies are warranted to explore the optimal follow-up strategy of those patients.

To our knowledge, this is the first study to compare the dynamic recurrence risk of patients with pStage II and pStage III GC through international multicenter studies. Moreover, proposing an optimal follow-up strategy that is more suitable for patients with pStage II and pStage III GC not only improves follow-up effectiveness but also optimizes the use of medical resources.

Conclusion

Patients with pStage II and pStage III GC showed significant differences in any type of recurrence. Individualized risk-based surveillance strategies are more applicable to patients with pStage II GC and are more cost-effective than the traditional guidelines. However, pStage III GC patients may be more suitable to the most intensive surveillance strategy recommended by the NCCN guidelines. This study is expected to further optimize the postoperative individualized follow-up strategies for patients with pStage II and pStage III GC to enable rational allocation of medical resources and reduce the medical burden on patients and society. Further prospective, large-scale studies are needed.

Ethical approval

This study passed the Institutional Review Board (IRB) of the FJMUUH (IRB number: 2022KY106).

Consent

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1964 and later versions. Informed consent or substitute for it was obtained from all patients for being included in the study.

Consent for publication: This article does not report an individual participant's data in any form.

Sources of funding

This study was supported by Fujian Research and Training Grants for Young and Middle-aged Leaders in Healthcare (Jun Lu).

Conflicts of interest disclosure

The authors declare no financial, consultant, or institutional conflict of interest.

Author contributions

C.-M.H., D.W., M.J.T., J.L., Y.-B.M.: conception/design; D.W., J.L., J.L., J.-b.H., B.-b.X., Z.X., H.-L.Z., G.-s.L., L.-l.S., J.-W.X., J.-B.W., J.-X.L., Q.-Y.C., L.-L.C., P.L., C.-H.Z., M.J.T., C.-M.H., Y.-B.M.: collection and/or assembly of data; D.W., J.L., J.L., J.-W.X., M.J.T., C.-M.H.: manuscript writing.

Research registration unique identifying number (UIN)

1. Name of the registry: Clinical Trials.gov.
2. Unique Identifying number or registration ID: NCT05465 993.
3. Hyperlink to your specific registration (must be publicly accessible and will be checked): https://clinicaltrials.gov/ct2/show/NCT05465993?term=NCT05465993&draw=2&rank=1.

Guarantor

Chang-ming Huang.

Data availability statement

Availability of data and materials: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

The authors thank all colleagues who provided patient care. The authors also thank all the family members (Chun-gui Ke *et al.*) who are supporting us. The authors hope our contribution can bring hope to patients with advanced gastric cancer.

References

[1] Sung H, Ferlay J, Siegel RL, *et al*. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71:209–49.
[2] Ohtsu A, Shah MA, Van Cutsem E, *et al*. Bevacizumab in combination with chemotherapy as first-line therapy in advanced gastric cancer: a randomized, double-blind, placebo-controlled phase III study. J Clin Oncol 2011;29:3968–76.
[3] Noh SH, Park SR, Yang H-K, *et al*. Adjuvant capecitabine plus oxaliplatin for gastric cancer after D2 gastrectomy (CLASSIC): 5-year follow-up of an open-label, randomised phase 3 trial. Lancet Oncol 2014;15:1389–96.
[4] Ajani, D'Amico TA, Bentrem DJ, *et al*. Gastric cancer, version 2.2022, NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw 2022;20:167–92.
[5] Japanese Gastric Cancer Association. Japanese gastric cancer treatment guidelines 2018. Gastric Cancer 2021;24:1–21.
[6] Mathew G, Agha R. for the STROCSS Group. STROCSS 2021: strengthening the reporting of cohort, cross-sectional and case-control studies in surgery. In: for the STROCSS Group, editors. Int J Surg 2021; 96:106165.
[7] Lee JH, Chang KK, Yoon C, *et al*. Lauren histologic type is the most important factor associated with pattern of recurrence following resection of gastric adenocarcinoma. Ann Surg 2018;267:105–13.
[8] Lu J, Wu D, Chen S, *et al*. A novel hematological classifier predicting chemotherapy benefit and recurrence hazard for locally advanced gastric cancer A multicenter IPTW analysis. Eur J Surg Oncol 2022; 48:1768–77.
[9] Japanese Gastric Cancer Association. Japanese classification of gastric carcinoma 3rd English edition. Gastric Cancer 2011;14:101–12.
[10] Amin MB, Edge S, Greene F, *et al*. AJCC Cancer Staging Manual, 8th ed. Springer International Publishing; 2017.
[11] Ishwaran H, Kogalur UB, Blackstone EH, *et al*. Random survival forests. Ann Appl Stat 2008;2:841–60.
[12] Ishwaran H, Kogalur U. Random Forest SRC: random forests for survival, regression and classification. https://CRAN.R-project.org/package=randomForestSRC
[13] Zhou GQ, Wu CF, Deng B, *et al*. An optimal posttreatment surveillance strategy for cancer survivors based on an individualized risk-based approach. Nat Commun 2020;11:3872.
[14] White C, Nimeh T, Gazelle GS, *et al*. A decision analysis comparing 3 active surveillance protocols for the treatment of patients with low-risk prostate cancer. Cancer 2019;125:952–62.
[15] Zhang Y, Denton BT, Nielsen ME. Comparison of surveillance strategies for low-risk bladder cancer patients. Med Decis Making 2013;33: 198–214.
[16] Wu B, Li T, Cai J, *et al*. Cost-effectiveness analysis of adjuvant chemotherapies in patients presenting with gastric cancer after D2 gastrectomy. BMC Cancer 2014;14:984.
[17] Hisashige A, Sasako M, Nakajima T. Cost-effectiveness of adjuvant chemotherapy for curatively resected gastric cancer with S-1. BMC Cancer 2013;13:443.
[18] Glatz T, Verst R, Kuvendjiska J, *et al*. Pattern of recurrence and patient survival after perioperative chemotherapy with 5-FU, leucovorin, oxaliplatin and docetaxel (FLOT) for locally advanced esophagogastric adenocarcinoma in patients treated outside clinical trials. J Clin Med 2020;9: 2654.
[19] Huang HL, Leung CY, Saito E, *et al*. Effect and cost-effectiveness of national gastric cancer screening in Japan: a microsimulation modeling study. BMC Med 2020;18:257.
[20] Xia R, Zeng H, Liu Q, *et al*. Health-related quality of life and health utility score of patients with gastric cancer: a multi-centre cross-sectional survey in China. Eur J Cancer Care (Engl) 2020;29:e13283.
[21] Yu J, Huang C, Sun Y, *et al*. Effect of laparoscopic vs open distal gastrectomy on 3-year disease-free survival in patients with locally advanced gastric cancer: the CLASS-01 randomized clinical trial. JAMA 2019;321: 1983–92.
[22] Cui J, Li F, Wang G, *et al*. Gene-expression signatures can distinguish gastric cancer grades and stages. PLoS One 2011;6:e17819.
[23] Sano Takeshi, Coit Daniel G, Kim Hyung Ho, *et al*. Proposal of a new stage grouping of gastric cancer for TNM classification: international gastric cancer association staging project. Gastric Cancer 2017;20: 217–25.
[24] Lu J, Wu D, Xu BB, *et al*. A matched cohort study of the failure pattern after laparoscopic and open gastrectomy for locally advanced gastric cancer: does the operative approach matter? Surg Endosc 2022;36: 689–700.
[25] Qiu WW, Chen QY, Zheng WZ, *et al*. Postoperative follow-up for gastric cancer needs to be individualized according to age, tumour recurrence pattern, and recurrence time. Eur J Surg Oncol 2022;48:1790–8.
[26] Chen QY, Zhong Q, Zhou JF, *et al*. Conditional survival and recurrence of remnant gastric cancer after surgical resection: a multi-institutional study. Cancer Sci 2020;111:502–12.
[27] Kang Jung Hun, Lee Soon Il, Lim Do Hyoung, *et al*. Salvage chemotherapy for pretreated gastric cancer: a randomized phase III trial comparing chemotherapy plus best supportive care with best supportive care alone. J Clin Oncol 2012;30:1513–8.
[28] Janowitz Tobias, Thuss-Patience Peter, Marshall Andrea, *et al*. Chemotherapy vs supportive care alone for relapsed gastric, gastro-esophageal junction, and oesophageal adenocarcinoma: a meta-analysis of patient-level data. Br J Cancer 2016;114:381–7.
[29] GBD 2017 Stomach Cancer Collaborators. The global, regional, and national burden of stomach cancer in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease study 2017. Lancet Gastroenterol Hepatol 2020;5:42–54.
[30] Altice CK, Banegas MP, Tucker-Seeley RD, *et al*. Financial hardships experienced by cancer survivors: a systematic review. J Natl Cancer Inst 2016;109:djw205.

[31] Kent EE, Forsythe LP, Yabroff KR, *et al*. Are survivors who report cancer-related financial problems more likely to forgo or delay medical care? Cancer 2013;119:3710–7.
[32] Cadham Christopher J, Cao Pianpian, Jayasekera Jinani, *et al*. Cost-effectiveness of smoking cessation interventions in the lung cancer screening setting: a simulation study. J Natl Cancer Inst 2021;113:1065–73.
[33] Hammer Mark M, Eckel Andrew L, Palazzo, *et al*. Cost-effectiveness of treatment thresholds for subsolid pulmonary nodules in ct lung cancer screening. Radiology 2021;300:586–93.
[34] Lee Ju-Hee, Kim Hyun-Il, Kim Min Gyu, *et al*. Recurrence of gastric cancer in patients who are disease-free for more than 5 years after primary resection. Surgery 2016;159:1090–8.