Comparison of short term surgical outcomes of male and female gastrointestinal surgeons in Japan: retrospective cohort study

Kae Okoshi,[1,2,*] Hideki Endo,[3,*] Sachiyo Nomura,[4] Emiko Kono,[5] Yusuke Fujita,[2] Itaru Yasufuku,[6] Koya Hida,[2] Hiroyuki Yamamoto,[3] Hiroaki Miyata,[3] Kazuhiro Yoshida,[6] Yoshihiro Kakeji,[7] Yuko Kitagawa[8]

For numbered affiliations see end of the article

*Contributed equally

Correspondence to: S Nomura sachiyo.nomura1012@gmail.com (ORCID 0000-0003-4293-6205)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2022;378:e070568 http://dx.doi.org/10.1136/bmj-2022-070568

Accepted: 27 July 2022

ABSTRACT

OBJECTIVE

To compare short term surgical outcomes between male and female gastrointestinal surgeons in Japan.

DESIGN

Retrospective cohort study.

SETTING

Data from the Japanese National Clinical Database (includes data on >95% of surgeries performed in Japan) (2013-17) and the Japanese Society of Gastroenterological Surgery.

PARTICIPANTS

Male and female surgeons who performed distal gastrectomy, total gastrectomy, and low anterior resection.

MAIN OUTCOME MEASURES

Surgical mortality, surgical mortality combined with postoperative complications, pancreatic fistula (distal gastrectomy/total gastrectomy only), and anastomotic leakage (low anterior resection only). The association of surgeons' gender with surgery related mortality and surgical complications was examined using multivariable logistic regression models adjusted for patient, surgeon, and hospital characteristics.

RESULTS

A total of 149 193 distal gastrectomy surgeries (male surgeons: 140 971 (94.5%); female surgeons: 8222 (5.5%)); 63 417 gastrectomy surgeries (male surgeons: 59 915 (94.5%); female surgeons: 3502 (5.5%)); and 81 593 low anterior resection procedures (male surgeons: 77 864 (95.4%);female surgeons: 3729 (4.6%)) were done. On average, female surgeons had fewer post-registration years, operated on patients at higher risk, and did fewer laparoscopic surgeries than male surgeons. No significant difference was found between male and female surgeons in the adjusted risk for surgical mortality (adjusted odds ratio 0.98 (95% confidence interval 0.74 to 1.29) for distal gastrectomy; 0.83 (0.57 to 1.19) for total gastrectomy; 0.56 (0.30 to 1.05) for low anterior resection), surgical mortality combined with Clavien-Dindo grade ≥3 complications (adjusted odds ratio 1.03 (0.93 to 1.14) for distal gastrectomy; 0.92 (0.81 to 1.05) for total gastrectomy; 1.02 (0.91 to 1.15) for low anterior resection), pancreatic fistula (adjusted odds ratio 1.16 (0.97 to 1.38) for distal gastrectomy; 1.02 (0.84 to 1.23) for total gastrectomy), and anastomotic leakage (adjusted odds ratio 1.04 (0.92 to 1.18) for low anterior resection).

CONCLUSION

This study found no significant adjusted risk difference in the outcomes of surgeries performed by male versus female gastrointestinal surgeons. Despite disadvantages, female surgeons take on patients at high risk. Greater access to surgical training for female physicians is warranted in Japan.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Women remain a minority in the surgical field, particularly in Japan

In the US and Canada, the proficiency of female physicians and surgeons was equal to or better than that of their male counterparts

WHAT THIS STUDY ADDS

This study included large numbers of distal gastrectomy, total gastrectomy, and low anterior resection surgeries in Japan

No overall significant differences existed in surgical mortality or Clavien-Dindo ≥3 complication rates associated with the three procedures performed by male and female surgeons.

More opportunities and encouragement should be provided to female surgeons to reduce the gender based inequity in the field of surgery

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 21/26 countries in 2018.[1] Despite this increase, women remain a minority in the surgical field. Female general surgeons accounted for 27.9% (in 2019), 22.0% (in 2019), and 32.5% (in 2017) of surgeons in Canada, the United States, and the United Kingdom, respectively.[2-4]

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 in surgery than are posed in other listed countries. In this unique social environment, comparing the outcomes of female and male surgeons is important to encourage women's choice of a career in surgery and to propose more effective training for female surgeons in Japan.

Previous studies in the US and Canada showed that the proficiency of female physicians and surgeons was equal to or better than that of their male counterparts. Tsugawa and colleagues reported that the mortality and readmission rates of older patients admitted to

hospital and 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, no difference was found in the complication rates of surgeries performed by male and female general surgeons in the US.[8] The postoperative mortality of patients operated on by female surgeons in Canada 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 propose more effective training for female surgeons in Japan, we used the Japanese National Clinical Database (NCD), which is the most extensive surgical database in Japan, to compare the surgical outcomes of female and male surgeons. We also examined the relation between postoperative mortality and surgical complication rates and the surgeon's licensing terms.

Methods

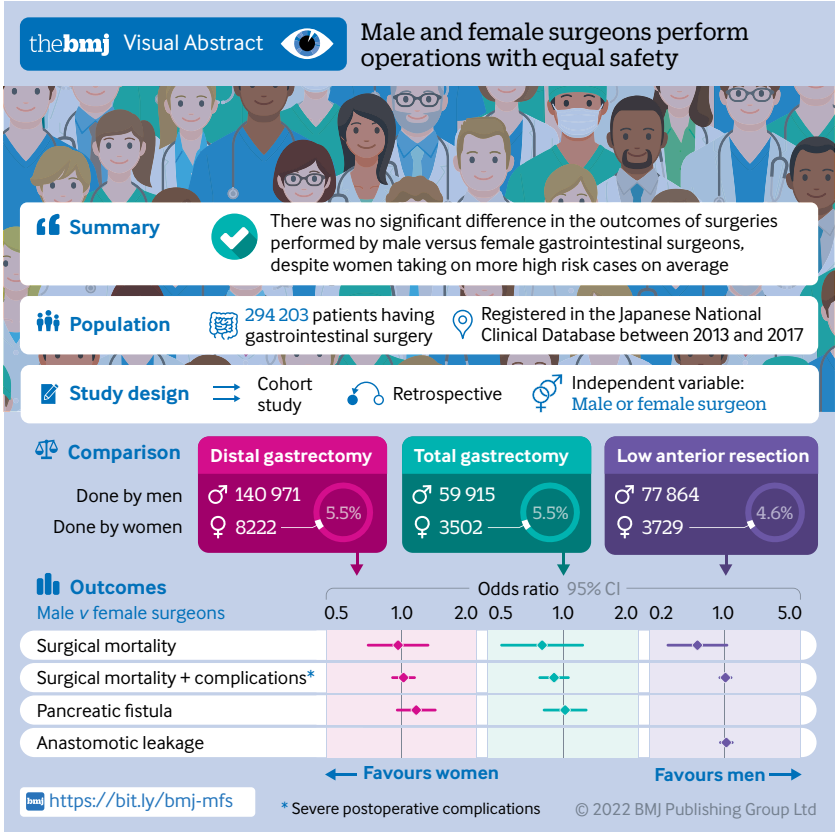
Study design and data source

We did a retrospective cohort study using data from the gastroenterological surgery section of the NCD. The NCD started data registration for surgical procedures in 2011.[10] By December 2019, 5276 facilities were registered with the NCD. Approximately 1.5 million surgical cases are registered in this database each year, which is equivalent to more than 95% of all surgeries in Japan.[11] The eligibility criteria for the NCD are accessible online (https://www.ncd.or.jp/). The NCD data entry system does not allow missing values except for laboratory data that were not taken for the patient. Validity of the data entries is evaluated through site visits and audits every year and has been shown 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—distal gastrectomy, total gastrectomy, and low anterior resection. We chose these three procedures because the number of female surgeons who did these surgeries was sufficient for analysis without the individual surgeon being identified. Other procedures among the aforementioned eight were difficult to analyse because fewer female surgeons did these procedures. The NCD does not directly contain information on surgeons' gender or the number of years since the registration of licensed doctors, but it does contain the licence number of the surgeons. We used these licence numbers to link the NCD information with the gender profile and the year of licensing registration for Japanese Society of Gastroenterological Surgery (JSGS) members.

We included surgeries done between 1 January 2013 and 31 December 2017. We excluded surgeries done by non-JSGS members because we assumed them to be doctors specialising in other surgical fields, such as cardiovascular surgery. In Japan, these doctors need to complete a general surgery programme, which includes performing gastroenterological surgery, to enter a subspeciality programme. Therefore, we considered them to be separate from doctors who specialise in gastroenterological surgery, and we also considered the effect on outcome to be different for surgeries done by these doctors. We excluded distal gastrectomy or total gastrectomy surgeries that were not for gastric cancer and low anterior resection surgeries that were not for colorectal cancer. We also excluded patients younger than 18 years, emergency surgery cases, those with unknown T/N factor in the TNM classification, and patients with metastasis because we aimed to assess the quality of surgery done as standard or major procedures, 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. We followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for this study.

Outcomes

Primary outcomes were surgical mortality, surgical mortality combined with severe postoperative complications, pancreatic fistula (in distal gastrectomy/total gastrectomy only), and anastomotic leakage (in low anterior resection only). In this study, we defined surgical mortality as all cause death up to 30 days



doi: 10.1136/bmj-2022-070568 | BMJ 2022;378:e070568 | thebmj

postoperatively, including death that occurred after discharge, and deaths that occurred within 90 days postoperatively during the index hospital admission. The extended timeframe for mortality during the index hospital admission 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 do within 30 days.[13] This measure has been commonly used in previous NCD based research to evaluate surgical outcomes.[13 14] We defined severe postoperative complications as any postoperative surgical and medical complications with a Clavien-Dindo classification of ≥3 that occurred within 30 days postoperatively.[15] Dindo and colleagues proposed the Clavien-Dindo classification for evaluating postoperative complications and comparing them among different hospitals; a Clavien-Dindo grade of ≥3 indicates that surgical, endoscopic, or radiological procedures are needed for the treatment of the complication.[15] We defined pancreatic fistula as a fistula of grade B or C according to the grading system proposed by the International Study Group of Pancreatic Fistula.[16] We defined anastomotic leakage 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; we considered these as intraoperative outcomes.

Statistical analysis

We used the χ^2 test for categorical variables and the Mann-Whitney U test for continuous variables when comparing baseline characteristics and short term outcomes. We constructed a multilevel multivariable logistic regression model for each surgical procedure, with adjustment for characteristics of the patient, surgeon, and hospital, to examine the association between surgeons' gender and surgery related mortality or surgical complications. We used a multilevel model to account for unmeasured hospital level characteristics.[17] We used hospital identification as a random intercept. An adjusted odds ratio of >1 indicated a higher risk and an adjusted odds ratio of <1 indicated a lower risk of the analysed outcome.

Patients' characteristics included age (<70 *v* ≥70 years), sex (male *v* female), body mass index (≤18.5 *v* >18.5, <25 *v* ≥25), American Society of Anesthesiologists Physical Status (ASA-PS; 1-2 *v* ≥3), clinical T factor (T1-2 *v* T3-4) and N factor (0 *v* 1-3) of tumours (N factor was included only for distal gastrectomy and total gastrectomy; based on the Union for International Cancer Control TNM classification, 7th edition), haemoglobin concentration (male: <13.5 *v* ≥13.5 g/dL; female: <11.5 *v* ≥11.5 g/dL), aspartate aminotransferase concentration (<35 *v* ≥35 IU/L; included in distal gastrectomy and total gastrectomy), albumin concentration (<3.5 *v* ≥3.5 g/dL), blood urea nitrogen concentration (<8.0 *v* ≥8.0 mg/dL), creatinine concentration (<1.2 *v* ≥1.2 mg/dL), presence/absence of diabetes mellitus, smoking status, habitual drinking status (only in low anterior resection), dependence in activities of daily living, history of chronic obstructive pulmonary disease, dialysis, ischaemic heart disease, congestive heart failure, long term steroid use, history of cerebrovascular diseases (only in low anterior resection), 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 relation between the variable and outcome. The surgical approach (open or laparoscopic) was included as an intraoperative factor.

Surgeons' characteristics included gender and years since licence registration in five year increments. We categorised years after medical licence registration on the basis of the following assumptions to account for their acquired surgical skills in the Japanese board certification and surgery training system: surgeons with an experience of five years or less were considered not to have completed the general surgery training programme; those with an experience of six to 10 years were assumed to be board certified general surgeons;

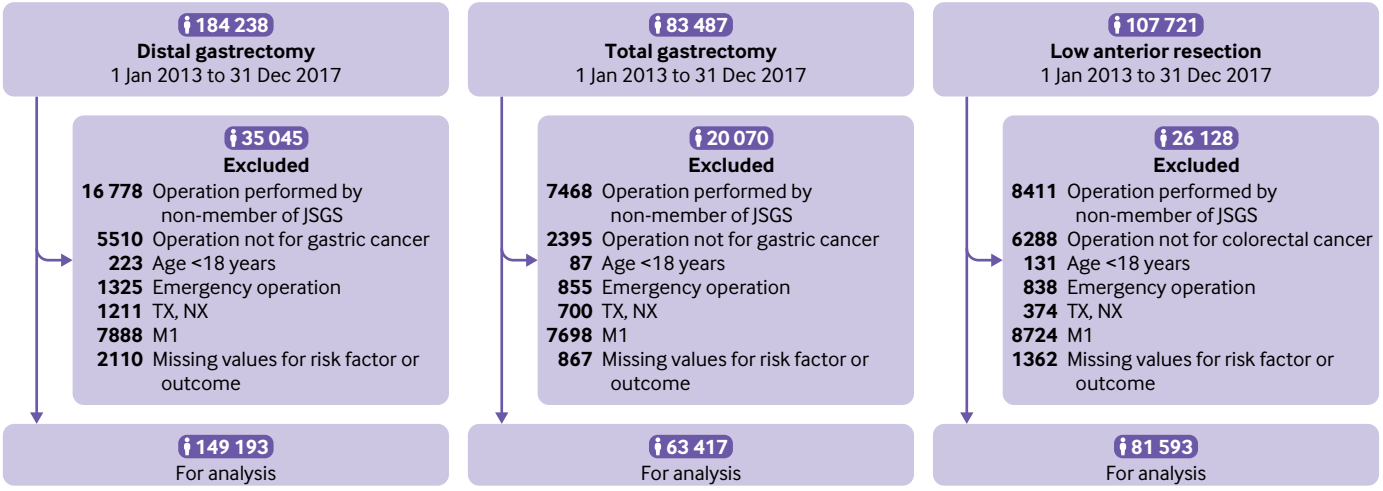


Fig 1 | Flow diagram for patient selection. JSGS=Japanese Society of Gastroenterological Surgery; M1=positive for distant metastasis; NX=N unknown; TX=T unknown

Table 1 | Surgeons' characteristics by gender according to number of cases. Values are numbers (percentages) unless stated otherwise

Characteristics	Male surgeons	Female surgeons
Distal gastrectomy		
Total cases of operation	140 971 (94.5)	8222 (5.5)
Median (IQR) years since registration of medical licence	16 (9-22)	9 (5-13)
Cases of operation according to years since registration of medical licence:		
≤5	19 246 (88.4)	2534 (11.6)
6-10	21 526 (89.9)	2430 (10.1)
11-15	27 084 (93.5)	1898 (6.5)
16-20	28 609 (97.0)	881 (3.0)
≥21	44 506 (98.9)	479 (1.1)
Total gastrectomy		
Total cases of operation	59 915 (94.5)	3502 (5.5)
Median (IQR) years since registration of medical licence	16 (9-23)	9 (5-14)
Cases of operation according to years since registration of medical licence:		
≤5	7959 (87.7)	1115 (12.3)
6-10	9097 (90.2)	989 (9.8)
11-15	11 204 (93.7)	749 (6.3)
16-20	11 956 (96.4)	441 (3.6)
≥21	19 699 (99.0)	208 (1.0)
Low anterior resection		
Total cases of operation	77 864 (95.4)	3729 (4.6)
Median (IQR) years since registration of medical licence	17 (11-23)	9 (6-15)
Cases of operation according to years since registration of medical licence:		
≤5	7066 (88.9)	885 (11.1)
6-10	10 576 (89.8)	1198 (10.2)
11-15	15 643 (94.8)	853 (5.2)
16-20	17 698 (96.9)	562 (3.1)
≥21	26 881 (99.1)	231 (0.9)

IQR=interquartile range.

11-15 years, board certified gastroenterological surgeons; 16-20, board certified trainers; and ≥21 years, directors (or a similar position) of surgical departments.

We categorised hospitals into quarters according to the annual number of cases of each procedure so that each category contained approximately the same number of cases to increase statistical power: very low, low, high, and very high (we defined very low, low, high, and very high as <15, 15 to <30, 30 to <50, and ≥50 for distal gastrectomy; <7, 7 to <13, 13 to <21, and ≥21 for total gastrectomy; and <8, 8 to <16, 16 to <29, and ≥29 for low anterior resection). On the basis of previous research on the volume-outcome relation, we assumed a non-linear association.[18]

Subsequently, we did an additional analysis to examine whether an interaction effect existed between gender and years since medical licence registration. We incorporated an interaction term of gender and years of experience post-medical licence registration, instead of including them individually as in the previous regression model. We excluded patients with missing data from this study because the proportion of cases with missing values was low in all three surgical procedures (distal gastrectomy 1.39%, total gastrectomy 1.35%, low anterior resection 1.64%).

Post hoc sensitivity analysis

To assess the robustness of the results, we did several analyses after the completion of the main analysis as a post hoc sensitivity analysis. Firstly, although the proportion of missing values was low and a complete case analysis was conducted, we compared cases with and without missing values and repeated the main analysis with a multiple imputed dataset. We assumed the mechanism of missingness to be at random.[19] We did imputation with a chained equation, and the number of imputed datasets was set to five.[20 21]

Secondly, we included patient's age, patient's body mass index, number of years after medical licence registration, and hospital case volume in the regression analysis as continuous variables instead of categorical variables. We included this analysis to explore confounding effects that might vary from previous studies depending on how the non-linear relation between the variable and outcome is modelled. We used a generalised additive model to account for the assumed non-linearity between the variables and the outcome.[22]

Thirdly, we added surgeon's case volume and region of the hospital to the regression model. We assumed that case volume is a surrogate of surgical experience that significantly affects outcome. For hospitals, hospital case volume would reflect surgical experience. For surgeons, we considered years since licence registration to be a more accurate measure of surgical experience than annual case volume of the individual surgeon because the years since 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's case volume may be a confounder, we included it as an additional variable in the regression model. We treated it as a continuous variable, and applied a smooth term of a generalised additive model to model a non-linear relation.[22] We additionally included the region of the hospital as a variable to partly account for the socioeconomic status of a patient. Socioeconomic status is not available in the NCD, and research on the relation between socioeconomic status and surgical outcomes in Japan is scarce. One study in Japan found no significant association between regional average income, which was considered to be one aspect of socioeconomic status, and outcome in cardiovascular surgery,[23] but whether regional mean household income reflects an individual's socioeconomic status and whether the results can be applied in gastroenterological surgery are unknown. Therefore, the considerable magnitude of socioeconomic status as a confounder could not be denied. We categorised the region of a hospital into urban or rural areas on the basis of those used in a previous Japanese study, which distinguished urban areas from rural ones according to the OECD definition.[24] Thirteen 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.

Table 2 | Surgeons' characteristics by gender according to number of surgeons. Values are numbers (percentages) unless stated otherwise

Characteristic	Male surgeons	Female surgeons
Distal gastrectomy		
Number of surgeons	9433 (92.3)	788 (7.7)
Median (IQR) years since registration of medical licence*	16 (8-25)	7 (4-13)
Number of surgeons according to years since registration of medical licence†:		
≤5	1430 (83.0)	292 (17.0)
6-10	1679 (87.4)	243 (12.6)
11-15	1381 (91.5)	129 (8.5)
16-20	1415 (94.9)	76 (5.1)
≥21	3528 (98.7)	48 (1.3)
Total gastrectomy		
Number of surgeons	8238 (92.8)	640 (7.2)
Median (IQR) years since registration of medical licence*	16 (8-24)	7 (5-13)
Number of surgeons according to years since registration of medical licence†:		
≤5	1286 (84.6)	234 (15.4)
6-10	1466 (88.4)	192 (11.6)
11-15	1258 (92.3)	105 (7.7)
16-20	1275 (94.8)	70 (5.2)
≥21	2953 (98.7)	39 (1.3)
Low anterior resection		
Number of surgeons	8200 (92.9)	627 (7.1)
Median (IQR) years since registration of medical licence*	16 (8-24)	8 (5-13)
Number of surgeons according to years since registration of medical licence†:		
≤5	1200 (85.2)	209 (14.8)
6-10	1380 (87.5)	198 (12.5)
11-15	1294 (92.0)	112 (8.0)
16-20	1326 (95.3)	65 (4.7)
≥21	3000 (98.6)	43 (1.4)

IQR=interquartile range.
*Highest number of years since registration of medical licence during 5 year study period was assigned to each surgeon. Years since registration for individual surgeon may increase during study period.
†Highest category (in terms of seniority) was selected to categorise each surgeon. Surgeons may move to higher category during 5 year study period.

Fourthly, as the number of surgeries done 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, we combined the study population for distal gastrectomy, total gastrectomy, and low anterior resection and analysed them as a single population. We included the type of surgical procedure as a covariate and repeated the main analysis. We assessed the relations between a surgeon's gender and surgical mortality, surgical mortality or postoperative complication with Clavien-Dindo classification ≥3, and anastomotic leakage.

Fifthly, as we found female surgeons to be more likely to be assigned to patients at higher risk, we compared surgical outcomes between male and female surgeons within the predicted risk strata. We calculated the predicted risk on the basis of the regression analysis; we excluded the doctor's gender as a variable. We categorised the predicted risk into five strata, from low to high risk, using the fifth of predicted risk.

All P values were two sided, and we considered P values <0.05 to be significant. We used R software (version 3.6.3) for statistical analyses.

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 of the surgical outcomes of female surgeons being comparable to those of their male counterparts. In Japanese society, the fact that women spend more time engaged in housework and childcare, making working in a profession such as surgery difficult for them to, has been a concern. 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 had distal gastrectomy, total gastrectomy, and low anterior resection, respectively, at Japanese institutes and were registered in the Japanese NCD between 2013 and 2017. Figure 1 shows the flow diagram for surgical case selection. Finally, 149 193 distal gastrectomy, 63 417 total gastrectomy, and 81 593 low anterior resection surgeries were eligible. A total of 140 971 (94.5%) eligible distal gastrectomy surgeries were done by male surgeons and 8222 (5.5%) by female surgeons; 59 915 (94.5%) eligible total gastrectomy surgeries were done by male surgeons and 3502 (5.5%) by female surgeons; and 77 864 (95.4%) eligible low anterior resection procedures were done by male surgeons and 3729 (4.6%) by female surgeons (table 1). The numbers of male surgeons who participated in distal gastrectomy, total gastrectomy, and low anterior resection were 9433 (92.3%), 8238 (92.8%), and 8200 (92.9%), respectively, and those of female surgeons were 788 (7.7%), 640 (7.2%), and 627 (7.1%), respectively (table 2). Female surgeons had fewer years of experience after licence registration than did male surgeons (9 *v* 16 years for distal gastrectomy/total gastrectomy; 9 *v* 17 years for low anterior resection).

Characteristics of institutions and patients

Table 3, table 4, and table 5 show the institutional factors, preoperative and intraoperative factors, intraoperative outcomes, and postoperative outcomes of distal gastrectomy, total gastrectomy, and low anterior resection, respectively. For distal gastrectomy, female surgeons were more distributed in hospitals with low (28.4%) and high (27.1%) case numbers than in those with very low (22.7%) or very high (21.8%) case numbers. For total gastrectomy, female surgeons were less distributed in hospitals with very high case numbers (20.7%) than in those in other categories. For low anterior resection, female surgeons were more typically distributed in hospitals with low (29.5%) numbers than in those with very low (23.3%), high (24.0%), or very high (23.2%) numbers.

Female surgeons did surgeries on patients at relatively high risk. Importantly, female surgeons

Table 3 | Institutional and operative characteristics by surgeon's gender in distal gastrectomy. Values are numbers (percentages) unless stated otherwise

	Male surgeons (n=140 971)	Female surgeons (n=8222)	P value
Institutional characteristics			
Median (IQR) No of surgeries per year	30 (15-54)	29 (16-52)	0.73
No of surgeries per year:			<0.001
<15	34 733 (24.6)	1867 (22.7)	
≥15, <30	35 826 (25.4)	2337 (28.4)	
≥30, <50	36 092 (25.6)	2226 (27.1)	
≥50	34 320 (24.3)	1792 (21.8)	
Preoperative factors			
Median (IQR) age, years	71 (64-78)	72 (65-78)	<0.001
Age ≥70 years	78 418 (55.6)	4840 (58.9)	<0.001
Female sex	46 798 (33.2)	2820 (34.3)	0.04
Median (IQR) body mass index	22.2 (20.0-24.4)	22.0 (19.8-24.3)	<0.001
Body mass index:			<0.001
≥18.5, <25	95 141 (67.5)	5556 (67.6)	
<18.5	17 118 (12.1)	1119 (13.6)	
≥25	28 712 (20.4)	1547 (18.8)	
Diabetes mellitus	25 484 (18.1)	1579 (19.2)	0.01
Smoking	63 731 (45.2)	3777 (45.9)	0.20
Dependence in activities of daily living	5965 (4.2)	401 (4.9)	0.005
Chronic obstructive pulmonary disease	6822 (4.8)	412 (5.0)	0.50
Dialysis	1062 (0.8)	77 (0.9)	0.07
History of ischaemic heart disease	5260 (3.7)	332 (4.0)	0.16
Congestive heart failure (within 30 days)	976 (0.7)	67 (0.8)	0.22
Long 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
Haemoglobin <13.5 g/dL in males, <11.5 g/dL in females	39 344 (27.9)	2459 (29.9)	<0.001
Albumin <3.5 g/dL	21 128 (15.0)	1519 (18.5)	<0.001
Blood urea nitrogen <8 mg/dL	19 371 (13.7)	1158 (14.1)	0.39
Creatinine >1.2 mg/dL	9961 (7.1)	626 (7.6)	0.06
Aspartate aminotransferase >35 IU/L	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
≥T3 (in TNM classification)	42 441 (30.1)	2939 (35.7)	<0.001
≥N1 (in TNM classification)	44 193 (31.3)	3056 (37.2)	<0.001
ASA-PS 3-5	15 563 (11.0)	1006 (12.2)	0.001
Intraoperative characteristics			
Surgical approach laparoscopic	74 282 (52.7)	2944 (35.8)	<0.001
Intraoperative outcomes			
Median (IQR) operating time, min	259 (205-320)	261 (209-322)	0.001
Median (IQR) estimated blood loss, mL	100 (25-250)	150 (50-327)	<0.001
Postoperative outcomes			
Median (IQR) postoperative hospital stay, days	13 (10-19)	14 (10-20)	<0.001

ASA-PS=American Society of Anesthesiologists Physical Status; IQR=interquartile range.

operated on older patients (58.9% *v* 55.6% for distal gastrectomy; 60.4% *v* 56.4% for total gastrectomy; 45.9% *v* 43.8% for low anterior resection) and on patients with diabetes mellitus (19.2% *v* 18.1%; 21.2% *v* 18.6%; 19.4% *v* 18.0%), dependence in activities of daily living (4.9% *v* 4.2%; 4.8% *v* 3.8%; 4.6% *v* 3.5%), lower haemoglobin (29.9% *v* 27.9%; 37.6% *v* 35.2%; 28.7% *v* 27.0%) and serum albumin (18.5% *v* 15.0%; 22.5% *v* 19.2%; 14.3% *v* 12.1%) concentrations, and higher T factors (35.7% *v* 30.1%; 58.3% *v* 55.1%; 63.7% *v* 60.0%) in all three procedures at a higher rate than their male counterparts (table 3; table 4; table 5). Additionally, female surgeons did distal gastrectomy in patients with long term steroid use (1.3% *v* 1.0%), weight loss (4.7% *v* 3.6%), preoperative blood transfusion (3.1% *v* 2.0%), a higher N factor (37.2% *v* 31.3%), and a worse ASA-PS (12.2% *v* 11.0%); total gastrectomy in patients who smoked (52.0% *v* 49.2%), were receiving dialysis (1.0% *v* 0.6%), had weight loss (6.7% *v* 5.8%), and had a higher N factor (51.4% *v* 48.9%); and low anterior resection in patients with a history of cerebrovascular disease (3.9% *v* 3.1%) at a higher rate than their male counterparts. By contrast, male surgeons did surgeries on patients who had had preoperative chemotherapy (2.2% *v* 1.9% for distal gastrectomy; 6.9% *v* 5.5% for total gastrectomy; 6.5% *v* 4.2% for low anterior resection) or radiotherapy (3.1% *v* 1.7% for low anterior resection) at a higher rate than their female counterparts.

Intraoperative factors and outcomes

Female surgeons did fewer laparoscopic surgeries (35.8% *v* 52.7% for distal gastrectomy; 13.0% *v* 26.3% for total gastrectomy; 60.4% *v* 69.6% for low anterior resection) than male surgeons. Significantly more blood loss was observed in all three procedures performed by female surgeons (150 *v* 100 mL for distal gastrectomy; 320 *v* 260 mL for total gastrectomy; 80 *v* 52 mL for low anterior resection) than by male surgeons (table 3; table 4; table 5).

Postoperative outcomes

After adjustment for characteristics of patients, surgeons, and hospitals, we observed no significant difference in the risk for surgical mortality in distal gastrectomy, total gastrectomy, and low anterior resection between male and female surgeons, as shown in figure 2 (risk adjusted odds ratio 0.98 (95% confidence interval 0.74 to 1.29) for distal gastrectomy; 0.83 (0.57 to 1.19) for total gastrectomy; 0.56 (0.30 to 1.05) for low anterior resection). The adjusted risk for surgical mortality or postoperative complication rated Clavien-Dindo ≥3 were similar for distal gastrectomy, total gastrectomy, and low anterior resection (risk adjusted odds ratio 1.03 (0.93 to 1.14) for distal gastrectomy; 0.92 (0.81 to 1.05) for total gastrectomy; 1.02 (0.91 to 1.15) for low anterior resection), pancreatic fistula for distal gastrectomy and total gastrectomy (risk adjusted odds ratio 1.16 (0.97 to 1.38) for distal gastrectomy; 1.02 (0.84 to 1.23) for total gastrectomy), and anastomotic leakage for low anterior resection (risk adjusted odds ratio 1.04, 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 licensing categories.

Distal gastrectomy

The adjusted risk for surgical mortality was higher for female surgeons than for male surgeons with five years or less experience after registration (risk adjusted odds ratio 1.64, 1.07 to 2.52) (fig 3). For surgery related death or postoperative adverse events rated

Table 4 | Institutional and operative characteristics by surgeon's gender in total gastrectomy. Values are numbers (percentages) unless stated otherwise

	Male surgeons (n=59 915)	Female surgeons (n=3502)	P value
Institutional characteristics			
Median (IQR) No of surgeries per year	13 (7-22)	13 (7-20)	0.09
No of surgeries per year:			<0.001
<7	15 790 (26.4)	906 (25.9)	
≥7, <13	14 037 (23.4)	913 (26.1)	
≥13, <21	14 379 (24.0)	957 (27.3)	
≥21	15 709 (26.2)	726 (20.7)	
Preoperative factors			
Median (IQR) age, years	71 (64-77)	72 (66-78)	<0.001
Age ≥70 years	33 821 (56.4)	2115 (60.4)	<0.001
Female sex	15 127 (25.2)	906 (25.9)	0.41
Median (IQR) body mass index	21.9 (19.7-24.2)	21.8 (19.6-24.0)	0.04
Body mass index:			0.08
≥18.5, <25	40 293 (67.3)	2377 (67.9)	
<18.5	8680 (14.5)	534 (15.2)	
≥25	10 942 (18.3)	591 (16.9)	
Diabetes mellitus	11 133 (18.6)	743 (21.2)	<0.001
Smoking	29 485 (49.2)	1821 (52.0)	0.001
Dependence in activities of daily living	2298 (3.8)	169 (4.8)	0.003
Chronic obstructive pulmonary disease	3135 (5.2)	199 (5.7)	0.25
Dialysis	331 (0.6)	34 (1.0)	0.001
History of ischaemic heart disease	2335 (3.9)	147 (4.2)	0.37
Congestive heart failure (within 30 days)	356 (0.6)	21 (0.6)	0.97
Long 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
Haemoglobin <13.5 g/dL in males, <11.5 g/dL in females	21 117 (35.2)	1316 (37.6)	0.005
Albumin <3.5 g/dL	11 513 (19.2)	788 (22.5)	<0.001
Blood urea nitrogen <8 mg/dL	8223 (13.7)	520 (14.8)	0.061
Creatinine >1.2 mg/dL	4191 (7.0)	269 (7.7)	0.12
Aspartate aminotransferase >35 IU/L	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
≥T3 (in TNM classification)	33 028 (55.1)	2040 (58.3)	<0.001
≥N1 (in TNM classification)	29 307 (48.9)	1799 (51.4)	0.005
ASA-PS 3-5	6694 (11.2)	421 (12.0)	0.12
Intraoperative characteristics			
Surgical approach laparoscopic	15 762 (26.3)	456 (13.0)	<0.001
Intraoperative outcomes			
Median (IQR) operating time, min	282 (221-354)	279 (225-347)	0.38
Median (IQR) estimated blood loss, mL	260 (100-521)	320 (150-595)	<0.001
Postoperative outcomes			
Median (IQR) postoperative hospital stay, days	16 (12-24)	16 (12-23)	0.18

ASA-PS=American Society of Anesthesiologists Physical Status; IQR=interquartile range.

Clavien-Dindo ≥3, female surgeons with five years or less experience after registration had a higher odds ratio (risk adjusted odds ratio 1.19, 1.01 to 1.41), whereas those with six to 10 years of experience after registration had a lower odds ratio (0.79, 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.

Total gastrectomy

The adjusted risk for surgery related death, postoperative adverse events rated Clavien-Dindo ≥3, and pancreatic fistula showed no significant differences between male and female surgeons at any year category after registration (fig 4).

Low anterior resection

The adjusted odds ratio for surgical mortality did not differ significantly between male and female surgeons at any year since licensing category (fig 5). The adjusted risk for surgical mortality or postoperative adverse events rated Clavien-Dindo ≥3 was higher for female surgeons than for male surgeons at the 16-20 years of experience category (risk adjusted odds ratio 1.41, 1.07 to 1.86). The adjusted risk for anastomotic leakage was lower for female surgeons with five years or less experience (risk adjusted odds ratio 0.71, 0.53 to 0.94).

Sensitivity analysis

The results of the sensitivity analyses are summarised in supplementary figures A-D and supplementary tables A-G. The proportion of missing values was relatively higher 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, patient's body mass index, hospital case volume, and surgeon's years after medical licence registration changed to the original continuous scale; and additional covariates (surgeon's case volume and urban-rural status) included, the changes in the point estimate and 95% confidence intervals were minimal compared with those in the main analysis (supplementary figures A-C), except for one of the analyses in low anterior resection that included surgeon's case volume and region of the hospital as additional covariates. As shown in supplementary figure C, this analysis showed a significant decrease in adjusted odds ratio for surgical mortality for female surgeons (adjusted odds ratio 0.54, 0.29 to 0.996). We found no significant differences between male and female surgeons when stratified with predicted risks of the outcome except for the highest fifth of predicted risk for pancreatic fistula in distal gastrectomy (1365 (4.9%) outcomes for male surgeons versus 115 (6.1%) for female surgeons; P=0.02; supplementary table E) and the second fifth of predicted risk for anastomotic leakage in low anterior resection (676 (4.3%) outcomes for male surgeons versus 49 (6.4%) for female surgeons; P=0.008; supplementary table G). Finally, the analysis of the three surgical procedures combined showed no significant association between female surgeons and surgical outcomes (supplementary figure D).

Discussion

Using the NCD data for 2013-17, we found no overall significant difference in the risk for surgical mortality in the three procedures performed by male and female surgeons after adjustment for confounders. We also found no significant difference between male and female surgeons in the adjusted risk for surgical mortality or Clavien-Dindo ≥3 complications in distal gastrectomy, total gastrectomy, and low anterior resection; pancreatic fistula in distal gastrectomy and total gastrectomy, and anastomotic leakage in low

Table 5 | Institutional and operative characteristics by surgeon's gender in low anterior resection. Values are numbers (percentages) unless stated otherwise

	Male surgeons (n=77 864)	Female surgeons (n=3729)	P value
Institutional characteristics			
Median (IQR) No of surgeries per year	16 (9-27)	15 (8-26)	0.007
No of surgeries per year:			0.01
<8	17 655 (22.7)	870 (23.3)	
≥8, <16	21 468 (27.6)	1100 (29.5)	
≥16, <29	20 112 (25.8)	895 (24.0)	
≥29	18 629 (23.9)	864 (23.2)	
Preoperative factors			
Median (IQR) age, years	68 (61-75)	68 (62-75)	0.004
Age ≥70 years	34 077 (43.8)	1711 (45.9)	0.01
Female sex	26 958 (34.6)	1353 (36.3)	0.04
Median (IQR) body mass index	22.3 (20.1-24.7)	22.2 (20.0-24.6)	0.01
Body mass index:			0.27
≥18.5, <25	51 808 (66.5)	2471 (66.3)	
<18.5	8838 (11.4)	454 (12.2)	
≥25	17 218 (22.1)	804 (21.6)	
Diabetes mellitus	14 049 (18.0)	722 (19.4)	0.04
Smoking	33 997 (43.7)	1620 (43.4)	0.81
Habitual drinking	41 677 (53.5)	1937 (51.9)	0.06
Dependence in activities of daily living	2725 (3.5)	170 (4.6)	0.001
Chronic obstructive pulmonary disease	2800 (3.6)	104 (2.8)	0.01
Dialysis	391 (0.5)	23 (0.6)	0.40
History of ischaemic heart disease	2277 (2.9)	122 (3.3)	0.24
Congestive heart failure (within 30 days)	411 (0.5)	25 (0.7)	0.29
Long term steroid use	611 (0.8)	28 (0.8)	0.89
History of cerebrovascular disease	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
Haemoglobin <13.5 g/dL in males, <11.5 g/dL in females	21 036 (27.0)	1072 (28.7)	0.02
Albumin <3.5 g/dL	9417 (12.1)	533 (14.3)	<0.001
Blood urea nitrogen <8 mg/dL	9306 (12.0)	447 (12.0)	0.97
Creatinine >1.2 mg/dL	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
≥T3 (in TNM classification)	46 697 (60.0)	2375 (63.7)	<0.001
ASA-PS 3-5	7155 (9.2)	344 (9.2)	0.96
Intraoperative characteristics			
Surgical approach laparoscopic	54 199 (69.6)	2252 (60.4)	<0.001
Intraoperative outcomes			
Median (IQR) operating time, min	265 (204-345)	269 (210-343)	0.04
Median (IQR) estimated blood loss, mL	52 (10-206)	80 (15-271)	<0.001
Postoperative outcomes			
Median (IQR) postoperative hospital stay, days	15 (11-23)	15 (11-23)	0.74

ASA-PS=American Society of Anesthesiologists Physical Status; IQR=interquartile range.

anterior resection. More blood loss was recorded in all three procedures performed by female surgeons, probably because they did a significantly larger proportion of open surgeries than did male surgeons. Importantly, we found that female gastrointestinal surgeons were more often responsible for patients with comorbid conditions (for example, diabetes mellitus, anaemia, and dependence in activities of daily living), even though female gastrointestinal surgeons were responsible for fewer surgeries than male surgeons, as described by Altieri and colleagues.[25] Data from a large institution in a western country indicated that female surgeons did not do more complex cases than male surgeons, even after subspeciality and seniority were accounted for.[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 an extremely 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 distal gastrectomy performed by female surgeons with an experience of five years or less since registration, the adjusted odds ratios for surgical mortality and surgical mortality with a complication grade of Clavien-Dindo ≥3 were statistically higher than those for male surgeons of the same category. For low anterior resection, female surgeons with 16-20 years' experience had a statistically higher adjusted risk for surgical mortality combined with a complication of Clavien-Dindo grade ≥3 than did male surgeons with the same surgical experience. However, the adjusted risks for surgical mortality or a complication grade of Clavien-Dindo ≥3 in distal gastrectomy performed by female surgeons with six to 10 years of experience were lower than those for male surgeons, and the rate of leakage in low anterior resection performed by female surgeons with five years or less experience was lower than that for male surgeons. Female surgeons in other subgroups for distal gastrectomy and low anterior resection and in all subgroups for total gastrectomy tended to have comparable surgical outcomes to their male counterparts.

Furthermore, in the category of surgeons with ≥21 years of experience, we observed no significant difference in outcomes between male and female surgeons in all three surgical procedures. Tsugawa and colleagues reported that the risk adjusted mortality rate in surgeries done by female surgeons aged >50 years was the lowest; however, they mentioned that evaluating the outcome of female surgeons aged >60 years was difficult because this group was very small.[7] These findings are consistent with ours. Wallis and colleagues reported a lower 30 day mortality rate for surgeries done by female surgeons.[9] Sharoky and colleagues reported no difference in mortality or complication rates for surgeries done by male and female surgeons assessed using cardinality matching with a refined balance.[8] However, these authors did not compare surgeons by age. Further research is needed to examine how gender and age affect surgical outcomes, but we 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 surgeons' case volume and urban-rural status in low anterior resection, we observed a significant decrease in the adjusted odds ratio for surgical mortality for female surgeons. The significant difference in surgical outcomes between female and male surgeons after adjustment 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 showed non-

Outcome	Doctor's gender	No of outcomes (%)	Unadjusted odds ratio (95% CI)	P value	Adjusted odds ratio (95% CI)	P value
Distal gastrectomy						
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						
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						
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	345 (9.3)	1.04 (0.93 to 1.17)	0.50	1.04 (0.92 to 1.18)	0.49

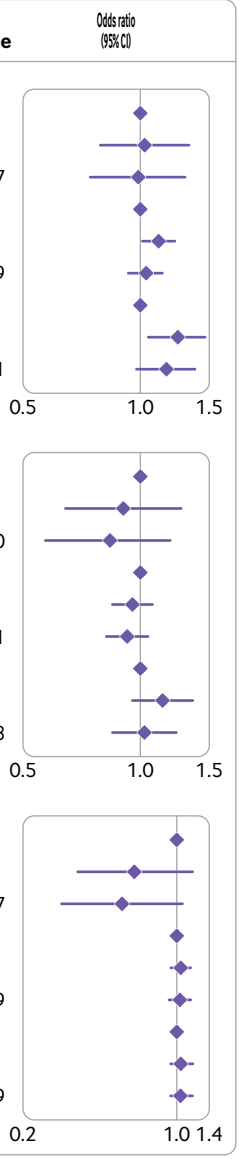


Fig 2 | Association between female surgeon and surgical outcome. CI=confidence interval; CDC=Clavien-Dindo classification; OR=odds ratio

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

Previous studies have suggested several problems related to the careers of female surgeons. Firstly, 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, female surgeons find attaining leadership positions difficult.[29]

Secondly, previous reports have shown a bias in the number of surgical cases assigned to male versus female surgeons during their training.[30] Foley and colleagues reported gender differences in the experience with robotic surgery in colorectal surgery training programmes, with female trainees having fewer opportunities to participate in the use of consoles and to complete the procedures.[31] They also reported that male supervisors provided fewer opportunities for console participation 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; thus, the process for case assignment to female surgeons by supervisors is essential in the training process for female surgeons.

Years after medical licence registration	Doctor's gender	No of outcomes (%)	Unadjusted odds ratio (95% CI)	P value	Adjusted odds ratio (95% CI)	P value
Surgical mortality in distal gastrectomy						
≤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						
≤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						
≤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	643 (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

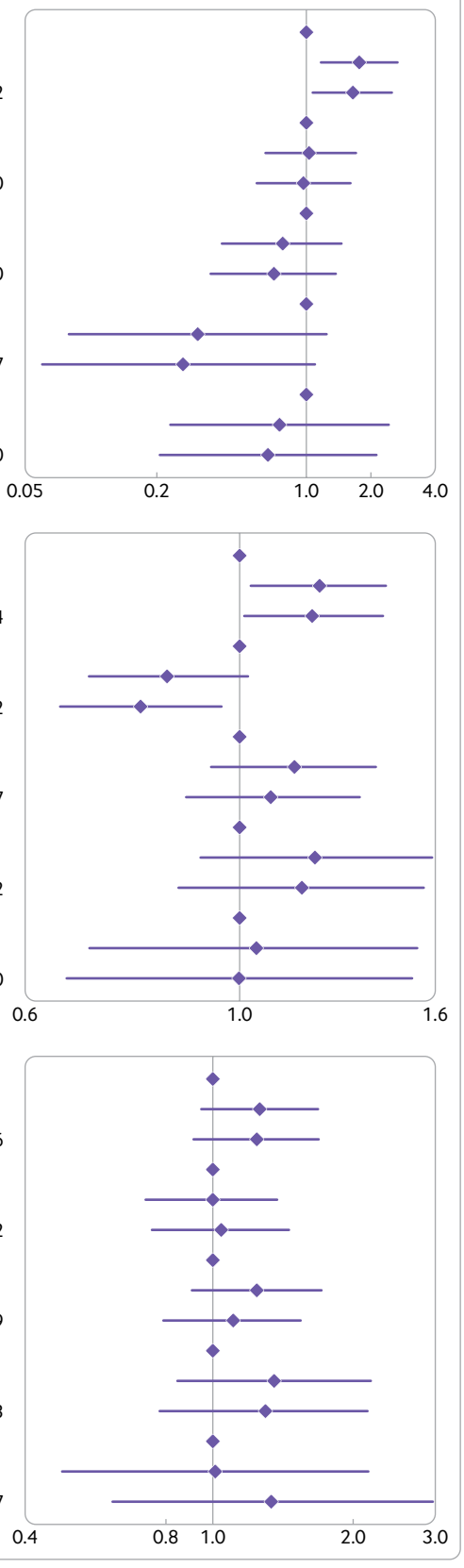


Fig 3 | Association between female surgeon and surgical outcome according to years after medical licence registration in distal gastrectomy. CI=confidence interval; CDC=Clavien-Dindo classification; OR=odds ratio

Years after medical licence registration	Doctor's gender	No of outcomes (%)	Unadjusted odds ratio (95% CI)	P value	Adjusted odds ratio (95% CI)	P value
Surgical mortality in total gastrectomy						
≤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						
≤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						
≤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

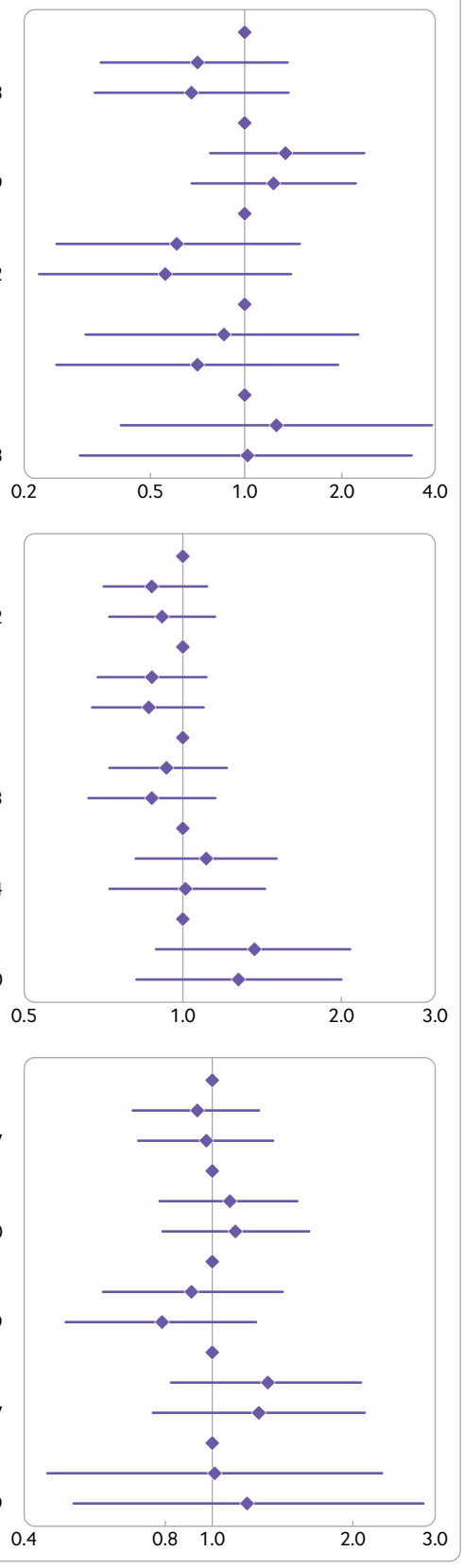


Fig 4 | Association between female surgeon and surgical outcome according to years after medical licence registration in total gastrectomy. CI=confidence interval; CDC=Clavien-Dindo classification; OR=odds ratio

Years after medical licence registration	Doctor's gender	No of outcomes (%)	Unadjusted odds ratio (95% CI)	P value	Adjusted odds ratio (95% CI)	P value
Surgical mortality in low anterior resection						
≤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						
≤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						
≤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

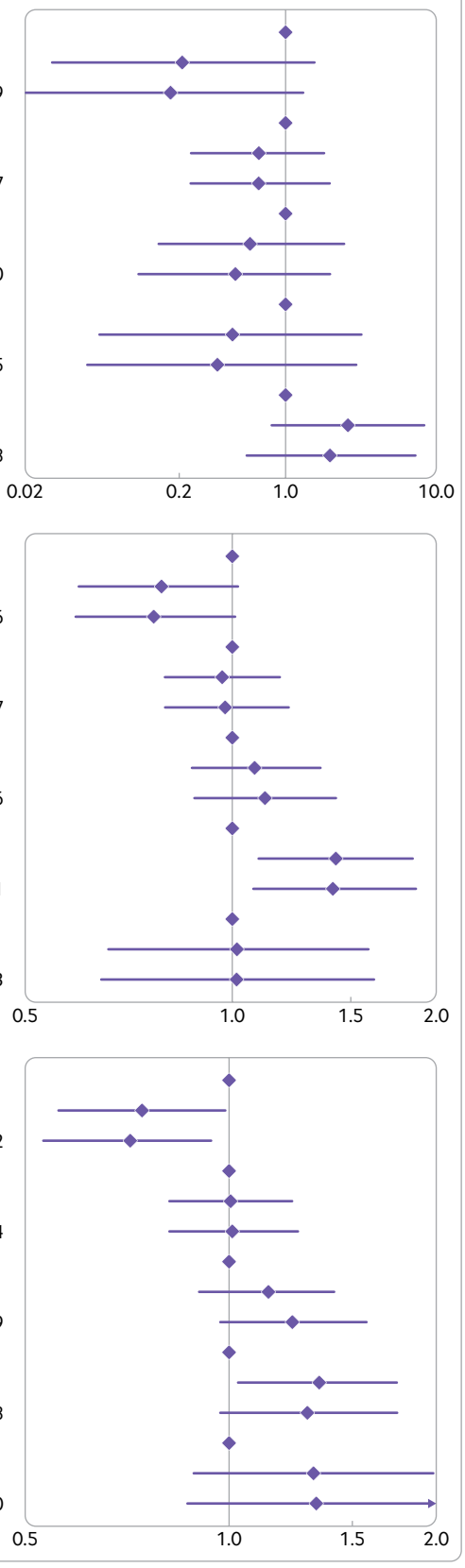


Fig 5 | Association between female surgeon and surgical outcome according to years after medical licence registration in low anterior resection. CI=confidence interval; CDC=Clavien-Dindo classification; OR=odds ratio

Thirdly, 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, as women often leave clinical practice owing to marriage, pregnancy, or childcare.[32] In traditional Japanese culture, women have often been considered unsuitable for doing surgery and are unwelcome in the field. We believed that showing that no differences existed in the results of surgical procedures performed by men and women would make it easier for women to be accepted as surgeons and professionals. Fourthly, work-family conflict is more pronounced among female surgeons, and they may experience burnout.[33 34]

Many aspects can impair the successful development of female surgeons. Nevertheless, in this analysis, no significant difference existed in the mortality or complication rates of surgeries done by female and male surgeons, suggesting that they are equally successful in developing their surgical skills. Notably, female surgeons did a lower percentage of laparoscopic procedures in all three procedures than male surgeons did. A tendency may have existed 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. Surgical teams need to welcome women as members, and gender equality needs to be 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 that is highly accurate in terms of patients' preoperative condition and surgical outcomes.

This study had some limitations. Firstly, this was an observational study, and we could not adjust for unmeasured confounders. Certain data—for example, regarding the socioeconomic status of a patient—were not available in the NCD. Secondly, because the number of female surgeons was smaller than that of male surgeons, a bias may have existed in that the outcomes of one female surgeon had a large effect on the overall outcomes. When interpreting the results, because there are so few female surgeons, a single adverse event can significantly affect the entire result; this is not the case for male surgeons. Thirdly, the study lacks details on surgeons' work and personal life conditions (part time or full time, family structure, and so on). Fourthly, as we intended to include only patients with relatively standard procedures performed 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

On the basis of our results, female Japanese surgeons took on high risk cases, and no significant differences existed in surgical mortality or Clavien-Dindo ≥3 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.

AUTHOR AFFILIATIONS

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

We sincerely appreciate all the participants in the National Clinical Database (NCD) project for their extraordinary efforts in data registration, and we are grateful to the Japanese Society of Gastrointestinal Surgery (JSGS) members who participated in this study. We also thank James R Goldenring for his proofreading assistance and Editage (www.editage.com) for English language editing.

Contributors: 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. KO and HE are equal contributors to this work and designated as co-first authors. All authors read the drafted manuscript, provided feedback, and approved the final submitted version. SN is the guarantor. 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: no support from any organisation for the submitted work; HE, HY, and HM have received grants from the National Clinical Database, Johnson & Johnson, and Nipro; KY has received 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 has received 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 and Company, Nippon Covidien, Ethicon, Ono Pharmaceutical, Olympus, Bristol-Myers Squibb, AstraZeneca, MSD, Smith and Nephew, and Kaken Pharmaceutical; 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 needed 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 JSGS were provided with the opportunity to opt out via email messages and through a website.

Data sharing: Data on individual surgical cases and surgeons reported in this study are not publicly available. Aggregate data, including data reported in this study, can be accessed by submitting a research plan to the NCD Office and requesting access, 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.

The corresponding author 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 (and, if relevant, registered) have been explained.

Dissemination to participants and related patient and public communities: Two plans exist for dissemination of the findings of this study. We plan to use formal scientific communication platforms for clinicians and professional societies, including sharing publication notices of scientific articles and promoting them through social media. In addition, research results will be shared in national and international forums, including specialty societies. For the general public, we will pursue engagement with the available media through press releases of research results and meetings with relevant government agencies.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Organisation for Economic Co-Operation and Development. Health care resources: Physicians by age and gender. https://stats.oecd.org/index.aspx?queryid=30172#.
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.
3 Association of American Medical Colleges. Active Physicians by Sex and Specialty, 2019. https://www.aamc.org/data-reports/workforce/interactive-data/active-physicians-sex-and-specialty-2019.
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.
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.
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. doi:10.1001/jamainternmed.2016.7875
7 Tsugawa Y, Jena AB, Orav EJ. Age and sex of surgeons and mortality of older surgical patients: observational study. *BMJ* 2018;361:k1343. doi:10.1136/bmj.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. doi:10.1097/SLA.0000000000002460
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. doi:10.1136/bmj.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. doi:10.1002/ags3.12026
11 Kakeji Y, Takahashi A, Hasegawa H, et al, National Clinical Database. Surgical outcomes in gastroenterological surgery in Japan: Report of the National Clinical Database 2011-2018. *Ann Gastroenterol Surg* 2020;4:250-74. doi:10.1002/ags3.12324
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* 2020;5:296-303. doi:10.1002/ags3.12419
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. doi:10.1007/s00595-015-1146-y
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. doi:10.1002/ags3.12353
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. doi:10.1097/01.sla.0000133083.54934.ae
16 Bassi C, Marchegiani G, Dervenis C, et al, International Study Group on Pancreatic Surgery (ISGPS). The 2016 update of the International Study Group (ISGPS) definition and grading of postoperative pancreatic fistula: 11 Years After. *Surgery* 2017;161:584-91. doi:10.1016/j.surg.2016.11.014
17 Finch WH, Bolin JE, Kelly K. *Multilevel modeling using R.* 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. doi:10.1007/s10120-018-0883-1
19 Little RJ, Rubin DB. *Statistical Analysis with Missing Data.* 2nd ed. John Wiley & Sons, 2002. doi:10.1002/9781119013563.
20 van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw* 2011;45:1-67. doi:10.18637/jss.v045.i03.
21 Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999;8:3-15. doi:10.1177/096228029900800102
22 Wood SN. *Generalized additive models: an introduction with R.* 2nd ed. CRC press, 2017. doi:10.1201/9781315370279.
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. doi:10.1371/journal.pone.0111071
24 Sakai R, Wang W, Yamaguchi N, Tamura H, Goto R, Kawachi I. The impact of Japan's 2004 postgraduate training program on intra-prefectural distribution of pediatricians in Japan. *PLoS One* 2013;8:e77045. doi:10.1371/journal.pone.0077045
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. doi:10.1016/j.amjsurg.2018.07.018
26 Chen YW, Westfal ML, Chang DC, Kelleher CM. Underemployment of female surgeons?*Ann Surg* 2021;273:197-201. doi:10.1097/SLA.0000000000004497
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. doi:10.1016/j.jamcollsurg.2004.05.281
28 Dossett LA, Vitous CA, Lindquist K, Jagsi R, Telem DA. Women surgeons' experiences of interprofessional workplace conflict. *JAMA Netw Open* 2020;3:e2019843. doi:10.1001/jamanetworkopen.2020.19843
29 Leschber G. From female surgical resident to academic leaders: challenges and pathways forward. *J Thorac Dis* 2021;13:480-4. doi:10.21037/jtd-2020-wts-08
30 Gill HK, Niederer RL, Danesh-Meyer HV. Gender differences in surgical case volume among ophthalmology trainees. *Clin Exp Ophthalmol* 2021;49:664-71. doi:10.1111/ceo.13969
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. doi:10.1097/DCR.0000000000001625
32 Oshima K, Ozaki A, Mori J, Takita M, Tanimoto T. Entrance examination misogyny in Japanese medical schools. *Lancet* 2019;393:1416. doi:10.1016/S0140-6736(18)33180-5
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. doi:10.1001/archsurg.2010.310
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. doi:10.1016/j.jss.2019.10.045

Web appendix: Supplementary materials