Can female surgeons break the glass ceiling? A comparison of short-term surgical outcomes of male and female surgeons in Japan: a retrospective cohort study

Journal:	*BMJ*
Manuscript ID	BMJ-2022-070568.R1
Article Type:	Original research
Date Submitted by the Author:	07-May-2022

Complete List of Authors:	Okoshi, Kae; Japan Baptist Hospital, Surgery Endo, Hideki; The University of Tokyo Graduate School of Medicine Faculty of Medicine, Healthcare Quality Assessment Nomura, Sachiyo; The University of Tokyo, Gastrointestinal Surgery Kono, Emiko; Osaka Medical and Pharmaceutical University, General and Gastroenterological Surgery Fujita, Yusuke; Kyoto University Graduate School of Medicine Faculty of Medicine, Surgery Yasufuku, Itaru; Gifu University School of Medicine Graduate School of Medicine Hida, Koya; Kyoto University Graduate School of Medicine Faculty of Medicine, Sugery Yamamoto, Hiroyuki; The University of Tokyo Graduate School of Medicine Faculty of Medicine, Healthcare Quality Assessment Miyata, Hiroaki; The University of Tokyo Graduate School of Medicine Faculty of Medicine, Healthcare Quality Assessment Yoshida, Kazuhiro; Gifu University School of Medicine Graduate School of Medicine, Surgical Oncology Kakeji, Yoshihiro ; The Japanese Society of Gastroenterological Surgery Kitagawa, Yuko; The Japanese Society of Gastroenterological Surgery

Keywords:	General surgery, Surgical procedures, operative, Gastrointestinal diseases

Can female surgeons break the glass ceiling? A comparison of short-term surgical outcomes of male and female surgeons in Japan: a retrospective cohort study

Kae Okoshi PhD[1*] (kae_md@kuhp.kyoto-u.ac.jp), Hideki Endo PhD[2*] (hidendo-thk@umin.ac.jp), Sachiyo Nomura PhD[3†] (sachiyo.nomura1012@gmail.com, ORCiD; 0000-0003-4293-6205), Emiko Kono MD[4] (konoemiko6@gmail.com), Yusuke Fujita MD[5] (yusuke0424@kuhp.kyoto-u.ac.jp), Itaru Yasufuku MD[6] (itaru.yasufuku1983@gmail.com), Koya Hida PhD[5] (hidakoya@kuhp.kyoto-u.ac.jp), Hiroyuki Yamamoto PhD[2] (yama-h@umin.ac.jp), Hiroaki Miyata PhD[2] (h-m@umin.ac.jp), Kazuhiro Yoshida PhD[6] (kyoshida@gifu-u.ac.jp), Yoshihiro Kakeji PhD[7] (kakeji@med.kobe-u.ac.jp), Yuko Kitagawa PhD[8] (kitagawa.a3@keio.jp)

[1]Department of Surgery, Japan Baptist Hospital, Sakyo-ku, Kyoto, Japan, [2]Department of Healthcare Quality Assessment, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan, [3]Department of Gastrointestinal Surgery, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan, [4]Department of General and Gastroenterological Surgery, Osaka Medical and Pharmaceutical University Takatsuki, Japan, [5]Department of Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan, [6]Department of Surgical Oncology, Gifu University School of Medicine, Gifu, Japan, [7]Database Committee, The Japanese Society of Gastroenterological Surgery, Tokyo, Japan, [8]The Japanese Society of Gastroenterological Surgery, Tokyo, Japan

*Contributed equally

†*Corresponding author*:
Sachiyo Nomura, MD, PhD
Department of Gastrointestinal Surgery
Graduate School of Medicine, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan
Telephone: +81-3-3815-5411
Fax: +81-3-5800-9734
E-mail: sachiyo.nomura1012@gmail.com

ABSTRACT

Objectives

To compare the short-term surgical outcomes between female and male surgeons in Japan with a large gender gap.

Design

Retrospective cohort study.

Setting

Data from the Japanese National Clinical Database (2013–2017) and the Japanese Society of Gastroenterological Surgery were used.

Participants

The National Clinical Database (2013–2017) includes data pertaining to >95% of surgeries performed in Japan and data from this database were used to analyse the outcomes of distal gastrectomy (DG), total gastrectomy (TG), and low anterior resection (LAR) performed by male and female surgeons. Cases with missing data were excluded from this study.

Main outcome measures

Primary outcomes included surgical mortality, surgical mortality combined with postoperative complications, pancreatic fistula (DG/TG only), and anastomotic leakage (LAR only). We examined the association of surgeons' gender with the number of years after the registration of licenced doctors, surgical complications, and surgery-related mortality using multivariable logistic regression models, adjusting for the characteristics of the patient, surgeon, and hospital.

Results

On average, female surgeons had fewer post-registration years of experience than male surgeons (DG/TG; median 9 vs. 16 years, LAR; median 9 vs. 17 years, respectively), operated on higher-risk patients, and performed fewer laparoscopic surgeries than male surgeons (DG; 52.7% vs. 35.8%, TG; 26.3% vs 13.0, LAR; 69.6% vs. 60.4%; respectively). There was no significant difference between male and female surgeons in the adjusted risk for surgical mortality, surgical mortality combined with Clavien–Dindo grade ≥3 complications in DG, TG, and LAR, pancreatic fistula in DG and TG, or anastomotic leakage in LAR.

Conclusion

There was no significant adjusted risk difference in the outcomes of surgeries performed by male vs. female surgeons. Despite disadvantages, female surgeons take on high-risk patients and strive to improve their skills. Greater access to surgical training for female physicians is warranted in Japan.

Key words

gastrointestinal surgery; surgical outcomes; surgeon; gender equality

INTRODUCTION

According to the Organisation for Economic Co-operation and Development (OECD), the number of female physicians has been increasing worldwide in recent years. The percentage of female physicians was ≥40% in seven of the 27 OECD member countries in 2000, and in 21 of 26 countries in 2018.[1] Despite this increase, women remain a minority in the surgical field. Female general surgeons accounted for 27.9% (in 2019),[2] 22.0% (in 2019),[3] and 32.5% (in 2017)[4] of surgeons in Canada, the United States (US), and the United Kingdom, respectively.

In Japan, the proportion of female physicians is 21.8%, the lowest among the 27 countries listed in the Gender Gap Report,[1] and the proportion of female surgeons in general and gastrointestinal surgery is even lower, at 5.9%.[5] This suggests that the working environment in Japan poses more challenges for women looking to continue their careers and develop their skills for surgery than those posed by other listed countries. In this unique social environment, it is important to compare the outcomes of female and male surgeons to encourage women's choice of a career in surgery and/or to propose more effective training for female surgeons in Japan.

Previous studies in the US and Canada demonstrated that the proficiency of female physicians and surgeons was equal to or better than that of their male counterparts. Tsugawa et al. reported that, the mortality and readmission rates of older hospitalised patients treated by female physicians in the US were lower than those of such patients treated by male physicians.[6] In the US, no significant difference was found in postoperative mortality between female and male surgeons.[7] Moreover, there was no difference in the complication rates of surgeries performed by male vs female general surgeons in the US.[8] The postoperative mortality of patients operated on by Canadian female surgeons was slightly, but significantly, lower than that of patients operated on by male surgeons.[9]

To support the choice of surgical careers for women in Japan and to suggest more effective training for female surgeons in Japan, we compared the surgical outcomes of female and male surgeons using the Japanese National Clinical Database (NCD), which is the most extensive surgical database in Japan. We also examined the relationship between postoperative mortality and surgical complication rates and the surgeon's licencing terms.

METHODS

Study design and data source

We conducted a retrospective cohort study using data from the gastroenterological surgery section of the NCD. The NCD initiated data registration for surgical procedures in 2011.[10] By December 2019, 5,276 facilities were registered with the NCD. Approximately 1.5 million surgical cases are registered in this database each year, which is equivalent to over 95% of all surgeries in Japan.[11] The eligibility criteria for the NCD are accessible online (http://www.ncd.or.jp/). The NCD data entry system does not allow missing values except for laboratory data that were not taken from the patient. Validity of the data entries is evaluated through site visits and audits every year and has been proven to be high.[12] In addition to collecting data on all types of gastroenterological surgery, the NCD evaluates the quality of surgery for eight commonly performed surgical procedures with detailed data on preoperative, intraoperative, and postoperative factors. We analysed the outcomes of three of these eight surgical procedures, namely, distal gastrectomy (DG), total gastrectomy (TG), and low anterior resection (LAR). These three procedures were chosen because the number of female surgeons who performed these surgeries was sufficient for analysis. Other procedures among the aforementioned eight were considered difficult to analyse because fewer female surgeons performed these procedures. The NCD does not directly contain information regarding surgeons' gender or the number of years since the registration of licenced doctors, but it does contain the licence number of the surgeons. Thus, using these licence numbers, analysis was conducted by linking the NCD information with the gender profile and the year of licencing registration for the JSGS members.
Surgeries performed between 1 January 2013 and 31 December 2017 were included. Surgeries performed by non-JSGS members were excluded because non-JSGS members were assumed to be doctors specialising in other surgical fields, such as cardiovascular surgery. In Japan, these doctors need to complete a general surgery program, which includes performing gastroenterological surgery, to enter a subspeciality program. Therefore, they are considered to be separate from doctors who specialise in gastroenterological surgery, and the effect on outcome was also considered to be different for surgeries performed by these doctors. DG or TG surgeries not for gastric cancer and LAR surgeries not for colon cancer were excluded. Patients younger than 18 years, emergency surgery cases, those with unknown T/N factor in the TNM classification, and patients with metastasis were also excluded, because we aimed to assess the quality of surgery performed as standard or major procedures, which was considered to improve comparability. In addition, non-standard procedures may have complicated confounders

such as the treatment preferences of the patients and doctors, which are not available in the NCD.

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines were followed for this study.

Outcomes

Primary outcomes were surgical mortality, surgical mortality combined with severe postoperative complications, pancreatic fistula (in DG/TG only), and anastomotic leakage (in LAR only). In this study, surgical mortality was defined as all-cause death up to 30 days postoperatively, including death that occurred after discharge, and deaths that occurred within 90 days postoperatively during the index hospitalisation. The extended time frame for mortality during index hospitalisation was intended to provide sufficient time for the outcome to be captured, because nearly the same number of patients die between 30 and 90 days after surgery as those within 30 days.[13] This measure has been commonly used in previous NCD-based research to evaluate surgical outcomes.[13,14] Severe postoperative complications were defined as any postoperative surgical and medical complications with a Clavien–Dindo (CD) classification of ≥3 that occurred within 30 days postoperatively.[15] The CD classification was proposed by Dindo et al. for evaluating postoperative complications and comparing them among different hospitals, and a CD grade of ≥3 indicates that surgical, endoscopic, or radiological procedures are required for the treatment of the complication.[15] Pancreatic fistula was defined as a fistula of grade B or C according to the grading system proposed by the International Study Group of Pancreatic Fistula.[16] Anastomotic leakage was defined as leakage of luminal content observed in the drain, leakage requiring drainage, or leakage proven with images. Other outcomes included operation time and blood loss; these were considered intraoperative outcomes.

Statistical analysis

We used the chi-square test for categorical variables and the Mann–Whitney U test for continuous variables when comparing baseline characteristics and short-term outcomes. A multilevel multivariable logistic regression model for each surgical procedure was constructed, adjusting for patient, surgeon, and hospital characteristics, to examine the association between the surgeon's gender and surgery-related mortality or surgical complications. A multilevel model was used to account for unmeasured hospital-level characteristics.[17] Hospital identification (ID) was used as a random intercept. An adjusted odds ratio (OR) of >1 indicated a higher risk and an adjusted OR of <0 indicated a lower risk of the analysed outcome.
Patient characteristics included age (<70 vs ≥70 years), sex (male vs female), body mass index (≤18.5 vs >18.5 kg/m^2, <25 vs ≥25 kg/m^2), American Society of Anesthesiologists Physical Status (ASA-PS, 1–2 vs ≥3), clinical T factor (T1–2 vs T3–4) and N (0 vs 1–3) of tumours (N factor was included only for DG and TG; based on the Union for International Cancer Control–TNM classification, 7th edition), haemoglobin (male: <13.5 g/dL vs ≥13.5 g/dL, female: <11.5 g/dL vs ≥11.5 g/dL), aspartate aminotransferase level (<35 IU/L vs ≥35 IU/L; included in DG and TG), albumin level (<3.5 g/dL vs ≥3.5 g/dL), blood urea nitrogen level (<8 mg/dL vs ≥8 mg/dL), creatinine level (<1.2 mg/dL vs ≥1.2 mg/dL), presence/absence of diabetes mellitus, smoking status, habitual drinking status (only in LAR), dependence in activities of daily living (ADL), history of chronic obstructive pulmonary disease, dialysis, ischaemic heart disease, congestive heart failure, long-term steroid use, history of cardiovascular diseases (only in LAR), weight loss, preoperative blood transfusion, preoperative chemotherapy, and preoperative radiotherapy. These variables and categorisation were based on previous research and risk models using the NCD.[13,14] Continuous variables were categorised to account for a non-linear relationship between the variable and outcome. The surgical approach (open or laparoscopic) was included as an intraoperative factor.
Surgeon's characteristics included gender and years since licence registration in five-year increments. Years after medical licence registration were categorised based on the following assumptions to account for their acquired surgical skills in the Japanese board certification and surgery training system: surgeons with an experience of 5 years or less were considered to not have completed the general surgery training program; those with an experience of 6-10 years were assumed to be board certified general surgeons; 11-15 years, board certified gastroenterological surgeons; 16-20, board certified trainers; and 21 years or more, directors (or a similar position) of surgical departments.
Hospitals were categorised into quartiles according to the annual number of cases of each procedure so that each category contained approximately the same number of cases in order to increase statistical power: very low (VL), low (L), high (H), and very high (VH) (VL, L, H, and VH were defined for DG as: <15, 15 to <30, 30 to <50, and ≥50; TG: <7, 7 to <13, 13 to <21, and ≥21; LAR: <8, 8 to <16, 16 to <29, and ≥29, respectively). Based on previous research on the volume-outcome relationship, a non-linear association was assumed.[18]
Subsequently, additional analysis was conducted to examine whether an interaction effect existed between gender and years after medical licence registration. An interaction term of gender and years of experience post-medical licence registration was incorporated, instead of including them individually in the previous regression model.

Patients with missing data were excluded from this study because the proportion of cases with missing values was low in all three surgical procedures (DG 1.39%, TG 1.35%, LAR 1.64%).

Sensitivity analysis

To assess the robustness of the results, a number of analyses were performed after the completion of the main analysis.

First, although the proportion of missing values was low and a complete case analysis was conducted, cases with and without missing values were compared and the main analysis was repeated with a multiple-imputed dataset. The mechanism of missingness was assumed to be at random.[19] Imputation with a chained equation was conducted, and the number of imputed datasets was set to five.[20,21]

Second, patient's age, patient's body mass index, number of years after medical licence registration, and hospital case volume were included in the regression analysis as continuous variables instead of categorical variables. A generalised additive model was used to account for the assumed non-linearity between the variables and the outcome.[22]

Third, surgeon case volume and region of the hospital were added to the regression model. We assume that case volume is a surrogate of surgical experience that significantly affects outcome. For hospitals, hospital case volume would reflect surgical experience. Meanwhile, for surgeons, years after licence registration were considered to be a more accurate measure of surgical experience than annual case volume of the individual surgeon because the years after licence registration account for surgical experience during the entire professional career, not just for the surgical experience of that year. However, considering that the surgeon case volume may be a confounder, it was included as an additional variable in the regression model. It was treated as a continuous variable, and to model a non-linear relationship, a smooth term of a generalised additive model was applied.[22] The region of the hospital was additionally included as a variable to partly account for the socioeconomic status (SES) of a patient. SES is not available in the NCD, and research on the relationship between SES and surgical outcome in Japan is scarce. One study in Japan found no significant association between regional average income, which was considered to be one aspect of SES, and outcome in cardiovascular surgery,[23] but it is unknown whether regional mean household income reflects an individual's SES and whether the results can be applied in gastroenterological surgery. Therefore, the considerable magnitude of SES as a confounder could not be denied. The region of a hospital was categorised into urban or rural areas based on those used in a previous Japanese study, which distinguished urban areas from rural ones according to the OECD definition.[24] Thirteen out of 47 prefectures were categorised as urban. Additionally, this factor could serve partly as a hospital-level characteristic that affects the assignment of surgeons based on gender and surgical outcome.

Fourth, as the number of surgeries performed by female surgeons was low and because a small number of female surgeons may have an extreme effect on the outcome or on the results, the study population for DG, TG, and LAR were combined and analysed as a single population. The type of surgical procedure was included as a covariate and the main analysis was repeated. The relationships between a surgeon's gender and surgical mortality, surgical mortality or postoperative complication with CD classification $\geq$3, and anastomotic leakage were assessed.

Fifth, since female surgeons were found to be more likely to be assigned to higher risk patients, surgical outcomes were compared between male and female surgeons within the predicted risk strata. The predicted risk was calculated based on the regression analysis; the doctor's gender was excluded as a variable. The predicted risk was categorised into five strata, from low to high risk, using the quintile of predicted risk.

All p-values were two-sided, and p-values <0.05 were considered significant. Statistical analyses were performed using R software (version 3.6.3, 2020; R Foundation for Statistical Computing, Vienna, Austria).

Patient and public involvement

Although patients and the public were not involved in the conception, design, or implementation of this study, we wish to publicise the study results among patients and the public to raise awareness regarding the surgical outcomes of female surgeons being comparable to those of their male counterparts. In the Japanese society, it has been a concern that women spend more time engaged in housework and childcare, making it difficult for them to work in a profession such as surgery. We would like to widely publicise these results through the media and public symposiums to encourage women's participation in professional fields, including surgery.

RESULTS

Study population

This study investigated 184,238, 83,487, and 107,721 patients who underwent DG, TG, and LAR, respectively, at Japanese institutes and were registered in the Japanese NCD between 2013 and 2017. The flow diagram for surgical case selection is shown in Fig 1. Finally, 149,193 DG, 63,417 TG, and 81,593 LAR surgeries were eligible.

A total of 140,971 (94.5%) eligible DG surgeries were performed by male surgeons and 8,222 (5.5%) by female surgeons; 59,915 (94.5%) eligible TG surgeries were performed by male surgeons and 3,502 (5.5%) by female surgeons; and 77,864 (95.4%) eligible LAR procedures were performed by male surgeons and 3,729 (4.6%) by female surgeons. The numbers of male surgeons who participated in DG, TG, and LAR were 9,433 (92.3%), 8,238 (92.8%), and 8,200 (92.9%), respectively, and those of female surgeons were 788 (7.7%), 640 (7.2%), and 627 (7.1%), respectively (Table 1). Female surgeons had fewer years of experience post-licence registration than male surgeons (9 vs. 16 years in DG/TG, and 9 vs. 17 years in LAR).

Characteristics of institutions and patients

The institutional factors, preoperative and intraoperative factors, intraoperative outcomes, and postoperative outcome of DG, TG, and LAR are presented in Tables 2, 3, and 4, respectively. Regarding DG, female surgeons were more distributed in hospitals with L (28.4%) and H (27.1%) than in those with VL (22.7%) or VH (21.8%) case numbers. Regarding TG, female surgeons were less distributed in hospitals with VH case numbers (20.7%) than in those in other categories. Regarding LAR, female surgeons were more typically distributed in hospitals with L (29.5%) numbers than in those with VL (23.3%), H (24.0%), or VH (23.2%) numbers.

Female surgeons performed surgeries on relatively high-risk patients. Importantly, female surgeons performed surgeries on older patients (DG, 58.9% vs. 55.6%; TG, 60.4% vs. 56.4%; LAR, 45.9% vs. 43.8%) and patients with diabetes mellitus (DG, 19.2% vs. 18.1%; TG, 21.2% vs. 18.6%; LAR, 19.4% vs. 18.0%), dependence in ADL (DG, 4.9% vs. 4.2%; TG, 4.8% vs. 3.8%; LAR, 4.6% vs. 3.5%), lower haemoglobin (DG, 29.9% vs. 27.9%; TG, 37.6% vs. 35.2%; LAR, 28.7% vs. 27.0%) and serum albumin (DG, 18.5% vs. 15.0%; TG, 22.5% vs. 19.2%; LAR, 14.3% vs. 12.1%) levels, and higher T factors (DG, 35.7% vs. 30.1%; TG, 58.3% vs. 55.1%; LAR, 63.7% vs. 60.0%) in all three procedures at a higher rate than their male counterparts (Tables 2–4). Additionally, female surgeons performed DG in patients with long-term steroid use (1.3% vs. 1.0%), weight loss (4.7% vs. 3.6%), preoperative blood transfusion (3.1% vs. 2.0%), a higher N factor (37.2% vs. 31.3%), and a worse ASA-PS (12.2% vs. 11.0%); TG for patients that smoked (52.0% vs. 49.2%), under dialysis (1.0% vs. 0.6%), weight loss (6.7% vs. 5.8%), and a higher N factor (51.4% vs. 48.9%); and LAR for patients with a history of cardiovascular disease (3.9% vs. 3.1%) at a higher rate than their male counterparts. In contrast, male surgeons performed surgeries on patients who had undergone preoperative chemotherapy (DG, 2.2% vs. 1.9%; TG, 6.9% vs. 5.5%; LAR, 6.5% vs. 4.2%) in DG, TG, and LAR or radiotherapy (3.1% vs. 1.7%, in LAR) at a higher rate than their female counterparts.

Intraoperative factors and outcomes

Female surgeons performed fewer laparoscopic surgeries (DG, 35.8% vs. 52.7%; TG, 13.0% vs. 26.3%; LAR, 60.4% vs. 69.6%) than male surgeons. There was significantly more blood loss observed in all three procedures performed by female surgeons (DG, 150 vs. 100 mL; TG, 320 vs. 260 mL; LAR 80 vs. 52 mL) than by male surgeons (Tables 2–4).

Postoperative outcomes

After adjusting for patient characteristics, surgeon characteristics, and hospital characteristics, no significant difference was noted in the risk for surgical mortality in DG, TG, and LAR between male and female surgeons as shown in Fig 2 (DG, risk-adjusted OR 0.98, 95% confidence interval [CI] 0.74 to 1.29; TG, risk-adjusted OR 0.83, 95% CI 0.57 to 1.19; LAR, risk-adjusted OR 0.56, 95% CI 0.30 to 1.05). The adjusted risk for surgical mortality or postoperative complication rated CD-3 or higher were similar for DG, TG, and LAR (DG, risk-adjusted OR 1.03, 95% CI 0.93 to 1.14; TG, risk-adjusted OR 0.92, 95% CI 0.81 to 1.05; LAR, risk-adjusted OR 1.02, 95% CI 0.91 to 1.15), pancreatic fistula for DG and TG (DG, risk-adjusted OR 1.16, 95% CI 0.97 to 1.38; TG, risk-adjusted OR 1.02, 95% CI 0.84 to 1.23), and anastomotic leakage for LAR (risk-adjusted OR 1.04, 95% CI 0.92 to 1.18) between male and female surgeons.

Interaction between surgeons' gender and years since registration of medical licence

For the sub-analysis, we compared surgical outcomes between male and female surgeons in the year-since-licencing categories (Fig 3–5).

DG

The adjusted risk for surgical mortality was higher for female surgeons than for male surgeons with ≤5 years of experience after registration (risk-adjusted OR 1.64, 95% CI 1.07 to 2.52). For surgery-related death or postoperative adverse events rated CD-3 or higher, female surgeons with ≤5 years of experience after registration had a higher OR (risk-adjusted OR 1.19, 95% CI 1.01 to 1.41), whereas those with 6–10 years of experience after registration had a lower OR (risk-adjusted OR 0.79, 95% CI 0.65 to 0.96) than male surgeons. The adjusted risk for pancreatic fistula showed no significant difference between male and female surgeons at any year category after registration.

TG

The adjusted risk for surgery-related death, postoperative adverse events rated CD-3 or higher, and pancreatic fistula showed no significant differences between male and female surgeons at any year category after registration.

LAR

The adjusted OR for surgical mortality did not differ significantly between male and female surgeons at any year-since-licencing category. The adjusted risk for surgical mortality or postoperative adverse events rated CD-3 or higher were higher for female surgeons than for male surgeons at the 16–20 years of experience category (risk-adjusted OR 1.41, 95% CI 1.07 to 1.86). The adjusted risk for anastomotic leakage was lower for female surgeons with ≤5 years of experience (risk-adjusted OR 0.71, 95% CI 0.53 to 0.94).

Sensitivity analysis

The results of the sensitivity analyses are summarised in supplementary Fig 1–4 and supplementary tables 1–7. A relatively higher proportion of missing values was observed among laboratory data; however, the proportion of missingness for all factors was below 1%. Female surgeons had a lower case volume and tended to work at hospitals in urban areas for all three surgical procedures. In the analyses with missing values imputed, patient's age, body mass index, hospital case volume, and years after medical licence registration changed to the original continuous scale, and additional covariates, i.e. surgeon case volume and urban-rural status included, the changes in the point estimate and 95% CI were minimal compared with those in the main analysis (supplementary Fig 1–3), except for one of the analyses in LAR that included surgeon case volume and region of the hospital as additional covariates. As shown in supplementary Fig 3, this analysis revealed a significant decrease in adjusted OR for surgical mortality for female surgeons (adjusted OR 0.54, 95% CI 0.29 to 0.996). There were no significant differences between male and female surgeons when stratified with predicted risks of the outcome except for the 5th quintile of predicted risk in pancreatic fistula in DG (number of outcomes for male surgeons, 1,365 (4.9%) vs. female surgeon, 115 (6.1%), p=0.02; supplementary table 5) and the 2nd quintile of predicted risk in anastomotic leakage in LAR (number of outcomes for male surgeons, 676 (4.3%) vs. female surgeons, 49 (6.4%), p=0.008; supplementary table 7). Finally, the analysis with the three surgical procedures combined showed no significant association between female surgeons and surgical outcomes (supplementary Fig 4).

DISCUSSION

Principal findings

Using the NCD data for 2013–2017, we found no overall significant difference in the risk after confounder adjustment for surgical mortality in the three procedures performed by male and female surgeons. There was also no significant difference in the adjusted risk for surgical mortality or CD-3 or higher complications in DG, TG, and LAR, pancreatic fistula in DG and TG, and anastomotic leakage in LAR between male and female surgeons. More blood loss was recorded in all three procedures performed by female surgeons, probably because they performed a significantly larger proportion of open surgeries than male surgeons. Importantly, we found that female gastrointestinal surgeons were more often responsible for patients with comorbid conditions (e.g., diabetes mellitus, anaemia, dependence in ADL, etc), even though, female gastrointestinal surgeons were responsible for fewer surgeries than male surgeons, as Altieri et al. has described.[25] Data from a large institution in a Western country indicated that female surgeons did not perform more complex cases than male surgeons, even after accounting for subspecialty and seniority.[26] This situation is different from that in Japan, as reported in our analysis. The number of surgeries performed per surgeon will be analysed more -precisely in our subsequent report, as it is a very crucial problem in the Japanese surgical society.

As a subgroup analysis, we compared the post-registration years and found differences in the risk for surgical outcomes between male and female surgeons. For DG performed by female surgeons with an experience of ≤5 years post-registration, the adjusted odds ratio for 'surgical mortality' and 'surgical mortality with a complication grade of CD-3 or higher' were statistically higher than those for male surgeons of the same category. For LAR, females with an experience of 16-20 years had a statistically higher adjusted risk for 'surgical mortality' than males with the same surgical experience. However, the adjusted risks for 'surgical mortality or a complication grade of CD-3 or higher' in DG performed by female surgeons with 6-10 years of experience was lower than those for males, and the rate of leakage in LAR performed by female surgeons with ≤5 years of experience was lower than that for males. Female surgeons in other subgroups in DG and LAR and in all subgroups in TG tended to have comparable surgical outcomes to male counterparts.

Further, in the category of surgeons with ≥21years of experience, no significant difference in outcomes was observed between male and female surgeons in all three surgical procedures. Tsugawa et al. reported that the risk-adjusted mortality rate in surgeries performed by female surgeons aged >50 years was the lowest; however, they mentioned that it was difficult to evaluate the outcome of female surgeons aged >60 years because this group was very small.[7] These findings are consistent with ours. Wallis et al. reported a lower 30-day mortality rate for surgeries performed by female surgeons.[9] Sharoky et al. reported no difference in mortality or complication rates for surgeries performed by male and female surgeons using cardinality matching with a refined balance.[8] However, these authors did not compare surgeons by age. Further research is required to examine how gender and age affect surgical outcomes, but it is necessary to note that the low volume of senior female surgeons is a particular concern.

The results of the sensitivity analyses differed minimally from those of the main analyses. In the analysis that additionally adjusted for the confounding effects of surgeon case volume and urban-rural status in LAR, a significant decrease in the adjusted OR for surgical mortality was observed for female surgeons. The significant difference in surgical outcomes between female and male surgeons after adjusting for the small number of procedures performed by female surgeons suggests that women may improve their outcomes further as they gain surgical experience. The risk-stratified comparison between male and female surgeons revealed non-significant differences in almost all stratified risk groups of the three surgical procedures. Two significant results favoured male surgeons in terms of better outcomes. Considering the multiple comparisons in this analysis, a type I error is likely to occur; therefore, the result would not alter the conclusion in the main analysis regarding the lack of significant differences in surgical outcomes between male and female surgeons.

Comparison with other studies

The proportion of female gastrointestinal surgeons in Japan is small; this was 7.1% in 2021 and 6.0% in 2015, the middle of the period covered by this study. In 2011, no gender data were available from the Office of the JSGS. First, the lack of role models is often pointed out as a barrier to female surgeons' careers,[27] and female surgeons experience interprofessional conflict due to breakdowns in communication.[28] Moreover, it is difficult for female surgeons to attain leadership positions.[29] Second, previous reports have shown a bias in the number of surgical cases assigned to male vs female surgeons during their training.[30] Foley et al.[31] reported gender differences in the robotic surgery experience in colorectal surgery training programmes, with female trainees having fewer opportunities to participate in the use of consoles and to complete the procedures. They also reported that male supervisors provided fewer console participation opportunities to female residents than to male residents, but female supervisors provided the same number of console-use opportunities to both female and male trainees. Female surgeons, as supervisors, may provide female residents with equitable training opportunities. Generally, in Japan, patients cannot nominate a primary surgeon, and primary surgeons are assigned to each surgery at random or at the discretion of the department head discretion; thus, the process for case assignment to female surgeons by supervisors is essential in the training process for female surgeons.

Third, in Japanese society, women are often viewed from a biased perspective. In 2018, gender discrimination was reported in admission tests for several medical schools, which had manipulated the scores of female applicants to interfere with their admissions. The admissions committees of these medical schools wanted to enrol more men, since women often leave clinical practice due to marriage, pregnancy, or childcare.[32] In traditional Japanese culture, women have often been considered unsuitable for performing surgery and are unwelcome in the field. We believed that showing that there were no differences in the results of surgical procedures performed by men and women would make it easier for women to be accepted as surgeons and professionals.

Fourth, work-family conflict is more pronounced among female surgeons, and they may experience burn-out.[33,34] Many aspects can impair the successful development of female surgeons. Nevertheless, in the present analysis, there was no significant difference in the mortality or complication rates of surgeries performed by female and male surgeons, suggesting that they are equally successful in developing their surgical skills. Notably, female surgeons performed a lower percentage of laparoscopic procedures in all three procedures than male surgeons did. There may have been a tendency for male surgeons to be assigned to laparoscopic procedures, which may require more time to develop experience. The percentage of women in the JSGS is gradually increasing. It is warranted that surgical teams welcome women as members and that gender equality is achieved in Japanese gastrointestinal surgery training. The three surgical procedures we analysed are only representative, but we believe that equality in training, inclusion, mentoring, and practice across the genders would produce better outcomes in medicine.

Strengths and limitations of study

The primary strength of our study is that we used the NCD, a comprehensive database, and adjusted for confounders with patient-related factors for the individual procedures selected. Many previous studies have used the Medicare claims database. By contrast, we used a clinical database such as NCD, which is highly accurate in terms of patients' preoperative condition and surgical outcomes.

This study had some limitations. First, this was an observational study, and we could not adjust for unmeasured confounders. Certain data, such as that regarding the SES of a patient, were not available in the NCD. Second, because the number of female surgeons was smaller than that of male surgeons, there may be a bias in that the outcomes of one female surgeon had a large effect on the overall outcomes. When interpreting the results, it is important to note that because there are so few female surgeons, a single adverse event can significantly impact the entire result; this is not the case for male surgeons. Third, the study included in this research paper lacks details regarding surgeons' work and personal life conditions (part-time or full-time, family structure, etc.). Fourth, since we intended to include only patients with relatively standard procedures operated by gastroenterological surgeons, our findings may not be applicable to non-standard procedures, emergency surgeries, surgeries performed by surgeons with other specialties, or other types of surgical procedures.

Conclusions

Based on our study, Japanese female surgeons took on high-risk cases, and there were no significant differences in surgical mortality or CD-3 or higher complication rates between patients operated on by male or female surgeons. We found that female surgeons were successful in developing their technical skills. More appropriate and effective surgical training for female surgeons could further improve surgical outcomes.

CONTRIBUTORS AND GUARANTOR

KO, SN, EK, YF, and KH designed the study. KY, IY, YK, and YK collected the data. HE, HY, and HM analysed the data. KO wrote the first draft of the manuscript. All authors read the drafted manuscript, provided feedback, and approved the final submitted version. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

FUNDING

This study received no financial support.

COMPETING INTERESTS

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare the following: HE, HY, and HM report grants from the National Clinical Database, Johnson & Johnson, and Nipro. KY reports grants from Abbott, Asahi Kasei Pharma, Astellas Pharma, Chugai Pharmaceutical, Covidien Japan, Daiichi Sankyo, Eisai, Eli Lilly Japan, Johnson & Johnson, Kaken Pharmaceutical, Kyowa Kirin, Nippon Kayaku, Otsuka Pharmaceutical, Sanofi, Taiho Pharmaceutical, Takeda Pharmaceutical, TERUMO, Tsumura, and Yakult Honsha; research funding from Abbvie, Biogen Japan, Celgene, EP-CRSU, EPS Corporation, FUJIFILM, GlaxoSmithKline, Meiji Seika Pharma, MSD, Novartis, Ono Pharmaceutical, Philips Japan, and ShiftZero; and personal fees from AstraZeneca, Bristol-Myers Squibb, Chugai Pharmaceutical, and Covidien Japan. YK reports grants from Takeda Pharmaceutical, Chugai Pharmaceutical, Yakult Honsha, Asahi Kasei Pharma, Otsuka Pharmaceutical, Ono Pharmaceutical, Tsumura, Kyowa Hakkkou Kirin, Dainippon Sumitomo Pharma, EA Pharma, Astellas Pharma, Toyama Chemical, MEDICON, Kaken Pharmaceutical, Eisai, Otsuka Pharmaceutical, Teijin Pharma, Nihon Pharmaceutical, and Nippon Covidien; and personal fees from Ono Pharmaceutical, Bristol-Myers Squibb, Chugai Pharmaceutical, Taiho Pharmaceutical, Asahi Kasei Pharma, Otsuka Pharmaceutical, Shionogi & Company, Nippon Covidien, Ethicon, Ono Pharmaceutical, Olympus, Bristol-Myers Squibb, AstraZeneca, MSD, Smith & Nephew, and Kaken Pharmaceutical. All remaining authors declare no support from any organisation for the submitted work, no financial relationships in the past three years with any organisations that might have an interest in the submitted work, and no other relationships or activities that could appear to have influenced the submitted work.

ETHICAL APPROVAL

This study was approved by the Ethics Committee of Japan Baptist Hospital (approval no. 19-1 Apr 2019), and written informed consent was not required because of the anonymous nature of the data. Regarding patient data registration in the NCD, each participating institution provided patients with the opportunity to opt-out of the study after their respective ethical committee review and approval. Regarding the use of data related to surgeons, members of the Japanese Society of Gastroenterological Surgery (JSGS) were provided with the opportunity to opt out via e-mail messages and through a website.

ACKNOWLEDGMENTS

We sincerely appreciate all the participants in the NCD project for their extraordinary efforts in data registration, and grateful to the JSGS members who participated in this study. We also thank Prof. James R. Goldenring for his proofreading assistance and Editage (www.editage.com) for English language editing.

DATA SHARING

Data on individual surgical cases and surgeons reported in this study are not publicly available. To access the aggregate data, including data reported in this study, please submit a research plan and request access to the NCD Office, usually through an NCD-related society (such as the JSGS). If the proposal is approved, the de-identified data (including participant and related data, if necessary) can be assessed by a statistics specialist affiliated with the NCD.

TRANSPARENCY STATEMENT

The corresponding author (SN) affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

COPYRIGHT FOR PUBLICATION

REFERENCES

1 Organisation for Economic Co-Operation and Development. Health care resources: Physicians by age and gender; https://stats.oecd.org/index.aspx?queryid=30172# (accessed 20 Jan 2022).
.

2 Canadian Medical Association. Number and percent distribution of physicians by specialty and gender, Canada 2019; https://www.cma.ca/sites/default/files/2019-11/2019-06-spec-sex_0.pdf (accessed 20 Jan 2022).

3 Association of American medical colleges. Active Physicians by Sex and Specialty, 2019; 2019; https://www.aamc.org/data-reports/workforce/interactive-data/active-physicians-sex-and-specialty-2019 (accessed 20 Jan 2022).

4 NHS Digital. Analysis of the representation of women across the hospital and community health services workforce, 2018; https://digital.nhs.uk/data-and-information/find-data-and-publications/supplementary-information/2018-supplementary-information-files/analysis-of-the-representation-of-women-across-the-hospital-and-community-health-services-workforce (accessed 20 Jan 2022).

5 Ministry of Health, Labour and Welfare (Japan). Survey of physicians, dentists, and pharmacists; 2018, https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00450026&tstat=000001030962&cycle=7&tclass1=000001109395&tclass2=000001109396&stat_infid=000031653231&tclass3val=0 (accessed 20 Jan 2022).

6 Tsugawa Y, Jena AB, Figueroa JF, Orav EJ, Blumenthal DM, Jha AK. Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs female physicians. *JAMA Intern Med* 2017;177:206–13.

7 Tsugawa Y, Jena AB, Orav EJ, et al. Age and sex of surgeons and mortality of older surgical patients: observational study. *BMJ* 2018;361:k1343.

8 Sharoky CE, Sellers MM, Keele LJ, et al. Does surgeon sex matter?: Practice patterns and outcomes of female and male surgeons. *Ann Surg* 2018;267:1069–76.

9 Wallis CJ, Ravi B, Coburn N, Nam RK, Detsky AS, Satkunasivam R. Comparison of postoperative outcomes among patients treated by male and female surgeons: A population based matched cohort study. *BMJ* 2017;359:j4366.

10 Seto Y, Kakeji Y, Miyata H, Iwanaka T. National Clinical Database (NCD) in Japan for gastroenterological surgery: Brief introduction. *Ann Gastroenterol Surg* 2017;1:80–1.

11 Kakeji Y, Takahashi A, Hasegawa H, et al. Surgical outcomes in gastroenterological surgery in Japan: Report of the National Clinical Database 2011–2018. *Ann Gastroenterol Surg* 2020;4:250–74.

12 Hasegawa H, Takahashi A, Kanaji S, et al. Validation of data quality in a nationwide gastroenterological surgical database: The National Clinical Database site-visit and remote audits, 2016-2018. *Ann Gastroenterol Surg* 2021;5:296–303.

13 Gotoh M, Miyata H, Hashimoto H, et al. National Clinical Database feedback implementation for quality improvement of cancer treatment in Japan: from good to great through transparency. *Surg Today* 2016;46:38–47.

14 Hoshino N, Endo H, Hida K, et al. Emergency surgery for gastrointestinal cancer: A nationwide study in Japan based on the National Clinical Database. *Ann Gastroenterol Surg* 2020;4:549–61.

15 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;240:205–13.

16 Bassi C, Marchegiani G, Dervenis C, et al. The 2016 update of the International Study Group (ISGPS) definition and grading of postoperative pancreatic fistula: 11 years After. *Surgery* 2017;161:584–91.

17 Finch WH, Bolin JE, Kelly K. Multilevel modeling using R. Boca Raton (FL): CRC Press; 2014.

18 Iwatsuki M, Yamamoto H, Miyata H, et al. Effect of hospital and surgeon volume on postoperative outcomes after distal gastrectomy for gastric cancer based on data from 145,523 Japanese patients collected from a nationwide web-based data entry system. *Gastric Cancer* 2019;22:190–201.

19 Little RJ, Rubin DB. Statistical Analysis with Missing Data. 2nd ed. New York: John Wiley & Sons, Inc; 2002.

20 van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw* 2011;45:1–67.

21 Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999;8:3–15.

22 Wood SN. Generalized additive models: an introduction with R. 2nd ed. Boca Raton: CRC press; 2017.

23 Lee SL, Hashimoto H, Kohro T, et al. Influence of municipality-level mean income on access to aortic valve surgery: a cross-sectional observational study under Japan's universal health-care coverage. *PLoS ONE* 2014;9:e111071.

24 Sakai R, Wang W, Yamaguchi N, et al. The impact of Japan's 2004 postgraduate training program on intra-prefectural distribution of pediatricians in Japan. *PLoS ONE* 2013;8:e77045.

25 Altieri MS, Yang J, Bevilacqua L, Zhu C, Talamini M, Pryor AD. Women are underrepresented as top surgical performers as assessed by case volumes in the state of New York. *Am J Surg* 2018;216:666–71.
26 Chen YW, Westfal ML, Chang DC, Kelleher CM. Underemployment of female surgeons?. *Ann Surg* 2021;273:197–201.
27 Barshes NR, Vavra AK, Miller A, Brunicardi FC, Goss JA, Sweeney JF. General surgery as a career: A contemporary review of factors central to medical student specialty choice. *J Am Coll Surg* 2004;199:792–9.
28 Dossett LA, Vitous CA, Lindquist K, Jagsi R, Telem DA. Women surgeons' experiences of interprofessional workplace conflict. *JAMA Network Open* 2020;3:e2019843.
29 Leschber G. From female surgical resident to academic leaders: Challenges and pathways forward. *J Thorac Dis* 2021;13:480–4.
30 Gill HK, Niederer RL, Danesh-Meyer HV. Gender differences in surgical case volume among ophthalmology trainees. *Clin Exp Ophthalmol* 2021;49:664-71.
31 Foley KE, Izquierdo KM, von Muchow MG, Bastawrous AL, Cleary RK, Soliman MK. Colon and rectal surgery robotic training programs: An evaluation of gender disparities. *Dis Colon Rectum* 2020;63:974–9.
32 Oshima K, Ozaki A, Mori J, Takita M, Tanimoto T. Entrance examination misogyny in Japanese medical schools. *Lancet* 2019;393:1416.
33 Dyrbye LN, Shanafelt TD, Balch CM, Satele D, Sloan J, Freischlag J. Relationship between work-home conflicts and burnout among American surgeons: A comparison by sex. *Arch Surg* 2011;146:211–7.
34 Lu PW, Columbus AB, Fields AC, Melnitchouk N, Cho NL. Gender differences in surgeon burnout and barriers to career satisfaction: A qualitative exploration. *J Surg Res* 2020;247:28–33.

Figure legends

Fig 1 Flow diagram for patient selection. DG, distal gastrectomy; TG, total gastrectomy; LAR, low anterior resection; JSGS, The Japanese Society of Gastroenterological Surgery; TX, Unknown; T, NX; Unknown N, M1, Positive for distant metastasis.

Fig 2 Association between female surgeon and surgical outcome. OR, odds ratio; CI, confidence interval; CDC, Clavien-Dindo classification.

Fig 3 Association between female surgeon and surgical outcome according to years after medical licence registration in distal gastrectomy. OR, odds ratio; CI, confidence interval; CDC, Clavien-Dindo classification.

Fig 4 Association between female surgeon and surgical outcome according to years after medical licence registration in total gastrectomy. OR, odds ratio; CI, confidence interval; CDC, Clavien-Dindo classification.

Fig 5 Association between female surgeon and surgical outcome according to years after medical licence registration in low anterior resection. OR, odds ratio; CI, confidence interval; CDC, Clavien-Dindo classification.

Summary boxes

What is already known on this topic:

- Women remain a minority in the surgical field, particularly in Japan
- In the United States and Canada, the proficiency of female physicians and surgeons was equal to or better than that of their male counterparts.

What this study adds:

- We found no overall significant differences in surgical mortality or CD-3 or higher complication rates associated with the three procedures (distal gastrectomy, total gastrectomy, and low anterior resection) performed by Japanese male and female surgeons.
- More opportunities and encouragement should be provided to female surgeons to address the gender-based inequity in the field of surgery.

Table 1 Surgeon characteristics by gender

Distal gastrectomy								
		Male surgeon	Female surgeon				Number of male surgeons	Number of female surgeons
Total cases of operation (%)		140971 (94.5)	8222 (5.5)		Number of surgeons (%)		9433 (92.3)	788 (7.7)
Years since registration of licensed doctors median [IQR]		16 [9, 22]	9 [5, 13]					
Years since registration of licensed doctors (%)					Years since registration of licensed doctors (%)			
	- 5	19246 (88.4)	2534 (11.6)			- 5	2461 (85.1)	432 (14.9)
	6 - 10	21526 (89.9)	2430 (10.1)			6 - 10	2432 (88.5)	315 (11.5)
	11 - 15	27084 (93.5)	1898 (6.5)			11 - 15	2273 (92.6)	181 (7.4)
	16 - 20	28609 (97.0)	881 (3.0)			16 - 20	2286 (96.0)	96 (4.0)
	21 -	44506 (98.9)	479 (1.1)			21 -	3528 (98.7)	48 (1.3)
Total gastrectomy								
		Male surgeon	Female surgeon				Number of male surgeons	Number of female surgeons
Total cases of operation (%)		59915 (94.5)	3502 (5.5)		Number of surgeons (%)		8238 (92.8)	640 (7.2)
Years since registration of licensed doctors median [IQR]		16 [9, 23]	9 [5, 14]					
Years since registration of licensed doctors (%)					Years since registration of licensed doctors (%)			
	- 5	7959 (87.7)	1115 (12.3)			- 5	2025 (86.1)	328 (13.9)
	6 - 10	9097 (90.2)	989 (9.8)			6 - 10	2026 (89.3)	244 (10.7)
	11 - 15	11204 (93.7)	749 (6.3)			11 - 15	1931 (93.2)	142 (6.8)
	16 - 20	11956 (96.4)	441 (3.6)			16 - 20	1924 (95.8)	85 (4.2)
	21 -	19699 (99.0)	208 (1.0)			21 -	2953 (98.7)	39 (1.3)
Low anterior resection								
		Male surgeon	Female surgeon				Number of male surgeons	Number of female surgeons

Total cases of operation (%)		77864 (95.4)	3729 (4.6)		Number of surgeons (%)		8200 (92.9)	627 (7.1)
Years since registration of licensed doctors median [IQR]		17 [11, 23]	9 [6, 15]					
Years since registration of licensed doctors (%)					Years since registration of licensed doctors (%)			
	- 5	7066 (88.9)	885 (11.1)			- 5	1864 (86.3)	296 (13.7)
	6 - 10	10576 (89.8)	1198 (10.2)			6 - 10	2007 (88.9)	251 (11.1)
	11 - 15	15643 (94.8)	853 (5.2)			11 - 15	2038 (93.1)	152 (6.9)
	16 - 20	17698 (96.9)	562 (3.1)			16 - 20	2072 (96.1)	83 (3.9)
	21 -	26881 (99.1)	231 (0.9)			21 -	3000 (98.6)	43 (1.4)

Note: The number of surgeons in each category does not add up to the total number of surgeons in the study population because some surgeons moved to a higher category (in terms of seniority) during the study period.

Table 2 Institutional and operative characteristics by surgeon's gender in DG

		Male surgeon	Female surgeon	P value
Total cases of operation		140971	8222	
Factor	**Category**			
Institutional factor				
Number of surgeries per year		30 [15, 54]	29 [16, 52]	0.73
Number of surgeries per year (%)				<0.001
	<15	34733 (24.6)	1867 (22.7)	
	15≤, <30	35826 (25.4)	2337 (28.4)	
	30≤, <50	36092 (25.6)	2226 (27.1)	
	50≤	34320 (24.3)	1792 (21.8)	
Preoperative factor				
Age, median [IQR]		71 [64, 78]	72 [65, 78]	<0.001
Age (%)	70≤	78418 (55.6)	4840 (58.9)	<0.001
Sex (%)	Female	46798 (33.2)	2820 (34.3)	0.04
Body mass index (kg/m2), median [IQR]		22.2 [20.0, 24.4]	22.0 [19.8, 24.3]	<0.001
Body mass index (kg/m2) (%)				<0.001
	≥18.5 <25	95141 (67.5)	5556 (67.6)	
	<18.5	17118 (12.1)	1119 (13.6)	
	≥ 25	28712 (20.4)	1547 (18.8)	
Diabetes mellitus (%)	+	25484 (18.1)	1579 (19.2)	0.01
Smoking (%)	+	63731 (45.2)	3777 (45.9)	0.20
Dependence in ADL (%)	+	5965 (4.2)	401 (4.9)	0.005
COPD (%)	+	6822 (4.8)	412 (5.0)	0.50
Dialysis (%)	+	1062 (0.8)	77 (0.9)	0.07
History of IHD (%)	+	5260 (3.7)	332 (4.0)	0.16
Congestive heart failure (Within 30 days)(%)	+	976 (0.7)	67 (0.8)	0.22
Lomg-term steroid use (%)	+	1424 (1.0)	105 (1.3)	0.02

Weight loss (%)	+	5046 (3.6)	386 (4.7)	<0.001
Preoperative blood transfusion (%)	+	2859 (2.0)	251 (3.1)	<0.001
Hemoglobin (%)	Male: < 13.5, Female: < 11.5	39344 (27.9)	2459 (29.9)	<0.001
Albumin (%)	<3.5	21128 (15.0)	1519 (18.5)	<0.001
BUN (%)	<8	19371 (13.7)	1158 (14.1)	0.39
Creatinine (%)	>1.2	9961 (7.1)	626 (7.6)	0.06
AST > 35 (%)		9542 (6.8)	581 (7.1)	0.31
Preoperative chemotherapy (%)		3092 (2.2)	153 (1.9)	0.049
Preoperative radiotherapy (%)		151 (0.1)	6 (0.1)	0.45
T factor (in the TNM cllasification) (%)	T3≤	42441 (30.1)	2939 (35.7)	<0.001
N factor (in the TNM cllasification) (%)	N1≤	44193 (31.3)	3056 (37.2)	<0.001
ASA-PS (%)	3, 4, 5	15563 (11.0)	1006 (12.2)	0.001
Intraoperative factor				
Surgical approach (open or laparoscopic) (%)	Laparoscopic surgery	74282 (52.7)	2944 (35.8)	<0.001
Intraoperative outcomes				
Operating time (min), median [IQR]		259 [205, 320]	261 [209, 322]	0.001
Estimated blood loss (mL), median [IQR]		100 [25, 250]	150 [50, 327]	<0.001
Postoperative outcomes				
Postoperative hospital stay (days), median [IQR]		13 [10, 19]	14 [10, 20]	<0.001

Abbreviations: DG, distal gastrectomy; IQR, interquartile range; ADL, activities of daily living; IHD, ischemic heart disease; COPD, chronic obstructive pulmonary disease; BUN, blood urea nitrogen; AST, aspartate aminotransferase; ASA-PS, American Society of Anesthesiologists Physical Status

Table 3 Institutional and operative characteristics by surgeon's gender in TG

		Male surgeon	Female surgeon	P value
Total cases of operation		59915	3502	
Factor	**Category**			
Institutional factor				
Number of surgeries per year		13 [7, 22]	13 [7, 20]	0.09
Number of surgeries per year (%)				<0.001
	<7	15790 (26.4)	906 (25.9)	
	7≤, <13	14037 (23.4)	913 (26.1)	
	13≤, <21	14379 (24.0)	957 (27.3)	
	21≤	15709 (26.2)	726 (20.7)	
Preoperative factor				
Age, median [IQR]		71 [64, 77]	72 [66, 78]	<0.001
Age (%)	70≤	33821 (56.4)	2115 (60.4)	<0.001
Sex (%)	Female	15127 (25.2)	906 (25.9)	0.41
Body mass index (kg/m2), median [IQR]		21.9 [19.7, 24.2]	21.8 [19.6, 24.0]	0.04
Body mass index (kg/m2) (%)				0.08
	≥18.5 <25	40293 (67.3)	2377 (67.9)	
	<18.5	8680 (14.5)	534 (15.2)	
	≥ 25	10942 (18.3)	591 (16.9)	
Diabetes mellitus (%)	+	11133 (18.6)	743 (21.2)	<0.001
Smoking (%)	+	29485 (49.2)	1821 (52.0)	0.001
Dependence in ADL (%)	+	2298 (3.8)	169 (4.8)	0.003
COPD (%)	+	3135 (5.2)	199 (5.7)	0.25
Dialysis (%)	+	331 (0.6)	34 (1.0)	0.001
History of IHD (%)	+	2335 (3.9)	147 (4.2)	0.37
Congestive heart failure (Within 30 days)(%)	+	356 (0.6)	21 (0.6)	0.97
Lomg-term steroid use (%)	+	512 (0.9)	33 (0.9)	0.58

Weight loss (%)	+	3460 (5.8)	235 (6.7)	0.02
Preoperative blood transfusion (%)	+	1552 (2.6)	92 (2.6)	0.89
Hemoglobin (%)	Male: < 13.5, Female: < 11.5	21117 (35.2)	1316 (37.6)	0.005
Albumin (%)	<3.5	11513 (19.2)	788 (22.5)	<0.001
BUN (%)	<8	8223 (13.7)	520 (14.8)	0.061
Creatinine (%)	>1.2	4191 (7.0)	269 (7.7)	0.12
AST > 35 (%)		4223 (7.0)	265 (7.6)	0.25
Preoperative chemotherapy (%)		4123 (6.9)	193 (5.5)	0.002
Preoperative radiotherapy (%)		100 (0.2)	7 (0.2)	0.64
T factor (in the TNM cllasification) (%)	T3≤	33028 (55.1)	2040 (58.3)	<0.001
N factor (in the TNM cllasification) (%)	N1≤	29307 (48.9)	1799 (51.4)	0.005
ASA-PS (%)	3, 4, 5	6694 (11.2)	421 (12.0)	0.12
Intraoperative factor				
Surgical approach (open or laparoscopic) (%)	Laparoscopic surgery	15762 (26.3)	456 (13.0)	<0.001
Intraoperative outcomes				
Operating time (min), median [IQR]		282 [221, 354]	279 [225, 347]	0.38
Estimated blood loss (mL), median [IQR]		260 [100, 521]	320 [150, 595]	<0.001
Postoperative outcomes				
Postoperative hospital stay (days), median [IQR]		16 [12, 24]	16 [12, 23]	0.18

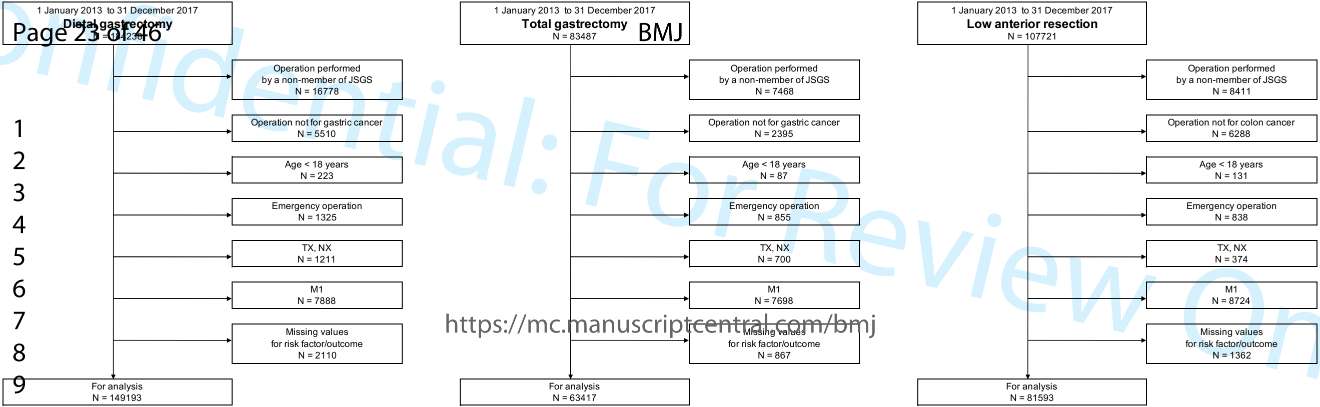
Abbreviations: TG, total gastrectomy; IQR, interquartile range; ADL, activities of daily living; IHD, ischemic heart disease; COPD, chronic obstructive pulmonary disease; BUN, blood urea nitrogen; AST, aspartate aminotransferase; ASA-PS, American Society of Anesthesiologists Physical Status

Table 4 Institutional and operative characteristics by surgeon's gender in LAR

		Male surgeon	Female surgeon	P value
Total cases of operation		77864	3729	
Factor	**Category**			
Institutional factor				
Number of surgeries per year		16 [9, 27]	15 [8, 26]	0.007
Number of surgeries per year (%)				0.01
	<8	17655 (22.7)	870 (23.3)	
	8≤, <16	21468 (27.6)	1100 (29.5)	
	16≤, <29	20112 (25.8)	895 (24.0)	
	29≤	18629 (23.9)	864 (23.2)	
Preoperative factor				
Age, median [IQR]		68 [61, 75]	68 [62, 75]	0.004
Age (%)	70≤	34077 (43.8)	1711 (45.9)	0.01
Sex (%)	Female	26958 (34.6)	1353 (36.3)	0.04
Body mass index (kg/m2), median [IQR]		22.3 [20.1, 24.7]	22.2 [20.0, 24.6]	0.01
Body mass index (kg/m2) (%)				0.27
	≥18.5 <25	51808 (66.5)	2471 (66.3)	
	<18.5	8838 (11.4)	454 (12.2)	
	≥ 25	17218 (22.1)	804 (21.6)	
Diabetes mellitus (%)	+	14049 (18.0)	722 (19.4)	0.04
Smoking (%)	+	33997 (43.7)	1620 (43.4)	0.81
Habitual drinking (%)	+	41677 (53.5)	1937 (51.9)	0.06
Dependence in ADL (%)	+	2725 (3.5)	170 (4.6)	0.001
COPD (%)	+	2800 (3.6)	104 (2.8)	0.01
Dialysis (%)	+	391 (0.5)	23 (0.6)	0.40
History of IHD (%)	+	2277 (2.9)	122 (3.3)	0.24
Congestive heart failure (Within 30 days)(%)	+	411 (0.5)	25 (0.7)	0.29

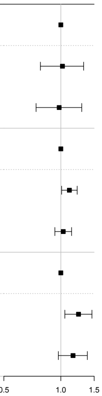
Lomg-term steroid use (%)	+	611 (0.8)	28 (0.8)	0.89
History of CVD (%)	+	2385 (3.1)	147 (3.9)	0.003
Weight loss (%)	+	1805 (2.3)	102 (2.7)	0.11
Preoperative blood transfusion (%)	+	710 (0.9)	45 (1.2)	0.08
Hemoglobin (%)	Male: < 13.5, Female: < 11.5	21036 (27.0)	1072 (28.7)	0.02
Albumin (%)	<3.5	9417 (12.1)	533 (14.3)	<0.001
BUN (%)	<8	9306 (12.0)	447 (12.0)	0.97
Creatinine (%)	>1.2	4350 (5.6)	221 (5.9)	0.40
Preoperative chemotherapy (%)		5032 (6.5)	156 (4.2)	<0.001
Preoperative radiotherapy (%)		2450 (3.1)	62 (1.7)	<0.001
T factor (in the TNM cllasification) (%)	T3≤	46697 (60.0)	2375 (63.7)	<0.001
ASA-PS	3, 4, 5	7155 (9.2)	344 (9.2)	0.96
Intraoperative factor				
Surgical approach (open or laparoscopic) (%)	Laparoscopic surgery	54199 (69.6)	2252 (60.4)	<0.001
Intraoperative outcomes				
Operating time (min), median [IQR]		265 [204, 345]	269 [210, 343]	0.04
Estimated blood loss (mL), median [IQR]		52 [10, 206]	80 [15, 271]	<0.001
Postoperative outcomes				
Postoperative hospital stay (days), median [IQR]		15 [11, 23]	15 [11, 23]	0.74

Abbreviations: LAR, low anterior resection; IQR, interquartile range; ADL, activities of daily living; IHD, ischemic heart disease; CVD, cardiovascular disease; COPD, chronic obstructive pulmonary disease; BUN, blood urea nitrogen; AST, aspartate aminotransferase; ASA-PS, American Society of Anesthesiologists Physical Status


1 January 2013 to 31 December 2017
Distal gastrectomy
N = 184238
Operation performed by a non-member of JSGS
N = 16778
Operation not for gastric cancer
N = 5510
Age < 18 years
N = 223
Emergency operation
N = 1325
TX, NX
N = 1211
M1
N = 7888
Missing values for risk factor/outcome
N = 2110
For analysis
N = 149193
1 January 2013 to 31 December 2017
Total gastrectomy
N = 83487
Operation performed by a non-member of JSGS
N = 7468
Operation not for gastric cancer
N = 2395
Age < 18 years
N = 87
Emergency operation
N = 855
TX, NX
N = 700
M1
N = 7698
Missing values for risk factor/outcome
N = 867
For analysis
N = 63417
1 January 2013 to 31 December 2017
Low anterior resection
N = 107721
Operation performed by a non-member of JSGS
N = 8411
Operation not for colon cancer
N = 6288
Age < 18 years
N = 131
Emergency operation
N = 838
TX, NX
N = 374
M1
N = 8724
Missing values for risk factor/outcome
N = 1362
For analysis
N = 81593

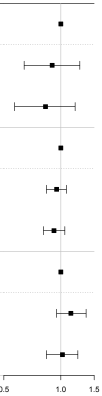
Distal gastrectomy

Outcome	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
Surgical mortality	Male	1030 (0.7)	Reference		Reference	
	Female	61 (0.7)	1.02 (0.78 to 1.32)	0.91		
					0.98 (0.74 to 1.29)	0.87
Surgical mortality or complication with CDC ≥3	Male	7817 (5.5)	Reference		Reference	
	Female	504 (6.1)	1.11 (1.01 to 1.22)	0.02		
					1.03 (0.93 to 1.14)	0.59
Pancreatic fistula	Male	2251 (1.6)	Reference		Reference	
	Female	162 (2.0)	1.24 (1.05 to 1.46)	0.009		
					1.16 (0.97 to 1.38)	0.11



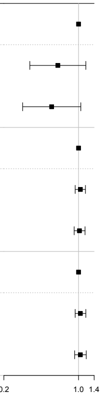
Total gastrectomy

Outcome	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
Surgical mortality	Male	667 (1.1)	Reference		Reference	
	Female	35 (1.0)	0.90 (0.64 to 1.26)	0.53		
					0.83 (0.57 to 1.19)	0.30
Surgical mortality or complication with CDC ≥3	Male	5569 (9.3)	Reference		Reference	
	Female	310 (8.9)	0.95 (0.84 to 1.07)	0.38		
					0.92 (0.81 to 1.05)	0.21
Pancreatic fistula	Male	1999 (3.3)	Reference		Reference	
	Female	132 (3.8)	1.13 (0.95 to 1.36)	0.17		
					1.02 (0.84 to 1.23)	0.88



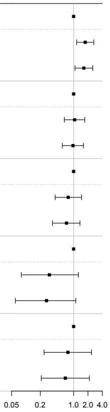
Low anterior resection

Outcome	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
Surgical mortality	Male	356 (0.5)	Reference		Reference	
	Female	11 (0.3)	0.64 (0.35 to 1.17)	0.15		
					0.56 (0.30 to 1.05)	0.07
Surgical mortality or complication with CDC ≥3	Male	7661 (9.8)	Reference		Reference	
	Female	380 (10.2)	1.04 (0.93 to 1.16)	0.48		
					1.02 (0.91 to 1.15)	0.69
Anastomotic leakage	Male	6950 (8.9)	Reference		Reference	
	Female	[illegible] (9.3)	1.04 ([illegible] to 1.17)	0.50		
					1.04 (0.92 to 1.18)	0.49



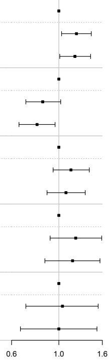
Surgical mortality in distal gastrectomy

Years after medical licence registration	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
- 5	Male	122 (0.6)	Reference		Reference	
	Female	28 (1.1)	1.75 (1.16 to 2.65)	0.008		
					1.64 (1.07 to 2.52)	0.02
6 - 10	Male	152 (0.7)	Reference		Reference	
	Female	18 (0.7)	1.05 (0.64 to 1.71)	0.85		
					0.97 (0.58 to 1.61)	0.90
11 - 15	Male	184 (0.7)	Reference		Reference	
	Female	10 (0.5)	0.77 (0.41 to 1.47)	0.43		
					0.71 (0.36 to 1.36)	0.30
16 - 20	Male	207 (0.7)	Reference		Reference	
	Female	2 (0.2)	0.31 (0.08 to 1.26)	0.10		
					0.27 (0.06 to 1.11)	0.07
21 -	Male	365 (0.8)	Reference		Reference	
	Female	3 (0.6)	0.76 (0.24 to 2.38)	0.64		
					0.67 (0.21 to 2.14)	0.50



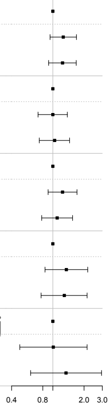
Surgical mortality or postoperative complication CDC ≥3 in distal gastrectomy

Years after medical licence registration	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
- 5	Male	1176 (6.1)	Reference		Reference	
	Female	185 (7.3)	1.21 (1.03 to 1.42)	0.02		
					1.19 (1.01 to 1.41)	0.04
6 - 10	Male	1269 (5.9)	Reference		Reference	
	Female	122 (5.0)	0.84 (0.70 to 1.02)	0.08		
					0.79 (0.65 to 0.96)	0.02
11 - 15	Male	1445 (5.3)	Reference		Reference	
	Female	115 (6.1)	1.14 (0.94 to 1.39)	0.18		
					1.08 (0.88 to 1.33)	0.47
16 - 20	Male	1501 (5.2)	Reference		Reference	
	Female	55 (6.2)	1.20 (0.91 to 1.59)	0.19		
					1.16 (0.86 to 1.56)	0.32
21 -	Male	2426 (5.5)	Reference		Reference	
	Female	27 (5.6)	1.04 (0.70 to 1.53)	0.86		
					1.00 (0.66 to 1.51)	1.00



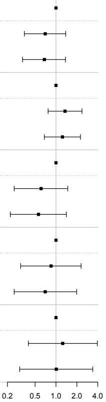
Pancreatic fistula in distal gastrectomy

Years after medical licence registration	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
- 5	Male	327 (1.7)	Reference		Reference	
	Female	54 (2.1)	1.26 (0.94 to 1.69)	0.12		
					1.24 (0.91 to 1.68)	0.16
6 - 10	Male	372 (1.7)	Reference		Reference	
	Female	42 (1.7)	1.00 (0.72 to 1.38)	1.00		
					1.04 (0.74 to 1.45)	0.82
11 - 15	Male	473 (1.7)	Reference		Reference	
	Female	41 (2.2)	1.24 (0.90 to 1.71)	0.19		
					1.10 (0.78 to 1.55)	0.59
16 - 20	Male	436 (1.5)	Reference		Reference	
	Female	18 (2.0)	1.35 (0.84 to 2.17)	0.22		
					1.29 (0.77 to 2.14)	0.33
21 -	Male	543 (1.4)	Reference		Reference	
	Female	7 (1.5)	1.01 (0.48 to 2.14)	0.98		
					1.34 (0.61 to 2.95)	0.47



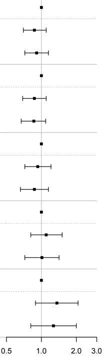
Surgical mortality in total gastrectomy

Years after medical licence registration	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
- 5	Male	92 (1.2)	Reference		Reference	
	Female	9 (0.8)	0.70 (0.35 to 1.38)	0.30		
					0.68 (0.33 to 1.37)	0.28
6 - 10	Male	96 (1.1)	Reference		Reference	
	Female	14 (1.4)	1.35 (0.77 to 2.37)	0.30		
					1.24 (0.68 to 2.25)	0.49
11 - 15	Male	123 (1.1)	Reference		Reference	
	Female	5 (0.7)	0.61 (0.25 to 1.48)	0.27		
					0.56 (0.22 to 1.41)	0.22
16 - 20	Male	128 (1.1)	Reference		Reference	
	Female	4 (0.9)	0.85 (0.31 to 2.30)	0.74		
					0.70 (0.25 to 1.99)	0.51
21 -	Male	228 (1.2)	Reference		Reference	
	Female	3 (1.4)	1.25 (0.40 to 3.94)	0.70		
					1.01 (0.30 to 3.39)	0.98



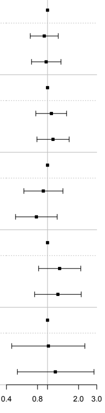
Surgical mortality or postoperative complication CDC ≥3 in total gastrectomy

Years after medical licence registration	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
- 5	Male	742 (9.3)	Reference		Reference	
	Female	92 (8.3)	0.87 (0.70 to 1.10)	0.25		
					0.91 (0.72 to 1.15)	0.42
6 - 10	Male	855 (9.4)	Reference		Reference	
	Female	82 (8.3)	0.87 (0.69 to 1.10)	0.25		
					0.86 (0.67 to 1.09)	0.21
11 - 15	Male	1054 (9.4)	Reference		Reference	
	Female	66 (8.8)	0.93 (0.72 to 1.21)	0.59		
					0.87 (0.66 to 1.15)	0.33
16 - 20	Male	1117 (9.3)	Reference		Reference	
	Female	45 (10.2)	1.10 (0.81 to 1.51)	0.54		
					1.01 (0.72 to 1.42)	0.94
21 -	Male	1801 (9.1)	Reference		Reference	
	Female	25 (12.0)	1.36 (0.89 to 2.07)	0.15		
					1.27 (0.81 to 2.00)	0.30



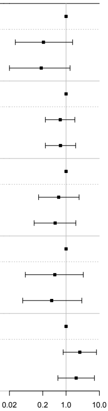
Pancreatic fistula in total gastrectomy

Years after medical licence registration	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
- 5	Male	359 (4.5)	Reference		Reference	
	Female	47 (4.2)	0.93 (0.68 to 1.27)	0.66		
					0.97 (0.70 to 1.35)	0.87
6 - 10	Male	322 (3.5)	Reference		Reference	
	Female	38 (3.8)	1.09 (0.77 to 1.53)	0.63		
					1.13 (0.79 to 1.62)	0.50
11 - 15	Male	361 (3.2)	Reference		Reference	
	Female	22 (2.9)	0.91 (0.59 to 1.41)	0.67		
					0.78 (0.49 to 1.24)	0.29
16 - 20	Male	397 (3.3)	Reference		Reference	
	Female	19 (4.3)	1.31 (0.82 to 2.10)	0.26		
					1.26 (0.75 to 2.12)	0.37
21 -	Male	560 (2.8)	Reference		Reference	
	Female	6 (2.9)	1.02 (0.45 to 2.30)	0.97		
					1.19 (0.51 to 2.82)	0.69



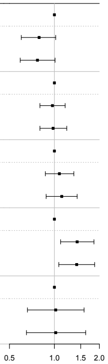
Surgical mortality in low anterior resection

Years after medical licence registration	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
- 5	Male	37 (0.5)	Reference		Reference	
	Female	1 (0.1)	0.21 (0.03 to 1.57)	0.13		
					0.18 (0.02 to 1.32)	0.09
6 - 10	Male	53 (0.5)	Reference		Reference	
	Female	4 (0.3)	0.67 (0.24 to 1.84)	0.43		
					0.68 (0.24 to 1.94)	0.47
11 - 15	Male	61 (0.4)	Reference		Reference	
	Female	2 (0.2)	0.60 (0.15 to 2.46)	0.48		
					0.47 (0.11 to 1.97)	0.30
16 - 20	Male	69 (0.4)	Reference		Reference	
	Female	1 (0.2)	0.46 (0.06 to 3.29)	0.44		
					0.37 (0.05 to 2.97)	0.35
21 -	Male	136 (0.5)	Reference		Reference	
	Female	3 (1.3)	2.59 (0.82 to 8.18)	0.11		
					2.01 (0.57 to 7.10)	0.28



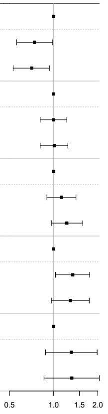
Surgical mortality or postoperative complication CDC ≥3 in low anterior resection

Years after medical licence registration	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
- 5	Male	667 (9.4)	Reference		Reference	
	Female	67 (7.6)	0.79 (0.60 to 1.02)	0.07		
					0.77 (0.59 to 1.01)	0.06
6 - 10	Male	1135 (10.7)	Reference		Reference	
	Female	125 (10.4)	0.97 (0.80 to 1.18)	0.75		
					0.98 (0.80 to 1.21)	0.87
11 - 15	Male	1607 (10.3)	Reference		Reference	
	Female	94 (11.0)	1.08 (0.87 to 1.35)	0.48		
					1.12 (0.88 to 1.42)	0.36
16 - 20	Male	1631 (9.2)	Reference		Reference	
	Female	71 (12.6)	1.42 (1.10 to 1.84)	0.01		
					1.41 (1.07 to 1.86)	0.01
21 -	Male	2621 (9.8)	Reference		Reference	
	Female	23 (10.0)	1.02 (0.66 to 1.58)	0.92		
					1.02 (0.65 to 1.62)	0.93

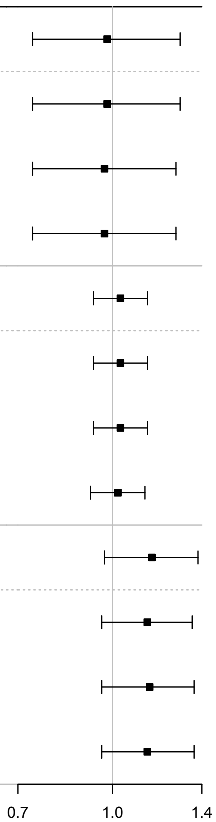


Anastomotic leakage in low anterior resection

Years after medical licence registration	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
- 5	Male	610 (8.6)	Reference		Reference	
	Female	58 (6.6)	0.74 (0.56 to 0.98)	0.04		
					0.71 (0.53 to 0.94)	0.02
6 - 10	Male	989 (9.4)	Reference		Reference	
	Female	112 (9.3)	1.00 (0.81 to 1.23)	1.00		
					1.01 (0.81 to 1.25)	0.94
11 - 15	Male	1410 (9.0)	Reference		Reference	
	Female	86 (10.1)	1.13 (0.90 to 1.42)	0.29		
					1.23 (0.97 to 1.58)	0.09
16 - 20	Male	1492 (8.4)	Reference		Reference	
	Female	62 (11.0)	1.35 (1.03 to 1.76)	0.03		
					1.30 (0.97 to 1.75)	0.08
21 -	Male	2449 (9.1)	Reference		Reference	
	Female	27 (11.7)	1.32 (0.88 to 1.98)	0.18		
					1.33 (0.86 to 2.05)	0.20

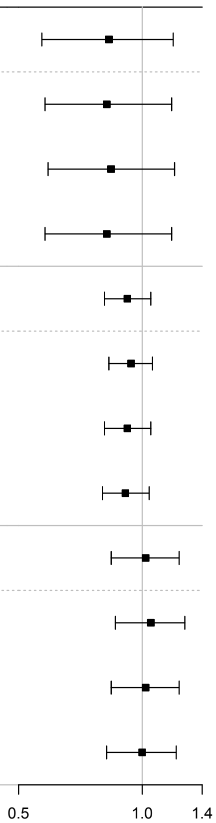


Outcome	Analysis	Adjusted OR (95% CI)	p value
Surgical mortality	Main analysis	0.98 (0.74 to 1.29)	0.87
	After multiple imputation	0.98 (0.74 to 1.29)	0.88
	Adding age, body mass index, hospital case volume, doctor's year after licence registration as continuous variables	0.97 (0.74 to 1.27)	0.81
	Adding sugeon case volume and rural-urban status as covariates	0.97 (0.74 to 1.27)	0.81
Surgical mortality or complication with CDC ≥3	Main analysis	1.03 (0.93 to 1.14)	0.59
	After multiple imputation	1.03 (0.93 to 1.14)	0.55
	Adding age, body mass index, hospital case volume, doctor's year after licence registration as continuous variables	1.03 (0.93 to 1.14)	0.59
	Adding sugeon case volume and rural-urban status as covariates	1.02 (0.92 to 1.13)	0.72
Pancreatic fistula	Main analysis	1.16 (0.97 to 1.38)	0.11
	After multiple imputation	1.14 (0.96 to 1.35)	0.15
	Adding age, body mass index, hospital case volume, doctor's year after licence registration as continuous variables	1.15 (0.96 to 1.36)	0.12
	Adding sugeon case volume and rural-urban status as covariates	1.14 (0.96 to 1.36)	0.13



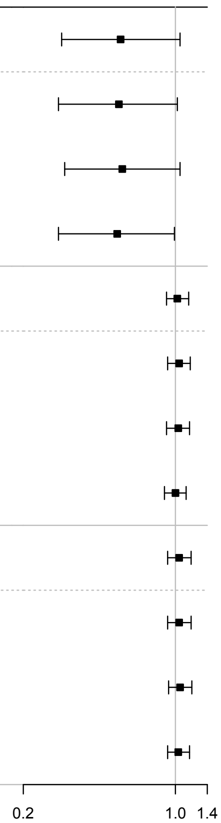
Supplementary Fig 1 Sensitivity analyses on the association between female surgeons and surgical outcomes in distal gastrectomy. OR, odds ratio; CI, confidence interval; CDC, Clavien-Dindo classification.

Outcome	Analysis	Adjusted OR (95% CI)	p value
Surgical mortality	Main analysis	0.83 (0.57 to 1.19)	0.30
	After multiple imputation	0.82 (0.58 to 1.18)	0.29
	Adding age, body mass index, hospital case volume, doctor's year after licence registration as continuous variables	0.84 (0.59 to 1.20)	0.35
	Adding sugeon case volume and rural-urban status as covariates	0.82 (0.58 to 1.18)	0.29
Surgical mortality or complication with CDC ≥3	Main analysis	0.92 (0.81 to 1.05)	0.21
	After multiple imputation	0.94 (0.83 to 1.06)	0.32
	Adding age, body mass index, hospital case volume, doctor's year after licence registration as continuous variables	0.92 (0.81 to 1.05)	0.21
	Adding sugeon case volume and rural-urban status as covariates	0.91 (0.80 to 1.04)	0.16
Pancreatic fistula	Main analysis	1.02 (0.84 to 1.23)	0.88
	After multiple imputation	1.05 (0.86 to 1.27)	0.64
	Adding age, body mass index, hospital case volume, doctor's year after licence registration as continuous variables	1.02 (0.84 to 1.23)	0.87
	Adding sugeon case volume and rural-urban status as covariates	1.00 (0.82 to 1.21)	1.00



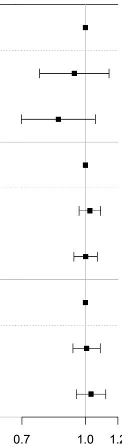
Supplementary Fig 2 Sensitivity analyses on the association between female surgeons and surgical outcomes in total gastrectomy. OR, odds ratio; CI, confidence interval; CDC, Clavien-Dindo classification.

Outcome	Analysis	Adjusted OR (95% CI)	p value
Surgical mortality	Main analysis	0.56 (0.30 to 1.05)	0.07
	After multiple imputation	0.55 (0.29 to 1.02)	0.06
	Adding age, body mass index, hospital case volume, doctor's year after licence registration as continuous variables	0.57 (0.31 to 1.05)	0.07
	Adding sugeon case volume and rural-urban status as covariates	0.54 (0.29 to 0.996)	0.049
Surgical mortality or complication with CDC ≥3	Main analysis	1.02 (0.91 to 1.15)	0.69
	After multiple imputation	1.04 (0.92 to 1.17)	0.54
	Adding age, body mass index, hospital case volume, doctor's year after licence registration as continuous variables	1.03 (0.91 to 1.16)	0.66
	Adding sugeon case volume and rural-urban status as covariates	1.00 (0.89 to 1.12)	0.99
Anastomotic leakage	Main analysis	1.04 (0.92 to 1.18)	0.49
	After multiple imputation	1.04 (0.92 to 1.18)	0.50
	Adding age, body mass index, hospital case volume, doctor's year after licence registration as continuous variables	1.05 (0.93 to 1.19)	0.41
	Adding sugeon case volume and rural-urban status as covariates	1.03 (0.92 to 1.16)	0.59



Supplementary Fig 3 Sensitivity analyses on the association between female surgeons and surgical outcomes in low anterior resection. OR, odds ratio; CI, confidence interval; CDC, Clavien-Dindo classification.

Outcome	Doctor's gender	Number of outcomes (%)	Unadjusted OR (95% CI)	p value	Adjusted OR (95% CI)	p value
Surgical mortality	Male	2053 (0.7)	Reference		Reference	
	Female	107 (0.7)	0.94 (0.77 to 1.14)	0.53		
					0.86 (0.70 to 1.06)	0.15
Surgical mortality or complication with CDC ≥3	Male	21046 (7.6)	Reference		Reference	
	Female	1194 (7.7)	1.03 (0.96 to 1.09)	0.42		
					1.00 (0.94 to 1.07)	0.97
Anastomotic leakage	Male	12978 (4.7)	Reference		Reference	
	Female	724 (4.7)	1.01 (0.93 to 1.09)	0.87		
					1.03 (0.95 to 1.12)	0.45



Supplementary Fig 4 Association between female surgeons and surgical outcomes in distal gastrectomy, total gastrectomy, and low anterior resection, combined. OR, odds ratio; CI, confidence interval; CDC, Clavien-Dindo classification.

Supplementary Table 1 Missing data regarding surgeon, institutional, and operative characteristics in DG

		Cases with no missing data	Cases with missing data	P value	Missing (N)	Missing (%)
Total cases of operation		149193	2110			
Factor	**Category**					
Surgeon's factor						
Female surgeon (%)		8222 (5.5)	153 (7.7)	<0.001	135	0.089225
Years since registration of licensed doctors, median [IQR]		15 [9, 22]	16 [9, 21]	0.08	146	0.096495
Years since registration of licensed doctors (%)				0.04	146	0.096495
	- 5	21780 (14.6)	315 (16.0)			
	6 - 10	23956 (16.1)	298 (15.2)			
	11 - 15	28982 (19.4)	366 (18.6)			
	16 - 20	29490 (19.8)	426 (21.7)			
	21 -	44985 (30.2)	559 (28.5)			
Number of surgeries per year, median [IQR]		8 [4, 15]	6 [3, 10]	<0.001	0	0
Institutional factor						
Number of surgeries per year, median [IQR]		30 [15, 54]	19 [9, 37]	<0.001	0	0
Number of surgeries per year (%)				<0.001	0	0
	<15	35753 (24.0)	840 (39.8)			
	15≤, <30	38198 (25.6)	608 (28.8)			
	30≤, <50	38856 (26.0)	470 (22.3)			
	50≤	36386 (24.4)	192 (9.1)			
Urban-rual status	Urban	85312 (57.2)	1183 (56.1)	0.30	0	0
Preoperative factor						
Age, median [IQR]		71 [64, 78]	72 [65, 79]	0.001	0	0
Age (%)	70≤	83258 (55.8)	1254 (59.4)	0.001	0	0
Sex (%)	Female	49618 (33.3)	687 (32.6)	0.51	1	0.000661
Body mass index (kg/m2), median [IQR]		22.2 [20.0, 24.4]	22.0 [19.8, 24.3]	0.009	11	0.007270
Body mass index (kg/m2) (%)				0.005	11	0.007270
	≥18.5 <25	100697 (67.5)	1391 (66.3)			
	<18.5	18237 (12.2)	305 (14.5)			
	≥ 25	30259 (20.3)	403 (19.2)			
Diabetes mellitus (%)	+	27063 (18.1)	327 (15.5)	0.002	0	0

Smoking (%)	+	67508 (45.2)	640 (30.3)	<0.001	1	0.000661
Dependence in ADL (%)	+	6366 (4.3)	119 (5.6)	0.002	0	0
COPD (%)	+	7234 (4.8)	61 (2.9)	<0.001	0	0
Dialysis (%)	+	1139 (0.8)	21 (1.0)	0.23	0	0
History of IHD (%)	+	5592 (3.7)	65 (3.1)	0.11	0	0
Congestive heart failure (Within 30 days)(%)	+	1043 (0.7)	9 (0.4)	0.14	0	0
Lomg-term steroid use (%)	+	1529 (1.0)	12 (0.6)	0.04	0	0
Weight loss (%)	+	5432 (3.6)	67 (3.2)	0.26	0	0
Preoperative blood transfusion (%)	+	3110 (2.1)	35 (1.7)	0.17	0	0
Hemoglobin (%)	Male: < 13.5, Female: < 11.5	41803 (28.0)	399 (24.2)	0.001	463	0.306008
Albumin (%)	<3.5	22647 (15.2)	214 (12.9)	0.01	453	0.299399
BUN (%)	<8	20529 (13.8)	135 (10.6)	0.001	838	0.553856
Creatinine (%)	>1.2	10587 (7.1)	91 (8.0)	0.22	977	0.645724
AST > 35 (%)		10123 (6.8)	100 (6.3)	0.49	533	0.352273
Preoperative chemotherapy (%)		3245 (2.2)	46 (2.2)	0.95	14	0.009253
Preoperative radiotherapy (%)		157 (0.1)	2 (0.1)	0.88	0	0
T factor (in the TNM cllasification) (%)	T3≤	45380 (30.4)	658 (31.2)	0.45	0	0
N factor (in the TNM cllasification) (%)	N1≤	47249 (31.7)	696 (33.0)	0.20	0	0
ASA-PS (%)	3, 4, 5	16569 (11.1)	175 (8.3)	<0.001	0	0
Intraoperative factor						
Surgical approach (open or laparoscopic) (%)	Laparoscopic surgery	77226 (51.8)	922 (43.7)	<0.001	0	0
Intraoperative outcomes						
Operating time (min), median [IQR]		259 [205, 320]	251 [195, 310]	<0.001	15	0.009914
Estimated blood loss (mL), median [IQR]		100 [25, 251]	108 [34, 272]	<0.001	0	0
Postoperative outcomes						
Surgical mortality (%)	+	1091 (0.7)	20 (0.9)	0.24	4	0.002644
Surgical mortality or complication with CD clasification of III or more (%)	+	8321 (5.6)	109 (5.2)	0.43	4	0.002644
Pancreatic leakage (%)	+	2413 (1.6)	25 (1.2)	0.12	0	0
Postoperative hospital stay (days), median [IQR]		13 [10, 19]	14 [11, 21]	<0.001	4	0.002644

Abbreviations: DG, distal gastrectomy; IQR, interquartile range; ADL, activities of daily living; IHD, ischemic heart disease; COPD, chronic obstructive pulmonary disease; BUN, blood urea nitrogen; AST, aspartate aminotransferase; ASA-PS, American Society of Anesthesiologists Physical Status.

Supplementary Table 2 Missing data regarding surgeon, institutional, and operative characteristics in TG

		Cases with no missing data	Cases with missing data	P value	Missing (N)	Missing (%)
Total cases of operation		63417	867			
Factor	**Category**					
Surgeon's factor						
Female surgeon (%)		3502 (5.5)	80 (9.8)	<0.001	47	0.073113
Years since registration of licensed doctors median [IQR]		16 [9, 22]	15 [9, 22]	0.31	65	0.101114
Years since registration of licensed doctors (%)				0.51	65	0.101114
	- 5	9074 (14.3)	113 (14.1)			
	6 - 10	10086 (15.9)	141 (17.6)			
	11 - 15	11953 (18.8)	159 (19.8)			
	16 - 20	12397 (19.5)	142 (17.7)			
	21 -	19907 (31.4)	247 (30.8)			
Number of surgeries per year, median [IQR]		4 [2, 7]	3 [2, 5]	<0.001	0	0
Institutional factor						
Number of surgeries per year, median [IQR]		13 [7, 22]	8 [4, 14]	<0.001	0	0
Number of surgeries per year (%)				<0.001	0	0
	<7	16375 (25.8)	381 (43.9)			
	7≤, <13	14757 (23.3)	186 (21.5)			
	13≤, <21	15339 (24.2)	174 (20.1)			
	21≤	16946 (26.7)	126 (14.5)			
Urban-rual status	Urban	35757 (56.4)	473 (54.6)	0.28	0	0
Preoperative factor						
Age, median [IQR]		71 [65, 77]	72 [65, 78]	0.03	0	0
Age (%)	70≤	35936 (56.7)	520 (60.0)	0.051	0	0
Sex (%)	Female	16033 (25.3)	211 (24.3)	0.53	0	0
Body mass index (kg/m2), median [IQR]		21.9 [19.7, 24.2]	21.8 [19.6, 24.2]	0.72	5	0.007778
Body mass index (kg/m2) (%)				0.71	5	0.007778
	≥18.5 <25	42670 (67.3)	570 (66.1)			
	<18.5	9214 (14.5)	133 (15.4)			
	≥ 25	11533 (18.2)	159 (18.4)			
Diabetes mellitus (%)	+	11876 (18.7)	127 (14.6)	0.002	0	0

Smoking (%)	+	31306 (49.4)	283 (32.7)	<0.001	1	0.001556
Dependence in ADL (%)	+	2467 (3.9)	31 (3.6)	0.63	0	0
COPD (%)	+	3334 (5.3)	41 (4.7)	0.49	0	0
Dialysis (%)	+	365 (0.6)	5 (0.6)	1.00	0	0
History of IHD (%)	+	2482 (3.9)	32 (3.7)	0.74	0	0
Congestive heart failure (Within 30 days)(%)	+	377 (0.6)	6 (0.7)	0.71	0	0
Lomg-term steroid use (%)	+	545 (0.9)	2 (0.2)	0.045	0	0
Weight loss (%)	+	3695 (5.8)	34 (3.9)	0.02	0	0
Preoperative blood transfusion (%)	+	1644 (2.6)	28 (3.2)	0.24	0	0
Hemoglobin (%)	Male: < 13.5, Female: < 11.5	22433 (35.4)	204 (30.4)	0.008	197	0.306453
Albumin (%)	<3.5	12301 (19.4)	113 (16.7)	0.08	192	0.298675
BUN (%)	<8	8743 (13.8)	48 (10.1)	0.02	391	0.608238
Creatinine (%)	>1.2	4460 (7.0)	36 (7.9)	0.47	411	0.639350
AST > 35 (%)		4488 (7.1)	48 (7.3)	0.80	213	0.331342
Preoperative chemotherapy (%)		4316 (6.8)	36 (4.2)	0.002	8	0.012445
Preoperative radiotherapy (%)		107 (0.2)	4 (0.5)	0.04	0	0
T factor (in the TNM cllasification) (%)	T3≤	35068 (55.3)	465 (53.6)	0.33	0	0
N factor (in the TNM cllasification) (%)	N1≤	31106 (49.0)	405 (46.7)	0.17	0	0
ASA-PS (%)	3, 4, 5	7115 (11.2)	79 (9.1)	0.051	0	0
Intraoperative factor						
Surgical approach (open or laparoscopic) (%)	Laparoscopic surgery	16218 (25.6)	183 (21.1)	0.003	0	0
Intraoperative outcomes						
Operating time (min), median [IQR]		282 [221, 353]	266 [210, 339]	<0.001	9	0.014000
Estimated blood loss (mL), median [IQR]		265 [105, 530]	250 [110, 496]	0.43	0	0
Postoperative outcomes						
Surgical mortality (%)	+	702 (1.1)	14 (1.6)	0.16	0	0
Surgical mortality or complication with CD clasification of III or more (%)	+	5879 (9.3)	54 (6.2)	0.002	0	0
Pancreatic leakage (%)	+	2131 (3.4)	8 (0.9)	<0.001	0	0
Postoperative hospital stay (days), median [IQR]		16 [12, 24]	17 [13, 26]	0.01	0	0

Abbreviations: TG, total gastrectomy; IQR, interquartile range; ADL, activities of daily living; IHD, ischemic heart disease; COPD, chronic obstructive pulmonary disease; BUN, blood urea nitrogen; AST, aspartate aminotransferase; ASA-PS, American Society of Anesthesiologists Physical Status.

Supplementary Table 3 Missing data regarding surgeon, institutional, and operative characteristics in LAR

		Cases with no missing data	Cases with missing data	P value	Missing (N)	Missing (%)
Total cases of operation		81593	1362			
Factor	**Category**					
Surgeon's factor						
Female surgeon (%)		3729 (4.6)	46 (3.6)	0.10	85	0.102465
Years since registration of licensed doctors median [IQR]		17 [11, 23]	17 [11, 23]	0.06	42	0.050630
Years since registration of licensed doctors (%)				0.34	42	0.050630
	- 5	7951 (9.7)	114 (8.6)			
	6 - 10	11774 (14.4)	188 (14.2)			
	11 - 15	16496 (20.2)	252 (19.1)			
	16 - 20	18260 (22.4)	297 (22.5)			
	21 -	27112 (33.2)	469 (35.5)			
Number of surgeries per year, median [IQR]		5 [3, 10]	4 [2, 8]	<0.001	0	0
Institutional factor						
Number of surgeries per year, median [IQR]		16 [9, 28]	13 [5, 21]	<0.001	0	0
Number of surgeries per year (%)				<0.001	0	0
	<8	18012 (22.1)	469 (34.4)			
	8≤, <16	22606 (27.7)	331 (24.3)			
	16≤, <29	21208 (26.0)	290 (21.3)			
	29≤	19767 (24.2)	272 (20.0)			
Urban-rual status	Urban	48751 (59.7)	874 (64.2)	0.001	0	0
Preoperative factor						
Age, median [IQR]		68 [61, 75]	68 [61, 75]	0.58	0	0
Age (%)	70≤	35788 (43.9)	607 (44.6)	0.60	0	0
Sex (%)	Female	28311 (34.7)	459 (33.7)	0.44	0	0
Body mass index (kg/m2), median [IQR]		22.3 [20.1, 24.7]	22.3 [20.0, 24.6]	0.71	8	0.009644
Body mass index (kg/m2) (%)				0.20	8	0.009644
	≥18.5 <25	54279 (66.5)	881 (65.1)		1	0.001205
	<18.5	9292 (11.4)	175 (12.9)		1	0.001205
	≥ 25	18022 (22.1)	298 (22.0)		1	0.001205
Diabetes mellitus (%)	+	14771 (18.1)	226 (16.6)	0.15	1	0.001205

Smoking (%)	+	35617 (43.7)	423 (31.1)	<0.001	1	0.001205
Habitual drinking (%)	+	43614 (53.5)	577 (42.4)	<0.001	1	0.001205
Dependence in ADL (%)	+	2895 (3.5)	40 (2.9)	0.23	1	0.001205
COPD (%)	+	2904 (3.6)	24 (1.8)	<0.001	1	0.001205
Dialysis (%)	+	414 (0.5)	8 (0.6)	0.68	1	0.001205
History of IHD (%)	+	2399 (2.9)	31 (2.3)	0.15	1	0.001205
Congestive heart failure (Within 30 days)(%)	+	436 (0.5)	6 (0.4)	0.64	1	0.001205
Lomg-term steroid use (%)	+	639 (0.8)	9 (0.7)	0.61	1	0.001205
History of CVD (%)	+	2532 (3.1)	32 (2.4)	0.11	395	0.476162
Weight loss (%)	+	1907 (2.3)	21 (1.5)	0.054	422	0.508710
Preoperative blood transfusion (%)	+	755 (0.9)	20 (1.5)	0.04	572	0.689530
Hemoglobin (%)	Male: < 13.5, Female: < 11.5	22108 (27.1)	203 (21.0)	<0.001	742	0.894461
Albumin (%)	<3.5	9950 (12.2)	107 (11.4)	0.45	5	0.006027
BUN (%)	<8	9753 (12.0)	73 (9.2)	0.02	0	0
Creatinine (%)	>1.2	4571 (5.6)	49 (7.9)	0.01	0	0
Preoperative chemotherapy (%)		5188 (6.4)	58 (4.3)	0.002	0	0
Preoperative radiotherapy (%)		2512 (3.1)	17 (1.2)	<0.001	0	0
T factor (in the TNM cllasification) (%)	T3≤	49072 (60.1)	841 (61.7)	0.23	0	0
ASA-PS	3, 4, 5	7499 (9.2)	111 (8.1)	0.19	0	0
Intraoperative factor						
Surgical approach (open or laparoscopic) (%)	Laparoscopic surgery	56451 (69.2)	809 (59.4)	<0.001	0	0
Intraoperative outcomes						
Operating time (min), median [IQR]		265 [205, 345]	259 [200, 335]	0.009	7	0.008438
Estimated blood loss (mL), median [IQR]		55 [10, 210]	75 [15, 239]	0.001	0	0
Postoperative outcomes						
Surgical mortality (%)	+	367 (0.4)	4 (0.3)	0.39	3	0.003616
Surgical mortality or complication with CD clasification of III or more (%)	+	8041 (9.9)	92 (6.8)	<0.001	2	0.002411
Anastomotic leakage (%)	+	7295 (8.9)	81 (6.0)	<0.001	1	0.001205
Postoperative hospital stay (days), median [IQR]		15 [11, 23]	15 [11, 25]	0.01	3	0.003616

Abbreviations: LAR, low anterior resection; IQR, interquartile range; ADL, activities of daily living; IHD, ischemic heart disease; CVD, cardiovascular disease; COPD, chronic obstructive pulmonary disease; BUN, blood urea nitrogen; AST, aspartate aminotransferase; ASA-PS, American Society of Anesthesiologists Physical Status.

Supplementary Table 4 Surgeon case volume and urban-rural status

Distal gastrectomy				
		Male surgeon	Female surgeon	P value
Total cases of operation		140971	8222	
Surgeon's factor				
Number of surgeries per year, median [IQR]		8 [4, 16]	6 [3, 11]	<0.001
Institutional factor				
Urban-rural status	Urban	79730 (56.6)	5582 (67.9)	<0.001
Total gastrectomy				
		Male surgeon	Female surgeon	P value
Total cases of operation		59915	3502	
Surgeon's factor				
Number of surgeries per year, median [IQR]		4 [2, 7]	3 [2, 5]	<0.001
Institutional factor				
Urban-rural status	Urban	33257 (55.5)	2500 (71.4)	<0.001
Low anterior resection				
		Male surgeon	Female surgeon	P value
Total cases of operation		77864	3729	
Surgeon's factor				
Number of surgeries per year, median [IQR]		5 [3, 10]	3 [2, 6]	<0.001
Institutional factor				
Urban-rual status	Urban	46132 (59.2)	2619 (70.2)	<0.001

Abbreviations: IQR, interquartile range.

Supplementary Table 5 Comparison of surgical outcomes between male and female surgeons according to stratified predicted risk in distal gastrectomy

					Surgeon's gender		
				Total	Male	Female	P value
	Total cases				140971	8222	
	Surgical mortality						
		Predicted risk range (%)			Male	Female	P value
Predicted risk from low to high risk	1st quintile	0.0187≤, <0.1075	Number of cases	29839	28353	1486	
			Number of outcomes (%)	7 (0.0)	7 (0.0)	0 (0.0)	1.00
	2nd quintile	0.1075≤, <0.1971	Number of cases	29838	28218	1620	
			Number of outcomes (%)	21 (0.1)	20 (0.1)	1 (0.1)	1.00
	3rd quintile	0.1971≤, <0.3542	Number of cases	29839	28202	1637	
			Number of outcomes (%)	69 (0.2)	66 (0.2)	3 (0.2)	1.00
	4th quintile	0.3542≤, <0.7358	Number of cases	29838	28086	1752	
			Number of outcomes (%)	159 (0.5)	151 (0.5)	8 (0.5)	0.65
	5th quintile	0.7358≤, ≤60.7759	Number of cases	29839	28112	1727	
			Number of outcomes (%)	835 (2.8)	786 (2.8)	49 (2.8)	0.92
	Surgical mortality or complication with a CD classification of III or more						
		Predicted risk range (%)			Male	Female	P value
Predicted risk from low to high risk	1st quintile	0.6789≤, <2.6167	Number of cases	29836	28488	1348	
			Number of outcomes (%)	381 (1.3)	364 (1.3)	17 (1.3)	0.96
	2nd quintile	2.6167≤, <3.7432	Number of cases	29841	28300	1541	
			Number of outcomes (%)	779 (2.6)	734 (2.6)	45 (2.9)	0.43
	3rd quintile	3.7432≤, <5.1228	Number of cases	29839	28205	1634	
			Number of outcomes (%)	1280 (4.3)	1217 (4.3)	63 (3.9)	0.37
	4th quintile	5.1228≤, <7.4709	Number of cases	29838	28034	1804	
			Number of outcomes (%)	1982 (6.6)	1859 (6.6)	123 (6.8)	0.76
	5th quintile	7.4709≤, ≤64.2078	Number of cases	29839	27944	1895	
			Number of outcomes (%)	3899 (13.1)	3643 (13.0)	256 (13.5)	0.56
	Pancreatic leakage						
		Predicted risk range (%)			Male	Female	P value
Predicted risk from low to high risk	1st quintile	0.0590≤, <0.5065	Number of cases	29839	28380	1459	
			Number of outcomes (%)	48 (0.2)	46 (0.2)	2 (0.1)	1.00
	2nd quintile	0.5065≤, <0.8214	Number of cases	29838	28287	1551	
			Number of outcomes (%)	119 (0.4)	117 (0.4)	2 (0.1)	0.10
	3rd quintile	0.8214≤, <1.2589	Number of cases	29839	28172	1667	
			Number of outcomes (%)	232 (0.8)	223 (0.8)	9 (0.5)	0.26
	4th quintile	1.2589≤, <2.1520	Number of cases	29838	28192	1646	
			Number of outcomes (%)	534 (1.8)	500 (1.8)	34 (2.1)	0.39
	5th quintile	2.1520≤, ≤34.2016	Number of cases	29839	27940	1899	
			Number of outcomes (%)	1480 (5.0)	1365 (4.9)	115 (6.1)	0.02

Abbreviation: CD, Clavien-Dindo.

Supplementary Table 6 Comparison of surgical outcomes between male and female surgeons according to stratified predicted risk in total gastrectomy

					Surgeon's gender		
				Total	Male	Female	P value
	Total cases				59915	3502	
	Surgical mortality						
		Predicted risk range (%)			Male	Female	P value
Predicted risk from low to high risk	1st quintile	0.0505≤, <0.2363	Number of cases	12684	12055	629	
			Number of outcomes (%)	8 (0.1)	8 (0.1)	0 (0.0)	1.00
	2nd quintile	0.2363≤, <0.3859	Number of cases	12683	12005	678	
			Number of outcomes (%)	20 (0.2)	19 (0.2)	1 (0.1)	1.00
	3rd quintile	0.3859≤, <0.6233	Number of cases	12683	11994	689	
			Number of outcomes (%)	57 (0.4)	55 (0.5)	2 (0.3)	0.77
	4th quintile	0.6233≤, <1.1523	Number of cases	12683	11919	764	
			Number of outcomes (%)	97 (0.8)	92 (0.8)	5 (0.7)	0.72
	5th quintile	1.1523≤, ≤50.0575	Number of cases	12684	11942	742	
			Number of outcomes (%)	520 (4.1)	493 (4.1)	27 (3.6)	0.51
	Surgical mortality or complication with a CD classification of III or more						
		Predicted risk range (%)			Male	Female	P value
Predicted risk from low to high risk	1st quintile	1.5804≤, <5.3523	Number of cases	12684	11989	695	
			Number of outcomes (%)	317 (2.5)	302 (2.5)	15 (2.2)	0.55
	2nd quintile	5.3523≤, <7.0254	Number of cases	12683	11983	700	
			Number of outcomes (%)	656 (5.2)	628 (5.2)	28 (4.0)	0.15
	3rd quintile	7.0254≤, <8.9722	Number of cases	12683	11985	698	
			Number of outcomes (%)	974 (7.7)	919 (7.7)	55 (7.9)	0.84
	4th quintile	8.9722≤, <12.0658	Number of cases	12683	11995	688	
			Number of outcomes (%)	1416 (11.2)	1341 (11.2)	75 (10.9)	0.82
	5th quintile	12.0658≤, ≤53.1935	Number of cases	12684	11963	721	
			Number of outcomes (%)	2516 (19.8)	2379 (19.9)	137 (19.0)	0.56
	Pancreatic leakage						
		Predicted risk range (%)			Male	Female	P value
Predicted risk from low to high risk	1st quintile	0.0758≤, <1.1022	Number of cases	12684	12166	518	
			Number of outcomes (%)	40 (0.3)	37 (0.3)	3 (0.6)	0.22
	2nd quintile	1.1022≤, <1.7600	Number of cases	12683	12013	670	
			Number of outcomes (%)	106 (0.8)	100 (0.8)	6 (0.9)	0.86
	3rd quintile	1.7600≤, <2.6742	Number of cases	12683	12001	682	
			Number of outcomes (%)	229 (1.8)	221 (1.8)	8 (1.2)	0.20
	4th quintile	2.6742≤, <4.5195	Number of cases	12683	11899	784	
			Number of outcomes (%)	483 (3.8)	444 (3.7)	39 (5.0)	0.08
	5th quintile	4.5195≤, ≤34.3401	Number of cases	12684	11836	848	
			Number of outcomes (%)	1273 (10.0)	1197 (10.1)	76 (9.0)	0.28

Abbreviation: CD, Clavien-Dindo.

Supplementary Table 7 Comparison of surgical outcomes between male and female surgeons according to stratified predicted risk in low anterior resection

					Surgeon's gender		
				Total	Male	Female	P value
	Total cases			81593	77864	3729	
	Surgical mortality						
		Predicted risk range (%)			Male	Female	P value
Predicted risk from low to high risk	1st quintile	0.0037≤, <0.0605	Number of cases	16319	15643	676	
			Number of outcomes (%)	1 (0.0)	1 (0.0)	0 (0.0)	1.00
	2nd quintile	0.0605≤, <0.1089	Number of cases	16318	15579	739	
			Number of outcomes (%)	4 (0.0)	4 (0.0)	0 (0.0)	1.00
	3rd quintile	0.1089≤, <0.1910	Number of cases	16319	15583	736	
			Number of outcomes (%)	17 (0.1)	16 (0.1)	1 (0.1)	0.54
	4th quintile	0.1910≤, <0.3871	Number of cases	16318	15522	796	
			Number of outcomes (%)	38 (0.2)	37 (0.2)	1 (0.1)	1.00
	5th quintile	0.3871≤, ≤68.2070	Number of cases	16319	15537	782	
			Number of outcomes (%)	307 (1.9)	298 (1.9)	9 (1.2)	0.12
	Surgical mortality or complication with a CD classification of III or more						
		Predicted risk range (%)			Male	Female	P value
Predicted risk from low to high risk	1st quintile	0.9348≤, <5.0464	Number of cases	16319	15594	725	
			Number of outcomes (%)	436 (2.7)	415 (2.7)	21 (2.9)	0.70
	2nd quintile	5.0464≤, <7.3290	Number of cases	16318	15575	743	
			Number of outcomes (%)	852 (5.2)	808 (5.2)	44 (5.9)	0.38
	3rd quintile	7.3290≤, <9.8525	Number of cases	16319	15587	732	
			Number of outcomes (%)	1272 (7.8)	1212 (7.8)	60 (8.2)	0.68
	4th quintile	9.8525≤, <13.5345	Number of cases	16318	15568	750	
			Number of outcomes (%)	2016 (12.4)	1920 (12.3)	96 (12.8)	0.70
	5th quintile	13.5345≤, ≤62.7916	Number of cases	16319	15540	779	
			Number of outcomes (%)	3465 (21.2)	3306 (21.3)	159 (20.4)	0.57
	Anastomotic leakage						
		Predicted risk range (%)			Male	Female	P value
Predicted risk from low to high risk	1st quintile	0.6513≤, <4.1243	Number of cases	16319	15598	721	
			Number of outcomes (%)	349 (2.1)	331 (2.1)	18 (2.5)	0.50
	2nd quintile	4.1243≤, <6.4416	Number of cases	16318	15548	770	
			Number of outcomes (%)	725 (4.4)	676 (4.3)	49 (6.4)	0.008
	3rd quintile	6.4416≤, <9.1038	Number of cases	16319	15584	735	
			Number of outcomes (%)	1133 (6.9)	1078 (6.9)	55 (7.5)	0.56
	4th quintile	9.1038≤, <12.6965	Number of cases	16318	15535	783	
			Number of outcomes (%)	1823 (11.2)	1741 (11.2)	82 (10.5)	0.52
	5th quintile	12.6965≤, ≤46.1444	Number of cases	16319	15599	720	
			Number of outcomes (%)	3265 (20.0)	3124 (20.0)	141 (19.6)	0.77

Abbreviation: CD, Clavien-Dindo.

Protocol for the comparative study of short-term surgical outcomes between male and female surgeons in Japan

19th March 2019

Introduction

In Japan, the proportion of female physicians is 21.1% and that of female surgeons in general and gastrointestinal surgery is even lower, at 5.9%.[1] This suggests that the working environment in Japan poses more challenges for women looking to continue their careers and develop their skills for surgery than those posed by other listed countries. In this unique social environment, it is important to compare the outcomes of female and male surgeons to encourage women's choice of a career in surgery and/or to propose more effective training for female surgeons in Japan.

Previous studies in the US and Canada demonstrated that the proficiency of female physicians and surgeons was equal to or better than that of their male counterparts. Tsugawa et al. reported that the mortality and readmission rates of older hospitalised patients treated by female physicians in the US were lower than those of such patients treated by male physicians.[2] In the US, no significant difference was found in postoperative mortality between female and male surgeons.[3] Moreover, there was no difference in the complication rates of surgeries performed by male and female general surgeons in the US.[4] The postoperative mortality of patients operated on by Canadian female surgeons was slightly, but significantly, lower than that of patients operated on by male surgeons.[5]

Objective

The objective of this study is to compare the short-term surgical outcomes between female and male surgeons in Japan with a large gender gap.

Study design

This study is a retrospective, nationwide, observational study including a multivariable logistic analysis adjusting for patient factors.

Resources

This study is supported by the Japanese Society of Gastroenterological Surgery (JSGS) and the National Clinical Database (NCD), however, is not funded by any research funds.

Study design and data source

This study is a retrospective observational study using data from the NCD. We will analyse data related to surgeons' gender and experience (years of clinical practice after licencing) and classify hospitals according to the number of cases of each procedure in one year.

The NCD initiated data registration in 2011.[6] By December 2016, almost 5,000 facilities were registered with the NCD. Over 9 100 000 cases are registered in this database each year, which is equivalent to more than 95% of all surgeries in Japan.[7] Using data from the NCD, we will analyse the outcomes of elective distal gastrectomy (DG), total gastrectomy (TG), and low anterior resection (LAR) performed by male and female surgeons between 2013 and 2017. These three procedures, which are commonly performed in Japan, will be statistically compared. The NCD does not contain direct information on surgeons' gender or the number of years since the registration of licenced doctors, but it does contain the licence number of the surgeons. Using these licence numbers, an analysis will be is conducted by linking the NCD information with the gender profile and the year of licencing registration for the JSGS members.

Endpoints

The primary endpoints will be surgical mortality, severe postoperative complications, pancreatic fistula (in DG/ TG only), and anastomotic leakage (in LAR only). In this study, surgical mortality will be defined as in-hospital deaths that occurred within 90 days postoperatively and any death up to 30 days postoperatively. Other primary endpoints include severe postoperative complications, which are defined as any postoperative surgical and medical complications with a Clavien–Dindo classification ≥3,[8] pancreatic fistulas (only in DG/TG), and anastomotic leakage (only in LAR). The operation time and blood loss are considered intraoperative outcomes.

Eligibility criteria

Inclusion criteria

Patients will be required to fulfill the following criteria for inclusion in this study:

Underwent DG, TG and LAR of rectum from 1 January to 2013 to 31 December 2017.

Exclusion criteria

Patients will be excluded if they meet any of the following criteria:

Patients operated on by a non-member of the JSGS.

Patients not operated for gastric cancer/rectal cancer

Aged under 18 years

Emergency surgery

Unknown T or N factor (DG, TG) and T factor (LAR) in TNM classification

Patients with any other organ metastasis (M1)

Patients with missing values for risk factor/outcome

Adjustment variables

Surgeon's characteristics include sex and years since registration of licenced doctors in five-year increments. Patient characteristics include age (<70 vs ≥70 years), sex (male vs female), body mass index (≤18.5 vs >18.5 kg/m^2, <25 vs ≥25 kg/m^2), American Society of Anesthesiologists Physical Status classification (ASA-PS, 1–2 vs ≥3), clinical T factor (T1–2 vs T3–4) and N (0 vs 1–3) of tumours (N factor was included only for DG and TG; based on the Union for International Cancer Control–TNM classification, 7th edition), haemoglobin (male: <13.5 g/dL vs ≥13.5 g/dL, female: <11.5 g/dL vs ≥11.5 g/dL), aspartate aminotransferase (<35 IU/L vs ≥35 IU/L; included in DG and TG), albumin (<3.5 g/dL vs ≥3.5 g/dL), blood urea nitrogen (<8 mg/dL vs ≥8 mg/dL), creatinine (<1.2 mg/dL vs ≥1.2 mg/dL), absence/presence of diabetes mellitus, smoking status, habitual drinking status (only in LAR), dependence in activities of daily living (ADL), history of chronic obstructive pulmonary disease, dialysis, ischaemic heart disease, or congestive heart failure, long-term steroid use, history of cardiovascular diseases (only in LAR), weight loss, preoperative blood transfusion, preoperative chemotherapy, and preoperative radiotherapy.

We will categorise hospitals into quartiles according to the number of cases of each procedure: very low (VL), low (L), high (H), and very high (VH). The surgical approach (open or laparoscopic) is included as an intraoperative factor.

Statistical analysis

We will use the chi-square test for categorical variables and the Mann–Whitney U test for continuous variables when comparing baseline characteristics and short-term outcomes. A multivariable logistic regression model will be constructed, adjusting for patient characteristics, surgeon characteristics, and hospital characteristics, to examine the association between the surgeon's gender, surgical complications, and surgery-related mortality. We will use a random-effects model to account for hospital-level characteristics. Hospital identification (ID) will be used as a random intercept. Subsequently, additional analysis will be conducted to examine whether an interaction effect existed between sex and years after medical licence registration. An interaction term of sex and years of experience post-medical licence registration will be incorporated, instead of including them individually in the previous regression model. All p-values will be two-sided, and p-values <0.05 will be considered significant. Statistical analyses will be performed using R software (R Foundation for Statistical Computing, Vienna, Austria).

Discussion

This study aims to compare surgical outcomes between female and male surgeons using the NCD, the most extensive surgical database in Japan. We will also examine the relationship between postoperative mortality and surgical complication rates and the number of years a surgeon has been licensed. From these results, we will support women's career choices as surgeons in Japan and discuss a more comfortable working environment and more effective training for female gastrointestinal surgeons in Japan.

1 Ministry of Health, Labour and Welfare (Japan). Survey of physicians, dentists, and pharmacists, 2018; https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00450026&tstat=000001030962&cycle=7&tclass1=000001109395&tclass2=000001109396&stat_infid=000031653231&tclass3val=0.

2 Tsugawa Y, Jena AB, Figueroa JF, Orav EJ, Blumenthal DM, Jha AK. Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs female physicians. *JAMA Intern Med* 2017;177:206–13.

3 Tsugawa Y, Jena AB, Orav EJ, et al. Age and sex of surgeons and mortality of older surgical patients: observational study. *BMJ* 2018;361:k1343.

4 Sharoky CE, Sellers MM, Keele LJ, et al. Does surgeon sex matter?: Practice patterns and outcomes of female and male surgeons. *Ann Surg* 2018;267:1069–76.

5 Wallis CJ, Ravi B, Coburn N, Nam RK, Detsky AS, Satkunasivam R. Comparison of postoperative outcomes among patients treated by male and female surgeons: A population based matched cohort study. *BMJ* 2017;359:j4366.

6 Seto Y, Kakeji Y, Miyata H, Iwanaka T. National Clinical Database (NCD) in Japan for gastroenterological surgery: Brief introduction. *Ann Gastroenterol Surg* 2017;1:80–1.

7 Kakeji, Y, Takahashi, A, Udagawa, H, et al. Surgical outcomes in gastroenterological surgery in Japan: Report of National Clinical database 2011–2016. *Ann Gastroenterol Surg* 2018;2:37–54.

8 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;240:205–13.

Print Abstract

Study question

Is there a difference in short-term outcomes of gastrointestinal surgery performed by female and male surgeons in Japan?

Methods

The National Clinical Database (NCD) (2013-2017), which includes >95% of surgeries in Japan, was used to analyse the outcomes of distal gastrectomy (DG), total gastrectomy (TG), and low anterior resection (LAR) performed by male and female surgeons. Primary endpoints included surgical mortality, postoperative complications, pancreatic fistulae (DG/TG), and anastomotic leakage (LAR). We examined the association of surgeons' gender and post-registration years with surgical complications and mortality using multivariable logistic regression models, adjusting for patient and hospital characteristics. Female surgeons had fewer post-registration years than males (DG/TG; median 9 vs 16 years, LAR; median 9 vs 17 years), operated on higher-risk patients, and performed fewer laparoscopic surgeries (DG; 52.7% vs 35.8%, TG; 26.3% vs 13.0, LAR; 69.6% vs 60.4%). There was no significant difference between male and female surgeons in the adjusted risk for surgical mortality and surgical mortality combined with the Clavien–Dindo grade ≥3 complications (DG/TG/LAR), pancreatic fistula (DG/TG), or anastomotic leakage (LAR).

Study answer and limitations

There was no statistical difference in the short-term outcomes of DG, TG, and LAR performed by female and male surgeons in Japan. However, one limitation is that there are far fewer female surgeons; therefore, a single adverse event can significantly impact the overall outcome.

What this study adds

Japanese female surgeons were responsible for relatively high-risk cases. There was no significant difference in surgical mortality or complication rates between male and female surgeons.

Funding, competing interests, data sharing

This study received no financial support. Several authors have competing interests in surgical instruments and pharmaceutical companies. To access the NCD, it is necessary to submit a proposal through an NCD-related society and request it to the NCD Secretariat.