Long-term Outcomes of Laparoscopic Versus Open Surgery for Clinical Stage I Gastric Cancer: The LOC-1 Study

Michitaka Honda, MD, PhD,* Naoki Hiki, MD, PhD,* Takahiro Kinoshita, MD, PhD,†
Hiroshi Yabusaki, MD, PhD,‡ Takayuki Abe, PhD,§ Souya Nunobe, MD, PhD,* Mitsumi Terada, MD,†
Atsushi Matsuki, MD, PhD,‡ Hideki Sunagawa, MD,† Masaki Aizawa, MD, PhD,‡ Mark A. Healy, MD,¶
Manabu Iwasaki, PhD,‖ and Toshi A. Furukawa, MD, PhD**

Background: Clinical trials comparing laparoscopic gastrectomy (LG) versus traditional open gastrectomy (OG) have been planned, their surgical outcomes reported but their oncologic outcomes are still pending. Consequently, we have conducted this large-scale historical cohort study to provide relevant information rapidly to guide our current practice.

Methods: Through a consensus meeting involving surgeons, biostatisticians, and epidemiologists, 30 variables of preoperative information possibly influencing surgeons' choice between LG versus OG and potentially associating with outcomes were identified to enable rigorous estimation of propensity scores. A total of 4235 consecutive patients who underwent gastrectomy for gastric adenocarcinoma were identified and their relevant data were gathered from the participating hospitals. After propensity score matching, 1848 patients (924 each for LG and OG) were selected for comparison of long-term outcomes.

Results: In the propensity-matched population, the 5-year overall survival was 96.3% [95% confidence interval (CI) 95.0–97.6] in the OG as compared with 97.1% (95% CI, 95.9–98.3) in LG. The number of all-cause death was 33/924 in the OG and 24/924 in the LG through the entire period, and the hazard ratio (LG/OG) for overall death was 0.75 (95% CI, 0.44–1.27; $P = 0.290$). The 3-year recurrence-free survival was 97.4% (95% CI, 96.4–98.5) in the OG and 97.7% (95% CI, 96.5–98.8) in the LG. The number of recurrence was 22/924 in the OG and 21/924 in the LG through the entire period, and the hazard ratio was 1.01 (95% CI, 0.55–1.84; $P = 0.981$).

Conclusions: This observational study adjusted for all-known confounding factors seems to provide strong enough evidence to suggest that LG is oncologically comparable to OG for gastric cancer.

Laparoscopic surgery for patients with gastric cancer has now been employed for more than 2 decades,[1] and its use has steadily increased worldwide. A number of randomized controlled trials (RCT's) with small sample sizes,[2–6] interim reports from planned large trials[7,8] and several meta-analyses[9–15] including retrospective observational studies[16–18] have indicated that surgical outcomes of laparoscopic gastrectomy (LG) in terms of factors, such as postoperative complications, patient recovery, blood loss, and the number of harvested lymph nodes, are generally comparable to open gastrectomy (OG) and acceptable. Thus, the surgical safety of LG is now well established. It, however, is generally believed that the issue of oncological outcomes, in terms of relapse-free and overall survival, recurrence rate, and pattern of recurrence will not be settled until results of large-scale randomized trials currently underway are published.[7,8]

Clinical decision making must ideally be based on medical evidence that considers the results of RCTs and meta-analyses. Laparoscopic gastrectomy, however, has already been accepted in many countries. In Japan, the number of patients who undergo LG has increased from 1823 in 2003 to 9168 in 2013, according to a survey conducted by the Japanese Society of Endoscopic Surgery,[19] thus representing a 5-fold increase over the last decade. In the real world, progress in new techniques continues to advance without waiting for the results of clinical trials, as exemplified by the increasing use of robotic surgery by many surgeons. How should we reconcile this time lag between daily practice and the availability of randomized evidence?

To fill in these evidence gaps while awaiting the outcomes of clinical trials, some investigators have been shifting emphasis to analytic observational approaches using large-scale databases.[20,21] Although this can shed light on simple associations between surgical treatments and outcomes, appropriate adjustment for confounding factors is essential in evaluating the effectiveness of interventions. Retrospective studies using propensity score matching (PSM) represent 1 notable approach for such confounder adjustment.[22–25] In most previous research to evaluate the effectiveness of surgical interventions, adjustment for confounding factors using propensity score, however, has been less than satisfactory, because factors used for calculating propensity scores were inadequate (only a small subset of probable confounders available were used) and/or inappropriate (post-intervention variables such as histopathological information that could be known only after surgery was used). To estimate propensity scores, all preoperative information influencing surgeons' clinical judgment of whether LG or OG is indicated should be isolated; otherwise, PSM

would not be able to adjust for confounding by indication. We have repeatedly pointed out these concerns about inadequate and inappropriate use of the PSM technique.[26–28]

In view of the current state of evidence surrounding LG, we have organized a consensus meeting involving expert laparoscopic surgeons, epidemiologists, and biostatisticians, and established a large-scale multicenter database to estimate propensity scores more precisely through multidisciplinary discussion. We herein report the details of this historical cohort study, "Laparoscopy versus Open Surgery for Clinical Stage I Gastric Cancer" (LOC-1), with the aim of establishing a more precise estimation of propensity scores and examining the effectiveness and safety of LG over OG, while adjusting as completely as possible for possible confounding by indication.

METHODS

Hypothesis and Cohort Development

The aim of this study was to verify our hypothesis that LG is not inferior to OG, in terms of overall survival. We designed the entire study protocol a priori, with consideration of the sample size necessary to evaluate noninferiority. Three Japanese cancer-specialized institutions were selected to participate on the basis of the following criteria: being able to offer both open and laparoscopic surgery to patients with gastric cancer; and having facilities for electronic storage of clinical data, including medical records, images, or laboratory data for all consecutive patients with gastric cancer who underwent gastrectomy from January 2006 through December 2012. All patients at these institutions who met the inclusion criteria below were enrolled. This study was conducted with the approval of the institutional review boards at all the participating hospitals. Cancer staging was based on the seventh Union for International Cancer Control (UICC) TNM classification.[29]

PARTICIPANTS

The patients enrolled in this study had histologically confirmed gastric adenocarcinoma, were diagnosed as clinical stage I (T1N0, T2N0, or T1N1), and had undergone gastrectomy, including total, subtotal, proximal, or pylorus-preserving gastrectomy. The exclusion criteria included carcinoma in the gastric stump (after previous gastrectomy), the presence of another primary malignancy, and a history of chemotherapy or chemo-radiotherapy.
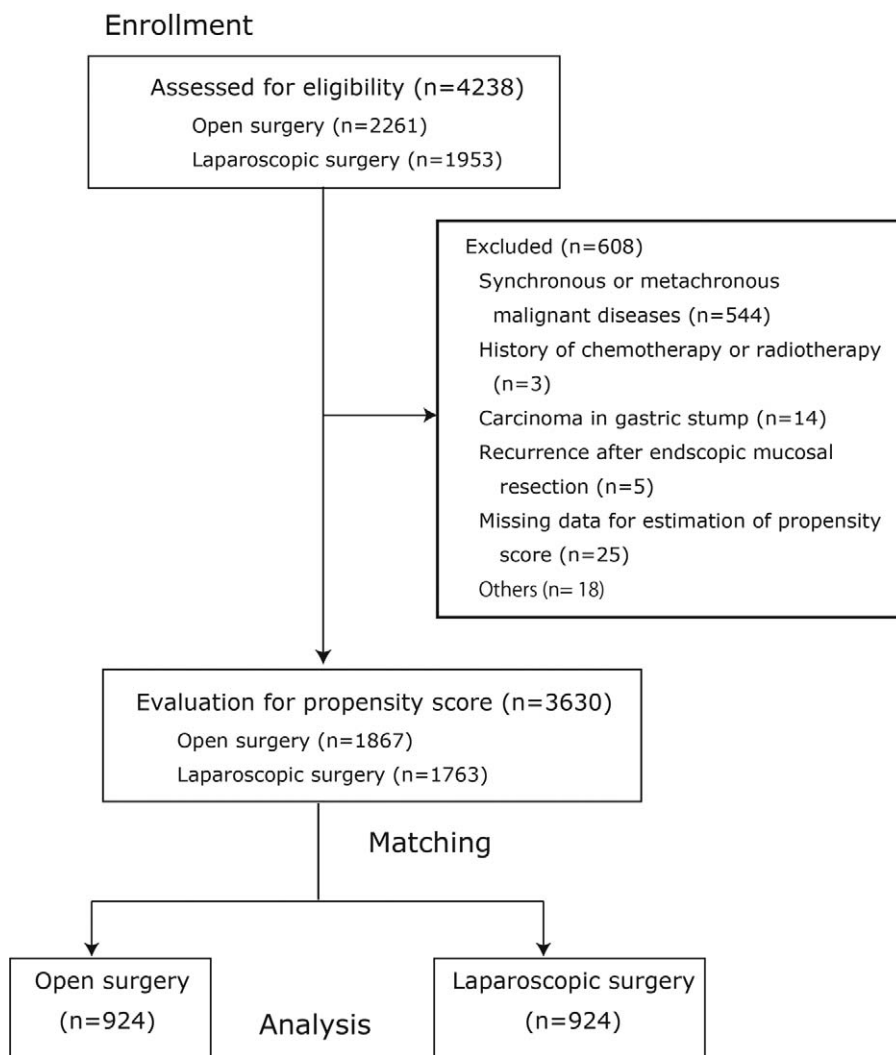
Enrollment

FIGURE 1. Study Population. Flowchart of patient enrollment. After matching, 924 patients in both groups were included in the final analyses.

Data Collection, Propensity Score Matching, and Sample Size Estimation

This observational study was not designed to be a formal noninferiority study because of the expected low incidence rate, but was designed to give a best possible interval estimate of hazard ratio (HR) using PSM based on 30 clinically relevant covariates. When designing this study, it was calculated that at least 592 patients per arm were necessary to show the noninferiority of the LG group to the OG group with 5% margin for proportion, 80% power, and a 2-tailed alpha of 5%. A total of 4235 patients were identified from the institutional databases: 2258 patients underwent open surgery and 1977 patients underwent LG. To optimize the accuracy of the propensity score, our study team clarified through their consensus meeting the preoperative information related to the choice made by the surgeon as to whether open surgery or laparoscopic surgery would be used. A total of 30 preoperative factors, including details of the patients' characteristics and tumor findings, were identified (Supplementary Table 1, http://links.lww.com/SLA/A974). To collect precise information on patient and tumor characteristics, inves-tigators who were blind to the outcome looked back on the medical records, stored images, and laboratory data for all patients. In all participating hospitals, upper gastrointestinal endoscopy and abdominal computed tomography were performed for every patient with gastric cancer, and barium swallow or endoscopic ultrasound were available and used for some cases. There was no performance of diagnostic laparoscopy for clinical stage I gastric cancer. Propensity score estimates and matching were derived by 2 biostatisticians (TA and MI) who were also blind to the outcome. The score was estimated using a logistic regression model and greedy matching (ratio = 1:1 without replacement) with a caliper of width 0.2 standard deviations of the logit of the estimated propensity score. In addition to the PSM, 5 factors (clinical T and N factor, esophageal invasion, duodenum invasion, and tumor location as a preoperative diagnosis) were exactly matched to achieve better balance.[30] We did not use statistical imputation for missing data because this was an issue in only 23 cases (0.53%). After matching, 924 patients each in the OG and LG groups were included in the final analyses. The balance of each covariate before and after the matching between the 2 groups was evaluated by

TABLE 1. Baseline Characteristics Before and After Propensity Score Matching

Characteristics		All Patients (n = 3630)					Propensity-Matched Patients (n = 1848)				
		Open Surgery (n = 1867)	%	Laparoscopic Surgery (n = 1763)	%	Standardized Difference	Open surgery (n = 924)	%	Laparoscopic surgery (n = 924)	%	Standardized Difference
Age	Mean	64.1		62.0		18.7	63.2		63.3		−0.7
Sex	Male	1245	66.7	1130	64.1	5.4	607	65.7	605	65.5	0.5
	Female	622	33.3	633	35.9	−5.4	317	34.3	319	34.5	−0.5
Year	2006	356	19.1	137	7.8	33.6	124	13.4	128	13.9	−1.3
	2007	312	16.7	185	10.5	18.2	159	17.2	153	16.6	1.7
	2008	289	15.5	201	11.4	12.0	135	14.6	146	15.8	−3.3
	2009	331	17.7	248	14.1	10.0	175	18.9	167	18.1	2.2
	2010	280	15.0	272	15.4	−1.2	155	16.8	147	15.9	2.3
	2011	168	9.0	364	20.6	−33.2	96	10.4	106	11.5	−3.5
	2012	131	7.0	356	20.2	−39.2	80	8.7	77	8.3	1.2
ASA-PS	1	698	37.4	884	50.1	−25.9	416	45.0	399	43.2	3.7
	2	1051	56.3	827	46.9	18.9	466	50.4	489	52.9	−5.0
	3	118	6.3	52	2.9	16.1	42	4.5	36	3.9	3.2
BMI, kg/m²	Mean (SD)	22.7 (3.2)		22.6 (3.2)		4.1	22.7 (3.1)		22.8 (3.2)		−5.1
History of abdominal surgery		84	4.5	64	3.6	4.4	39	4.2	41	4.4	−1.1
Combined surgery		5	0.3	2	0.1	3.5	1	0.1	2	0.2	−2.7
Site of lesion	Upper	364	19.5	293	16.6	7.5	159	17.2	159	17.2	0.0
	Upper to middle	90	4.8	40	2.3	13.8	22	2.4	22	2.4	0.0
	Middle	795	42.6	882	50.0	−15.0	447	48.4	447	48.4	0.0
	Middle to lower	58	3.1	20	1.1	13.7	12	1.3	12	1.3	0.0
	Lower	502	26.9	497	28.2	−2.9	269	29.1	269	29.1	0.0
	Entire	47	2.5	22	1.2	9.4	15	1.6	15	1.6	0.0
Esophageal invasion		39	2.1	13	0.7	11.5	7	0.8	7	0.8	0.0
Duodenum invasion		12	0.6	3	0.2	7.4	0	0.0	0	0.0	NA
Preoperative ER		136	7.3	257	14.6	−23.5	103	11.1	92	10.0	3.9
Clinical T	1a	293	15.7	497	28.2	−30.5	224	24.2	224	24.2	0.0
	1b	1002	53.7	1191	67.6	−28.7	632	68.4	632	68.4	0.0
	2	572	30.6	75	4.3	74.1	68	7.4	68	7.4	0.0
Clinical N	0	1758	94.2	1733	98.3	−21.8	904	97.8	904	97.8	0.0
	1	109	5.8	30	1.7	21.8	20	2.2	20	2.2	0.0
Size of tumor, mm	Mean (SD)	30.7 (15.2)		30.2 (15.6)		3.2	29.2 (14.7)		28.9 (14.5)		1.7
Histological findings of Biopsy specimen	Well	872	46.7	645	36.6	20.6	396	42.9	389	42.1	1.5
	Poor	640	34.3	786	44.6	−21.2	361	39.1	367	39.7	−1.3
	Mixed type	354	19.0	329	18.7	0.8	167	18.1	168	18.2	−0.3
	Others	0	0.0	2	0.1	−4.8	0	0.0	0	0.0	NA

ASA-PS indicates American Society of Anesthesiologists physical status; BMI, Body mass index; SD, Standard deviation.
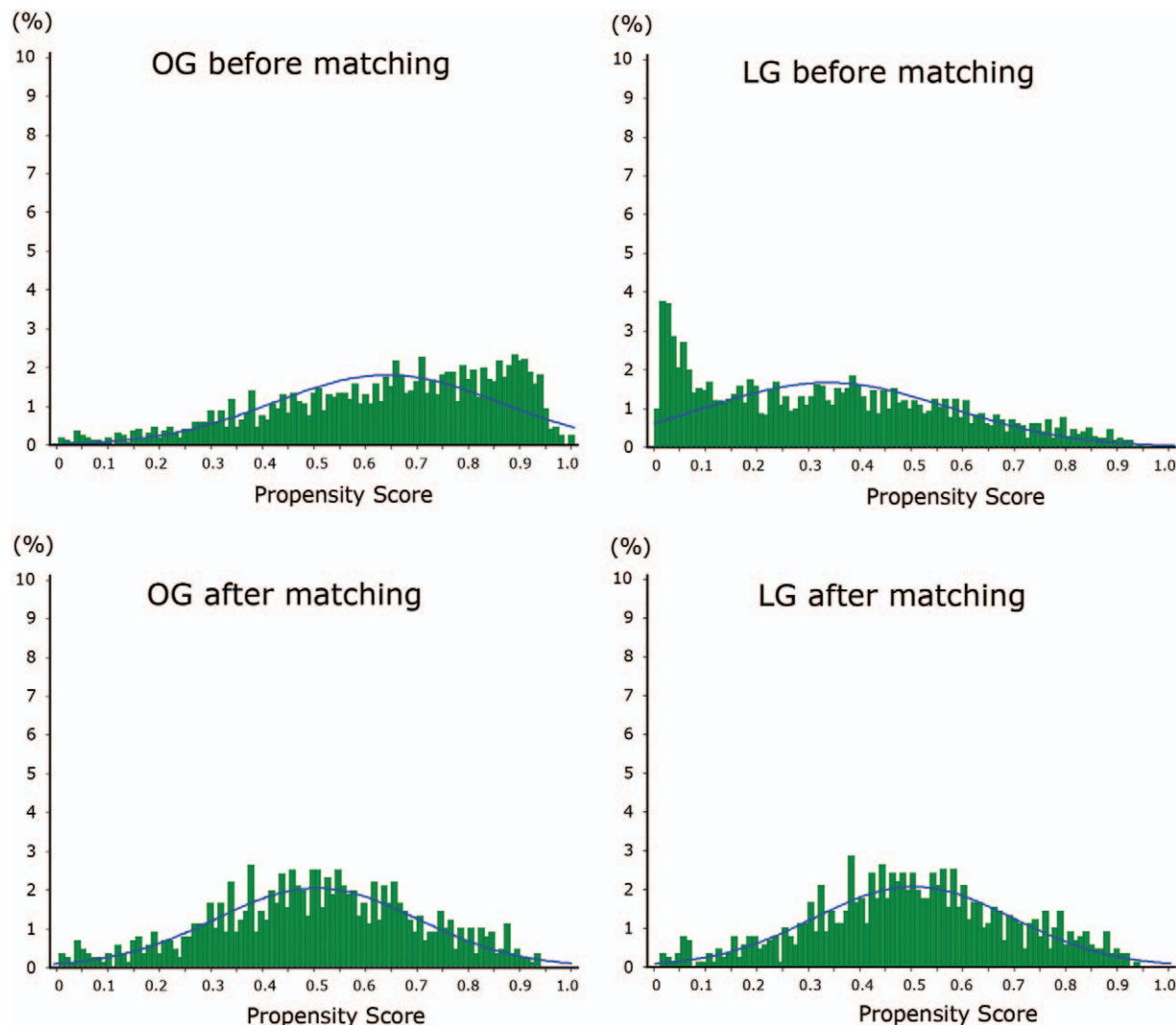
FIGURE 2. Distribution of propensity score before and after matching.

standardized differences.[31] Absolute value of standardized differences less than 10% was considered to be a relatively small imbalance. A flowchart of patient enrollment is shown in Figure 1.

TREATMENT METHODS

All institutions that participated in this study were specialized cancer hospitals and all enrolled patients received relatively homogeneous treatments, according to gastric cancer treatment guideline in Japan.[32] Although there were no surgeon-specific criteria in this study, considering the average number of patients undergoing gastrectomy was more 300 cases per year in each hospital during this period, all surgeons were considered to have enough experiences to perform both laparoscopic and open surgery. In Japanese specialized cancer hospitals, D2 lymphadenectomy,[33] which includes dissection of the region around the splenic artery, celiac artery and hepatic artery without distal pancreatectomy, is considered standard. In this regard, cases with lymphadenectomy around the proper hepatic artery (No. 12a) were omitted in some instances of clinical stage I disease. Modified D2 lymphadenectomy was defined as "D1 +" in this study. Adjuvant

chemotherapy with S-1 for 1 year was performed for most patients with curative resection and pathological stage II, IIIA, or IIIB tumors.[34]

Outcomes and Statistical Analysis

After matching and fixing the enrolled cases, investigators collected all outcome data. The main outcome was overall survival (OS). Secondary outcomes included relapse-free survival (RFS), disease-specific survival, recurrence pattern, the number of harvested lymph nodes, and the incidence of postoperative complications that were grade 3 or more severe according to the Clavien-Dindo classification.[35] The OS, RFS, and disease-specific survival were assessed using the Kaplan-Meier method and compared between the OG and LG groups. The hazard ratios (HR) and 95% confidence intervals (CI) were estimated using the unstratified Cox proportional hazards model as primary analyses.[36,37] The stratified Cox model for matched pairs was also fitted as a sensitivity analysis. The descriptive statistics were evaluated in other secondary outcomes, and as necessary, continuous variables were compared using Student t tests and categorical variables by Fisher exact test. All statistical tests were 2-sided, and P values of 0.05 or less were considered to indicate

TABLE 2. Operating and Pathological Findings

		Open Surgery (n = 924)	(%)	Laparoscopic Surgery (n = 924)	(%)	*P* Value
Procedure	TG	103	11.1	102	11.0	0.797
	DG	566	61.3	568	61.5	
	(B-I/B-II/RY)	(467/9/90)		(390/0/178)		
	PPG	187	20.2	196	21.2	
	PG	68	7.4	58	6.3	
Lymph node dissection	D1+	346	37.4	656	71.0	0.000
	D2	578	62.6	268	29.0	
Retrieved Lymph nodes	median (range)	33 (5–145)		35 (5–94)		—
Operating time	Mean (SD)	167.7 (47.0)		240.0 (58.2)		0.000
Blood loss	Mean (SD)	174.1 (181.2)		60.7 (111.7)		0.000
Switch to open surgery		—		12	1.3	—
Patological T	1	794	85.9	781	84.5	0.449
	2	67	7.3	86	9.3	
	3	38	4.1	42	4.5	
	4a	14	1.5	15	1.6	
	4b	1	0.1	0	0.0	
Pathological N	0	810	87.7	793	85.8	0.613
	1	76	8.2	96	10.4	
	2	22	2.4	20	2.2	
	3a	13	1.4	12	1.3	
	3b	3	0.3	3	0.3	
Pathological M	0	922	99.8	923	99.9	1.000
	1	2	0.2	1	0.1	
Pathological Stage	IA	741	80.2	712	77.1	0.680
	IB	92	10.0	104	11.3	
	IIA	45	4.9	60	6.5	
	IIB	23	2.5	28	3.0	
	IIIA	9	1.0	10	1.1	
	IIIB	9	1.0	7	0.8	
	IIIC	3	0.3	2	0.2	
	IV	2	0.2	1	0.1	
Recurrence site	Total occurrences	22	2.4	21	2.3	1.000
	Peritneum	8		7		
	Liver	6		8		
	Lung	1		1		
	Bone	2		0		
	Lymphnode	2		4		
	Local	2		1		
	Others	1		0		

B-1 indicates Billroth I reconstruction; B-2, Billroth II reconstruction; DG, distal gastrectomy; PG, proximal gastrectomy; PPG, pylorus preserving gastrectomy; RY, Roux-en-Y reconstruction; TG, total gastrectomy.

TABLE 3. Timing of Events Occurrence

		Postoperative Year							
Event	Group	1	2	3	4	5	6	7	Total
All death	OG	2	10	11	6	1	1	2	33
	(%)*	6.06	36.4	69.7	87.9	90.9	93.9	100	
	LG	1	9	8	4	1	1	0	24
	(%)*	4.17	41.7	75.0	91.7	95.8	100	100	
Disease specific death	OG	0	4	4	0	1	0	0	9
	(%)*	0.00	44.4	88.9	88.9	100	100	100	
	LG	1	6	5	1	0	0	0	13
	(%)*	7.69	53.8	92.3	100	100	100	100	
Recurrence	OG	6	11	3	1	0	1	0	22
	(%)*	27.3	77.3	90.9	95.5	95.5	100	100	
	LG	7	5	4	1	1	2	1	21
	(%)*	33.3	57.1	76.2	81.0	85.7	95.2	100	

*Cumulative percentage.
OG indicates open gastrectomy; LG, laparoscopic gastrectomy.

Overall Survival (%)

HR = 0.75
(95%CI 0.44 - 1.27; p = 0.290)

Red line = OG
Blue line = LG

Follow-up (Year)

Number of at risk

		0	1	2	3	4	5	6	7
OG	n=924	904	857	747	614	365	161	63
LG	n=924	918	896	724	535	372	92	21

A

Relapse-Free Survival (%)

HR = 1.01
(95%CI 0.55 - 1.84; p = 0.981)

Red line = OG
Blue line = LG

Follow-up (Year)

Number of at risk

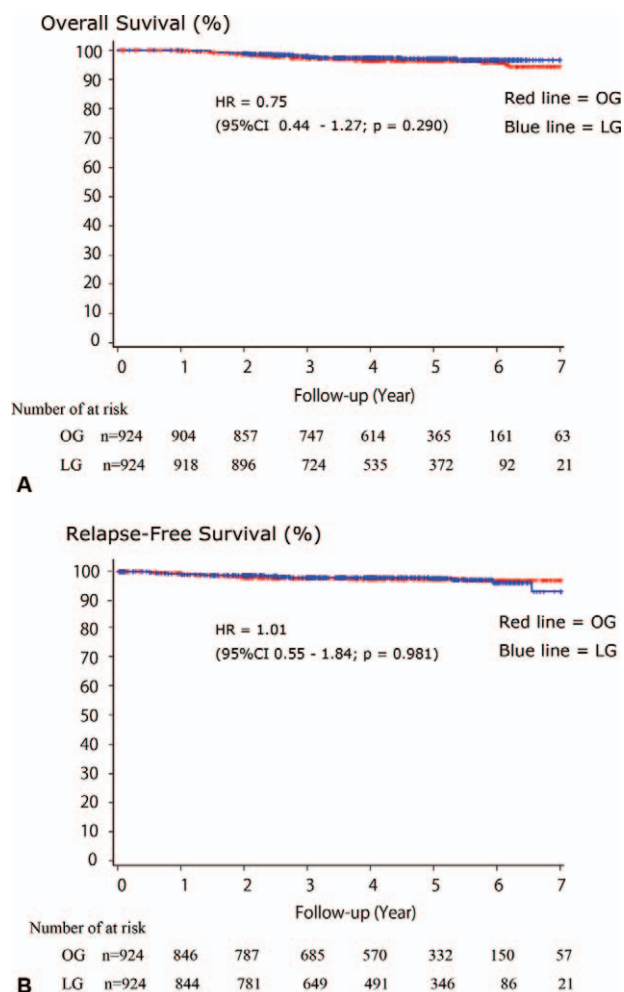|		0	1	2	3	4	5	6	7
OG	n=924	846	787	685	570	332	150	57
LG	n=924	844	781	649	491	346	86	21

B

FIGURE 3. Survival curves using the Kaplan-Meier method and the number of at-risk patients in the matched cohort. Overall survival rate. The 5-year overall survival was 95.3% in the open gastrectomy group and 94.8% in the laparoscopic gastrectomy group. Relapse-free survival rate.

statistical significance. All the analyses were performed with SAS version 9.3 (SAS Institute, Cary, NC).

Subgroup Analyses

To confirm the main outcome for all procedures, disease-free survival was compared for each resection type: total, subtotal, proximal or pylorus-preserving gastrectomy, and D1 + or D2 lymphadenectomy.

RESULTS

Table 1 shows the patient and tumor characteristics for both pre- and post-PSM. Propensity score distributions are shown in Figure 2.

Oncological Outcomes

The HR for overall mortality, recurrence, and disease-specific mortality in the LG group compared with OS group were 0.75 (95% CI 0.44–1.27; $P = 0.290$), 1.01 (95% CI 0.55–1.84; $P = 0.981$), and 1.38 (95% CI 0.61–3.34; $P = 0.411$), respectively. Estimated HR

with the stratified Cox model gave similar results (not shown). Figure 3 shows the survival curves obtained using the Kaplan-Meier method. The 5-year OS and 3-year RFS were 96.3% (95% CI 95.0–97.6) and 97.4% (95% CI 96.4–98.5) in the OG group, and 97.1% (95% CI 95.9–98.3) and 97.7% (95% CI 96.5–98.8) in the LG group, respectively. The risk differences of OS and RFS in the LG group were −0.82% and −0.32%. The number (proportion) of all death occurrence was 33 (3.57%) in the OG group and 24 (2.60%) in the LG group ($P = 0.230$) and recurrent cases was 22 (2.38%) in the OG group and 21 (2.27%) in the LG group ($P = 1.000$), respectively. The most common sites of recurrence were the peritoneum and liver (Table 2). The median observation period was 4.4 years (interquartile range: 3.1–5.2 years). In terms of oncological outcomes, there were no significant differences between groups. The timing of event occurrence was shown in Table 3. Most recurrences occurred within 3 postoperative years. There were no significant differences in oncological outcome among any surgical procedures in subgroup analyses; the 5-year OS in OG group versus LG group were 95.0% (95% CI 92.7–96.6) versus 95.1 (95% CI 92.5–96.8; $P = 0.4331$) in distal gastrectomy, 90.2% (95% CI 79.3–95.5) versus 95.6% (95% CI 88.4–98.4; $P = 0.441$) in total gastrectomy, 98.4% (95% CI 95.0–99.5) versus 96.8 (95% CI 93.0–98.6; $P = 0.387$) in pylorus-preserving gastrectomy, 95.0% (95% CI 85.1–98.4) versus 84.8% (95% CI 67.6–93.2; $P = 0.187$) in proximal gastrectomy, 95.6% (95% CI 92.5–97.5) versus 96.0% (95% CI 93.8–97.5; $P = 0.5931$) in D1 + lymphadenectomy and 95.1% (95% CI 92.8–96.6) versus 92.3% (88.0–95.1; $P = 0.2637$) in D2 lymphadenectomy.

Surgical Outcomes

Table 2 provides the details of surgical procedures and outcomes. Two intraoperative accidents occurred in the OG group: 1 unplanned splenectomy due to injury to the splenic artery and 1 fatal arrhythmia during surgery. In the LG group, 12 cases had to be converted to open surgery: for control of bleeding in 4 cases, and to secure safe oncological manipulation because of serosal invasion by the tumor in 8 cases. There were no significant differences in terms of surgical outcome, such as the number of harvested lymph nodes, pathological TNM stage, or histological type. Details of the postoperative complications are shown in Table 4. The incidence of postoperative complications more severe than grade III was 5.8% in the OG group and 5.1% in the LG group ($P = 0.539$). There were more cases of pancreatic fistula in the LG group than in the OG group, whereas the latter group showed a higher incidence of wound infection. With regard to long-term complications, the incidence of small bowel obstruction in the OG group was higher than in the LG group, whereas the latter group showed a higher incidence of internal hernia requiring reoperation. Neither length of postoperative hospitalization nor the proportion of readmission rate differed significantly between the groups. No death within 30 postoperative days occurred in either of the groups.

DISCUSSION

Our results suggest that there are no significant differences in OS, RFS, or site of recurrence between OG and LG, and that there are no differences in short-term postoperative complications. These results were also confirmed for all procedure subgroups, including similar oncological outcomes for laparoscopic total gastrectomy with a D1 + lymphadenectomy in early stage gastric cancer. Given the large sample size and the use of strict propensity score estimation and matching, the results of this study seem to establish that LG is neither oncologically nor surgically inferior to OG for stage I gastric cancer. Our study has many advantages over previous studies addressing the

TABLE 4. Postoperative Course

		Open Surgery (n = 924)	%	Laparoscopic surgery (n = 924)	%	P Value
In-hospital stay, median (range)		11 (4–77)		11 (7–144)		—
Reoperation		5	0.54	5	0.54	1.000
Operative mortality		0	0.00	0	0.00	—
Readmission within 60 postoperative day		59	6.39	66	7.14	0.579
Postoperative complication (> = Grade 3)		54	5.84	47	5.09	0.539
Short-term problems						
	Anastomotic leakage	12	1.30	12	1.30	
	Intra-abdominal bleeding	0	0.00	1	0.11	
	Intraluminal bleeding	1	0.11	3	0.32	
	Pancreatic fistula	8	0.87	12	1.30	
	Abdominal abscess or fluid collection	4	0.43	6	0.65	
	Wound infection	5	0.54	1	0.11	
	Stenosis	5	0.54	8	0.87	
	Enteroparalysis	1	0.11	1	0.11	
	Ascites	0	0.00	1	0.11	
	Pneumonia	2	0.22	3	0.32	
	Cardiac problem	2	0.22	1	0.11	
	Neurological problem	0	0.00	0	0.00	
	Others	2	0.22	0	0.00	
Long-term problems						
	Small bowel obstruction	12	1.30	0	0.00	
	Internal hernia	0	0.00	5	0.54	

same clinical question using PSM. In the present study, we assembled a team, including expert gastric cancer surgeons, epidemiologists, and biostatisticians, and tried to evaluate the propensity of procedural allocation as precisely as possible. This allowed us to finally identify 30 preoperative factors related to surgical decision making. These clinical data were collected by investigators who were blinded to the outcomes. Thanks to recent advances in electronic medical data storage, which has made it possible to store most clinical data, including unstructured data, such as CT or endoscopic images, without any loss or deterioration of quality, in our study, only 0.2% of data were missing. In the literature, 2 previous studies[22,24] have addressed the same research question, but they selected only 3 to 5 covariates, such as body mass index, TNM, or comorbidity to calculate the propensity score. It is clear that surgeons choose the most appropriate surgical approach on the basis of much more information, including for example, patient age, tumor size, location, histological findings of biopsy, history of abdominal surgery, performance status, patient selection, institution, and the year the operation was performed. Surgeons are expected to consider all available information to decide the optimal procedure for each patient. Another weakness of previous studies has been that the propensity score has often been estimated using pathological findings, which is postoperative information. Any decision regarding surgical approach can only be influenced by preoperative factors, such as clinical TNM stage. Adoption of pathological TNM stage would obviously violate the temporal sequence of cause-effect in propensity estimation. We have discussed the difficulty of propensity estimation for surgical interventions in previous articles.[27,28,38]

Our study also has some advantages over conventional RCTs, both in terms of internal and external validity. Through accurate estimation of the propensity scores based on our survey, all of the known confounding factors were much better adjusted for in our study than in an RCT. Therefore, this study can be regarded as an epidemiological attempt to come closer to the truth, encompassing the characteristics of a well-designed observational study, or RCT for which all confounding factors, including unknown ones, are adjusted automatically. Some recent review articles have evaluated the quality of PSM.[20,21,39,40] These reviews showed that PSM may sometimes overestimate the efficacy of interventions, while some studies using PSM have led to almost the same conclusions as RCTs. Propensity score matching accuracy seems to vary depending on the research question posed or the type of outcome investigated.[20] Because the results of RCTs aimed at addressing the same research question will be published later,[7,8] we should be able to assess the validity of our study through comparison with those results.

One notable feature of our study methodology is an increase in external validity, even over many RCT's. Because the patients we enrolled were acquired consecutively from all participating institutions, many patients who would not be entered into a clinical trial were included in our analyses, such as those with severe comorbidities, the elderly, or those requiring emergency surgery. Therefore we were able to secure a high degree of external validity. Randomized controlled trails aimed at establishing efficacy of a new intervention often need to sacrifice external validity to enhance their internal validity. As a result, patients actually enrolled into a trial tend to account for only a small proportion of the total candidate patients who would receive the intervention in question in the real world. The same would be true of the quality of the surgical interventions: in randomized surgical intervention trials, it is quite difficult to ensure homogeneity of the quality of surgical interventions, unlike clinical trials of drug therapies. This has been regarded as an inevitable limitation of RCTs of surgical interventions but this limitation would not apply to our retrospective cohort study because we have included all surgeons who use LG, OG, or both in the participating institutions.

Lastly, our study has ethical and pragmatic advantages over RCTs in certain respects. Randomized controlled trails are always hampered by a shortage of registered participants, and they tend to require a long period before oncological follow-up can be completed. Even while waiting for the results, advances in surgical techniques or mechanical devices will have been taking place.[41] Therefore there is a possibility that even if trials yield positive results, they may not lead to changes in standard treatment, because surgical techniques may have become more refined in the meantime. Indeed, LG has already been accepted as a standard procedure in many countries, even before publication of the results of clinical trials.[19] In this situation, the ethical problem of patients being allocated randomly without

consideration of technical or mechanical advances seems to be an important issue that needs to be addressed.

There are several important limitations to our study. First, there is no guarantee that all confounding factors were included in our database. It might be possible to overlook unmeasurable or unknown but important factors. In addition, we need to discuss the possibility that our results could be extrapolated to patients with gastric cancer worldwide. The epidemiology and treatment of gastric cancer in other East Asian countries is similar to our present study. In Western countries, particularly in North America, there, however, may be significant disadvantageous factors in surgical treatment. These include higher body mass index (BMI) in patients and a greater proportion of patients with advanced stage and upper third cancer, requiring total gastrectomy. It is important to note that it is not known if the results of our series will be able to be reproduced in patients with higher BMI, such as the many patients with severe obesity (BMI > 35 m^2/kg) seen in the Western world. Although higher BMI has been linked to a higher incidence of postoperative complications and to longer operative times in previous reports, this has not been shown to contribute to oncological prognosis.[42–45] We hypothesize though that, if similar oncologic principles are followed, patients with severe obesity may have similar results, as laparoscopic surgery has been applied safely and routinely in the Western world in these patients for other indications. Another clinical question is the potential use of LG for more advanced cases such as serosa-invasive tumors or patients with bulky lymph node metastasis. We plan to conduct future research to show outcomes of advanced cases using this same rigorous study design, once more data is available.

In conclusion, while it is known that LG has similar postoperative outcomes to open surgery, this is the first study of its kind to show that long-term oncological outcomes are also similar for these 2 procedures. Methods of advanced PSM, such as those used in our study, should be considered in future work to evaluate surgical interventions with multiple potential confounders or when clinical trials are not feasible.

REFERENCES

1. Kitano S, Iso Y, Moriyama M, et al. Laparoscopy-assisted Billroth I gastrectomy. *Surg Laparosc Endosc*. 1994;4:146–148.
2. Kim HH, Hyung WJ, Cho GS, et al. Morbidity and mortality of laparoscopic gastrectomy versus open gastrectomy for gastric cancer: an interim report: a phase III multicenter, prospective, randomized Trial (KLASS Trial). *Ann Surg*. 2010;251:417–420.
3. Takiguchi S, Fujiwara Y, Yamasaki M, et al. Laparoscopy-assisted distal gastrectomy versus open distal gastrectomy. A prospective randomized single-blind study. *World J Surg*. 2013;37:2379–2386.
4. Kitano S, Shiraishi N, Fujii K, et al. A randomized controlled trial comparing open vs laparoscopy-assisted distal gastrectomy for the treatment of early gastric cancer: an interim report. *Surgery*. 2002;131:S306–311.
5. Sakuramoto S, Yamashita K, Kikuchi S, et al. Laparoscopy versus open distal gastrectomy by expert surgeons for early gastric cancer in Japanese patients: short-term clinical outcomes of a randomized clinical trial. *Surg Endosc*. 2013;27:1695–1705.
6. Kim YW, Yoon HM, Yun YH, et al. Long-term outcomes of laparoscopy-assisted distal gastrectomy for early gastric cancer: result of a randomized controlled trial (COACT 0301). *Surg Endosc*. 2013;27:4267–4276.
7. Nakamura K, Katai H, Mizusawa J, et al. A phase III study of laparoscopy-assisted versus open distal gastrectomy with nodal dissection for clinical stage IA/IB gastric Cancer (JCOG0912). *Jpn J Clin Oncol*. 2013;43:324–327.
8. Kim HH, Han SU, Kim MC, et al. Prospective randomized controlled trial (phase III) to comparing laparoscopic distal gastrectomy with open distal gastrectomy for gastric adenocarcinoma (KLASS 01). *J Korean Surg Soc*. 2013;84:123–130.
9. Vinuela EF, Gonen M, Brennan MF, et al. Laparoscopic versus open distal gastrectomy for gastric cancer: a meta-analysis of randomized controlled trials and high-quality nonrandomized studies. *Ann Surg*. 2012;255:446–456.
10. Peng JS, Song H, Yang ZL, et al. Meta-analysis of laparoscopy-assisted distal gastrectomy and conventional open distal gastrectomy for early gastric cancer. *Chin J Cancer*. 2010;29:349–354.
11. Deng Y, Zhang Y, Guo TK. Laparoscopy-assisted versus open distal gastrectomy for early gastric cancer: a meta-analysis based on seven randomized controlled trials. *Surg Oncol*. 2015;24:71–77.
12. Chen XZ, Wen L, Rui YY, et al. Long-term survival outcomes of laparoscopic versus open gastrectomy for gastric cancer: a systematic review and meta-analysis. *Medicine (Baltimore)*. 2015;94:e454.
13. Zou ZH, Zhao LY, Mou TY, et al. Laparoscopic vs open D2 gastrectomy for locally advanced gastric cancer: a meta-analysis. *World J Gastroenterol*. 2014;20:16750–16764.
14. Huang YL, Lin HG, Yang JW, et al. Laparoscopy-assisted versus open gastrectomy with D2 lymph node dissection for advanced gastric cancer: a meta-analysis. *Int J Clin Exp Med*. 2014;7:1490–1499.
15. Cheng Q, Pang TC, Hollands MJ, et al. Systematic review and meta-analysis of laparoscopic versus open distal gastrectomy. *J Gastrointest Surg*. 2014;18:1087–1099.
16. Fukunaga T, Hiki N, Kubota T, et al. Oncologic outcomes of laparoscopy-assisted distal gastrectomy for gastric cancer. *Ann Surg Oncol*. 2013;20:2676–2682.
17. MacLellan SJ, MacKay HJ, Ringash J, et al. Laparoscopic gastrectomy for patients with advanced gastric cancer produces oncologic outcomes similar to those for open resection. *Surg Endosc*. 2012;26:1813–1821.
18. Hamabe A, Omori T, Tanaka K, et al. Comparison of long-term results between laparoscopy-assisted gastrectomy and open gastrectomy with D2 lymph node dissection for advanced gastric cancer. *Surg Endosc*. 2012;26:1702–1709.
19. Japan Society for Endoscopic Surgery. 12th Nationwide survey of endoscopic surgery in Japan. *JSES*. 2014;19:495–558.
20. Dahabreh IJ, Kent DM. Can the learning health care system be educated with observational data? *JAMA*. 2014;312:129–130.
21. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. *Cochrane Database Syst Rev*. 2014;4:MR000034.
22. Kim HH, Han SU, Kim MC, et al. Long-term results of laparoscopic gastrectomy for gastric cancer: a large-scale case-control and case-matched Korean multicenter study. *J Clin Oncol*. 2014;32:627–633.
23. Hasegawa S, Kunisaki C, Ono H, et al. Omentum-preserving gastrectomy for advanced gastric cancer: a propensity-matched retrospective cohort study. *Gastric Cancer*. 2013;16:383–388.
24. Kim KH, Kim MC, Jung GJ, et al. Comparative analysis of five-year survival results of laparoscopy-assisted gastrectomy versus open gastrectomy for advanced gastric cancer: a case-control study using a propensity score method. *Dig Surg*. 2012;29:165–171.
25. Hanna WC, de Valence M, Atenafu EG, et al. Is video-assisted lobectomy for non-small-cell lung cancer oncologically equivalent to open lobectomy? *Eur J Cardiothorac Surg*. 2013;43:1121–1125.
26. Honda M, Kuriyama A, Furukawa TA. Comments on "Omentum-preserving gastrectomy for advanced gastric cancer: a propensity-matched retrospective cohort study". *Gastric Cancer*. 2014;17:757–758.
27. Yamamoto K, Honda M, Furukawa TA. Video-assisted thoracoscopic surgery lobectomy or open lobectomy for non-small-cell lung cancer? Minimizing selection bias in observational studies. *Eur J Cardiothorac Surg*. 2013;46:150.
28. Honda M, Yamamoto K, Furukawa TA. Retrospective propensity score matching and the selection of surgical procedures: how precise can a propensity estimate be? *J Clin Oncol*. 2014;32:3200–3201.
29. Biondi A, Hyung WJ. Seventh edition of TNM classification for gastric cancer. *J Clin Oncol*. 2011;29:4338–4339. author reply 4340-4332.
30. Rubin DBT, Thomas N. Combining propensity score matching with additional adjustments for prognostic covariates. *J Amer Statist Assoc*. 2000;95:573–585.
31. Bangalore S, Guo Y, Samadashvili Z, et al. Everolimus-eluting stents or bypass surgery for multivessel coronary disease. *N Engl J Med*. 2015;372:1213–1222.
32. Nakajima T. Gastric cancer treatment guidelines in Japan. *Gastric Cancer*. 2002;5:1–5.
33. Japanese Gastric Cancer Association. Japanese gastric cancer treatment guidelines 2010 (ver. 3). *Gastric Cancer*. 2011;14:113–123.
34. Sakuramoto S, Sasako M, Yamaguchi T, et al. Adjuvant chemotherapy for gastric cancer with S-1, an oral fluoropyrimidine. *N Engl J Med*. 2007;357:1810–1820.
35. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg*. 2004;240:205–213.

36. Schafer JL, Kang J. Average causal effects from nonrandomized studies: a practical guide and simulated example. *Psychol Methods*. 2008;13:279–313.

37. Austin PC. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Stat Med*. 2008;27:2037–2049.

38. Honda M, Kuriyama A, Furukawa TA. Comments on ''Omentum-preserving gastrectomy for advanced gastric cancer: a propensity-matched retrospective cohort study''. *Gastric Cancer*. 2014;17:757–758.

39. Dahabreh IJ, Sheldrick RC, Paulus JK, et al. Do observational studies using propensity score methods agree with randomized trials? A systematic comparison of studies on acute coronary syndromes. *Eur Heart J*. 2012;33:1893–1901.

40. Lonjon G, Boutron I, Trinquart L, et al. Comparison of treatment effect estimates from prospective nonrandomized studies with propensity score analysis and randomized controlled trials of surgical procedures. *Ann Surg*. 2014;259:18–25.

41. Nunobe S, Hiki N, Tanimura S, et al. The clinical safety of performing laparoscopic gastrectomy for gastric cancer by trainees after sufficient experience in assisting. *World J Surg*. 2013;37:424–429.

42. Kim JH, Chin HM, Hwang SS, et al. Impact of intra-abdominal fat on surgical outcome and overall survival of patients with gastric cancer. *Int J Surg*. 2014;12:346–352.

43. Bickenbach KA, Denton B, Gonen M, et al. Impact of obesity on perioperative complications and long-term survival of patients with gastric cancer. *Ann Surg Oncol*. 2013;20:780–787.

44. Wong J, Rahman S, Saeed N, et al. Effect of body mass index in patients undergoing resection for gastric cancer: a single center US experience. *J Gastrointest Surg*. 2014;18:505–511.

45. Lee HJ, Kim HH, Kim MC, et al. The impact of a high body mass index on laparoscopy assisted gastrectomy for gastric cancer. *Surg Endosc*. 2009;23:2473–2479.