RESEARCH ARTICLE

Surgeon Gender and Post-Operative In-Hospital Mortality

Authors
Tara M. Barry, MD
Haroon Janjua, MS
Evelena Cousin-Peterson, MD
Mary Ottinger, MD
Paul C. Kuo, MD, MS, MBA

Affiliations
Department of Surgery University of South Florida, Tampa, FL USA

Corresponding Author:
Tara M. Barry
tbarry@usf.edu
(609) 658-6685 (mobile)
University of South Florida
Department of Surgery
2 Tampa General Circle
Tampa, FL 33606

Presentation
Presented in oral presentation format at the Academic Surgical Congress (Sponsored by the Society of University Surgeons) on 2/6/2020 in Orlando, FL...\AppData\Roaming\WebWork Tracker\images\1633932064.jpeg

Article Summary:
This study demonstrates lower in-hospital mortality rates for patients who underwent select procedures by female surgeons.

Acknowledgements:
The authors would like to thank the Research Methodology and Biostatistics Core at Morsani College of Medicine at the University of South Florida, particularly Jessica A. Garcia for her assistance in reviewing our analysis plan, completing a statistical review, and providing editorial assistance throughout the preparation of this manuscript.

Abstract:

Background: Prior studies report improved mortality rates for patients of female internists, few studies have evaluated the effect of surgeon gender on post-operative mortality.
Methods: A retrospective analysis using the Agency for Health Care Administration Florida database from 2010-2015 examined patients undergoing one of 25 selected surgical procedures. Surgeon gender (self-reported), number of Medicare beneficiaries and years of experience were imported from CMS Physician Compare/Provider Utilization & Payment Data set using NPI. 25 procedures of varying complexity from all subspecialties were selected. For each procedure, inverse probability of treatment weighting (IPTW) was used to match cases performed by male vs female surgeons to achieve maximum balance between the groups.
Results: There were 73,994 admissions for patients undergoing surgical procedures performed by 2,828 surgeons(361 females, 2467 males). Fewer patients who had emergent procedures performed by female surgeons died in the hospital(291/13957,2.08% vs 348/14017,2.48% p=0.026). Those who underwent the following procedures had significantly lower rates of in-hospital mortality if the surgeon was female: CABG(2/387,0.52% vs 8/387,2.07%), mastectomy(2/4797,0.04% vs 10/4797,0.21%), open cholecystectomy(7/955,0.73% vs 17/955, 1.78%) all p≤0.05.
Conclusion: Patients who underwent CABG, mastectomy and open cholecystectomy had lower rates of in-hospital mortality if the surgeon was female while male surgeons did not have a significant mortality advantage for any procedure. Further studies examining national data may provide additional insight regarding the effect of surgeon gender on patient outcomes.

INTRODUCTION

Females make up 52.4% of medical school matriculants but remain underrepresented in the field of surgery.[1] The number of women in surgical leadership positions is particularly lacking, with only 16 female department chairs in 2018 compared with 320 males.[2] The concept of the leadership pipeline does not hold true in the field of surgery. Despite a significant increase in the number of female trainees, the pace of advancement to leadership roles is almost stagnant.[3] This is disadvantageous for multiple reasons. Diversity has been shown to be financially beneficial and improve patient outcomes in the healthcare setting and should be a priority during recruitment.[4,5] Females in the field of medicine prefer gender concordant mentors, and mentorship plays a role in progression through the "pipeline".[6] Lack of mentorship is cited as one of many obstacles aspiring female physicians face when considering a career in surgery or deciding on a subspecialty.[7, 8] Strong mentorship also fosters professional advancement which in turn increases the representation of women in leadership roles.

Prior studies exploring the impact of physician gender on patient outcomes have shown no significant difference between male and female surgeons, but superior outcomes for Medicare patients treated by female hospitalists.[9-11] These studies highlight equivalent and even superior outcomes by female physicians, however comparing hospitalizations of medical patients may not be applicable to surgical patients. The studies specific to the field of surgery include Canadian data from one province with a single payer healthcare system, or multi state data with strict inclusion criteria that compromises generalizability.

The objective of this study is to determine if surgeon gender impacts patient outcomes, specifically post-operative in-hospital mortality. Comparing a greater number of surgeons from various subspecialties and controlling for procedure type will allow for a more accurate analysis of surgeon gender as a factor impacting post-operative mortality.

METHODS

Data Source

We conducted a retrospective observational study using a hybrid data set. We selected data from 2010-2015 to prevent discrepancies caused by conversion of ICD 9 to ICD 10 procedure codes after Q3 in 2015. Physician factors including physician gender, specialty, number of Medicare beneficiaries, and number of services were obtained from the Centers for Medicare and Medicaid Services (CMS) Physician Compare and Fee-For-Service Provider Utilization & Payment Data Physician and Other Supplier Public Use Files, and linked on the basis of National Provider Identification number (NPI) to the 2016 CMS Physician Compare file containing medical school graduation year.[12] Gender is defined as the gender self-reported by the physician to the National Plan & Provider Enumeration System (NPPES) when they enrolled for their NPI number. Specialty is defined as the provider specialty code reported on Medicare claims associated with the largest number of services for that provider. The number of Medicare beneficiaries is not representative of total patient volume for each surgeon because it only includes Medicare patients. The number of services refers only to the number of Medicare claims filed by the surgeon within that year. The aggregate data set containing provider information was linked to the Florida Agency for Health Care Administration (AHCA) yearly Hospital Inpatient Data files for 2010-2015 on the basis of operating physician NPI.[13] A new variable was created for "years of experience" by subtracting the provider's medical school graduation year from the year of procedure within the yearly files. There was no missing data for provider gender, however surgeons with missing graduation year were dropped (2 females, 25 males). Data was limited to adult patients only (all patients <18 years of age were excluded). The files for 2010-2015 were appended to create the master file, which was used for preliminary analysis of mortality by admission priority (elective, urgent, emergent, trauma) and provider gender as well as an overall breakdown of provider type by gender. Data for 25 procedures was identified and isolated using ICD 9 Procedure codes for the primary procedure (see supplemental file for complete list of codes used for each procedure). The procedures were chosen based on the modified Johns Hopkins Surgical Criteria to include operations from all major subspecialties from the moderate to significantly invasive procedure and highly invasive procedure groups (Grade II and III respectively).[14] Both CMS and Florida AHCA Data use agreements dictate no value <5 may be published to prevent potential identification of individuals, therefore all values <5 were represented with an asterix in the tables.

Matching

Inverse probability of treatment weighting (IPTW) propensity score matching was utilized to balance baseline covariates for cases done by male and female surgeons, for each of the 25 procedures.[15, 16] Cases performed by female surgeons were matched to cases performed by male surgeons based on patient characteristics including: age, sex, race, ethnicity, comorbidities (using Charlson Comorbidity Index/CCI), case factors including admission priority (urgent, emergent, elective, trauma), and surgeon factors including years of experience. The variables were selected for matching a priori and were felt to be prognostically important covariates that could impact the primary outcome (mortality) and have been used in other studies of a similar nature.[10, 15] Standardized mean differences were calculated before and after propensity matching to confirm improved balance between the groups and Kolmogorov–Smirnov testing was performed on the matched data, confirming the distribution of patient age and surgeon years of experience were similar for cases done by male and female surgeons. The "Matchit" package in R statistical software was used to perform matching, which improves robustness and decreases dependence of causal inferences on statistical modeling assumptions.[17]

The Stata CCI module was used to calculate the Charlson Index of Comorbidity from data containing ICD-9-CM, diagnoses codes, which groups the comorbidity score into low moderate and severe (CCI-0, CCI-1, CCI-2).[18] The matched data sets for individual procedures were appended to create a master matched data set.

Univariate Analysis

Chi square analysis in Stata was used to determine if post-operative in-hospital mortality was dependent on surgeon gender before and after IPTW matching, for each of the 25 procedures individually. Similar analysis was performed to address surgical subspecialty as a potential confounding factor. T-tests were performed on the aggregate matched data set using R to compare means of length of stay, number of Medicare beneficiaries, surgeon years of experience, patient age and total charges.

All analysis was conducted using Stata 15.1 statistical software (StataCorp, College Station, TX) and RStudio v1.2.1335 statistical software (*RStudio Team (2009-2019). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/*). The University of South Florida Institutional Review Board granted exemption for this research due to the use of publicly available de-identified data sets.

RESULTS

Provider and Patient Characteristics (Table 1)

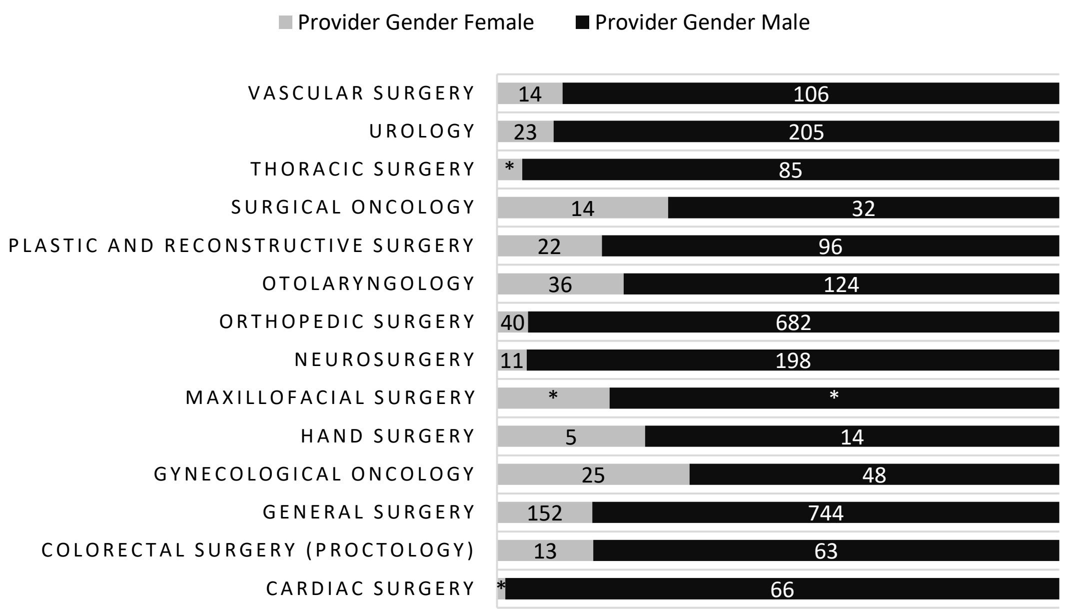
After the data set was limited to the 25 procedures of interest there was a total of 1,011,883 admissions for procedures performed by 4,092 surgeons (361 female, 3,731 male). (Table 1) In the unmatched data set, female surgeons had fewer years of experience compared to their male counterparts, they took care of fewer of Medicare beneficiaries and their patients had longer length of stay (LOS), (all $p<0.00$). After matching, there was no significant difference in length of stay or mean total charges (despite the fact that these variables were not used as covariates in during IPTW matching), however breakdown of payer type remained minimally but significantly different with female surgeons taking care of a higher percentage of Medicaid and Self-Pay patients. While provider and patient characteristics were included in IPTW matching for each procedure individually, the appended matched data set revealed minimal but significant differences between male and female providers. Patient age (58.29 vs 58.63 years), provider years of experience (16.53 vs 16.71 years) and number of Medicare beneficiaries (43.7 vs 39.8) were statistically different (all p<0.01), which can be attributed to the large sample size and is not functionally important.

Table 1. Patient and Provider Demographics for Unmatched and Matched Data

	Unmatched Data (1,011,883 procedures)				**Matched Data** (73,994 procedures)			
Providers	**Male Surgeons** (n = 3,731)	**Female Surgeons** (n = 361)	**95% CI**	**p**	**Male Surgeons** (n = 2,467)	**Female Surgeons** (n = 361)	**95% CI**	**p**
Mean Years experience	23.86	16.53	7.25-7.4	<0.001	16.71	16.53	0.07-0.28	<0.001
Mean # of Medicare Beneficiaries	71.73	43.7	27.21-28.84	<0.001	39.81	43.7	-4.98 - -2.81	<0.001
Patients	n = 974,886	n = 36,997			n = 36,997	n = 36,997		
Age (mean)	63.67	58.29	5.2-5.54	<0.001	58.63	58.29	0.103-0.586	<0.001
CCI (mean)	0.66	0.86	-0.22- -0.18	<0.001	0.84	0.86	-0.04-0.006	0.15
LOS (mean)	4.83	5.18	-0.42- -0.28	<0.001	5.15	5.18	-0.128-0.078	0.63
Mean Total charges	$91,679.19	$82,842.65	8005.4-9667.6	<0.001	$81,891.39	$82,842.65	-2170.9-268.4	0.13
Sex	n (%)	n (%)			n (%)	n (%)		
Male	449,195 (46.08)	11,368 (30.7)		<0.001	11,561 (31.25)	11,368 (30.7)		0.125
Female	525,691 (53.92)	25,629 (69.27)		<0.001	25,436 (68.75)	25,629 (69.27)		0.125
Race				<0.001				0.987
AI/Alaskan	1200 (0.12)	41 (0.11)			46 (0.12)	41 (0.11)		
Asian	5,852 (0.6)	373 (1.01)			375 (1.01)	373 (1.01)		
Black/AA	89,595 (9.19)	4,651 (12.57)			4,689 (12.67)	4,651 (12.57)		
Hawaiian/PI	388 (0.04)	25 (0.07)			25 (0.07)	25 (0.07)		
White	818,891 (84.0)	28,876 (78.05)			28,882 (78.07)	28,876 (78.05)		
Other	49,840 (5.11)	2,650 (7.16)			2,613 (7.06)	2,650 (7.16)		
Unknown	9,120 (0.94)	381 (1.03)			367 (0.99)	381 (1.03)		
Ethnicity				<0.001				0.762
Hisp/Latino	118,731(12.18)	5,549 (15)			5,610 (15.16)	5,549 (15)		
Non-Hispanic	836,379 (85.97)	30,720 (83.03)			30,675 (82.91)	30,720 (83.03)		
Unknown	19,776 (2.03)	728 (1.97)			712 (1.92)	728 (1.97)		
Payer				<0.001				<0.001
Medicare	553,889 (56.8)	15,778 (42.6)			16,245 (43.9)	15,778 (42.6)		
Medicaid	53,351 (5.5)	3,791 (10.2)			3,582 (9.7)	3,791 (10.2)		
Private	285,713 (29.3)	13,032 (35.2)			12,968 (35.1)	13,032 (35.2)		
Self-Pay	30,426 (3.1)	2,254 (6.1)			1,952 (5.3)	2,254 (6.1)		
Non-Payment	11,368 (1.2)	732 (2.0)			901 (2.4)	732 (2.0)		
Other	40,139 (4.1)	1,410 (3.8)			1,348 (3.6)	1,410 (3.8)		

(AI = American Indian, AA = African American, PI = Pacific Islander, Hisp/Latino = Hispanic/Latino)

FIGURE 1. SURGICAL SUBSPECIALTY AND GENDER AFTER MATCHING



*represents cell value <5 as mandated by data use agreement

Comparison of Post-operative Inpatient Mortality by Surgeon Gender

The matched data set contained 73,994 admissions for 25 procedures by 2,828 providers including 361 female surgeons and 2,467 male surgeons representing 14 surgical subspecialties (Figure 1). IPTW matching was performed on cases rather than surgeons to achieve maximal balance between the two groups by including patient, surgeon and admission factors. All variables chosen for matching were considered prognostically important with the potential to impact the outcome. We were able to retain all cases performed by the 361 female surgeons and find the most similar cases performed by male surgeons, however this resulted in a discordance between the number of female and male surgeons represented in the data set. Although female surgeons' overall mortality rates were higher in the raw data set, after matching cases based on patient factors, surgeon years of experience and admission priority there was no gender difference for elective, urgent, and trauma admissions. (Table 2) Female providers had lower mortality rates for emergent procedures ($p<0.02$), as well as for open cholecystectomy, CABG, and mastectomy, specifically (all $p\leq0.05$). Male surgeons did not demonstrate a mortality advantage for any procedures. (Table 3)

Table 2. Mortality by admission priority for Unmatched and Matched Data

	Mortality (%)			
	Male Surgeons	**Female Surgeons**	**p value**	**No. procedures**
Unmatched Data				
Overall	8,156/974,886 **(0.84)**	440/36,997 **(1.19)**	**0.000**	1,011,883
Admission Priority				
Emergent	5,002/245,956 (2.03)	291/13,957 (2.08)	0.677	259,913
Urgent	943/81,311 **(1.16)**	49/2,845 **(1.72)**	**0.006**	84,156
Elective	1,816/640,318 (0.28)	51/19,873 (0.26)	0.481	660,191
Trauma	395/7,301 **(5.41)**	49/322 **(15.22)**	**0.000**	7,623
Matched Data				
Overall	493/36,997 (1.33)	440/36,997 (1.19)	0.081	73,994
Admission Priority				
Emergent	348/14,017 **(2.48)**	291/13,957 **(2.08)**	**0.026**	27,974
Urgent	41/2,859 (1.43)	49/2,845 (1.72)	0.382	5,704
Elective	60/19,829 (0.3)	51/19,873 (0.26)	0.386	39,702
Trauma	44/292 (15.07)	49/322 (15.22)	0.959	614

Table 3. Number of Procedures and Mortality Rates by Surgeon Gender Before and After Matching

	Unmatched				**Matched**			
	n	Female	Male	p value	n	Female	Male	p value
Procedure	1,011,683	% mortality			73,994	% mortality		
laparoscopic cholecystectomy	133,044	0.23	0.29	0.275	19,980	0.23	0.28	0.483
open cholecystectomy	11,363	**0.73**	**1.95**	**0.008**	1,910	**0.73**	**1.78**	**0.04**
exploratory laparotomy	4,075	16.71	17.14	0.835	730	16.71	17.26	0.844
EVAR	14,678	3.23	2.65	0.658	310	3.23	6.45	0.186
CEA	36,484	0	0.25	0.478	402	0	0	NA
laparoscopic colectomy	29,900	**0.3**	**0.69**	**0.024**	4,726	0.3	0.55	0.179
open colectomy	50,039	**6.07**	**4.98**	**0.003**	7,742	6.07	6.28	0.706
CABG	56,773	**0.52**	**1.86**	**0.05**	774	**0.52**	**2.07**	**0.05**
total hip replacement	108,628	0.14	0.14	0.968	2,836	0.14	0.21	0.654
total knee replacement	193,615	0.1	0.06	0.376	5,720	0.1	0.07	0.655
lobectomy/pneumonectomy	14,889	**4.79**	**2.22**	**0.037**	292	4.79	3.42	0.555
prostatectomy	40,625	0	0.17	0.41	778	0	0	NA
thyroidectomy	9,133	0.11	0.27	0.362	1,836	0.11	0.11	1
pancreatectomy	5,430	2.68	3.35	0.532	596	2.68	2.68	1
LE amputation	13,847	2.76	3.85	0.287	724	2.76	5.25	0.088
hysterectomy	18,075	0.32	0.26	0.524	6,830	0.32	0.29	0.827
spinal fusion	143,747	0.46	0.22	0.085	2,164	0.46	0.46	1
nephrectomy	19,633	0.85	0.91	0.852	1,418	0.85	0.28	0.156

mastectomy	14,158	0.04	0.13	0.121	9,594	**0.04**	**0.21**	**0.021**
cystectomy	3,986	0	1.6	0.335	114	0	1.75	0.315
gastric bypass	31,126	0.14	0.11	0.773	1,420	0.14	0.14	1
ORIF femur	50,483	1.86	1.4	0.247	1,828	1.86	2.08	0.736
flap reconstruction	2,252	0	0	NA	570	0	0	NA
liver resection	2,406	2.5	2.67	0.928	160	2.5	2.5	1
splenectomy	3,294	**9.26**	**5.99**	**0.033**	540	9.26	8.52	0.762

Further analysis of the data for open cholecystectomy, CABG and mastectomy was performed, confirming that after matching, female surgeons represented a narrower range of subspecialties than the male surgeons for cases included in this data set. (Table 4) More male providers performed procedures outside the usual scope of their subspecialty as reported in this database. Mortality was not dependent on provider type for any of these three procedures, including subset analysis by provider gender (p = 0.97 open cholecystectomy, p = 0.95 mastectomy, p =0.74 CABG). All cases from the analysis resulting in death were reviewed in the original data set, and operating physician NPI confirmed to ensure there was no association between poor patient outcomes for any single surgeon.

Table 4. Provider Type and Gender for Selected Procedures

	Male Surgeons		Female Surgeons		
Provider Type	providers	procedures	providers	procedures	total
Open Cholecystectomy	n = 378	n = 955	n = 94	n = 955	n = 1,910
Colorectal Surgery	15	20	*	*	23
General Surgery	339	870	89	915	1,785
Surgical Oncology	9	35	*	37	72
Vascular Surgery	11	25	0	0	25
Other (Gyn Onc, PRS, Hand, Thoracic)	*	5	0	0	5
CABG	n = 127	n = 387	n = 5	n = 387	n = 774
Cardiac Surgery	59	180	0	0	180
Thoracic Surgery	63	198	*	354	552
Other (Vascular, Colorectal, General)	5	9	*	33	42
Mastectomy	n = 478	n = 4,797	n = 123	n = 4,797	n = 9594
Colorectal Surgery	7	180	0	0	180
General Surgery	378	3,896	95	3,873	7,769
Plastic & Reconstructive Surgery (PRS)	63	242	15	39	281
Surgical Oncology	18	429	11	883	1,312
Other (Vasc, Uro, Thoracic, Ortho, Gyn Onc)	12	50	*	*	52

*represents cell value <5 as mandated by data use agreements

Discussion

We were able to demonstrate surgeon gender as a significant factor affecting patient outcomes, with lower in-hospital mortality rates among patients treated by female surgeons for open cholecystectomy, mastectomy and CABG. Our study benefited from the inclusion of all adults >18 years of age rather than limiting the population to Medicare patients as is the case in prior studies. IPTW matching was done for each of the 25 procedures individually maximizing similarities between the groups and decreasing confounding caused by variation in procedure as reported by Sharoky et al.[9]

A 2017 JAMA study by Tsugawa et al. confirmed Medicare patients had lower 30-day mortality and readmission rates if they were cared for by female internists. The authors suggest differences in practice patterns may contribute to these improved outcomes, including the fact that female physicians are more likely to practice evidence-based medicine, perform better on standardized tests and prioritize patient centered care.[11] Results from a study of medical patients may not be directly applicable to surgical patients, and the direct impact of provider gender may be more pronounced when comparing patients who underwent a specific procedure rather than comparing hospitalizations. A follow up study in Ontario, Canada examined surgeon gender and post-operative mortality, highlighting a small but significant decreased 30-day mortality rate for patients treated by female surgeons.[10] These results may not be generalizable to the United States given the single payer healthcare system in Canada and heterogeneity of administrative data for individuals in a single province.[19]

Sharoky et al. studied practice patterns and post-operative outcomes for male vs. female surgeons in three states with strict inclusion criteria. The cohort was limited to Medical Doctors age 30-70 trained in the United States practicing General Surgery. They found no significant difference in post-operative complications, length of stay or inpatient mortality.[9] Our study includes all subspecialties and compares a greater number of surgeons than the previous study, which was limited to 152 male/female surgeon pairs at the same hospitals for the outcomes analysis. This allowed the authors to account for hospital level factors, however they compared each male/female surgeon pair without exact matching of procedures. We feel that treating surgeon gender as a factor while comparing patients who underwent the same procedure is a more meaningful assessment than a direct comparison of two individual surgeons. The inability to account for hospital level factors is a significant limitation, as differences in resources and post-operative care may have an impact on mortality. However, subset analysis for the three procedures confirmed that mortality was not dependent on the individual hospital as determined by hospital Medicare number).

The idea of increased risk tolerance among men specifically as it pertains to economics has been widely explored, and may be applicable to medicine and surgery as well[20]. Male surgeons may be more apt to take on “riskier” cases, however this is not discernable from administrative data sets. To address this as a potential confounder, we matched patients by admission priority in addition to comorbidities and procedure type to ensure an equal comparison of emergent, urgent, elective and trauma cases. Our results from the adjusted data revealed that female providers had significantly lower mortality rates for emergent procedures, debunking the myth that mortality rates may be higher for male surgeons because they take on more high-risk cases.

Females are underrepresented in all surgical subspecialties, across all practice settings. The gap is even more apparent in the world of academic surgery, where only 41% of faculty, specifically 24% of full professors and less than 15% of department chairs are women.[21] However, the number of women entering surgical subspecialties is rising, with a 5-10% increase in the number of female residents across all surgical disciplines from 2005-2015, the highest being 38.2% of General Surgery residents and lowest at 14.8% of Orthopedic Surgery residents.[22] This disparity in the surgical subspecialties was prominent in our data set, with 48 female Orthopedic Surgeons compared to 1,184 males, and less than 15 female surgeons in the state with subspecialty reported as either Cardiac, Thoracic, Neurosurgery, Hand Surgery, Maxillofacial Surgery and Surgical Oncology. Male subspecialists were more likely to perform procedures outside the scope of their specialty, despite multiple studies showing this is associated with worse outcomes for specific procedures.[23] (Table 4) This supports the rationale that females are more likely to follow evidence-based practice guidelines, which may contribute to improved outcomes.[11] Historically, General Surgery training was all encompassing with less sub-specialization and the majority of surgeons who trained during the era before fellowships were male. However, this is unlikely the reason for a higher number of male surgeons practicing outside their specialty in this data set because provider years of experience was included in propensity matching to eliminate any potential generational bias.

The primary outcome of post-operative in-hospital mortality cannot be directly attributed to a single individual as patient care involves a large team of people who work together. For the purpose of this study the gender of the operating physician of record was used, as the surgeon is ultimately responsible for the outcome of the operation and the patient's life.[24] It is possible that female leadership in team settings may contribute to superior outcomes. While intraoperative leadership has classically been authoritative and associated with male stereotypes, Minehart et al. explored the idea that inclusive leadership, a style more typical of females, may contribute to improved outcomes.[25] Inclusive leadership allows for improved communication and problem solving within a team and provides a safe environment where team members feel comfortable raising concerns, which is associated with fewer never events and preventable errors. A meta-analysis of emotional intelligence (EI) in medicine using ACGME competencies as a framework cited most studies found females to have superior EI to males from samples of medical students, medical school applicants and human service professionals. A high level of EI was found to correlate with increased team effectiveness and performance.[26]

With regards to selection of variables for matching, Rosenbaum suggests that one ask "which covariates do you wish to balance by matching on the propensity score?"[16] the goal of propensity score analyses should be to induce balance in measured baseline covariates between treatment groups. However, when considering balance, not all covariates are of equal importance. It is more important to balance prognostically important covariates than those covariates that influence treatment selection but have no effect on the outcome. In a similar vein, Myers et al. state that conditioning on instruments (i.e., variables that affect treatment-selection but not the outcome) can result in increased bias and variance of the treatment-effect estimate.[15, 16, 27]

Limitations include the retrospective nature of the study, and use of administrative data sets, from a single state. All data sets used are limited data sets, both AHCA and CMS disclaim responsibility for any analysis, interpretations, or conclusions that may be created because of the limited data set. The data set required us to use in-hospital mortality as a primary endpoint, resulting in significantly fewer deaths included in the data as compared to 30- or 90-day mortality. The variable created for surgeon years of experience does not account for time taken off from a surgical career for pregnancy, illness or other reasons. The volume-outcome relationship for hospitals and individual surgeons have been shown to impact mortality post-operative complications for certain procedures, specifically for vascular, cardiothoracic and trauma surgery. We did not include surgeon specific procedure volumes as a covariate which we acknowledge as a limitation. Not controlling for surgical subspecialty is another significant limitation. The impact of fellowship training on mortality for specific procedures has been well documented, however the data set utilized for this study does not report fellowship training, and provider type is dictated by the provider type submitted by the surgeon to CMS for the majority of their claims.[28] Therefore, we felt that the inclusion of provider type as a factor for matching would not be an accurate representation, and would further limit the number of observations available for comparison since women are disproportionately under-represented in many subspecialties. We were able to confirm that provider type did not affect our results by performing subset analysis of provider type and mortality for the three procedures of interest.

Conclusion

Our findings support the hypothesis that surgeon gender plays a role in patient outcomes for specific procedures. Female surgeons had significantly lower post-operative in-hospital mortality rates than male surgeons for open cholecystectomy, mastectomy, and CABG. These results provide objective evidence of competence that is critical to address the "confidence gap" impeding the development of female surgeon leaders, which is disadvantageous to surgical teams, patients and healthcare systems.[29] Additional studies on national or multi-state data, controlling for hospital factors may provide insight into why female surgeons had better outcomes for these three procedures specifically, and whether these results are applicable to a broader patient population.

COI/Disclosures: All authors have no conflicts of interest to declare and no direct or indirect financial support/funding was used for this project.

References

1. Colleges AoAM. Medical School Applications and Matriculants by School, State of Legal Residence, and Sex 2019-2020. 2020.
2. Colleges AoAM. AAMC Faculty Salary Survey MD or Equivalent Faculty by Gender. 2019.
3. Bernardi K, Shah P, Lyons NB, Olavarria OA, Alawadi ZM, Leal IM, et al. Perceptions on gender disparity in surgery and surgical leadership: A multicenter mixed methods study. Surgery. 2020;167(4):743-50. Epub 2020/01/26. doi: 10.1016/j.surg.2019.12.004. PubMed PMID: 31980138.
4. Gomez LE, Bernet P. Diversity improves performance and outcomes. J Natl Med Assoc. 2019;111(4):383-92. Epub 2019/02/16. doi: 10.1016/j.jnma.2019.01.006. PubMed PMID: 30765101.
5. West MA, Hwang S, Maier RV, Ahuja N, Angelos P, Bass BL, et al. Ensuring Equity, Diversity, and Inclusion in Academic Surgery: An American Surgical Association White Paper. Ann Surg. 2018;268(3):403-7. Epub 2018/07/14. doi: 10.1097/sla.0000000000002937. PubMed PMID: 30004923.
6. Carapinha R, Ortiz-Walters R, McCracken CM, Hill EV, Reede JY. Variability in Women Faculty's Preferences Regarding Mentor Similarity: A Multi-Institution Study in Academic Medicine. Academic medicine : journal of the Association of American Medical Colleges. 2016;91(8):1108-18. Epub 2016/06/23. doi: 10.1097/acm.0000000000001284. PubMed PMID: 27332871; PubMed Central PMCID: PMCPMC4965308.
7. Marks IH, Diaz A, Keem M, Ladi-Seyedian SS, Philipo GS, Munir H, et al. Barriers to Women Entering Surgical Careers: A Global Study into Medical Student Perceptions. World J Surg. 2020;44(1):37-44. Epub 2019/10/17. doi: 10.1007/s00268-019-05199-1. PubMed PMID: 31616970.
8. Reed CE, Vaporciyan AA, Erikson C, Dill MJ, Carpenter AJ, Guleserian KJ, et al. Factors dominating choice of surgical specialty. Journal of the American College of Surgeons. 2010;210(3):319-24. Epub 2010/03/03. doi: 10.1016/j.jamcollsurg.2009.11.016. PubMed PMID: 20193895.
9. Sharoky CE, Sellers MM, Keele LJ, Wirtalla CJ, Karakousis GC, Morris JB, et al. Does Surgeon Sex Matter?: Practice Patterns and Outcomes of Female and Male Surgeons. Ann Surg. 2018;267(6):1069-76. Epub 2017/07/26. doi: 10.1097/SLA.0000000000002460. PubMed PMID: 28742695.
10. 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.
11. 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 Internal Medicine. 2017;177(2):206-13. doi: 10.1001/jamainternmed.2016.7875.

12. [dataset] Physician Compare Data. Centers for Medicare & Medicaid Services; 2016.
13. [dataset] Hospital Inpatient Data File Agency for Health Care Administration State of Florida; 2010-2015.
14. Donati A, Ruzzi M, Adrario E, Pelaia P, Coluzzi F, Gabbanelli V, et al. A new and feasible model for predicting operative risk. BJA: British Journal of Anaesthesia. 2004;93(3):393-9. doi: 10.1093/bja/aeh210 %J BJA: British Journal of Anaesthesia.
15. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Statistics in medicine. 2015;34(28):3661-79. Epub 2015/08/05. doi: 10.1002/sim.6607. PubMed PMID: 26238958; PubMed Central PMCID: PMCPMC4626409.
16. Rosenbaum PR. Model-Based Direct Adjustment. Journal of the American Statistical Association. 1987;82(398):387-94. doi: 10.2307/2289440.
17. Ho D, Imai K, King G, Stuart EA. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. 2011. 2011;42(8):28 %J Journal of Statistical Software. Epub 2011-06-14. doi: 10.18637/jss.v042.i08.
18. Stagg V. CHARLSON: Stata module to calculate Charlson index of comorbidity. Statistical Software Components S456719, Boston College Department of Economics; revised 13 Sep 2017.
19. Doyle CM, Lix LM, Hemmelgarn BR, Paterson JM, Renoux C. Data variability across Canadian administrative health databases: Differences in content, coding, and completeness. 2020;29(S1):68-77. doi: 10.1002/pds.4889.
20. Nelson JA. Gender and Risk-Taking: Economics, Evidence, and Why the Answer Matters: Taylor & Francis; 2017.
21. Colleges. AoAM. Table 13: US medical school faculty by sex, rank, and department. 2018.
22. Colleges AAoM. Distribution of Residents by Specialty, 2005 Compared to 2015. 2016.
23. Dorrance HR, Docherty GM, O'Dwyer PJ. Effect of surgeon specialty interest on patient outcome after potentially curative colorectal cancer surgery. Diseases of the colon and rectum. 2000;43(4):492-8. Epub 2000/05/02. doi: 10.1007/bf02237192. PubMed PMID: 10789744.
24. Henrickson Parker S, Flin R, McKinley A, Yule S. The Surgeons' Leadership Inventory (SLI): a taxonomy and rating system for surgeons' intraoperative leadership skills. Am J Surg. 2013;205(6):745-51. Epub 2012/08/08. doi: 10.1016/j.amjsurg.2012.02.020. PubMed PMID: 22867725.
25. Minehart RD, Foldy EG, Long JA, Weller JM. Challenging gender stereotypes and advancing inclusive leadership in the operating theatre. Br J Anaesth. 2020;124(3):e148-e54. Epub 2020/02/01. doi: 10.1016/j.bja.2019.12.015. PubMed PMID: 32000976.
26. Arora S, Ashrafian H, Davis R, Athanasiou T, Darzi A, Sevdalis N. Emotional intelligence in medicine: a systematic review through the context of the ACGME competencies. Med Educ. 2010;44(8):749-64. Epub 2010/07/17. doi: 10.1111/j.1365-

2923.2010.03709.x. PubMed PMID: 20633215.
27. Myers JA, Rassen JA, Gagne JJ, Huybrechts KF, Schneeweiss S, Rothman KJ, et al. Effects of adjusting for instrumental variables on bias and precision of effect estimates. Am J Epidemiol. 2011;174(11):1213-22. Epub 2011/10/26. doi: 10.1093/aje/kwr364. PubMed PMID: 22025356; PubMed Central PMCID: PMCPMC3254160.
28. McKay A, You I, Bigam D, Lafreniere R, Sutherland F, Ghali W, et al. Impact of surgeon training on outcomes after resective hepatic surgery. Ann Surg Oncol. 2008;15(5):1348-55. Epub 2008/03/01. doi: 10.1245/s10434-008-9838-9. PubMed PMID: 18306973.
29. Stephens EH, Heisler CA, Temkin SM, Miller P. The Current Status of Women in Surgery: How to Affect the Future. JAMA Surgery. 2020. doi: 10.1001/jamasurg.2020.0312 %J JAMA Surgery.

