Journal of the American Heart Association

ORIGINAL RESEARCH

Sex Differences in Postoperative Recovery and Mortality After High-Risk Cardiac Surgery: A Propensity Score–Matched Post Hoc Analysis of the SUSTAIN-CSX Trial

Rashad Zayat, MD, PhD*; Alish Kolashov, MD*; Christian Stoppe, MD, PhD; Quirin Notz, MD, PhD; Yusuf Shieba, MD; Ajay Moza, MD, PhD; Daniel Catena, MD; Ghaith Mohsen, MD; Jan Beer, MD; Georg Daniel Duerr, MD, PhD; Ahmed F.A. Mohammed, MD

BACKGROUND: Sex-related differences after cardiac surgery remain controversial because women often present with higher baseline risk and complexity than men. We performed a post hoc propensity score–matched analysis of the SUSTAIN-CSX (Sodium Selenite Administration in Cardiac Surgery) trial to evaluate sex differences in mortality, postoperative complications, and recovery after high-risk cardiac surgery.

METHODS: Of 1394 trial participants, 1386 had complete data. Women were matched 1:1 to men using nearest-neighbor propensity score matching based on age and European System for Cardiac Operative Risk Evaluation II (EuroSCORE II), with exact matching on surgical category, yielding 327 female–male pairs. Prespecified sensitivity analyses adjusted for frailty, baseline hemoglobin, renal disease, left ventricular ejection fraction, previous myocardial infarction, preoperative medications, and baseline creatinine.

RESULTS: In the primary matched analysis, 180-day survival did not differ between women and men (log-rank P=0.086; unadjusted hazard ratio, 1.80 [95% CI, 0.91–3.55]; P=0.091). In descriptive matched comparisons, women had numerically longer intensive care unit stay (median, 3 days [quartile 1, quartile 3 (Q1, Q3)=1, 6 days] versus 2 days [Q1, Q3=1, 5 days]) and hospital stay (median, 10 days [Q1, Q3=7, 18 days] versus 9 days [Q1, Q3=6, 16 days]; P=0.292), whereas major postoperative complications were similar. In adjusted sensitivity analyses accounting for the matched design and residual imbalance, female sex remained associated with longer intensive care unit stay (adjusted incidence rate ratio [IRR], 1.8 [95% CI, 1.2–2.9]; P=0.009) and hospital stay (adjusted IRR, 1.4 [95% CI, 1.0–1.9]; P=0.031). Mortality sensitivity analyses were model-dependent.

CONCLUSIONS: In this propensity score–matched cohort of high-risk cardiac surgery patients, women showed a longer postoperative recovery trajectory in adjusted analyses, whereas mortality findings were sensitive to model specification and should be interpreted cautiously.

REGISTRATION: URL: https://www.clinicaltrials.gov; Unique identifier: NCT02002247.

Key Words: cardiac surgery ■ epidemiology ■ high-risk cardiac surgery ■ propensity score matching ■ quality and outcomes ■ SUSTAIN-CSX trial ■ women

Correspondence to: Rashad Zayat, MD, PhD, Department of Cardiac Surgery, RWTH University Hospital Aachen, Pauwelsstraße 30, 52074 Aachen, Germany. Email: rzayat@ukaachen.de

*R. Zayat and A. Kolashov contributed equally.

This article was sent to Monik C. Botero, SM, ScD, Associate Editor, for review by expert referees, editorial decision, and final disposition.

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.126.049243

For Sources of Funding and Disclosures, see page 11.

JAHA is available at: www.ahajournals.org/journal/jaha

CLINICAL PERSPECTIVE

What Is New?

- In a propensity score–matched cohort of high-risk cardiac surgery patients from the SUSTAIN-CSX (Sodium Selenite Administration in Cardiac Surgery) trial, descriptive matched comparisons of intensive care unit and hospital stay were attenuated, but adjusted sensitivity analyses consistently showed longer postoperative recovery in women.
- The association between female sex and prolonged postoperative recovery remained consistent in sensitivity analyses adjusting for frailty, baseline hemoglobin, renal disease, preoperative medications, and baseline creatinine.

What Are the Clinical Implications?

- In women undergoing high-risk cardiac surgery, perioperative counseling and postoperative planning should account for a potentially longer recovery trajectory, whereas short-term mortality differences should be interpreted cautiously because they were sensitive to model specification.

Nonstandard Abbreviations and Acronyms

CFS	Clinical Frailty Scale
EuroSCORE II	European System for Cardiac Operative Risk Evaluation II
IRR	incidence rate ratio
SMD	standardized mean difference
SUSTAIN-CSX	Sodium Selenite Administration in Cardiac Surgery

High-risk and complex cardiac surgery remains associated with substantial morbidity and mortality despite advances in perioperative care.[1] These procedures provoke systemic inflammation and ischemia–reperfusion injury, contributing to organ dysfunction, prolonged intensive care unit (ICU) stay, and delayed recovery.[2,3] The SUSTAIN-CSX (Sodium Selenite Administration in Cardiac Surgery) trial enrolled high-risk patients (European System for Cardiac Operative Risk Evaluation II [EuroSCORE II]–predicted mortality ≥5% or those undergoing planned complex procedures), providing a well-characterized cohort for evaluating postoperative outcomes.[4] Female sex has often been linked to worse outcomes after cardiac surgery, including higher operative mortality, but findings are inconsistent and confounded by baseline differences: women typically present older, with smaller body size and greater comorbidity burden.[5–7] It remains unclear whether sex is an independent risk factor or a surrogate for preoperative risk and procedural complexity.[8]

Using the SUSTAIN-CSX data set, we applied propensity score matching based on age and EuroSCORE II, with exact matching on surgical category, to reduce key baseline and procedural imbalances.[7,9] This post hoc analysis examined whether female sex was associated with postoperative mortality, major morbidity, and postoperative recovery, including ICU and hospital length of stay. We hypothesized that sex-related differences in this high-risk cohort, if present, would be more evident in postoperative recovery than in major postoperative complications or mortality.

METHODS

Data Availability Statement

The data, analytic methods, and study materials underlying this analysis are available from the corresponding author on reasonable request, subject to approval by the SUSTAIN-CSX trial steering committee and institutional regulations.

Study Design

This post hoc, propensity score–matched analysis derives from the multicenter, double-blind, placebo-controlled SUSTAIN-CSX randomized trial. The trial was conducted among 23 centers in Germany and Canada between 2015 and 2021 and investigated whether perioperative high-dose intravenous sodium selenite could reduce postoperative organ dysfunction and mortality in high-risk cardiac surgery patients.[4] Because the parent trial reported neutral primary results, both randomized treatment groups were pooled for the present sex-based analysis. A study flow diagram summarizing cohort selection and matching is shown in Figure 1.

Ethics

The parent trial was approved by the responsible institutional review boards/ethics committees at all participating centers, including Queen's University and Rheinisch-Westfälische Technische Hochschule Aachen University, and by the relevant regulatory authorities. The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines. All participants provided written informed consent. The parent trial was registered at ClinicalTrials.gov (NCT02002247).

Patient Population

Adults aged ≥18 years undergoing elective or urgent cardiac surgery requiring cardiopulmonary bypass were

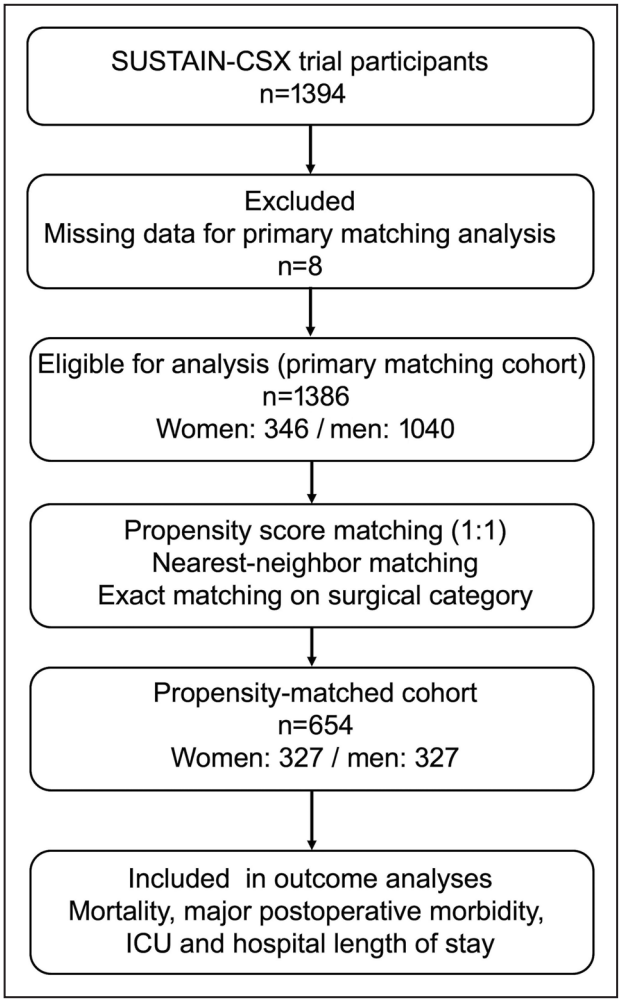

Figure 1. Study flow diagram of cohort selection and propensity score matching.
ICU indicates intensive care unit; and SUSTAIN-CSX, Sodium Selenite Administration in Cardiac Surgery.

eligible. High risk was defined as a EuroSCORE II–predicted mortality ≥5% or a planned complex or combined procedure, including combined coronary artery bypass grafting plus valve or multivalve surgery, coronary artery bypass grafting plus other nonvalvular procedures, multivalve surgery, valve plus other nonvalvular procedures, or thoracic aortic surgery. Key exclusion criteria in the parent trial included pregnancy, chronic corticosteroid therapy, immunosuppressive therapy, severe hepatic dysfunction, prior transplant, or planned ventricular assist device implantation.

Data Collection and Variables

Data were collected prospectively by trained research coordinators at each of the 23 participating centers using standardized electronic case report forms. Baseline and perioperative data were extracted from hospital electronic medical records, while the 180-day follow-up for mortality was conducted via telephone interviews or review of official death registries.

Collected variables included demographics, comorbidities, ethnicity, frailty, laboratory values, medications, and detailed procedural characteristics. Variables for the present post hoc analysis were selected a priori based on established cardiac surgery risk stratification frameworks, including factors reflected in EuroSCORE II, as well as their clinical relevance to sex-related differences in postoperative recovery and complications and their availability within the parent trial data set.[1,5,6,10–12] Ethnicity data were available in the parent trial and were recorded using standardized predefined categories in the electronic case report forms. Frailty was assessed using the Clinical Frailty Scale (CFS), a validated global clinical measure of fitness and frailty, with higher scores indicating greater frailty.[13] Frailty is also an established predictor of adverse outcomes after cardiac surgery.[14]

Outcomes

The primary clinical outcomes of interest for this sex-based analysis were 180-day mortality, major postoperative complications, and postoperative recovery. Major postoperative complications included stroke, postoperative dialysis, cardiac arrest, myocardial infarction, gastrointestinal ischemia, and major infective complications. Because the present analysis focused on recovery-related sex differences, ICU length of stay and total postoperative hospital length of stay were interpreted as key recovery outcomes. Additional postoperative outcomes included clinically significant atrial fibrillation, delirium in the ICU, hospital-acquired infections, pneumonia, invasive ventilation duration, vasopressor duration, circulatory support duration, and persistent organ dysfunction-free days.

Propensity Score Matching

For the primary analysis, women were matched 1:1 to men using nearest-neighbor propensity score matching based on age and EuroSCORE II, with a caliper of 0.2 SD of the logit of the propensity score, no replacement, and exact matching on surgical category. Age and EuroSCORE II were selected as clinically established baseline risk variables, whereas surgical category was exact-matched because procedural mix is a major determinant of postoperative outcomes and an important source of sex-related confounding in cardiac surgery.[1,7,9] This primary matching strategy yielded 327 female–male pairs.

Covariate balance before and after matching was assessed using absolute standardized mean differences (SMDs). Surgical category, which was exact-matched and represented by multiple indicator variables, is fully reported in Table S1 and omitted from Supplemental Figure S1 for visual clarity.

Although matching substantially improved balance in the primary matching variables, residual imbalance remained in selected clinical covariates, including baseline hemoglobin, CFS, severe renal disease, left ventricular ejection fraction (LVEF), previous myocardial infarction, and selected preoperative medications. These variables were therefore addressed in prespecified sensitivity analyses rather than forced into the primary propensity score model.

Statistical Analysis

Continuous variables are presented as median (quartile 1, quartile 3 [Q1, Q3]), and categorical variables as number (percentage). Distributional assumptions for continuous variables were assessed using Shapiro–Wilk testing together with visual inspection of histograms and Q–Q plots. In the matched cohort, several key continuous variables, particularly EuroSCORE II, operative times, creatinine, and the primary recovery end points (ICU and hospital length of stay), showed nonnormal within-pair difference distributions. Accordingly, rank-based nonparametric methods were used for descriptive continuous comparisons. Results of the normality assessment for selected paired differences are provided in Table S2.

In the unmatched cohort, groups were compared using Wilcoxon rank sum tests for continuous variables and Pearson χ^2 tests with Monte Carlo simulation for categorical variables. In the matched cohort, descriptive comparisons used Wilcoxon signed rank tests for continuous variables, McNemar or exact McNemar tests for paired binary categorical variables, and Stuart-Maxwell tests for paired categorical variables with >2 levels, as appropriate. Surgical category was exact-matched and is therefore reported descriptively without inferential testing. Formal inference after matching relied primarily on adjusted regression models that accounted for the matched design. Time-to-event outcomes were evaluated using Kaplan–Meier methods, log-rank testing, and Cox proportional hazards models. The proportional hazards assumption was assessed for all Cox models using scaled Schoenfeld residuals with global and covariate-specific testing. No evidence of violation of the proportional hazards assumption was detected; results are summarized in Table S3.

To address residual post-match imbalance, prespecified complete-case sensitivity analyses were performed in the matched cohort. Covariates included in the adjusted models were selected a priori based on residual imbalance after matching and established clinical relevance for postoperative mortality and recovery in cardiac surgery.[10,11,14–19] The primary adjusted mortality model was a pair-stratified Cox proportional hazards model including CFS, baseline hemoglobin, severe renal disease, LVEF, and previous myocardial infarction. A secondary Cox model with SEs clustered by matched pair was used to evaluate the robustness of the mortality findings under an alternative approach to handling the matched design. Recovery outcomes (ICU and hospital length of stay) were analyzed using negative binomial models with SEs clustered by matched pair because these outcomes represent overdispersed count-like data. Medication-adjusted sensitivity models additionally included preoperative angiotensin-converting enzyme inhibitor, angiotensin receptor blocker, and statin use because these variables also remained imbalanced after matching. Renal-marker sensitivity analyses replaced severe renal disease with baseline creatinine to test whether the findings were robust to an alternative representation of renal risk. A DAG-based variable-selection approach was not used in this post hoc analysis. Formal interaction testing between sex and other covariates was not prespecified for this post hoc analysis and was not pursued because of the limited number of outcome events, particularly for mortality, and the resulting risk of model overfitting. Absolute SMDs were used to assess covariate balance before and after matching. A threshold of <0.10 was considered indicative of acceptable balance for individual covariates, while acknowledging that several nonmatched variables remained partially imbalanced and required adjusted sensitivity analyses. All tests were 2-sided, and $P<0.05$ was considered statistically significant. Analyses were performed in R version 4.3.2 (R Foundation for Statistical Computing).

RESULTS

Study Cohort and Matching

Figure 1 summarizes cohort selection and matching. The primary matching cohort included 1386 patients with complete data for age, EuroSCORE II, and surgical category, of whom 346 were women and 1040 were men. Propensity score matching yielded 327 female–male pairs.

Baseline Characteristics

Baseline characteristics before and after matching are summarized in Table 1. Before matching, women were older than men (median, 71 years [Q1, Q3=64, 77 years] versus 70 years [Q1, Q3=61, 76 years]; $P=0.010$) and had higher EuroSCORE II values (median, 11 [Q1, Q3=7, 20] versus 8 [Q1, Q3=5, 13]; $P<0.001$). Women also had lower baseline creatinine concentrations (median, 71 μmol/L [Q1, Q3=62, 84 μmol/L] versus 88 μmol/L [Q1, Q3=76, 106 μmol/L]; $P<0.001$) and lower baseline hemoglobin concentrations (women, 122 g/L [Q1,

Table 1. Baseline Patient Characteristics Before and After Propensity Score Matching

Variable	Male before matching (n=1040)	Female before matching (n=346)	*P* value before matching	Male after matching (n=327)	Female after matching (n=327)	*P* value after matching
Age, y	70 (61, 76)	71 (64, 77)	0.010	71 (65, 77)	71 (64, 77)	0.300
BMI, kg/m^2	28.0 (25.3, 31.5)	27.4 (24.1, 31.3)	0.087	28.0 (25.3, 31.9)	27.7 (24.2, 31.6)	0.671
White race, n (%)	1003 (96)	334 (97)	>0.990	315 (96)	318 (97)	0.541
Cardiac and laboratory status						
LVEF, %	55 (50, 60)	55 (50, 60)	0.039	55 (45, 60)	56 (50, 60)	<0.001
Baseline platelets (10^9/L, lowest)	140 (111, 176)	144 (114, 181)	0.214	142 (107, 173)	145 (115, 182)	0.167
Baseline WBC count (10^3/μL, highest)	7.20 (6.00, 8.90)	7.30 (5.80, 8.80)	0.639	7.19 (5.90, 8.50)	7.30 (5.90, 8.80)	0.542
Baseline creatinine, μmol/L, highest	88 (76, 106)	71 (62, 84)	<0.001	95 (80, 113)	71 (62, 84)	<0.001
Baseline hemoglobin, g/L, lowest	134 (120, 146)	122 (107, 132)	<0.001	133 (119, 145)	122 (108, 132)	<0.001
Risk scores						
EuroSCORE II	8 (5, 13)	11 (7, 20)	<0.001	11 (7, 17)	11 (7, 19)	0.055
Clinical Frailty Scale, n (%)			<0.001			0.005
1	91 (8.9)	17 (5.0)		28 (8.6)	17 (5.3)	
2	308 (30)	72 (21)		85 (26)	71 (22)	
3	439 (43)	141 (41)		143 (44)	134 (41)	
4	155 (15)	96 (28)		55 (17)	86 (27)	
5	19 (1.9)	10 (2.9)		8 (2.5)	9 (2.8)	
6	9 (0.9)	6 (1.8)		3 (0.9)	6 (1.9)	
7	3 (0.3)	0 (0)		2 (0.6)	0 (0)	
Charlson comorbidity score, n (%)			0.109			0.390
0	384 (37)	138 (40)		95 (29)	129 (39)	
1	185 (18)	38 (11)		43 (13)	36 (11)	
2	264 (25)	91 (26)		105 (32)	87 (27)	
3	128 (12)	52 (15)		49 (15)	51 (16)	
4	51 (4.9)	18 (5.2)		24 (7.3)	17 (5.2)	
5	23 (2.2)	6 (1.7)		8 (2.4)	4 (1.2)	
6	3 (0.3)	1 (0.3)		1 (0.3)	1 (0.3)	
8	2 (0.2)	1 (0.3)		2 (0.6)	1 (0.3)	
9	0 (0)	1 (0.3)		0 (0)	1 (0.3)	
Comorbidities						
Moderate renal disease, n (%)	317 (30)	111 (32)	0.623	123 (38)	108 (33)	0.193
Severe renal disease, n (%)	51 (4.9)	34 (9.8)	0.001	23 (7.0)	30 (9.2)	0.157
CKD on dialysis, n (%)	1 (<0.1)	0 (0)	>0.990	1 (0.3)	0 (0)	1.000
Diabetes, n (%)	279 (27)	80 (23)	0.196	90 (28)	76 (23)	0.409
Preexisting AF, n (%)	208 (20)	87 (25)	0.051	80 (24)	80 (24)	0.649
Previous MI, n (%)	185 (18)	42 (12)	0.018	69 (21)	39 (12)	0.010
Previous cardiac surgery, n (%)	103 (9.9)	41 (12)	0.355	44 (13)	35 (11)	0.674

Data are median (Q1, Q3) or n (%). *P* values before matching were derived from Wilcoxon rank sum tests for continuous variables and Pearson chi-square tests with Monte Carlo simulation for categorical variables. *P* values after matching are descriptive and were derived from Wilcoxon signed rank tests for continuous variables, McNemar or exact McNemar tests for paired binary categorical variables, and Stuart-Maxwell tests for paired categorical variables with >2 levels, as appropriate. Surgical category was exact-matched and is presented descriptively without inferential testing. AF indicates atrial fibrillation; BMI, body mass index; CKD, chronic kidney disease; EuroSCORE II, European System for Cardiac Operative Risk Evaluation II; LVEF, left ventricular ejection fraction; MI, myocardial infarction; and WBC, white blood cell.

Q3=107, 132 g/L] versus men 134 g/L [Q1, Q3=120, 146 g/L]; *P*<0.001). Frailty distribution also differed before matching (*P*<0.001). After matching, age was similar between groups (median, 71 years [Q1, Q3=64, 77 years] versus 71 years [Q1, Q3=65, 77 years]; *P*=0.300), and EuroSCORE II was closely aligned although not perfectly identical (median, 11 [Q1, Q3=7, 19] versus 11 [Q1, Q3=7, 17]; *P*=0.055). White race also

remained similarly distributed (97% in women versus 96% in men; *P*=0.541). However, residual differences remained in selected nonmatched variables, including baseline creatinine (71 µmol/L [Q1, Q3=62, 84 µmol/L] versus 95 µmol/L [Q1, Q3=80, 113 µmol/L]; *P*<0.001), baseline hemoglobin (122 g/L [Q1, Q3=108, 132 g/L] versus 133 g/L [Q1, Q3=119, 145 g/L]; *P*<0.001), CFS distribution (*P*=0.005), LVEF (median, 56% [Q1, Q3=50, 60%] versus 55% [Q1, Q3=45, 60%]; *P*<0.001), and previous myocardial infarction (12% versus 21%; *P*=0.010). Balance diagnostics are summarized in Table S1 and Figure S1.

Cardiac Medications and Operative Characteristics

Cardiac medications and operative data before and after matching are shown in Table 2. After matching, women remained less likely to receive angiotensin-converting enzyme inhibitor therapy (28% versus 38%; *P*=0.027). Statin use also remained lower in women (54% versus 61%; *P*=0.031), whereas the difference in corticosteroid exposure was attenuated and no longer statistically significant after matching (4.6% versus 1.5%; *P*=0.189). Amiodarone use was numerically lower in women but did not differ significantly after matching (1.8% versus 2.8%; *P*=0.057). Operatively, women had shorter cardiopulmonary bypass times (median, 128 minutes [Q1, Q3=99, 163 minutes) versus 138 minutes [Q1, Q3=110, 171 minutes]; *P*=0.022) and shorter cross-clamp times (median, 95 minutes [Q1, Q3=72, 123 minutes] versus 105 minutes [Q1, Q3=83, 135 minutes]; *P*=0.003). Urgency distribution did not differ significantly after matching (*P*=0.198). Surgical category was balanced by exact matching and is therefore presented descriptively without inferential testing.

Postoperative Outcomes

Postoperative outcomes in the matched cohort are summarized in Table 3, and length-of-stay distributions are shown in Figure 2. In descriptive matched comparisons, women had numerically longer postoperative hospital stay (median, 10 days [Q1, Q3=7, 18 days] versus 9 [Q1, Q3=6, 16 days]; *P*=0.292) and longer time to discharge alive from the ICU (median, 3 days [Q1, Q3=1, 6 days] versus 2 days [Q1, Q3=1, 5 days]; *P*=0.067). In contrast, duration of invasive ventilation (median, 1.00 days [Q1, Q3=1.00, 2.00 days) versus 1.00 [Q1, Q3=1.00, 2.00 days]; *P*=0.480) and persistent organ dysfunction-free days (median, 29 [Q1, Q3=27, 30] versus 29 [Q1, Q3=28, 30]; *P*=0.729) were similar between groups.

Major postoperative complications did not differ significantly between sexes in the matched cohort, including clinically significant atrial fibrillation (16% versus 17%; *P*=0.756), ICU delirium (17% versus 19%; *P*=0.677), postoperative stroke (2.1% versus 2.4%; *P*=0.581), new-onset dialysis (5.8% versus 5.2%; *P*=0.739), hospital-acquired infections (9.2% versus 9.8%; *P*=0.686), and gastrointestinal complications. Biliary complications and pancreatitis were infrequent in both groups.

Sensitivity Analyses for Recovery Outcomes

The association between female sex and prolonged postoperative recovery remained consistent among adjusted negative binomial sensitivity analyses. For ICU length of stay, female sex remained associated with longer stay in the primary adjusted model (adjusted incidence rate ratio [IRR], 1.8 [95% CI, 1.2–2.9]; *P*=0.009), after additional adjustment for preoperative medications (adjusted IRR, 1.9 [95% CI, 1.2–3.0]; *P*=0.010), and in the creatinine-adjusted renal-marker sensitivity model (adjusted IRR, 2.2 [95% CI, 1.3–3.6]; *P*=0.002) (Table S4). For hospital length of stay, female sex remained associated with longer stay in the primary adjusted model (adjusted IRR, 1.4 [95% CI, 1.0–1.9]; *P*=0.031), after additional medication adjustment (adjusted IRR, 1.4 [95% CI, 1.0–1.9]; *P*=0.032), and in the creatinine-adjusted sensitivity model (adjusted IRR, 1.5 [95% CI, 1.1–2.2]; *P*=0.011) (Table S5). In adjusted negative binomial models, female sex remained independently associated with longer ICU and hospital length of stay, with consistent effect estimates among sensitivity analyses (Figure 2).

Survival Analysis

Kaplan–Meier estimates of 180-day survival for the matched cohort are shown in Figure 3. In the primary matched analysis, 180-day survival did not differ significantly between women and men (log-rank *P*=0.086). The corresponding unadjusted Cox model yielded a hazard ratio (HR) of 1.80 (95% CI, 0.91–3.55; *P*=0.091) for female versus male sex. By 180 days, 23 deaths had occurred in women and 13 in men. In the complete-case matched cohort used for adjusted mortality sensitivity analyses, 31 deaths occurred (21 in women and 10 in men). In a pair-stratified Cox model adjusted for CFS, baseline hemoglobin, severe renal disease, LVEF, and previous myocardial infarction, female sex was not independently associated with 180-day mortality (adjusted HR, 4.18 [95% CI, 0.80–21.78]; *P*=0.090). However, in a secondary Cox model with SEs clustered by matched pair and adjustment for the same covariates, female sex was associated with higher 180-day mortality (adjusted HR, 2.3 [95% CI, 1.1–4.8]; *P*=0.023). This association

Table 2. Preoperative Medications and Intraoperative Characteristics Before and After Matching

Variable	Male before matching (n=1040)	Female before matching (n=346)	*P* value before matching	Male after matching (n=327)	Female after matching (n=327)	*P* value after matching
Cardiac medications, n (%)						
ACEi	377 (36)	97 (28)	0.006	123 (38)	93 (28)	0.027
Amiodarone	28 (2.7)	6 (1.7)	0.425	9 (2.8)	6 (1.8)	0.057
ARBs	203 (20)	74 (21)	0.500	66 (20)	70 (21)	0.187
β-Blockers	573 (55)	182 (53)	0.456	175 (54)	172 (53)	0.167
Calcium channel blockers	262 (25)	91 (26)	0.735	85 (26)	85 (26)	0.409
Statins	640 (62)	187 (54)	0.017	200 (61)	178 (54)	0.031
Use of IV/oral corticosteroids, n (%)	19 (1.8)	17 (4.9)	0.003	5 (1.5)	15 (4.6)	0.189
Operative data						
CPB, min	134 (105, 171)	131 (99, 164)	0.057	138 (110, 171)	128 (99, 163)	0.022
CCT, min	104 (80, 134)	96 (72, 125)	0.009	105 (83, 135)	95 (72, 123)	0.003
Urgency, n (%)			0.964			0.198
Elective	858 (82)	285 (82)		258 (79)	271 (83)	
Emergency	4 (0.4)	1 (0.3)		0 (0)	1 (0.3)	
Urgent	178 (17)	60 (17)		69 (21)	55 (17)	
Type of surgery, n (%)			<0.001			—
CABG + valve/ multivalve	640 (62)	173 (50)		173 (53)	173 (53)	
CABG + other nonvalve	35 (3.4)	5 (1.4)		5 (1.5)	5 (1.5)	
Isolated CABG	37 (3.6)	17 (4.9)		15 (4.6)	15 (4.6)	
Isolated valve	64 (6.2)	26 (7.5)		23 (7.0)	23 (7.0)	
Isolated thoracic aorta	34 (3.3)	8 (2.3)		7 (2.1)	7 (2.1)	
Multivalve	94 (9.0)	66 (19)		57 (17)	57 (17)	
Valve + other nonvalve	136 (13)	51 (15)		47 (14)	47 (14)	

Data are median (Q1, Q3) or n (%). *P* values before matching were derived from Wilcoxon rank sum tests for continuous variables and Pearson χ^2 tests with Monte Carlo simulation for categorical variables. *P* values after matching were derived from Wilcoxon signed rank tests for continuous variables, McNemar or exact McNemar tests for paired binary categorical variables, and Stuart-Maxwell tests for paired categorical variables with >2 levels, as appropriate. The type of surgery was exact-matched and is therefore presented descriptively without inferential testing after matching. ACEi indicates angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; CABG, coronary artery bypass grafting; CCT, cross-clamp time; CPB, cardiopulmonary bypass; and IV, intravenous.

remained materially unchanged after additional adjustment for preoperative angiotensin-converting enzyme inhibitor, angiotensin receptor blocker, and statin use (adjusted HR, 2.3 [95% CI, 1.1–4.8]; *P*=0.025) and in a renal-marker sensitivity analysis replacing severe renal disease with baseline creatinine (adjusted HR, 2.6 [95% CI, 1.1–5.8]; *P*=0.022) (Table S6).

DISCUSSION

In this propensity score–matched post hoc analysis of high-risk cardiac surgery patients from the SUSTAIN-CSX trial, the most consistent finding was a sex-related difference in postoperative recovery rather than a uniform mortality effect. Although descriptive matched comparisons of ICU and hospital stay were attenuated, adjusted negative binomial sensitivity analyses consistently showed that female sex remained independently associated with longer ICU and hospital stay. In contrast, the mortality signal was more nuanced. The primary matched survival analysis did not show a statistically significant difference in 180-day survival between women and men, whereas adjusted mortality sensitivity analyses were model-dependent, with cluster-robust Cox models suggesting higher mortality in women and the pair-stratified model remaining imprecise because of sparse events. Taken together, these findings suggest that sex-related differences in this cohort were more consistently

Table 3. Postoperative Outcomes in the Propensity-Matched Cohort

Variable	Male (n=327)	Female (n=327)	*P* value
Postoperative recovery			
Postoperative length of stay, d	9 (6, 16)	10 (7, 18)	0.292
Duration to discharge alive from ICU, d	2 (1, 5)	3 (1, 6)	0.067
Length of invasive ventilation	1.00 (1.00, 2.00)	1.00 (1.00, 2.00)	0.480
Persistent organ dysfunction-free days	29 (28, 30)	29 (27, 30)	0.729
Cardiac and hemodynamic outcomes			
Clinically significant atrial fibrillation, n (%)	55 (17)	53 (16)	0.756
Duration of vasoactive support, d	1.00 (0.00, 2.00)	1.00 (0.00, 2.00)	0.758
Postoperative circulatory support, n (%)	12 (3.7)	8 (2.4)	0.648
Cardiac arrest, n (%)	5 (1.5)	8 (2.4)	0.581
Neurological and renal outcomes			
ICU delirium, n (%)	62 (19)	56 (17)	0.677
Postoperative stroke, n (%)	8 (2.4)	7 (2.1)	0.581
New-onset dialysis, n (%)	17 (5.2)	19 (5.8)	0.739
Gastrointestinal complications			
Postoperative GI ischemia, n (%)	2 (0.6)	2 (0.6)	1.000
Upper GI bleeding, n (%)	2 (0.6)	1 (0.3)	1.000
Lower GI bleeding, n (%)	1 (0.3)	2 (0.6)	1.000
Biliary complications, n (%)	1 (0.3)	0 (0)	1.000
Pancreatitis, n (%)	2 (0.6)	0 (0)	1.000
Infectious complications			
Hospital acquired infections, n (%)	32 (9.8)	30 (9.2)	0.686
Pneumonia, n (%)	1 (0.3)	2 (0.7)	1.000
Abdominal infection, n (%)	1 (0.3)	1 (0.3)	1.000
Urinary tract infection, n (%)	2 (0.6)	8 (2.4)	0.180
Surgical wound infection, n (%)	4 (1.2)	0 (0)	0.062
Other infections, n (%)	1 (0.3)	2 (0.6)	1.000

Values are median (Q1, Q3) or n (%). *P* values are descriptive and were derived from Wilcoxon signed rank tests for continuous variables, McNemar or exact McNemar tests for paired binary categorical variables, and Stuart-Maxwell tests for paired categorical variables with >2 levels, as appropriate. Adjusted matched-cohort sensitivity analyses for intensive care unit length of stay, hospital length of stay, and mortality are presented in Tables S4–S6. ICU indicates intensive care unit; and GI, gastrointestinal.

expressed through postoperative recovery than through a uniform mortality effect.

Mortality in the High-Risk Surgical Setting

Sex-related differences in mortality after cardiac surgery have been reported inconsistently among prior studies, with many analyses suggesting that the observed excess risk in women is strongly influenced by older age, greater comorbidity burden, smaller body size, coronary anatomy, and procedural complexity rather than sex alone.[5–9,20,21]

In our primary matched analysis, women and men were closely balanced with respect to age, EuroSCORE II, and surgical category, and the unadjusted matched analysis showed no statistically significant difference in 180-day survival. However, this neutral primary result should not be overinterpreted. The mortality findings in our sensitivity analyses were clearly model-dependent. Whereas the pair-stratified Cox model remained nonsignificant and yielded wide CIs, cluster-robust Cox models suggested a higher adjusted mortality hazard in women. This discrepancy likely reflects the limited number of deaths in the complete-case matched cohort (n=31), making the results sensitive to how the matched-pair structure and residual postmatch imbalance are handled analytically. Therefore, although our data do not support a definitive neutral mortality conclusion, they also do not establish a uniformly robust excess mortality effect; instead, they indicate that mortality differences in this high-risk cohort remain nuanced and require confirmation in larger studies.

Recovery Trajectory and Resource Utilization

In contrast to mortality, the association between female sex and prolonged postoperative recovery was highly consistent among all adjusted models. In descriptive paired comparisons, women showed numerically longer ICU and hospital stays that were no longer statistically significant after appropriate paired testing. However, in prespecified adjusted negative binomial sensitivity analyses that accounted for the matched design and residual postmatch imbalance, female sex remained consistently associated with longer ICU stay (adjusted IRR, 1.8–2.2) and hospital stay (adjusted IRR, 1.4–1.5) (all $P \leq 0.031$). This pattern suggests that the observed recovery gap is unlikely to be fully explained by residual imbalance in measured baseline variables alone and supports the presence of a sex-related difference in postoperative recovery trajectory after high-risk cardiac surgery.

At the same time, the present analysis does not establish the mechanism underlying the longer recovery trajectory. Frailty, anemia, and renal dysfunction are established predictors of adverse postoperative recovery and prolonged resource utilization after major surgery, including cardiac surgery.[10,11,14–19] Although women in our cohort had lower baseline hemoglobin, greater frailty burden, and persistent differences in selected clinical variables after matching, the adjusted models suggest that these factors do not fully account for the observed sex difference in resource use. In addition, several potentially relevant determinants of ICU or hospital discharge were not available in the trial data set, including postoperative nausea, early mobilization barriers, rehabilitation readiness, social support, discharge

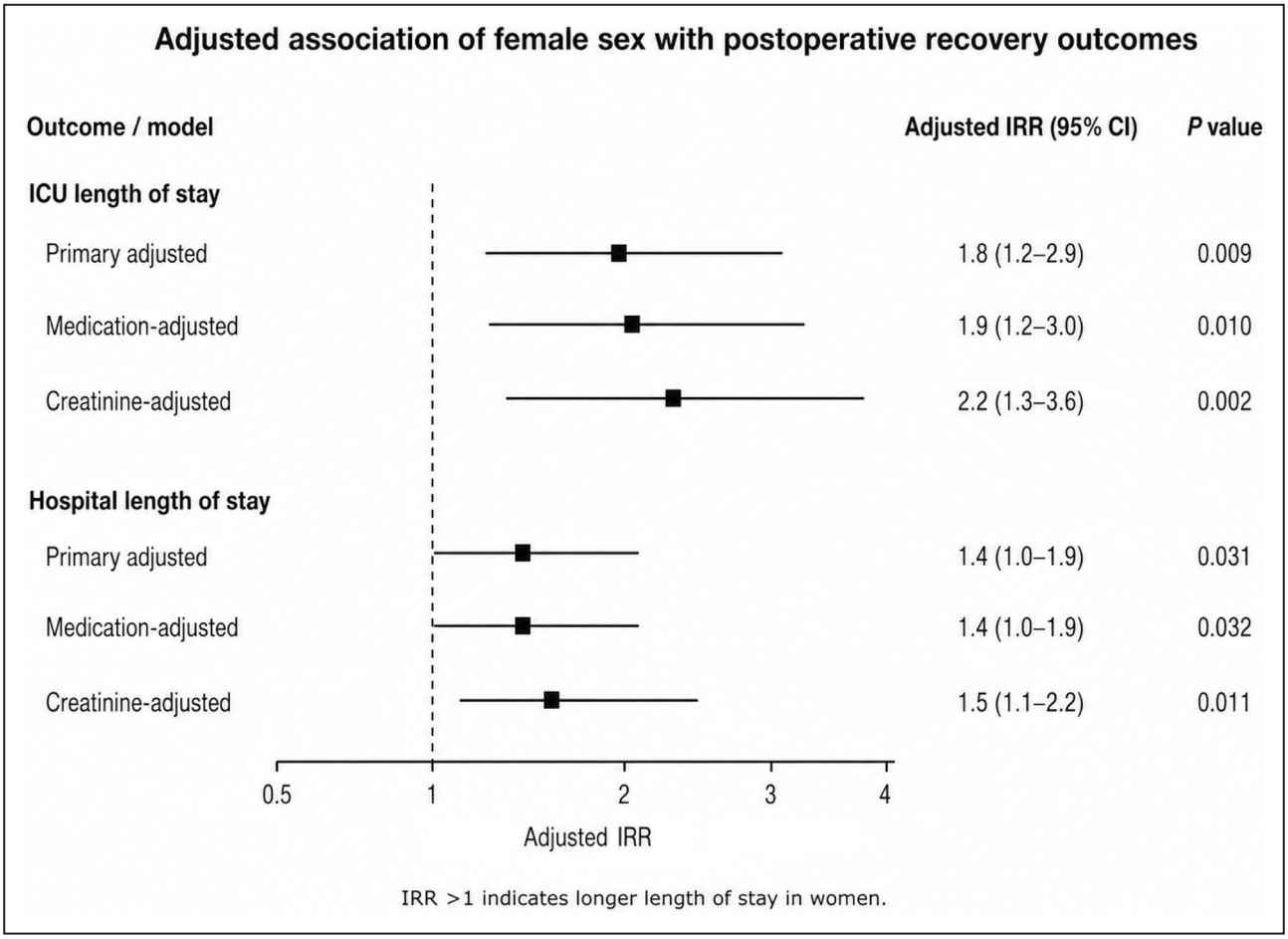

Figure 2. Adjusted association of female sex with postoperative recovery outcomes in the matched cohort.
Forest plot showing IRRs and 95% CIs for the association between female sex and postoperative ICU length of stay and hospital length of stay in negative binomial regression models. The primary model adjusted for D'Agostino RSCFS, baseline hemoglobin, severe renal disease, left ventricular ejection fraction, and previous myocardial infarction. Sensitivity models additionally adjusted for preoperative medication use or replaced severe renal disease with baseline creatinine. An IRR >1 indicates longer length of stay in women. CFS indicates Clinical Frailty Scale; ICU, intensive care unit; and IRR, incidence rate ratio.

logistics, and clinician-level thresholds for ICU or ward discharge. Therefore, the finding of prolonged recovery in women should be interpreted as clinically meaningful but mechanistically unresolved.

Preoperative Therapy and Residual Care Differences

Another relevant observation is that differences in preoperative medical therapy persisted even after matching, particularly for angiotensin-converting enzyme inhibitor use and corticosteroid exposure, with numerically lower statin use in women. Prior cardiovascular literature has repeatedly shown sex-related differences in the uptake of guideline-directed therapy, even in highly structured care settings.[22–26] In our study, however, the association between female sex and prolonged ICU and hospital stay remained materially unchanged after additional medication adjustment, indicating that preoperative treatment differences alone are unlikely to explain the recovery signal. These findings nonetheless highlight the importance of scrutinizing preoperative optimization pathways for potential sex-based disparities.

Clinical Interpretation

From a clinical perspective, these data suggest that sex-based perioperative counseling in high-risk cardiac surgery should place particular emphasis on postoperative recovery trajectory rather than relying exclusively on mortality estimates. Women in this cohort did not exhibit higher rates of major postoperative complications such as stroke, dialysis, or hospital-acquired infection, yet adjusted analyses consistently supported a longer postoperative recovery trajectory. This distinction is important because it shifts the clinical conversation away from framing female sex primarily as a mortality risk factor and toward anticipating a potentially more protracted

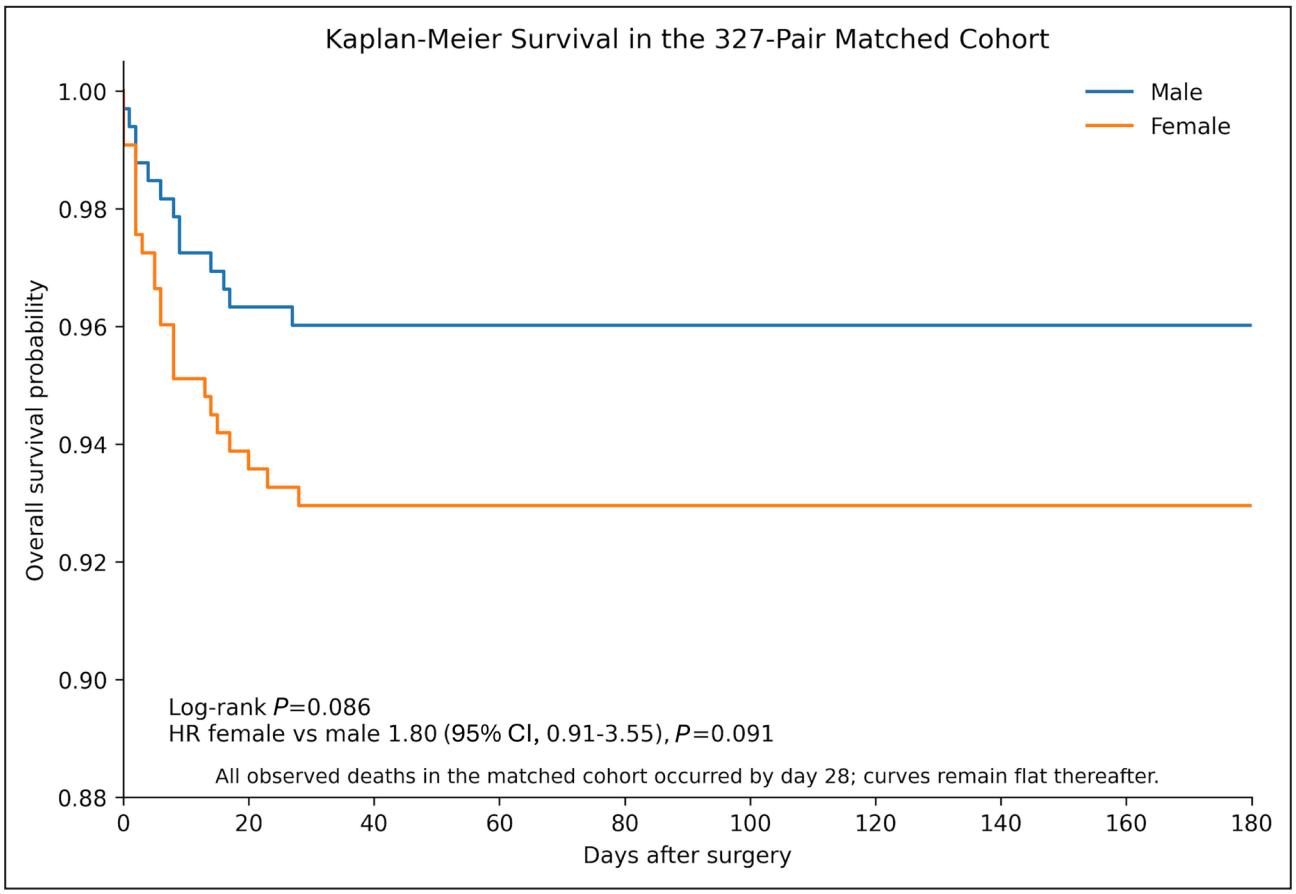

Figure 3. Kaplan–Meier estimates of 180-day survival in the propensity score–matched cohort.
Women and men were matched 1:1 on age and EuroSCORE II with exact matching on surgical category, yielding 327 pairs. All observed deaths in the matched cohort occurred by day 28; survival curves remained flat thereafter. Survival did not differ significantly between groups in the primary matched analysis (log-rank *P*=0.086). EuroSCORE II indicates European System for Cardiac Operative Risk Evaluation II; and HR, hazard ratio.

recovery course, greater resource utilization, and the possible need for more individualized discharge planning and rehabilitation support.

Strengths and Limitations

This study has several strengths. It was based on a multicenter trial data set with prospectively collected perioperative data, standardized case report forms, and detailed postoperative outcomes. The primary analysis also incorporated exact matching on surgical category, a key source of confounding in sex-based cardiac surgery comparisons, and used prespecified sensitivity analyses to address clinically important residual imbalance after matching.

Several limitations should also be acknowledged. First, although matching substantially improved balance in the primary matching variables, residual imbalance remained in selected covariates, including frailty, baseline hemoglobin, renal disease, LVEF, and previous myocardial infarction. Second, only 31 deaths occurred in the complete-case matched cohort used for adjusted mortality sensitivity analyses, limiting precision and contributing to model dependence of the mortality results. Potential interactions between sex and other covariates were not formally tested; such analyses would have been exploratory in this post hoc setting and were limited by sample size and sparse mortality events. Third, the adjusted recovery signal cannot be mechanistically explained by the present data set, because detailed chart-adjudicated reasons for prolonged ICU or hospital stay were not available. Fourth, although ethnicity data were available, the cohort was overwhelmingly White, which limited informative ethnicity-stratified analysis. Finally, as with any post hoc observational analysis, residual confounding cannot be fully excluded despite matching and multivariable sensitivity analyses.

Clinical Implications and Future Directions

The present findings support a more recovery-focused interpretation of sex differences after

high-risk cardiac surgery. Although descriptive matched comparisons of ICU and hospital stay were attenuated, women consistently showed longer recovery among multiple adjusted models, whereas mortality findings were more sensitive to model specification. Future studies should evaluate whether sex-specific prehabilitation, frailty-targeted optimization, anemia management, and recovery-oriented discharge pathways can reduce postoperative resource use and improve patient-centered recovery outcomes. Larger cohorts with greater event counts will also be needed to clarify whether the mortality signal observed in adjusted sensitivity analyses reflects true biological or care-related differences or residual confounding.

CONCLUSIONS

In this propensity score–matched cohort of high-risk cardiac surgery patients, descriptive paired comparisons showed numerically longer postoperative recovery in women that was attenuated after paired testing, whereas adjusted sensitivity analyses consistently demonstrated prolonged ICU and hospital length of stay. In contrast, mortality findings were model-dependent: the primary matched analysis was neutral, whereas adjusted cluster-robust sensitivity models suggested higher mortality in women and the pair-stratified model remained imprecise because of sparse events. These results support a sex-specific difference in postoperative recovery while underscoring the need for cautious interpretation of short-term mortality risk.

ARTICLE INFORMATION

Received January 15, 2026; accepted June 24, 2026.

Affiliations

Deparment of cardiac surgery, Medical Faculty, University Hospital RWTH Aachen, Aachen, Germany (R.Z., A.M., A.F.A.M.); Department of Cardiothoracic Surgery, Heart Centre Trier, Barmherzigen Brueder Hospital, Trier, Germany (A.K.); Department of Anaesthesiology, Intensive Care, Emergency and Pain Medicine, University Hospital, Würzburg, Germany (C.S., Q.N.); Department of Cardiac Anaesthesiology and Intensive Care Medicine, Deutsches Herzzentrum der Charité, Berlin, Germany (C.S., D.C., G.M.); Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin, Germany (C.S., D.C., G.M.); Department of Cardiothoracic Surgery, Faculty of Medicine, Qena University, Qena, Egypt (Y.S., A.F.A.M.); and Department of Cardiovascular Surgery, University Medical Center Mainz (Johannes Gutenberg-University Mainz), Mainz, Germany (J.B., G.D.D.).

Sources of Funding

The study has been funded by the Lotte and John Hecht Memorial Foundation. The investigational product, placebo, and initial financial support were provided by Biosyn Arzneimittel GmbH (Fellbach, Germany).

Disclosures

Dr Stoppe received support by the Deutsche Forschungsgemeinschaft (grants STO1099/10-11and STO1099/8–1). All other authors have no disclosures to report.

Supplemental Material

Tables S1–S6
Figure S1

REFERENCES

1. Bowdish ME, D'Agostino RS, Thourani VH, Schwann TA, Krohn C, Desai N. STS adult cardiac surgery database: 2021 update on outcomes, quality, and research. *Ann Thorac Surg*. 2021;111:1770–1780. doi: 10.1016/j.athoracsur.2021.03.043
2. Laffey J, Boylan J, Cheng D. The systemic inflammatory response to cardiac surgery. *Anesthesiology*. 2002;97:215–252.
3. Kraft F, Schmidt C, Van Aken H, Zarbock A. Inflammatory response and extracorporeal circulation. *Best Pract Res Clin Anaesthesiol*. 2015;29:113–123. doi: 10.1016/j.bpa.2015.03.001
4. Stoppe C, McDonald B, Meybohm P. Effect of high-dose selenium on postoperative organ dysfunction and mortality in cardiac surgery patients: the SUSTAIN CSX randomized clinical trial. *JAMA Surg*. 2023;158:235–244. doi: 10.1001/jamasurg.2022.6855
5. Leterrier J, Iung B, de Tymoski C. Sex differences and outcomes in surgical infective endocarditis. *Eur J Cardiothorac Surg*. 2024;65:ezae114.
6. Harik L, Perezgrovas-Olaria R, Soletti GJ. Sex differences in coronary artery bypass graft surgery outcomes: a narrative review. *J Thorac Dis*. 2023;15:5041–5054. doi: 10.21037/jtd-23-294
7. Dumitriu LaGrange D, Tessitore E, Reymond P, Mach F, Huber C. A systematic review and meta-analysis of differences between men and women in short-term outcomes following coronary artery bypass graft surgery. *Sci Rep*. 2024;14:20682. doi: 10.1038/s41598-024-71414-2
8. Bernard J, Dal Soglio S, Zhang B. Sex-related differences in outcomes according to surgical treatment approach in degenerative mitral regurgitation. *JACC Adv*. 2025;4:101897. doi: 10.1016/j.jacadv.2025.101897
9. Ren J, Bowyer A, Tian DH. Sex differences in long-term survival after total arterial coronary artery bypass grafting. *Eur J Cardiothorac Surg*. 2024;65:ezae106.
10. Neugarten J, Golestaneh L. Female sex reduces the risk of hospital-associated acute kidney injury: a meta-analysis. *BMC Nephrol*. 2018;19:314. doi: 10.1186/s12882-018-1122-z
11. Neugarten J, Golestaneh L. Sex differences in acute kidney injury. *Semin Nephrol*. 2022;42:208–218. doi: 10.1016/j.semnephrol.2022.04.010
12. Guida P, Mastro F, Scrascia G, Whitlock R, Paparella D. Performance of the European system for cardiac operative risk evaluation II: a meta-analysis of 22 studies involving 145,592 cardiac surgery procedures. *J Thorac Cardiovasc Surg*. 2014;148:3049–3057.e3041. doi: 10.1016/j.jtcvs.2014.07.039
13. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, Mitnitski A. A global clinical measure of fitness and frailty in elderly people. *CMAJ*. 2005;173:489–495. doi: 10.1503/cmaj.050051
14. Sepehri A, Beggs T, Hassan A. The impact of frailty on outcomes after cardiac surgery: a systematic review. *J Thorac Cardiovasc Surg*. 2014;148:3110–3117.
15. Karkouti K, Wijeysundera DN, Beattie WS. Risk associated with preoperative anemia in cardiac surgery: a multicenter cohort study. *Circulation*. 2008;117:478–484.
16. Musallam KM, Tamim HM, Richards T. Preoperative anaemia and postoperative outcomes in non-cardiac surgery: a retrospective cohort study. *Lancet*. 2011;378:1396–1407. doi: 10.1016/S0140-6736(11)61381-0
17. Hobson CE, Yavas S, Segal MS. Acute kidney injury is associated with increased long-term mortality after cardiothoracic surgery. *Circulation*. 2009;119:2444–2453. doi: 10.1161/CIRCULATIONAHA.108.800011
18. Muñoz M, Acheson AG, Auerbach M. International consensus statement on the peri-operative management of anaemia and iron deficiency. *Anaesthesia*. 2017;72:233–247. doi: 10.1111/anae.13773
19. Wang Y, Bellomo R. Cardiac surgery-associated acute kidney injury: risk factors, pathophysiology and treatment. *Nat Rev Nephrol*. 2017;13:697–711. doi: 10.1038/nrneph.2017.119
20. Ram E, Sternik L, Moshkovitz Y. Coronary artery bypass grafting following acute coronary syndrome: impact of gender. *Semin Thorac Cardiovasc Surg*. 2022;34:920–929. doi: 10.1053/j.semtcvs.2021.07.015
21. Chang FC, Chen SW, Chan YH. Sex differences in risks of in-hospital and late outcomes after cardiac surgery: a nationwide population-based cohort study. *BMJ Open*. 2022;12:e058538. doi: 10.1136/bmjopen-2021-058538

22. Vogel B, Acevedo M, Appelman Y. The lancet women and cardiovascular disease commission: reducing the global burden by 2030. *Lancet*. 2021;397:2385–2438.
23. Mehta LS, Beckie TM, DeVon HA. Acute myocardial infarction in women: a scientific statement from the American Heart Association. *Circulation*. 2016;133:916–947.
24. Kytö V, Sipilä J, Rautava P, Gunn J. Sex differences in outcomes following acute coronary syndrome treated with coronary artery bypass surgery. *Heart Lung Circ*. 2021;30:100–107. doi: 10.1016/j.hlc.2020.02.009
25. Bugiardini R, Ricci B, Cenko E. Delayed care and mortality among women and men with myocardial infarction. *J Am Heart Assoc*. 2017;6:e005968.
26. Zhao M, Woodward M, Vaartjes I. Sex differences in cardiovascular medication prescription in primary care: a systematic review and meta-analysis. *J Am Heart Assoc*. 2020;9:e014742.