Original Investigation

Effect of Pregnancy on Adverse Outcomes After General Surgery

Hunter B. Moore, MD; Elizabeth Juarez-Colunga, PhD; Michael Bronsert, PhD, MS; Karl E. Hammermeister, MD; William G. Henderson, MPH, PhD; Ernest E. Moore, MD; Robert A. Meguid, MD, MPH

IMPORTANCE The literature regarding the occurrence of adverse outcomes following nonobstetric surgery in pregnant compared with nonpregnant women has conflicting findings. Those differing conclusions may be the result of inadequate adjustment for differences between pregnant and nonpregnant women. It remains unclear whether pregnancy is a risk factor for postoperative morbidity and mortality of the woman after general surgery.

OBJECTIVE To compare the risk of postoperative complications in pregnant vs nonpregnant women undergoing similar general surgical procedures.

DESIGN, SETTING, AND PARTICIPANTS In this retrospective cohort study, data were obtained from the American College of Surgeons' National Surgical Quality Improvement Program participant user file from January 1, 2006, to December 31, 2011. Propensity-matched females based on 63 preoperative characteristics were matched 1:1 with nonpregnant women undergoing the same operations by general surgeons. Operations performed between January 1, 2006, and December 31, 2011, were analyzed for postoperative adverse events occurring within 30 days of surgery.

MAIN OUTCOMES AND MEASURES Rates of 30-day postoperative mortality, overall morbidity, and 21 individual postoperative complications were compared.

RESULTS The unmatched cohorts included 2764 pregnant women (50.5% underwent emergency surgery) and 516 705 nonpregnant women (13.2% underwent emergency surgery) undergoing general surgery. After propensity matching, there were no meaningful differences in all 63 preoperative characteristics between 2539 pregnant and 2539 nonpregnant patients (all standardized differences, <0.1). The 30-day mortality rates were similar (0.4% in pregnant women vs 0.3% in nonpregnant women; *P* = .82), and the rate of overall morbidity was also not significantly different between pregnant vs nonpregnant patients (6.6% vs 7.4%; *P* = .30).

CONCLUSIONS AND RELEVANCE There was no significant difference in overall morbidity or 30-day mortality rates in pregnant and nonpregnant propensity-matched women undergoing similar general surgical operations. General surgery appears to be as safe for pregnant women as it is for nonpregnant women.

JAMA Surg. doi:10.1001/jamasurg.2015.91
Published online May 13, 2015.

Supplemental content at jamasurgery.com

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Robert A. Meguid, MD, MPH, Surgical Outcomes and Applied Research Program, Department of Surgery, School of Medicine, University of Colorado, Mail Stop C310, 12631 E 17th Ave, Aurora, CO 80045 (robert.meguid@ucdenver.edu).

Historical data[1] suggest that 1 in 500 pregnant patients require nonobstetric surgery. Pregnancy is associated with physiologic changes in body habitus and the coagulation,[2] cardiovascular,[3] pulmonary,[4] and immune[5] systems. These changes pose a diagnostic and treatment challenge to surgeons because physical examination findings and laboratory test values are different from those routinely encountered.[6,7] Therefore, it might be expected that postoperative complications in pregnant patients are increased compared with those in nonpregnant patients. Several retrospective studies[8,9] have supported this hypothesis. However, equivalent complication rates between pregnant and nonpregnant patients have also been reported.[10,11] Review of these disparate studies suggests that the heterogeneity of outcomes is likely the result of differences in the types of operations studied and the inability to account for differences in patient characteristics of pregnant and nonpregnant women in the statistical analyses.

Although several investigators[9-11] report that they have adjusted for differences in patient baseline characteristics, none of these studies has attempted to do this using propensity matching. Because pregnancy usually occurs in younger women as well as considering the broad alterations across the body's physiology as a result of pregnancy, we anticipated differences in preoperative characteristics between nonpregnant and pregnant patients. Therefore, we used propensity matching, a technique recommended[12] for comparison of groups of interest with low event rates.

Using the American College of Surgeons' National Surgical Quality Improvement Program (ACS NSQIP) participant use file (PUF), we evaluated adverse operative outcome rates in general surgery contrasting pregnant with propensity-matched nonpregnant women. We hypothesized that pregnant women are at greater risk of complications than are comparable nonpregnant women undergoing similar general surgical operations.

Methods

Study Design and Population

This retrospective cohort study compared 30-day postoperative surgical outcomes of pregnant vs nonpregnant women undergoing nonobstetric operations by general surgeons. Patients were identified from the ACS NSQIP PUF, a database of surgical procedures performed in hospitals participating in the ACS NSQIP from January 1, 2006, to December 31, 2011. Patients were excluded if they were male, underwent obstetric surgery, or were missing one or more of the preoperative patient characteristics used in the study. Nonpregnant women were also excluded if they underwent operations that were not performed in the group of pregnant patients. The Colorado Multiple Institutional Review Board classified this study as not involving human subject research.

Primary Outcome

The primary outcome variables for the analyses in this study were death from any cause and overall morbidity (≥1 of the 21 ACS NSQIP perioperative complications) occurring within 30 days of the index operation. These complications included acute renal failure requiring dialysis or hemofiltration; progressive renal insufficiency; bleeding requiring transfusion of more than 4 U of packed red blood cells; cardiac arrest requiring cardiopulmonary resuscitation; Q-wave myocardial infarction; deep venous thrombosis or thrombophlebitis requiring treatment; pulmonary embolism; pneumonia; prolonged (>48 hours) intubation; unplanned intubation; septic shock; cerebrovascular accident (including trauma such as a fall, resulting in an injury to the head) or stroke with subsequent neurologic deficit; sepsis; superficial surgical site infection; deep incisional surgical site infection; organ or organ space surgical site infection; urinary tract infection; wound disruption; peripheral nerve injury; graft, prosthesis, or flap failure; and coma lasting longer than 24 hours.[13]

Statistical Analysis

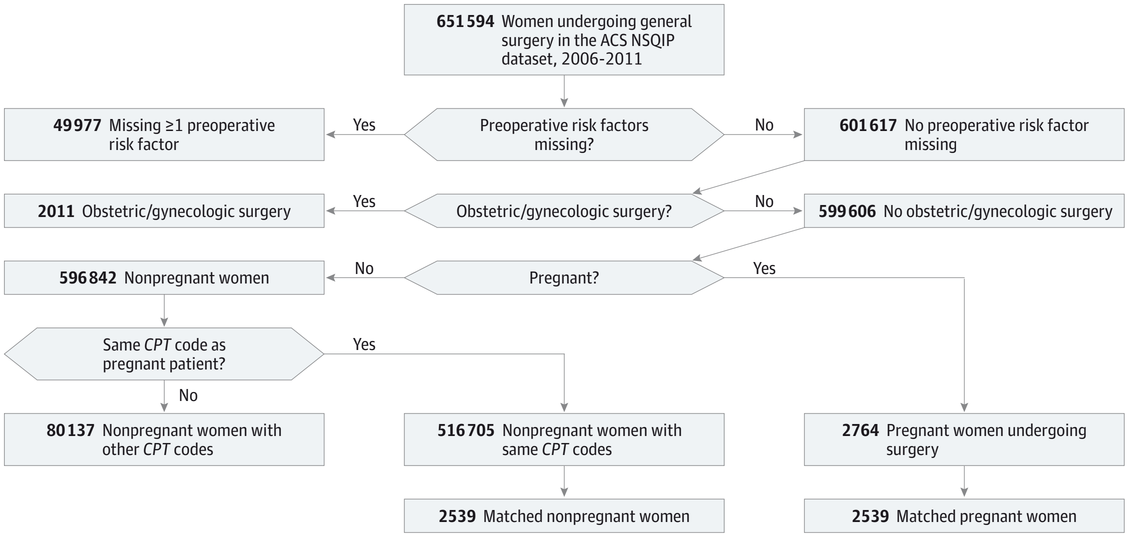
Differences between pregnant and nonpregnant patients were compared with χ^2 tests for categorical variables and 2-tailed independent t tests for continuous variables in the unmatched cohorts. Differences between the groups were also evaluated using the standardized differences[14] to enable comparison of covariate imbalance between the matched and unmatched cohorts. The absolute value of the standard difference of less than 0.1 indicates that the groups are well balanced for that characteristic; differences greater than 0.1 or less than −0.1 indicate some imbalance. The McNemar test, either in its large sample size approximation or exact form, was used to compare 30-day mortality and morbidity in the propensity-matched cohorts.

Propensity-score matching methods were used to reduce confounding related to nonrandom assignment of pregnancy.[15] A propensity score is the predicted probability, based on logistic regression, that a given woman will be pregnant. This approach was used because of its performance and simplicity. Pregnant patients were propensity matched 1:1 to controls with a greedy algorithm.[16] The propensity score logit model included 63 patient preoperative characteristics.

Despite the large sample of nonpregnant women, we conducted one-to-one matching to avoid the possible bias of many-to-one matching.[17] Each pregnant patient was matched to a single nonpregnant control patient if her predicted propensity scores were identical to 8 decimal places. If such a match was not found, the pregnant patient was matched to a nonpregnant patient on the basis of a 7-, 6-, 5-, 4-, 3-, 2-, or 1-decimal place match, tested sequentially. Missing values were treated as a separate category for the categorical variables of race/ethnicity, body mass index, and the 12 preoperative laboratory test values. Laboratory test values were coded as missing, abnormal low, normal, and abnormal high according to values presented in a widely used medical textbook.[18]

Differences in complications and mortality rates were also analyzed in subgroups of emergency and nonemergency operations. To retain high power, we included an interaction term for pregnancy and emergency in a conditional logistic regression model. Evidence of an interaction would indicate that the association between pregnancy and complication rates was dif-

Figure 1. Strengthening the Reporting of Observational Studies in Epidemiology Flow Diagram of Pregnant vs Nonpregnant Women Undergoing the Same General Surgical Operations



Data were obtained from the American College of Surgeons' National Surgical Quality Improvement Program (ACS NSQIP) (2006-2011). *CPT* indicates *Current Procedural Terminology*.

ferent in emergency vs nonemergency cases. The a priori level of statistical significance was set at α = .05 for all analyses, which were 2-tailed. Statistical analyses were performed with SAS, version 9.4 (SAS Institute Inc).

Results

There were 651 594 adult women undergoing operations by general surgeons in the ACS NSQIP PUF from 2006 to 2011. Exclusion criteria and sample size of patients in this analysis are demonstrated in the Strengthening the Reporting of Observational Studies in Epidemiology diagram (**Figure 1**). Of the 651 594 patients, 49 977 (7.7%) were excluded because they were missing 1 or more preoperative patient characteristic, and 2011 patients (0.3%) were excluded because they underwent an obstetric operation. An additional 80 137 nonpregnant patients (12.3%) were excluded because they underwent an operation that was not performed in the group of pregnant patients. A total of 519 469 patients remained: 2764 (0.5%) were pregnant and 516 705 (99.5%) were nonpregnant.

Characteristics of unmatched pregnant patients and nonpregnant patients are presented in eTable 1 in the Supplement. The unmatched pregnant patients were significantly younger than the nonpregnant patients (29.8 vs 53.0 years; $P < .001$). Compared with the nonpregnant patients, the pregnant women were more likely to undergo surgery as an inpatient (2074 [75.0%] vs 308 375 [59.7%]; $P < .001$) and undergo an emergency operation (1396 [50.5%] vs 68 156 [13.2%]; $P < .001$). Pregnant patients generally had lower rates of preoperative comorbidities but higher rates of abnormal laboratory test results (high white blood cell count and low blood urea nitrogen, hematocrit, serum creatinine, serum sodium, and serum albumin levels) compared with nonpregnant patients.

The standardized differences for the proportions or means between the pregnant and nonpregnant unmatched cohorts are reported in eTable 1 in the Supplement. The standardized differences were greater than 0.1 or less than −0.1 for 31 of the 63 preoperative patient characteristics (49.2%), indicating the expected, important imbalances between pregnant and nonpregnant patients in the unmatched cohorts.

In the unmatched cohort, 10 of 2764 pregnant patients (0.4%) died within 30 days of surgery compared with 5759 of 516 705 nonpregnant patients (1.1%) ($P < .001$) (**Table**). The overall morbidity rate was also lower for pregnant patients (183/2764 [6.6%] vs 48 394/516 705 [9.4%]; $P < .001$) than nonpregnant patients. Pregnant patients had significantly lower rates of superficial surgical site infection, urinary tract infection, bleeding requiring transfusion of more than 4 U of packed red blood cells, myocardial infarction, and unplanned intubation compared with nonpregnant patients in the unmatched cohort (P value range, .005-.049).

The propensity model is provided in eTable 2 in the Supplement. Thirty-seven of the preoperative patient characteristics were significant predictors of pregnancy, with the C statistic for the full model of 0.939. A total of 2539 of the 2764 pregnant patients (91.9%) were matched to 2539 of 516 705 (0.5%) nonpregnant patients. The standardized differences in baseline characteristics between the groups before and after matching on the propensity score are shown in **Figure 2**. In the propensity-matched cohort, none of the 63 patient characteristics had standardized differences greater than 0.1 or less than −0.1, indicating that the propensity-matched samples were well balanced.

As reported in the Table for the propensity-matched cohort, there was no significant difference in the 30-day mortality rates between pregnant and nonpregnant patients (0.4% vs 0.3%; $P = .82$) or in the overall morbidity rate in the preg-

Table. Bivariable Association of Complications With Pregnancy in the Unmatched and Matched Cohorts

	No. (%)					
	Unmatched			Matched		
Complication	Nonpregnant	Pregnant	*P* Value	Nonpregnant	Pregnant	*P* Value
No. (%)	516 705 (99.5)	2764 (0.5)		2539 (4.9)	2539 (91.9)	
Infection						
Superficial surgical site infection	12 975 (2.5)	46 (1.7)	.005	44 (1.7)	36 (1.4)	.43
Sepsis	7854 (1.5)	36 (1.3)	.35	38 (1.5)	35 (1.4)	.82
Urinary tract infection	8169 (1.6)	28 (1.0)	.02	29 (1.1)	27 (1.1)	.89
Organ/organ space surgical site infection	6072 (1.2)	30 (1.1)	.66	42 (1.7)	29 (1.1)	.15
Deep incisional surgical site infection	3246 (0.6)	14 (0.5)	.42	13 (0.5)	14 (0.6)	>.99
Wound disruption	2133 (0.4)	11 (0.4)	.90	6 (0.2)	9 (0.4)	.61
Cardiac						
Bleeding requiring transfusion of >4 U of PRBCs	7345 (1.4)	27 (1.0)	.049	19 (0.8)	26 (1.0)	.37
Cardiac arrest requiring CPR	1291 (0.2)	2 (0.1)	.06	1 (0.04)	2 (0.1)	>.99
Q-wave MI	901 (0.2)	0	.03	1 (0.04)	0	NA
Respiratory						
Prolonged intubation (>48 h)	8002 (1.5)	35 (1.3)	.23	28 (1.1)	35 (1.4)	.44
Pneumonia	5340 (1.0)	22 (0.8)	.22	14 (0.6)	21 (0.8)	.30
Unplanned intubation	4886 (0.9)	13 (0.5)	.01	11 (0.4)	13 (0.5)	.84
Septic shock	4276 (0.8)	14 (0.5)	.06	17 (0.7)	14 (0.6)	.72
Venous thromboembolism						
Deep venous thrombosis/thrombophlebitis	2617 (0.5)	10 (0.4)	.29	6 (0.2)	10 (0.4)	.45
Pulmonary embolism	1366 (0.3)	7 (0.3)	.91	1 (0.04)	7 (0.3)	.07
Renal						
Acute renal failure requiring dialysis or hemofiltration	1353 (0.3)	4 (0.1)	.23	3 (0.1)	4 (0.2)	>.99
Progressive renal insufficiency[a]	1016 (0.2)	3 (0.1)	.30	2 (0.1)	3 (0.1)	>.99
Stroke						
Stroke/cerebrovascular accident with neurologic deficit	621 (0.1)	0	.07	3 (0.1)	0	NA
Other						NA
Peripheral nerve injury	162 (0.03)	0	.35	1 (0.04)	0	NA
Graft/prosthesis/flap failure	337 (0.1)	0	.18	1 (0.04)	0	NA
Coma for >24 h	255 (0.05)	0	.24	0	0	
30-d Mortality	5759 (1.1)	10 (0.4)	<.001	8 (0.3)	10 (0.4)	.82
≥1 Complication	48 394 (9.4)	183 (6.6)	<.001	188 (7.4)	168 (6.6)	.30

Abbreviations: CPR, cardiopulmonary resuscitation; MI, myocardial infarction; NA, not available; PRBCs, packed red blood cells.

[a] Progression was indicated by an increase in creatinine of more than 2 mg/dL from the preoperative value. To convert creatinine to micromoles per liter, multiply by 88.4.

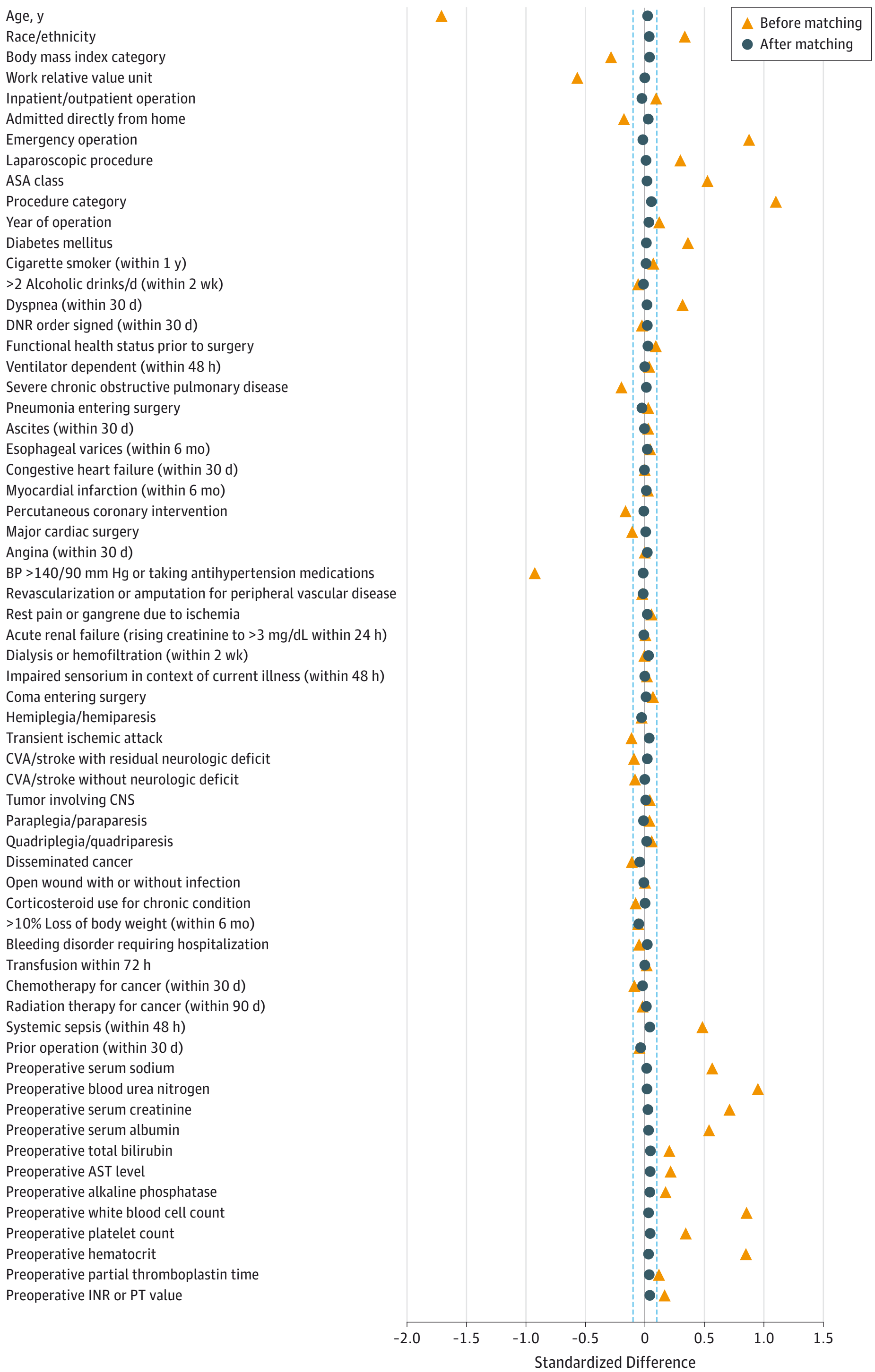
nant patients vs nonpregnant women (6.6% vs 7.4%; *P* = .30). No significant differences were found when we compared the rates of the 21 individual complications in the pregnant vs nonpregnant patients after propensity matching. There was no evidence of a different association between pregnancy and overall morbidity or mortality rates in the emergency and nonemergency subgroups (interaction *P* values: overall morbidity, *P* = .11; mortality, *P* = .74).

Discussion

We performed an analysis of pregnant women matched to nonpregnant women undergoing general surgical operations using the ACS NSQIP PUF to determine whether pregnancy was associated with an increased rate of postoperative adverse outcomes. We observed that pregnant patients had different preoperative risk factors than nonpregnant women: the pregnant women were younger, had fewer comorbidities, and more frequently had abnormal laboratory test values. Pregnant women were also more likely to undergo emergency procedures. Unbalanced preoperative risk factors between the groups were balanced after propensity matching, thereby minimizing bias in comparison of outcomes between the 2 patient populations. Analysis of matched cohorts showed no significant differences in 30-day mortality or the occurrence of 1 or more complications between the groups. Nonobstetric general surgery appears to be as safe in pregnant women as in nonpregnant women.

Prior studies reporting increased complication rates in pregnant vs nonpregnant women undergoing nonobstetric surgery come from analysis of the Health Care Utilization Proj-

Figure 2. Standardized Differences in Population Baseline Characteristics in Pregnant vs Nonpregnant Women Undergoing the Same General Surgical Operations Before and After Matching



To convert creatinine to micromoles per liter, multiply by 88.4. Solid vertical line indicates 0, dashed lines, 0.1 and −0.1. ASA indicates American Association of Anesthesiologists; AST, aspartate aminotransferase; BP, blood pressure; CNS, central nervous system; CVA, cerebrovascular accident; DNR, do not resuscitate; INR, international normalized ratio; and PT, prothrombin time.

ect Nationwide Inpatient Sample. Kuy et al[9] reported increased rates of complications, length of stay, and cost for pregnant women undergoing thyroid and parathyroid surgery despite risk adjustment with logistic regression analysis for the dichotomous outcome (complications) and linear regression for the continuous variables (length of stay and cost).

Using a similar time period, the same group[11] evaluated pregnant women undergoing cholecystectomy. Prior to regression analysis, the pregnant patients had an increased complication rate. However, after age and procedure matching, as well as adjustment for insurance, race, and surgeon case volume, pregnancy was not associated with an increased risk of surgical complications. A recent publication[8] using the Health Care Utilization Project Nationwide Inpatient Sample reported that postoperative complication rates following appendectomy were higher in pregnant vs nonpregnant women. In that study, Abbasi et al[8] matched more than 7000 pregnant women to 35 000 nonpregnant women based on age and then performed multivariable logistic regression on categories of race, obesity, income, and insurance type. Although postoperative complication rates were higher in the pregnant group, the most notable finding was that peritonitis on presentation was the highest predictor of postoperative complication rates. This study identified that pregnant women more frequently present with peritonitis than do nonpregnant women. The authors concluded that this factor was the causality for this discrepancy between the groups.

In contrast, some studies report low postoperative complication rates in pregnant patients. Erekson et al[19] analyzed the ACS NSQIP PUF. Their descriptive findings parallel our results of low maternal postoperative complication rates, but they did not contrast pregnant and nonpregnant women. Silvestri et al[10] found a similar rate of morbidity between pregnant and nonpregnant women undergoing cholecystectomy and appendectomy. McMaster et al[20] also found that pregnant patients had postoperative complication rates similar to those of nonpregnant women after breast surgery.

Because a prospective randomized clinical trial to identify whether pregnancy is a risk factor for postoperative complications is not feasible, only observational studies are available. The latter are dependent upon statistical adjustment to account for significant baseline differences between pregnant and nonpregnant patients. The ACS NQIP PUF during our study time frame contained data on more than 500 000 nonpregnant women undergoing operations similar to those of pregnant patients. Propensity matching is well suited for this type of observational study in which a "large reservoir" of potential controls is contrasted to a moderately sized group.[21] Propensity matching controls for measured baseline covariates before analysis of the outcomes.[22] This technique does not require the complexity of forming multiple strata to balance covariates and is superior in reducing bias.[12] Propensity matching is used frequently in medical studies because of its simplicity and robust performance, but it is not always reported appropriately.[15] Key elements of propensity modeling that are often neglected include reporting the model construct, assessment of prematching and postmatching differences between groups, and appropriate outcome analysis.

In our study, the maximum C statistic for the propensity model was 0.939 (eTable 2 in the Supplement) compared with the value of 1.0 for a perfect model. This finding supports the conclusion that we have developed a reasonable model to predict pregnancy based on preoperative characteristics. The model eliminated measured, unbalanced preoperative variables quantified by standardized differences (Figure 2), and outcomes were appropriately assessed with a paired analysis contrasting pregnant to nonpregnant women. Other approaches, such as double robust inverse probability weighting, requiring specification of an outcomes regression model would not have been a feasible approach with these data given the small number of outcome events.[12]

This study has strengths and limitations. Strengths include (1) a large sample from a broad range of hospitals, (2) a broad range of operations included in the database using a systematic sampling method, and (3) a standardized protocol for collection of the ACS NSQIP PUF, with central auditing of the data. The primary limitations of this study include (1) the observational design so that only association (ie, not causation) may be concluded and (2) a lack of data on fetal outcomes. There is clearly a risk to the fetus when a pregnant woman undergoes surgery. Fetal loss after appendectomy was found to be 4% in women with a normal appendix.[23] The increasing number of reports indicates that infectious surgical indications, such as appendicitis and cholecystitis, are associated with an unfavorable outcome for the fetus[24,25] and that advanced disease is a risk factor for fetal and maternal complications.[23,26] Attempting medical management of surgical diseases (eg, appendicitis and cholecystitis) is associated with a worse outcome compared with early operative management.[8,11] Therefore, the well-being of the fetus represents an additional risk-benefit factor to consider in pregnant women, and an unclear diagnosis may require further expeditious evaluation to minimize delay of definitive management.

Conclusions

Pregnant patients undergoing emergency and nonemergency general surgery do not appear to have elevated rates of mortality or morbidity. We did not account for fetal complications in this study and would not advocate that our findings be generalized to elective surgical situations that can be postponed until after delivery. Therefore, general surgery appears to be as safe in pregnant as it is in nonpregnant women. These findings support previous reports that pregnant patients who present with acute surgical diseases should undergo the procedure if delay in definitive care will lead to progression of disease.

ARTICLE INFORMATION

Accepted for Publication: December 1, 2014.

Published Online: May 13, 2015. doi:10.1001/jamasurg.2015.91.

Author Affiliations: Department of Surgery, School of Medicine, University of Colorado, Aurora (H. B. Moore, E. E. Moore, Meguid); Surgical Outcomes and Applied Research Program, School of Medicine, University of Colorado, Aurora (Juarez-Colunga, Bronsert, Hammermeister, Henderson, Meguid); Adult and Child Center for Health Outcomes Research and Delivery Science, School of Medicine, University of Colorado, Aurora (Juarez-Colunga, Bronsert, Hammermeister, Henderson); Department of Biostatistics and Informatics, Colorado School of Public Health, Aurora (Juarez-Colunga, Henderson); Division of Cardiology, Department of Medicine, School of Medicine, University of Colorado, Aurora (Hammermeister); Department of Surgery, Denver Health Medical Center, Denver, Colorado (E. E. Moore).

Author Contributions: Drs H. B. Moore and Juarez-Colunga had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: H. B. Moore, Hammermeister, Henderson, Meguid.
Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: H. B. Moore, Juarez-Colunga, Hammermeister, E. E. Moore, Meguid.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Juarez-Colunga, Bronsert, Henderson, Meguid.
Obtained funding: Meguid.
Administrative, technical, or material support: E. E. Moore, Meguid.
Study supervision: Hammermeister, Henderson, E. E. Moore, Meguid.

Conflict of Interest Disclosures: None reported.

Funding/Support: This project was supported by funding from the Department of Surgery, Adult and Child Center for Health Outcomes Research and Delivery Science Joint Surgical Outcomes and Applied Research Program at the University of Colorado.

Role of the Funder/Sponsor: The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES

1. Kammerer WS. Nonobstetric surgery during pregnancy. *Med Clin North Am*. 1979;63(6):1157-1164.

2. Della Rocca G, Dogareschi T, Cecconet T, et al. Coagulation assessment in normal pregnancy: thrombelastography with citrated non activated samples. *Minerva Anestesiol*. 2012;78(12):1357-1364.

3. Ouzounian JG, Elkayam U. Physiologic changes during normal pregnancy and delivery. *Cardiol Clin*. 2012;30(3):317-329.

4. Orzechowski KM, Miller RC. Common respiratory issues in ambulatory obstetrics. *Clin Obstet Gynecol*. 2012;55(3):798-809.

5. Munoz-Suano A, Hamilton AB, Betz AG. Gimme shelter: the immune system during pregnancy. *Immunol Rev*. 2011;241(1):20-38.

6. Coleman MT, Trianfo VA, Rund DA. Nonobstetric emergencies in pregnancy: trauma and surgical conditions. *Am J Obstet Gynecol*. 1997;177(3):497-502.

7. Augustin G, Majerovic M. Non-obstetrical acute abdomen during pregnancy. *Eur J Obstet Gynecol Reprod Biol*. 2007;131(1):4-12.

8. Abbasi N, Patenaude V, Abenhaim HA. Management and outcomes of acute appendicitis in pregnancy–population-based study of over 7000 cases. *BJOG*. 2014;121(12):1509-1514.

9. Kuy S, Roman SA, Desai R, Sosa JA. Outcomes following thyroid and parathyroid surgery in pregnant women. *Arch Surg*. 2009;144(5):399-406.

10. Silvestri MT, Pettker CM, Brousseau EC, Dick MA, Ciarleglio MM, Erekson EA. Morbidity of appendectomy and cholecystectomy in pregnant and nonpregnant women. *Obstet Gynecol*. 2011;118(6):1261-1270.

11. Kuy S, Roman SA, Desai R, Sosa JA. Outcomes following cholecystectomy in pregnant and nonpregnant women. *Surgery*. 2009;146(2):358-366.

12. Cepeda MS, Boston R, Farrar JT, Strom BL. Optimal matching with a variable number of controls vs a fixed number of controls for a cohort study: trade-offs. *J Clin Epidemiol*. 2003;56(3):230-237.

13. Cohen ME, Ko CY, Bilimoria KY, et al. Optimizing ACS NSQIP modeling for evaluation of surgical quality and risk: patient risk adjustment, procedure mix adjustment, shrinkage adjustment, and surgical focus. *J Am Coll Surg*. 2013;217(2):336-346.e1.

14. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*. 2009;28(25):3083-3107.

15. Austin PC. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Stat Med*. 2008;27(12):2037-2049.

16. Austin PC. A comparison of 12 algorithms for matching on the propensity score. *Stat Med*. 2014;33(6):1057-1069.

17. Austin PC. Statistical criteria for selecting the optimal number of untreated subjects matched to each treated subject when using many-to-one matching on the propensity score. *Am J Epidemiol*. 2010;172(9):1092-1097.

18. Kratz A, Pesce MA, Basner RC. Laboratory values of clinical importance. In: Longo DL, Fauci AS, Kasper DL, Hauser SL, Jameson JL, Loscalzo J, eds. *Harrison's Principles of Internal Medicine.* 18th ed. New York, NY: McGraw Hill; 2012:appendix.

19. Erekson EA, Brousseau EC, Dick-Biascoechea MA, Ciarleglio MM, Lockwood CJ, Pettker CM. Maternal postoperative complications after nonobstetric antenatal surgery. *J Matern Fetal Neonatal Med*. 2012;25(12):2639-2644.

20. McMaster J, Dua A, Desai SS, Kuy S, Kuy S. Short term outcomes following breast cancer surgery in pregnant women. *Gynecol Oncol*. 2014;135(3):539-541.

21. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41-44.

22. Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med*. 2007;26(1):20-36.

23. McGory ML, Zingmond DS, Tillou A, Hiatt JR, Ko CY, Cryer HM. Negative appendectomy in pregnant women is associated with a substantial risk of fetal loss. *J Am Coll Surg*. 2007;205(4):534-540.

24. Wei PL, Keller JJ, Liang HH, Lin HC. Acute appendicitis and adverse pregnancy outcomes: a nationwide population-based study. *J Gastrointest Surg*. 2012;16(6):1204-1211.

25. Swisher SG, Schmit PJ, Hunt KK, et al. Biliary disease during pregnancy. *Am J Surg*. 1994;168(6):576-579.

26. Mourad J, Elliott JP, Erickson L, Lisboa L. Appendicitis in pregnancy: new information that contradicts long-held clinical beliefs. *Am J Obstet Gynecol*. 2000;182(5):1027-1029.