E.2 All AAAs

Full citation	**Casey K, Hernandez-Boussard T, Mell MW, Lee JT. (2013) Differences in readmissions after open repair versus endovascular aneurysm repair. J Vasc Surg. 57(1):89-95**
Study details	Study design: retrospective cohort study

Full citation	**Casey K, Hernandez-Boussard T, Mell MW, Lee JT. (2013) Differences in readmissions after open repair versus endovascular aneurysm repair. J Vasc Surg. 57(1):89-95**
	Location(s): USA Study period: 2003 to 2008 Aim of the study: to determine reasons for all-cause readmissions within the first year after open repair and EVAR
Participants	Sample size: EVAR group, n=9,356; OSR group, n= 6,380 Inclusion criteria: adults who underwent open repair or endovascular repair of infrarenal AAA. Patients were identified using International Classification of Diseases, Ninth Revision (ICD-9) diagnosis codes for intact AAA. All patients at risk for 1-year readmission after elective AAA repair were included. Exclusion criteria: patients identified with ICD-9 codes for ruptured AAA. Age <40 years. Age >90 years. Baseline characteristics: Mean age (SD): EVAR group, 75.0 (8.3) years; OSR group, 72.1 (9.0) years Gender: EVAR group, 84.7% male; OSR group, 75.6% male Mean aneurysm diameter: not reported Private insurance: EVAR group, 15.3%; OSR group, 20.8% Charlson Index ≥2: EVAR group, 51.7%; OSR group, 55.5% Obesity: EVAR group, 7.2%; OSR group, 7.0% Congestive heart failure: EVAR group, 9.8%; OSR group, 14.1% Complicated DM: EVAR group, 1.7%; OSR group, 1.2% Peripheral vascular disease: EVAR group, 98.8%; OSR group, 98.5% End-stage renal disease: EVAR group, 8.6%; OSR group, 9.1%
Methods	Data collection: The State Inpatient Database (SID), contains a range of data collected from discharge inpatient hospital records, including demographics, ICD-9-Clinical Modification codes for primary and secondary diagnoses and procedures, admission source, length of stay, discharge disposition, inpatient mortality, and hospital characteristics. This database also allows for identification of patient readmissions ≤1 full year in California and attempts to capture patient data characteristics from the readmission, including primary and secondary diagnoses. Multiple readmissions from the same patient ≤1 year counted toward the total number of readmissions for that cohort. Analysis: Modified Cox proportional hazards modelling with adjustment for patient and hospital characteristics was used to adjust for patient mix on readmission rates between the two surgical procedures. Patient characteristics included age, sex, race, insurance status, obesity, complicated diabetes mellitus (DM), complicated hypertension, peripheral vascular disease (PVD), chronic obstructive pulmonary disease (COPD), and end-stage renal disease.
Intervention	EVAR
Comparator	OSR

Full citation	**Casey K, Hernandez-Boussard T, Mell MW, Lee JT. (2013) Differences in readmissions after open repair versus endovascular aneurysm repair. J Vasc Surg. 57(1):89-95**
Outcomes	Readmission rates
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Moderate risk - ≥5-yr recruitment with no adjustment for year of operation 1.2. Were cohorts from the same place? Low risk – same country 1.3. Is the definition of AAA the same across cohorts? Moderate risk – recruitment methods would have led to all infrarenal and complex AAA included. This is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk - good range of individual comorbidities 2.3. Does study control appropriately for AAA characteristics? High risk - none 2.4. Could any adjustment variables have been affected by the intervention? Low risk **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – administrative registry 3.2. Is method of data collection likely to record perioperative outcomes accurately? N/A 3.3. Is method of data collection likely to record long-term outcomes accurately? Low risk – the administrative registry is likely to record readmission rates **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – none reported 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – no details reported, and long-term nature of readmission outcome raise risk of loss to follow-up 4.3. Have different methods been compared within the study? High risk – none reported **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? Low risk 6.2 Were interactions between treatment and other covariates considered? High risk - no **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Chang DC, Parina RP, Wilson SE et al. (2015) Survival After Endovascular vs Open Aortic Aneurysm Repairs. JAMA Surg. 150 (12): 1160-6.**
Study details	Study design: Retrospective cohort study Location(s): USA Study period: 2001 to 2009 Aim of the study: to determine long-term outcomes of EVAR vs open repair on a population level.
Participants	Sample size: EVAR group, n=12,239; OSR group, n= 11,431 Inclusion criteria: all people who underwent EVAR or OSR of unruptured AAA at non-federal hospitals in California were included. Note: there is no indication that selection as limited to people with infrarenal AAA, and people with complex aneurysms may have been included. Exclusion criteria: people with concomitant thoracic aneurysm repairs or diagnosis of syphilitic, traumatic, thoracoabdominal, ruptured, or unspecified aortic aneurysms were excluded. Baseline characteristics: Mean age: EVAR group, 75.1 years; OSR group, 72.3 years Gender: EVAR group, 84.4% male; OSR group, 77.5% male Mean aneurysm diameter: not reported Private insurance: EVAR group, 15.4%; OSR group, 20.6% Charlson Comorbidity Index score of ≥2: EVAR group, 57.1%; OSR group, 62.7%
Methods	Data collection: participants were identified and data were collected by querying an administrative database maintained by California state authorities using ICD9 diagnosis and procedure codes. The administrative database was also inked to the Social Security Death Index to obtain mortality records. Analysis: Cox proportional hazards analyses were performed, controlling for repair type (open vs EVAR), age, race/ethnicity, gender, insurance types, Charlson Comorbidity Index score, calendar year, admission type (scheduled vs unscheduled), and hospital type
Intervention	EVAR
Comparator	OSR
Outcomes	Overall long-term mortality, and reinterventions
Study Appraisal using NICE's bespoke risk of bias	**Selection** 1.1. Were cohorts from the same time period? Low risk – ≥5-yr recruitment with adjustments made for year of operation 1.2. Were cohorts from the same place? Low risk – all participants came from California

Full citation	Chang DC, Parina RP, Wilson SE et al. (2015) Survival After Endovascular vs Open Aortic Aneurysm Repairs. JAMA Surg. 150 (12): 1160-6.
assessment tool	1.3. Is the definition of AAA the same across cohorts? Moderate risk – recruitment methods would have led to all infrarenal, complex, and possibly some symptomatic unruptured AAA being included. This is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – study controlled for age, gender and ethnicity 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – a validated comorbidity index (Charlson Comorbidity score) was considered in the regression model. 2.3. Does study control appropriately for AAA characteristics? High risk – AAA characteristics were not controlled for 2.4. Could any adjustment variables have been affected by the intervention? Low risk – no post-intervention variables which could mediate the treatment effect were controlled for **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – an administrative database with high-level diagnosis and procedure codes was used to identify participants 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – an administrative database was used to assess outcomes 3.3. Is method of data collection likely to record long-term outcomes accurately? Moderate risk – an administrative database was used with linkage to a reliable mortality registry **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – no details of any checks for model fit were reported. 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – no details provided. 4.3. Have different methods been compared within the study? High risk – different methods were not compared. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? Low risk – number of events is ≥10 times greater than number of variables considered 6.2 Were interactions between treatment and other covariates considered? High risk – no interactions were considered **Overall risk of bias:** High risk **Directness:** Directly applicable

Full citation	**Davenport DL, and Xenos ES., et al. (2013) Deep venous thrombosis after repair of nonruptured abdominal aneurysm. J Vasc Surg. 2013 Mar;57(3):678-683**
Study details	Study design: retrospective cohort study Location(s): USA Study period: 2005 to 2009 Aim of the study: to examine venous thromboembolism (VTE) rates, timing, and risk factors after nonruptured open or endoluminal abdominal aortic aneurysm (AAA) repair.
Participants	Sample size: EVAR group, n=8,502; OSR group, n= 3,967 Inclusion criteria: people undergoing elective AAA repair were included. Note: there is no indication that selection as only limited to people with infrarenal AAA. Exclusion criteria: people <18 years old, people with ruptured AAA and people with admissions for trauma were excluded. Baseline characteristics (for all participants): Mean age (SD): 73.2 (8.7) years Gender: 80.2% male Mean aneurysm diameter: not reported Medically treated diabetes: 14.0% Medically treated hypertension:80.8% Prior cardiac stent or operation: 38.3% Chronic heart failure, cardiac arrest or infarct: 3.9% History of stroke, TIA, or hemiplegia: 15.3%
Methods	Data collection: data were extracted from a detailed national surgical registry (the National Surgical Quality Improvement Program; NSQIP). Patients who underwent repair procedures for complex AAAs were identified using ICD9 diagnostic codes and CPT procedure codes. Analysis: Stepwise multivariate logistic regression of the occurrence of VTE (in hospital) was performed considering over 40 preoperative variables contained in the NSQIP database. Intraoperative variables (including procedure type, operative duration, wound class, and intraoperative transfusion) were also considered in the regression model.
Intervention	EVAR
Comparator	OSR
Outcomes	Deep vein thrombosis
Study Appraisal	**Selection**

Full citation	Davenport DL, and Xenos ES., et al. (2013) Deep venous thrombosis after repair of nonruptured abdominal aneurysm. J Vasc Surg. 2013 Mar;57(3):678-683
using NICE's bespoke risk of bias assessment tool	1.1. Were cohorts from the same time period? Low risk – cohorts were drawn from the same time period. 1.2. Were cohorts from the same place? Low risk – all study participants were identified using the same national surgical registry. 1.3. Is the definition of AAA the same across cohorts? High risk – the definition of OSR includes juxta-/pararenal AAA whereas EVAR codes are limited to infrarenal. **Confounding** 2.1. Does study control appropriately for demographics? High risk – the study did not control for demographic variables. 2.2. Does study control appropriately for comorbidity and/or fitness? Moderate risk – a limited number of individual comorbidities were controlled for. 2.3. Does study control appropriately for AAA characteristics? High risk – AAA characteristics were not controlled for. 2.4. Could any adjustment variables have been affected by the intervention? High risk – intra operative variables that may mediate the treatment effect (Intraoperative transfusion) were controlled for. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? Moderate risk – detailed surgical registries were used with diagnosis and procedure codes specified. 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – administrative codes with high level diagnosis and procedure codes were used to record outcomes. 3.3. Is method of data collection likely to record long-term outcomes accurately? Low risk – no long-term outcomes were assessed. **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – no checks for model specification/fit were reported. 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – authors briefly mention that there was some missing data but do not provide information as to how the missing data was dealt with. 4.3. Have different methods been compared within the study? High risk – different methods were not compared. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? High risk – number of events is <10 times greater than number of variables considered 6.2 Were interactions between treatment and other covariates considered? High risk – no interactions were considered **Overall risk of bias:** High risk

Full citation	Davenport DL, and Xenos ES., et al. (2013) Deep venous thrombosis after repair of nonruptured abdominal aneurysm. J Vasc Surg. 2013 Mar;57(3):678-683
	Directness: directly applicable

Full citation	Egorova NN, Vouyouka AG, McKinsey JF, et al. (2011) Effect of gender on long-term survival after abdominal aortic aneurysm repair based on results from the Medicare national database. J Vasc Surg. 54(1):1-12.e6
Study details	Study design: retrospective cohort study Location(s): USA Study period: 1995 to 2006 Aim of the study: to determine the effect of gender on long-term survival differences after EVAR and OSR of AAA under elective and emergency circumstances
Participants	Sample size of matched cohort: EVAR group, n=42,320; OSR group, n=42,320 Inclusion criteria: all people who underwent EVAR or OSR for ruptured or unruptured AAA were included. NB: it is possible that the study sample included people with complex AAA morphologies. Exclusion criteria: the study excluded people with thoracic or thoracoabdominal aneurysms. Baseline characteristics: Mean age: EVAR group - men 76.2 years, women 77.45 years; OSR group - men 74.6 years, women 75.5 years Gender: EVAR group, 82.4% male; OSR group, 76.0% male Mean aneurysm diameter: not reported Coronary comorbidities: EVAR group - men 37.6%, women 30.6%; OSR group - men 49.4%, women 38.1%; Arrhythmia: EVAR group - men 26.6%, women 20.6%; OSR group - men 28.3%, women 22.9%; Coronary heart failure: EVAR group - men 15.0%, women 16.1%; OSR group - men 16.0%, women 17.9%; Renal failure: EVAR group - men 6.3%, women 6.1%; OSR group - men 4.3%, women 4.1%; Diabetes: EVAR group - men 16.9%, women 14.9%; OSR group - men 11.3%, women 10.0%; Cancer: EVAR group - men 8.2%, women 4.8%; OSR group - men 7.1%, women 3.9%; Hypertension: EVAR group - men 72.6%, women 78.2%; OSR group - men 60.3%, women 68.2%; Hyperlipidaemia: EVAR group - men 42.5%, women 41.3%; OSR group - men 23.1%, women 24.5%;
Methods	Data collection: data were retrospectively collected from the Medicare Inpatient Standard Analytical and Denominator databases. Participants were selected through a combination of diagnosis codes (ICD9). The hospitalisation with the first AAA procedure was identified as the index hospitalisation. Pre-index and index hospitalisations were used to assess baseline comorbidities; including all those identified in prior hospitalisations and the chronic conditions reported at the index hospitalisation.

Full citation	**Egorova NN, Vouyouka AG, McKinsey JF, et al. (2011) Effect of gender on long-term survival after abdominal aortic aneurysm repair based on results from the Medicare national database. J Vasc Surg. 54(1):1-12.e6**
	Analysis: Propensity matching was performed. To determine the propensity score, a logistic regression model was developed where the dependent variable was the type of procedure or gender. All baseline confounders, including gender, race, comorbidities age, year of surgery, hospital, and surgeon volume (as continuous variables) were included in the model as independent variables. The patients were matched by greedy algorithm using the 8- to 1-digit matching scheme without replacement. Perioperative complications and 30-day mortality were compared in matched groups using the McNemar test. Survival curves were constructed with Cox models. Propensity score and type of repair or gender were included in this regression analysis.
Intervention	EVAR
Comparator	OSR
Outcomes	30-day mortality, perioperative adverse events, long term survival, long-term adverse events, and reintervention.
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? High risk – all 1995 to 2006 hospitalisations for OSR and hospitalisations from 2000 through 2006 for EVAR were included. 1.2. Were cohorts from the same place? Low risk – all participants were selected from the same databases of a nationwide health insurance provider. 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA were included. This is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – demographic variables including age and gender were controlled. 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – a good range of individual comorbidities were controlled for. 2.3. Does study control appropriately for AAA characteristics? High risk – no indication that AAA characteristics were adjusted for. 2.4. Could any adjustment variables have been affected by the intervention? Low risk – no post-intervention variables which could mediate the treatment effect were controlled for. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – an administrative (healthcare insurance provider) database with high-level diagnosis and procedure codes was used in this study. 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – an administrative (healthcare insurance provider) database with high-level diagnosis and procedure codes was used in this study. 3.3. Is method of data collection likely to record long-term outcomes accurately? High risk – an administrative (healthcare insurance provider) database with high-level diagnosis and procedure codes was used in this study. **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – none reported.

Full citation	**Egorova NN, Vouyouka AG, McKinsey JF, et al. (2011) Effect of gender on long-term survival after abdominal aortic aneurysm repair based on results from the Medicare national database. J Vasc Surg. 54(1):1-12.e6**
	4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – authors did not demonstrate that missing data were considered in the analyses. 4.3. Have different methods been compared within the study? Moderate risk – different methods were compared but they relied on the same assumption about selection. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? Low risk – a greedy algorithm that used an 8 to 1 matching scheme without replacement was used. 5.2. Was overlap / common support appropriately assessed? High risk – no assessment reported. 5.3. Has balancing of the covariates been demonstrated? Moderate risk – Conventional hypothesis tests, with no evidence of significant differences. **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? N/A 6.2 Were interactions between treatment and other covariates considered? N/A **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Eslami MH, Rybin DV, Doros G, Farber A. (2017) Description of a risk predictive model of 30-day postoperative mortality after elective abdominal aortic aneurysm repair. J Vasc Surg. 65:65-75**
Study details	Study design: retrospective cohort study Location(s): USA Study period: 2005 to 2011 Aim of the study: to describe such a 30-day postoperative risk prediction model using American College of Surgeons National Surgical Quality Improvement Project (NSQIP) data.
Participants	Sample size: EVAR group, n=4,635; OSR group, n= 14,282 Inclusion criteria: patients who underwent elective AAA repair Exclusion criteria: Patients who had a non-elective admission for AAA repair or were missing information for age, sex, discharge status (dead, discharged to home, or discharged to rehabilitation facility) or procedure type were excluded. We also excluded patients who had secondary procedures. Any missing data was treated as missing completely at random and not included in the analyses, and no data imputation was considered.

Full citation	**Eslami MH, Rybin DV, Doros G, Farber A. (2017) Description of a risk predictive model of 30-day postoperative mortality after elective abdominal aortic aneurysm repair. J Vasc Surg. 65:65-75**
	Baseline characteristics: Mean age: data not provided for the EVAR and OSR groups separately Gender: data not provided for the EVAR and OSR groups separately Mean aneurysm diameter: not reported
Methods	Data collection: The NSQIP database was queried. This quality improvement database that is risk-adjusted and includes information about both index admission and subsequent 30-day postoperative follow-up after the index operation. These risk adjustments are performed for the hospitals. Currently, over 500 hospitals in the United States participate in NSQIP, including university centres and private hospitals. These institutions select for analysis a random set of surgical cases, including those from general, vascular, plastic, and cardiac surgery. Cases are gathered on an 8-day cycle, trained staff abstract 135 pre-, intra-, and postoperative measures directly from the medical record and enter them as de-identified data in the database Analysis: multivariable logistic regression. In this parsimonious model, choice of operation significantly increased odds of mortality (OAR vs EVAR: odds ratio [OR], 2.71; 95% confidence interval [CI], 2.12-3.47; P<0.001). In this model, presence of chronic obstructive pulmonary disease (OR, 1.44; 95% CI, 1.10-1.90; P=0.008) and history of myocardial disease (OR, 1.36; 95% CI, 1.06-1.74; P=0.017) were significantly associated with increased mortality. Elevated creatinine also significantly increased the risk of mortality. End-stage renal disease (ESRD) was associated with POD in a univariate fashion, but this variable was not included in the final model. Age and sex also affected mortality; patients 70 years of age and older had significantly higher odds of death compared with those under 70 (OR, 2.24, 95% CI, 1.66-3.03; P<0.001) and women had increased odds of mortality compared with men (OR 1.46, 95% CI, 1.11-1.92; P=0.007). Functionally dependent patients had more than twice odds of mortality after elective AAA repair than functionally independent patients (OR, 2.29; 95% CI, 1.44-3.64; P<0.001). It was noted that weight loss, defined by ACS-NSQIP as ">10% loss of body weight in the last six months", was highly predictive of mortality after elective AAA repair.
Intervention	EVAR
Comparator	OSR
Outcomes	30-day mortality
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Moderate risk - ≥5-yr recruitment with no adjustment for year of operation 1.2. Were cohorts from the same place? Low risk – same country 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA were included. This is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk

Full citation	**Eslami MH, Rybin DV, Doros G, Farber A. (2017) Description of a risk predictive model of 30-day postoperative mortality after elective abdominal aortic aneurysm repair. J Vasc Surg. 65:65-75**
	2.3. Does study control appropriately for AAA characteristics? High risk - none 2.4. Could any adjustment variables have been affected by the intervention? Low risk **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – administrative database 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – administrative database 3.3. Is method of data collection likely to record long-term outcomes accurately? N/A **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? Low risk 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? Low risk 4.3. Have different methods been compared within the study? High risk – no different methods compared **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? Low risk (20 covariates; 315 events) 6.2 Were interactions between treatment and other covariates considered? High risk **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Giles KA, Landon BE, Cotterill P, et al. (2011) Thirty-day mortality and late survival with reinterventions and readmissions after open and endovascular aortic aneurysm repair in Medicare beneficiaries. J Vasc Surg. 53(1):6-12**
Study details	Study design: retrospective cohort study Location(s): USA Study period: 2001 to 2004 Aim of the study: to evaluate the impact of reinterventions and readmission after initial AAA repair on 30-day and long-term mortality
Participants	Sample size of matched cohort: EVAR group, n=22,826; OSR group, n=22,826

Full citation	Giles KA, Landon BE, Cotterill P, et al. (2011) Thirty-day mortality and late survival with reinterventions and readmissions after open and endovascular aortic aneurysm repair in Medicare beneficiaries. J Vasc Surg. 53(1):6-12
	Inclusion criteria: All Medicare beneficiaries who underwent EVAR or OSR for unruptured AAA repair were included. Note: there is no indication that selection was limited to only people with infrarenal AAA. Exclusion criteria: not reported Baseline characteristics (in participants who did not received any reintervention): Mean age (SD): EVAR group, 76.2 (5.4 years; OSR group, 75.9 (5.2) years Gender: EVAR group, 84.7% male; OSR group, 85.6% male Mean aneurysm diameter: not reported Myocardial disease within 6 months: EVAR group, 1.9%; OSR group, 1.8% Congestive heart failure: EVAR group, 13.2%; OSR group, 13.1% Peripheral vascular disease: EVAR group, 20.9%; OSR group, 20.3% Cerebrovascular disease: EVAR group, 16.5%; OSR group, 16.2% Hypertension: EVAR group, 65.9%; OSR group, 65.7% Diabetes: EVAR group, 15.8%; OSR group, 15.8% COPD: EVAR group, 29.7%; OSR group, 29.1%
Methods	Data collection: data were retrospectively collected from the Medicare databases. Patient demographic characteristics were identified from the Medicare denominator file database. Comorbidities were ascertained using inpatient and outpatient claims up to 2 years before but not including the admission for repair. Reinterventions and readmissions were determined from inpatient and outpatient claims by assessing ICD9 coding, and mortality was determined from the Medicare denominator file. Analysis: To control for the non-random assignment of patients to OSR vs EVAR, investigators created matched cohorts of patients using a logistic regression model from demographics and pre-existing comorbidities. Investigators measured the rates of coexisting conditions using an Elixhauser algorithm that was adapted to also include diagnoses made in the outpatient setting. They matched each beneficiary who underwent EVAR to the beneficiary who underwent OSR with the closest estimated propensity score. To ensure close matches, investigators required that the estimated log-odds scores for EVAR of a patient who underwent EVAR and one who underwent OSR were within 0.60 SD of one another. They stated that this requirement ensured the removal of approximately 90% of the bias in estimates of effects due to differences in covariate distributions between the EVAR group and the OSR group.
Intervention	EVAR
Comparator	OSR
Outcomes	30-day mortality and reinterventions up to 7 years.
Study Appraisal using NICE's	**Selection** 1.1. Were cohorts from the same time period? Low risk – cohorts were drawn from the same time period.

Full citation	**Giles KA, Landon BE, Cotterill P, et al. (2011) Thirty-day mortality and late survival with reinterventions and readmissions after open and endovascular aortic aneurysm repair in Medicare beneficiaries. J Vasc Surg. 53(1):6-12**
bespoke risk of bias assessment tool	1.2. Were cohorts from the same place? Low risk – all participants were selected from the same databases of a nationwide health insurance provider. 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA were included. This is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – although authors did not explicitly state that age and gender were adjusted for they stated that they used all available demographic and clinical characteristics for beneficiaries at baseline as explanatory variables. These are likely to have included age and gender. 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – authors stated that they used all available demographic and clinical characteristics for beneficiaries at baseline as explanatory variables. 2.3. Does study control appropriately for AAA characteristics? High risk – authors stated that they were not able to assess anatomic differences among patients. 2.4. Could any adjustment variables have been affected by the intervention? Low risk – no post-intervention variables that could mediate the treatment effect appear to be controlled for. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – an administrative (healthcare insurance provider) database with high-level diagnosis and procedure codes was used in this study. 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – an administrative (healthcare insurance provider) database was used. 3.3. Is method of data collection likely to record long-term outcomes accurately? High risk – an administrative (healthcare insurance provider) database was used. **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – no checks reported. 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – authors did not demonstrate that missing data were considered in the analyses. 4.3. Have different methods been compared within the study? High risk – different methods were not compared. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? Low risk – matching algorithm reported in previous study by the same authors and seems reasonable. 5.2. Was overlap / common support appropriately assessed? High risk – no indication that overlap was accurately assessed. 5.3. Has balancing of the covariates been demonstrated? High risk – no standardised differences reported.

Full citation	Giles KA, Landon BE, Cotterill P, et al. (2011) Thirty-day mortality and late survival with reinterventions and readmissions after open and endovascular aortic aneurysm repair in Medicare beneficiaries. J Vasc Surg. 53(1):6-12
	Analysis – simple multivariable models 6.1 Is sample size adequate relative to number of covariates considered? N/A 6.2 Were interactions between treatment and other covariates considered? N/A **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	Jackson RS, Chang DC, and Freischlag JA. (2012) Comparison of long-term survival after open vs endovascular repair of intact abdominal aortic aneurysm among Medicare beneficiaries. JAMA. 307(15):1621-8
Study details	Study design: retrospective cohort study Location(s): USA Study period: January 2003 to December 2007 Aim of the study: to compare long-term outcomes after EVAR and OSR of unruptured AAA
Participants	Sample size: EVAR group, n=3,826; OSR group, n=703 Inclusion criteria: people 65 and over with clinical diagnosis as well as procedure codes (ICD9 codes) corresponding to EVAR or OSR of unruptured AAA were included. Exclusion criteria: People younger than 65 years Baseline characteristics: Mean age (SD): EVAR group, 76.4 (6.3) years; OSR group, 75.2 (5.7) years Gender: EVAR group, 79.9% male; OSR group, 70.8% male Mean aneurysm diameter: not reported Diabetes: EVAR group, 18.8%; OSR group, 14.1% Hypertension: EVAR group, 59.1%; OSR group, 53.9% Chronic renal failure: EVAR group, 3.0%; OSR group, 4.1% Congestive heart failure: EVAR group, 12.3%; OSR group, 13.2% COPD: EVAR group, 24.3%; OSR group, 25.2%
Methods	Data collection: the study cohort was composed of eligible patients identified from the Medicare Standard Medicare Standard Analytic Files database using ICD9 codes. In the database, patient records contain longitudinal data from the index admission and subsequent hospital admissions, vital status, and date of death for deceased beneficiaries (based on linkage with social security administrative database). Comorbidities were assigned using Clinical Classifications Software (CCS): the software consolidates over 14,000 ICD9 codes and is developed by the Healthcare Cost and Utilization Project of the Agency for Healthcare Research and Quality. Analysis: Logistic regression was used to calculate a propensity score for each patient based on emergency presentation, age, calendar year of repair, gender, race, type of CCS diagnostic category, and number CCS diagnostic categories. Propensity scores were divided into quintiles by repair type, and propensity score quintile was treated as a categorical variable in the multivariable models. Multivariate logistic regression was used to assess 30-day mortality and hospital length of stay. Coz proportional hazards regression was used for all other outcomes. Two sensitivity analyses were performed to examine the effect of model assumptions on study results. The first sensitivity analysis examined the influence of possible misclassification of surgical complications as comorbidities in the portion of the cohort whose CCS categories were drawn from the same calendar years as the index AAA repair. The second sensitivity analysis examined the influence of emergency presentation on study results. In the primary analysis, emergency presentation was adjusted for as a component of the propensity score.
Intervention	EVAR

Full citation	**Jackson RS, Chang DC, and Freischlag JA. (2012) Comparison of long-term survival after open vs endovascular repair of intact abdominal aortic aneurysm among Medicare beneficiaries. JAMA. 307(15):1621-8**
Comparator	OSR
Outcomes	Mortality
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Low risk – all cohorts were drawn from the same time period. 1.2. Were cohorts from the same place? Low risk – both cohorts were derived from the same national healthcare provider database. 1.3. Is the definition of AAA the same across cohorts? High risk –all infrarenal and complex AAA included, which is likely to have led to implicitly different inclusion criteria across treatment arms. I In addition, a higher proportion of participants in the OSR group (23.3%) were admitted under emergency conditions compared to the EVAR (14.7%), possibly due to symptomatic aneurysms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – the study controlled for demographic variables including age and gender. 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – the study controlled for a broad set of comorbidities. 2.3. Does study control appropriately for AAA characteristics? High risk – there is no indication that AAA characteristics were controlled for. 2.4. Could any adjustment variables have been affected by the intervention? Low risk - no post-intervention variables which could mediate the treatment effect were controlled for. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – participants were identified using administrative registries with high-level diagnosis and procedure codes. 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – outcomes were assessed using administrative registries with high-level codes. 3.3. Is method of data collection likely to record long-term outcomes accurately? Moderate risk – surgical registries were used with linkage to a social security administrative database. **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – none reported. 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? Low risk – casewise deletion was used to exclude patients with missing data. 4.3. Have different methods been compared within the study? Moderate risk - different methods were compared but all relied on the same assumption about selection **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? Moderate risk –box plots of propensity scores stratified by AAA repair type were examined, and median propensity score and interquartile range were compared between groups.

Full citation	**Jackson RS, Chang DC, and Freischlag JA. (2012) Comparison of long-term survival after open vs endovascular repair of intact abdominal aortic aneurysm among Medicare beneficiaries. JAMA. 307(15):1621-8**
	5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? N/A 6.2 Were interactions between treatment and other covariates considered? High risk – no interactions appear to have been considered. **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Jetty P, Hebert P, van Walraven C, et al. (2010) Long-term outcomes and resource utilization of endovascular versus open repair of abdominal aortic aneurysms in Ontario. J Vasc Surg. 51(3):577-83**
Study details	Study design: retrospective cohort study Location(s): Canada Study period: April 2002 to March 2007 Aim of the study: to compare population-based clinical outcomes and resource utilisation for EVAR and OSR of all elective AAA repairs in Ontario
Participants	Sample size: EVAR group, n=888; OSR group, n= 5,573 Inclusion criteria: people who underwent EVAR or OSR of unruptured infrarenal AAA in the province of Ontario were included. Exclusion criteria: people with ruptured aneurysms, thoracic aneurysms, isolated iliac aneurysms, and pseudoaneurysms were excluded. Non-Ontarians who underwent surgical repair in an Ontario hospital were excluded because their follow-up period is not captured by the Ontario health databases used for the study. Baseline characteristics: Median age (range): EVAR group, 76 (70-81) years; OSR group, 72 (66-77) years Gender: EVAR group, 86.2% male; OSR group, 80.3% male Mean aneurysm diameter: not reported Peripheral vascular disease: EVAR group, 49.1%; OSR group, 32.3% Cerebrovascular disease: EVAR group, 13.7%; OSR group, 9.9% Congestive heart failure: EVAR group, 42.3%; OSR group, 28.6% Pulmonary disease: EVAR group, 33.6%; OSR group, 21.7% Diabetes: EVAR group, 23.8%; OSR group, 18.0% Hypertension: EVAR group, 72.9%; OSR group, 60.9%

Full citation	Jetty P, Hebert P, van Walraven C, et al. (2010) Long-term outcomes and resource utilization of endovascular versus open repair of abdominal aortic aneurysms in Ontario. J Vasc Surg. 51(3):577-83
	Cancer: EVAR group, 17.7%; OSR group, 12.7%
Methods	Data collection: All study data were obtained from several population based administrative databases. The primary database was the Canadian Institute for Health Information-Discharge Abstract Database (CIHI-DAD) which contains ICD10 diagnosis codes and Canadian Classification of Health Interventions procedure codes. Other databases used for the study included a diabetes database, the Registered Persons Database (which captures the death dates), a myocardial infarction database, a health insurance plan database, and a prescriptions database. These databases were linked by a patient identifier number Analysis: Investigators performed regression based on propensity scores. Propensity scores were calculated using multivariable logistic regression with preoperative patient factors outlined in the Charlson Comorbidity Index as the predictor variables and the assignment to EVAR as the outcome variable. Clinically important terms were included in the model that generated the probability that a patient would receive either treatment. Patients were then classified into quintiles by their propensity score. This allowed investigators to evaluate the quality of the model by confirming balance. Finally, Cox proportional hazard modelling was used for the adjusted survival analysis using the propensity quintiles as stratifying variables.
Intervention	EVAR
Comparator	OSR
Outcomes	30-day mortality, mortality at 5 years, length of stay, length of stay in ICU, discharge to long-term care
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Low risk – cohorts were drawn from the same time period. 1.2. Were cohorts from the same place? Low risk – all participants underwent AAA repair in hospitals from the Ontario region. 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included. This is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – study controlled for variables in the Charlson Comorbidity Index, including age and gender. 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – a broad range of comorbidities were controlled for. 2.3. Does study control appropriately for AAA characteristics? High risk – study did not control for AAA characteristics. 2.4. Could any adjustment variables have been affected by the intervention? Low risk – no post-intervention variables which could mediate the treatment effect were controlled for. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – administrative registries with high-level diagnosis and procedure codes were used.

Full citation	Jetty P, Hebert P, van Walraven C, et al. (2010) Long-term outcomes and resource utilization of endovascular versus open repair of abdominal aortic aneurysms in Ontario. J Vasc Surg. 51(3):577-83
	3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – administrative registries with high-level diagnosis and procedure codes were used. 3.3. Is method of data collection likely to record long-term outcomes accurately? Moderate risk – data were collected via linkage to routine registries. **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – none reported. 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – authors do not discuss what approach was taken for missing data. 4.3. Have different methods been compared within the study? Moderate risk – different methods compared but all rely on the same assumption about selection **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? High risk – no checks reported 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? N/A 6.2 Were interactions between treatment and other covariates considered? High risk – no interactions were explored. **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Johnson ML, Bush RL, Collins TC, et al. (2006) Propensity score analysis in observational studies: outcomes after abdominal aortic aneurysm repair. Am J Surg. 192(3):336-43.**
Study details	Study design: retrospective cohort study Location(s): USA Study period: May 2001 and September 2003 Aim of the study: to compare outcomes of EVAR and OSR of unruptured AAA.
Participants	Sample size of unmatched cohort: EVAR group, n=717; OSR group, n=1,187 Sample size of matched cohort: EVAR group, n=670; OSR group, n=670 Inclusion criteria: all people wo underwent EVAR or OSR of unruptured AAA were included. Exclusion criteria: people with ruptured AAA, thoracic or thoracoabdominal aortic aneurysm, or those who underwent conversion from EVAR to OSR were excluded. Baseline characteristics: Mean age (SD): EVAR group, 71.6 (7.8) years; OSR group,70.2 (7.9) years Gender: EVAR group, 99.6% male; OSR group, 99.1% male Mean aneurysm diameter: not reported COPD: EVAR group, 26.6%; OSR group, 26.0% Chronic heart failure: EVAR group, 3.6%; OSR group, 2.4% Renal insufficiency: EVAR group, 1.1%; OSR group, 1.0% Cerebrovascular accident with neuro-deficit: EVAR group, 5.3%; OSR group, 6.0% Diabetes: EVAR group, 14.4%; OSR group, 13.2% Cancer: EVAR group, 1.1%; OSR group, 1.0%
Methods	Data collection: data were extracted from a detailed surgical registry run by the run by the military Veterans Health Administration: (A Veterans Affairs component of the National Surgical Quality Improvement Program; NSQIP). The NSQIP database requires hospitals to provide complete 30-day follow-up on at least 95% of patients. To supplement the information in the NSQIP records investigators used unique identifiers to link records with other Veterans Affairs databases: including the patient treatment file (which contains abstracts of all patients discharged), the outpatient clinic file (which contains records for every outpatient visit), and the VA beneficiary identification record locator system death file Analysis: Propensity score matching was used with no additional analyses performed. The propensity score (predicted probability of receiving EVAR) was obtained by performing a multivariable logistic regression of 32 independent variables. One-to-one matched samples were created and matched pairs were subsequently categorised into 5 groups or strata of increasing score. Patients with the lowest propensity score were most likely to receive OSR based on their baseline risk factors, whereas patients with higher propensity scores were more likely to receive EVAR. The clinical outcomes were subsequently compared between strata.

Full citation	**Johnson ML, Bush RL, Collins TC, et al. (2006) Propensity score analysis in observational studies: outcomes after abdominal aortic aneurysm repair. Am J Surg. 192(3):336-43.**
Intervention	EVAR
Comparator	OSR
Outcomes	30-day mortality, 1 year mortality and postoperative adverse events (included adverse cardiac events, renal dysfunction, pulmonary complications, wound complications, neurologic complications, bleeding requiring blood transfusion, and graft failure)
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Low risk – both cohorts were drawn from the same time period. 1.2. Were cohorts from the same place? Low risk – both cohorts were drawn from the same time period. 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included. This is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk - patient demographic data including age, gender and age were controlled for in the analyses. 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – comorbidities, including those conditions known to have an influence on the risk of cardiovascular morbidity and mortality, were chosen for analysis. 2.3. Does study control appropriately for AAA characteristics? High risk – there is no indication that the study controlled for AAA characteristics. 2.4. Could any adjustment variables have been affected by the intervention? Low risk – no post-intervention variables which could mediate the treatment effect were controlled for. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? Moderate risk – a detailed surgical registry was used to identify participants with diagnosis and procedure codes specified. 3.2. Is method of data collection likely to record perioperative outcomes accurately? Low – a detailed surgical registry was used to collect data on outcomes. 3.3. Is method of data collection likely to record long-term outcomes accurately? Moderate – a surgical registry was used with linkage to routine data registries. **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – no checks were reported. 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? Low risk – authors state that given the robust nature of the NSQIP and other databases used the likelihood of missing essential covariates is low. 4.3. Have different methods been compared within the study? High risk – different methods were not compared. **Analysis – matching**

Full citation	Johnson ML, Bush RL, Collins TC, et al. (2006) Propensity score analysis in observational studies: outcomes after abdominal aortic aneurysm repair. Am J Surg. 192(3):336-43.
	5.1. Is the matching algorithm reported and reasonable? High risk – matching algorithm was discussed/reported. 5.2. Was overlap / common support appropriately assessed? High risk – insufficient checks were performed. 5.3. Has balancing of the covariates been demonstrated? Moderate risk – “baseline factors that were statistically significantly associated with type of surgery in unadjusted bivariate analyses were re-examined after propensity score stratification to confirm that baseline differences had been removed.” **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? N/A 6.2 Were interactions between treatment and other covariates considered? N/A **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	Karthikesalingam A, Holt PJ, Vidal-Diez A, et al. (2016) The impact of endovascular aneurysm repair on mortality for elective abdominal aortic aneurysm repair in England and the United States. J Vasc Surg. 64(2):321-327	
Study details	Study design: retrospective cohort study Location(s): UK and USA Study period: January 2005 to December 2010 Aim of the study: to compare in-hospital mortality between people who underwent EVAR and those who OSR for unruptured AAA repair, and subsequently compare outcomes between people in England and the United States.	
Participants	**<u>UK cohort</u>** Sample size: EVAR group, n=7,937; OSR group, n= 13,335 Inclusion criteria: all patients undergoing elective AAA repair. Exclusion criteria: ruptured AAA. Baseline characteristics: Median age (IQR): 74 (69-79) years Gender: 86.6% male Mean aneurysm diameter: not reported Diabetes: 11.3% Myocardial infarction: 5.2% Cerebrovascular disease: 2.0%	**<u>USA cohort</u>** Sample size: EVAR group, n=126,211; OSR group, n= 69,902 Inclusion criteria: all patients undergoing elective AAA repair. Exclusion criteria: ruptured AAA. Baseline characteristics: Median age (IQR): 72.7 (66.9-78.2) years Gender: 76.4% male Mean aneurysm diameter: not reported Diabetes: 17.8% Myocardial infarction: 14.7% Cerebrovascular disease: 5.7%

Full citation	**Karthikesalingam A, Holt PJ, Vidal-Diez A, et al. (2016) The impact of endovascular aneurysm repair on mortality for elective abdominal aortic aneurysm repair in England and the United States. J Vasc Surg. 64(2):321-327**	
	COPD: 15.8% Renal disease: 7.5% Cancer: 0.3%	COPD: 34.9% Renal disease: 10.0% Cancer: 0.4%
Methods	Data collection: Demographic and in-hospital outcome data were extracted from the UK Hospital Episode Statistics (HES) and the USA Nationwide Inpatient Sample (NIS) for all patients undergoing elective AAA repair. HES data captures the majority of people undergoing AAA repair in the UK whereas, NIS data is an anonymised, stratified sample of 20% of all discharges from U.S. hospitals. Participants were identified by examining diagnosis and procedure codes; such as ICD9 and OPCS4 codes. Analysis: logistic regression was performed adjusting for age, gender, social deprivation, comorbidity index scores hospital procedural volume (caseload), hospital bed capacity and teaching status. Backward selection procedures were used with comparison of models by the likelihood ratio test to ascertain whether individual covariates improved goodness-of-fit for prediction of in-hospital mortality. Note: Age- and gender-matched analyses were constructed to compare English and U.S. outcomes for in-hospital; however this was not considered relevant for this review.	
Intervention	EVAR	
Comparator	OSR	
Outcomes	In-hospital mortality	
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Low risk – recruitment over ≥5 yrs, but year of operation controlled for in analysis. 1.2. Were cohorts from the same place? Low risk – EVAR and OSR cohorts were compared in the context of the country in which they were performed. 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included. This is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – demographic variables including age and gender were controlled for. 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – investigators adjusted for comorbidities by considering Charlson scores (a comorbidity index). 2.3. Does study control appropriately for AAA characteristics? High risk – the study did not control for AAA characteristics. 2.4. Could any adjustment variables have been affected by the intervention? Low risk – no post-intervention variables which could mediate the treatment effect were controlled for. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – administrative registries with high-level diagnosis and procedure codes were used.	

Full citation	Karthikesalingam A, Holt PJ, Vidal-Diez A, et al. (2016) The impact of endovascular aneurysm repair on mortality for elective abdominal aortic aneurysm repair in England and the United States. J Vasc Surg. 64(2):321-327
	3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – administrative registries with high-level diagnosis and procedure codes were used. 3.3. Is method of data collection likely to record long-term outcomes accurately? Low risk – no long-term outcomes assessed. **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – no checks reported. 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – no demonstration that missing data was considered. 4.3. Have different methods been compared within the study? High risk – different methods were not compared. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? Low risk - number of events is ≥10 times greater than number of variables considered. 6.2 Were interactions between treatment and other covariates considered? High risk – no interactions were considered. **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	Lo RC, Bensley RP, Hamdan AD, et al. (2013) Gender differences in abdominal aortic aneurysm presentation, repair, and mortality in the Vascular Study Group of New England. J Vasc Surg. 57:1261-8
Study details	Study design: retrospective cohort study Location(s): USA Study period: 2003 to 2011 Aim of the study: to describe differences in the presentation, choice of repair, and mortality among men and women undergoing AAA repair
Participants	Sample size: EVAR group, n=2,159; OSR group, n=1,867 Inclusion criteria: Patient presentation was categorized as intact (including patients who were symptomatic as well as those undergoing elective repair) or ruptured. Four subgroup analyses were performed: intact EVAR, ruptured EVAR, intact open repair, and ruptured open repair. For the purposes of multivariable modelling, intact EVAR served as the referent group. Exclusion criteria: None mentioned Baseline characteristics: Mean age for men (range): EVAR group, 74 (67-80) years; OSR group, 71 (64-77) years. Mean age for women (range): EVAR group, 77 (71-81) years; OSR group, 73 (68-78) years Gender: EVAR group, 80.0% male; OSR group, 74.6% male Mean aneurysm diameter: not reported Diabetes mellitus: Males: EVAR group, 1.16%; OSR group, 1.08%. Females: EVAR group, 4.42%; OSR group, 2.74% Coronary artery disease: Males: EVAR group, 2.14%; OSR group, 2.51%. Females: EVAR group, 6.28%; OSR group, 5.7% Congestive heart failure: Males: EVAR group, 0.64%; OSR group, 0.53%. Females: EVAR group, 2.79%; OSR group, 1.43% COPD: Males: EVAR group, 1.97%; OSR group, 2.37%. Females: EVAR group, 9.53%; OSR group, 9.49%
Methods	Data collection: retrospective review of open and endovascular AAA repairs in the Vascular Study Group of New England (VSGNE) database, a voluntary collaboration among vascular surgeons, cardiologists, and radiologists from 30 academic and community hospitals in the six New England states. Formed as a quality improvement initiative, VSGNE represents a pool of clinical data related to several frequently performed vascular procedures that have been collected since 2003. Analysis: Multivariable logistic regression was used to determine predictors of 30-day mortality. Individual survival curves for each presentation and treatment group were evaluated for differences in survival between men and women. Cox proportional hazards modelling was used to determine predictors of long-term mortality. Multivariable predictors of 30-day mortality were age, coronary artery disease, preoperative dialysis dependence, and presentation/ treatment subgroup (intact vs ruptured and EVAR vs open repair). Age >80 years increased the odds of 30-day mortality by 10.9 and rupture increased the odds by 48.4 for EVAR and 83.8 for open repair. Female gender did not reach statistical significance as a predictor of 30-day mortality.
Intervention	EVAR
Comparator	OSR

Full citation	**Lo RC, Bensley RP, Hamdan AD, et al. (2013) Gender differences in abdominal aortic aneurysm presentation, repair, and mortality in the Vascular Study Group of New England. J Vasc Surg. 57:1261-8**
Outcomes	30-day mortality
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Moderate risk - ≥5-yr recruitment with no adjustment for year of operation 1.2. Were cohorts from the same place? Low risk 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included. This is likely to have led to implicitly different inclusion criteria across treatment arms.. **Confounding** 2.1. Does study control appropriately for demographics? Low risk 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk 2.3. Does study control appropriately for AAA characteristics? High risk - none 2.4. Could any adjustment variables have been affected by the intervention? Low risk **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? Moderate risk – detailed surgical registry 3.2. Is method of data collection likely to record perioperative outcomes accurately? Low risk 3.3. Is method of data collection likely to record long-term outcomes accurately? N/A **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk - none reported 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? Low risk 4.3. Have different methods been compared within the study? High risk – none reported **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? High risk 6.2 Were interactions between treatment and other covariates considered? High risk **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Locham S, Rizwan M, Dakour-Aridi H et al. (2018) Outcomes after elective abdominal aortic aneurysm repair in obese versus nonobese patients. J Vasc Surg. S0741-5214(18)30907-8**
Study details	Study design: retrospective cohort study Location(s): USA Study period: 2003 to 2017 Aim of the study: to use a nationally representative vascular database to compare in-hospital outcomes in obese versus non-obese patients undergoing elective EVAR or OSR.
Participants	Sample size: EVAR group, n=26,723; OSR group, n=6,359 Inclusion criteria: people who underwent elective OSR and EVAR were included. Note: people with complex aneurysm anatomies are likely to have been included. Exclusion criteria: people who underwent urgent or emergent AAA repair were excluded from the study. Baseline characteristics: Mean age: treatment specific ages not reported Gender: EVAR group, 81.4% male; OSR group, 73.8% male Mean aneurysm diameter: treatment specific ages not reported Diabetes: EVAR group, 20.5%; OSR group, 16.3% Hypertension: EVAR group, 83.3%; OSR group, 84.2% Coronary artery disease: EVAR group, 29.8%; OSR group, 27.0% Coronary heart failure: EVAR group, 11.5%; OSR group, 7.5% COPD: EVAR group, 32.5%; OSR group, 32.7% Dialysis: EVAR group, 1.0%; OSR group, 0.6% Prior bypass: EVAR group, 3.0%; OSR group, 4.4%
Methods	Data collection: investigators queried a large national database that contained detailed preoperative, operative, and postoperative characteristics of patients on several common vascular surgical procedures (the Vascular Quality Initiative: VQI) to identify people who underwent elective AAA repair and obtain information on their outcomes. Analysis: Multivariate logistic regression was performed. Covariates were chosen on the basis of clinical and statistical significance in univariate analysis (significance level not specified). The interaction between surgery and obesity was also evaluated. All models were tested using variation inflation factor, Hosmer-Lemeshow goodness of fit test, and area under the receiver operating characteristic curve (C-statistic).
Intervention	EVAR
Comparator	OSR
Outcomes	In-hospital mortality and adverse events

Full citation	Locham S, Rizwan M, Dakour-Aridi H et al. (2018) Outcomes after elective abdominal aortic aneurysm repair in obese versus nonobese patients. J Vasc Surg. S0741-5214(18)30907-8
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Moderate risk – ≥5-yr recruitment with no adjustment for year of operation. 1.2. Were cohorts from the same place? Low risk – all participants were selected from the same national vascular registry. 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included. This is likely to have led to implicitly different inclusion criteria across treatment arms.. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – study adjusted for demographics variables, including age and gender. 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – a broad range of demographic variables were controlled for. 2.3. Does study control appropriately for AAA characteristics? Moderate risk – the study controlled for aneurysm diameter. 2.4. Could any adjustment variables have been affected by the intervention? High risk – investigators controlled for blood loss, which could mediate the treatment effect. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? Moderate risk – a detailed vascular registry was used to identify participants. 3.2. Is method of data collection likely to record perioperative outcomes accurately? Low risk – outcomes were obtained from a detailed vascular registry. 3.3. Is method of data collection likely to record long-term outcomes accurately? Low risk – no long-term outcomes were assessed. **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? Low risk – model specification/fit was checked using the Hosmer–Lemeshow test as well as the C-statistic 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? Low risk – authors reported that the effect of missing data was minimal as the VQI performs thorough checks to maintain standards. All study hospitals are required to enter complete information if they are found to have a large number of missing data. 4.3. Have different methods been compared within the study? High risk – different methods were not compared. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? High risk – number of events is <10 times greater than number of variables considered.

Full citation	Locham S, Rizwan M, Dakour-Aridi H et al. (2018) Outcomes after elective abdominal aortic aneurysm repair in obese versus nonobese patients. J Vasc Surg. S0741-5214(18)30907-8
	6.2 Were interactions between treatment and other covariates considered? Low risk – the following interaction was considered OSR*obese **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	Mark TL, Lawrence W, Coffey RM, et al. (2013) The value of linking hospital discharge and mortality data for comparative effectiveness research. Journal of comparative effectiveness research 2(2), 175-84
Study details	Study design: retrospective cohort study Location(s): USA Study period: July 2000 to January 2006 Aim of the study: to demonstrate the value of linking state community hospital discharge data to vital statistics death files for research by conducting a comparative effectiveness analysis of elective EVAR versus OSR.
Participants	Sample size of unmatched cohort: EVAR group, n=6,046; OSR group, n=7,606 Sample size of matched cohort: EVAR group, n=4,483; OSR group, n=4,483 Inclusion criteria: people who underwent elective EVAR or OSR for unruptured AAA were included. Note: there is no indication that selection was limited to only infrarenal AAAs. Exclusion criteria: people who underwent emergency AAA repair, both EVAR and OSR during the same admission, revision of a previous AAA repair, and those with less than 1 year of follow-up data available were excluded. Furthermore, people with the clinical codes indicating the following conditions were excluded: thoracoabdominal aneurysms, and/or visceral or renal bypass, polyarteritis nodosa, coarctation of the aorta, Marfan syndrome and other congenital anomalies. Finally, individuals who had multiple death records, a death date prior to an admission date, or multiple codes for sex were excluded. Baseline characteristics (of matched cohort): % <65: EVAR group, 22%; OSR group, 24% Gender: EVAR group, 8% male; OSR group, 6% male Mean aneurysm diameter: not reported COPD: EVAR group, 29%; OSR group,31% Diabetes with or without complications: EVAR group, 15%; OSR group, 14% Renal failure: EVAR group, 5%; OSR group, 5% Lymphoma: EVAR group, 1%; OSR group, 1% Metastatic cancer: EVAR group, 1%; OSR group, 0%

Full citation	**Mark TL, Lawrence W, Coffey RM, et al. (2013) The value of linking hospital discharge and mortality data for comparative effectiveness research. Journal of comparative effectiveness research 2(2), 175-84**
	Solid tumour without metastasis: EVAR group, 7%; OSR group, 7% Ischaemic heart disease: EVAR group, 43%; OSR group, 43% Cerebral vascular occlusive disease: EVAR group, 5%; OSR group, 5%
Methods	Data collection: investigators identified participants and obtained data on their outcomes by linking (using social security numbers) 2 administrative databases managed by the California Office of State-wide Health Planning and Development. Analysis: Propensity score matching was performed. Logistic regression was performed to derive propensity scores. The predicted propensity score was then used to create a 1:1 match between cohorts using greedy matching with a caliper of 0.25 standard deviation of the propensity score. Balance on the covariates was assessed by computing standardized differences for each covariate; 'balanced' was defined as <10 standardised differences. Supplementary, analyses using simple multivariable models were also performed.
Intervention	EVAR
Comparator	OSR
Outcomes	In-hospital, 30-day, 1-year and 5-year mortality
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Low risk – recruitment over ≥5 yrs, but year of operation controlled for in analysis 1.2. Were cohorts from the same place? Low risk – all participants were selected from the same regional database. 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included. This is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? High risk – no details were provided. 2.2. Does study control appropriately for comorbidity and/or fitness? High risk – no details were provided. 2.3. Does study control appropriately for AAA characteristics? High risk – authors stated that information about aneurysm size or anatomical features was not available. 2.4. Could any adjustment variables have been affected by the intervention? High risk – no details were provided. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – participants were identified using an administrative database with high-level procedure codes. 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – outcomes were assessed using an administrative database with high-level procedure codes. 3.3. Is method of data collection likely to record long-term outcomes accurately? Moderate risk – an administrative patient discharge database was used with linkage to a reliable routine registry.

Full citation	**Mark TL, Lawrence W, Coffey RM, et al. (2013) The value of linking hospital discharge and mortality data for comparative effectiveness research. Journal of comparative effectiveness research 2(2), 175-84**
	Analysis – general 4.1. Were any checks conducted on model specification and/or fit? High risk – no checks were reported 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – authors stated that approximately 20-30% of patient discharge records did not have a valid social security number and therefore were linked between databases. 4.3. Have different methods been compared within the study? High risk – no different methods were compared. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? Low risk – matching was performed using greedy matching with a caliper of 0.25 standard deviation of the propensity score. 5.2. Was overlap / common support appropriately assessed? High risk – no assessment reported. 5.3. Has balancing of the covariates been demonstrated? Low risk – conventional hypothesis tests were performed. Furthermore, standardised differences in covariates were also reported and considered by investigators. **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? N/A 6.2 Were interactions between treatment and other covariates considered? N/A **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Quintana MJ, Gich I, Librero J, Bellmunt-Montoya S, Escudero JR, Bonfill X. Variation in the choice of elective surgical procedure for abdominal aortic aneurysm in spain. Vascular Health and Risk Management. 2019;15:69.**
Study details	Study design: retrospective cohort study Location(s): Spain Study period: 2002 to 2012 Aim of the study: to identify current preferences among Spanish hospitals for OSR or EVAR and to determine changes in these preferences over the course of the study period
Participants	Sample size: EVAR group, n=4,010; OSR group, n=4,010 Inclusion criteria: primary diagnosis of unruptured AAA, elective admission, treated with OSR or EVAR Exclusion criteria: thoracic or thoracoabdominal AA, aortic dissection, patients who underwent both EVAR and OSR during a single intervention, ruptured AAA and emergency admission Baseline characteristics (whole cohort):

Full citation	**Quintana MJ, Gich I, Librero J, Bellmunt-Montoya S, Escudero JR, Bonfill X. Variation in the choice of elective surgical procedure for abdominal aortic aneurysm in spain. Vascular Health and Risk Management. 2019;15:69.**
	Mean age 71.4 Gender: 96.7% male Mean aneurysm diameter: not reported Mean Charlson index: 0.7
Methods	Data collection: data in the database “Minimum Basic Dataset at Hospital Discharge” (in Spanish: “Conjunto Mínimo Básico de Datos al Alta Hospitalaria” [CMBDAH]). This database contains data on patients discharged from Spanish public hospitals. ICD9 codes used reported. Analysis:. Generalised linear mixed models (multivariate logistic regressions with hospital random effect) to examine interhospital variation. Three multivariate models that included temporal and hospital effects were developed. The first model included adjusted variables; the second model also considered a potential time trend (year of surgery); and the third model added surgical volume (ie, number of procedures) performed at each hospital. The results were adjusted for individual factors (age, gender, and comorbidities) to assess the influence of surgical volumes.
Intervention	EVAR
Comparator	OSR
Outcomes	In-hospital mortality
Study Appraisal using NICE’s bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Moderate risk – ≥5-yr recruitment with no adjustment for year of operation 1.2. Were cohorts from the same place? Low risk – all participants were selected from the same national surgical registry 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included, which is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – cohort controls for age and sex. 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – Charlson index. 2.3. Does study control appropriately for AAA characteristics? High risk – AAA characteristics were not controlled for. 2.4. Could any adjustment variables have been affected by the intervention? Low risk – no post-intervention variables controlled for that may mediate treatment effect. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – administrative registries with high-level diagnosis and procedure codes. 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – administrative registries with high-level diagnosis and procedure codes.

Full citation	**Quintana MJ, Gich I, Librero J, Bellmunt-Montoya S, Escudero JR, Bonfill X. Variation in the choice of elective surgical procedure for abdominal aortic aneurysm in spain. Vascular Health and Risk Management. 2019;15:69.**
	3.3. Is method of data collection likely to record long-term outcomes accurately? N/A – no long-term outcomes were assessed. **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? Low risk – Hosmer–Lemeshow test 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – no details. 4.3. Have different methods been compared within the study? High risk – no different methods were compared. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? Low risk – 601 events, 6 covariates. 6.2 Were interactions between treatment and other covariates considered? High risk – no interactions were considered **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Raval MV, and Eskandari MK. (2012) Outcomes of elective abdominal aortic aneurysm repair among the elderly: endovascular versus open repair. Surgery. 151(2):245-60.**
Study details	Study design: retrospective cohort study Location(s): USA Study period: January 2005 to December 2008 Aim of the study: to analyse outcomes of endovascular EVAR and OSR of elective AAA
Participants	Sample size: EVAR group, n=2,350; OSR group, n=5,586 Inclusion criteria: people who underwent EVAR or OSR for unruptured AAA were included. Note: people with complex aneurysms may also have been included. Exclusion criteria: people with ruptured AAA, wound infection, pneumonia, ventilator dependence/reintubation, renal failure, stroke, and coma prior to AAA repair were excluded. Baseline characteristics (whole cohort): Mean age not reported

Full citation	Raval MV, and Eskandari MK. (2012) Outcomes of elective abdominal aortic aneurysm repair among the elderly: endovascular versus open repair. Surgery. 151(2):245-60.
	Gender: 81.1% male Mean aneurysm diameter: not reported Diabetes: 13.9% COPD: 18.3% Congestive heart failure: 1.0% Hypertension requiring treatment: 80.2% Neurologic disease/event: 15.5% Cancer: 0.8%
Methods	Data collection: investigators obtained data by querying the American College of Surgeons version of the National Surgical Quality Improvement Program (ACS-NSQIP) database using procedure codes. Mortality was determined by examination of medical records, contacting patients, and querying social Security Death Index and the National Obituary Archives. Analysis: multivariate logistic regression was performed. Predictor variables were entered into the regression models by forward stepwise selection except operative approach (which was "forced" into the models). Final models were assessed for ''goodness-of-fit'' using Hosmer and Lemeshow tests and model discrimination was evaluated using C-index.
Intervention	EVAR
Comparator	OSR
Outcomes	People of any age (whole cohort): 30-day mortality and adverse events within 30 days (wound, pulmonary, cardiac, renal and infectious complications) People > 80 years: 30-day mortality, length of stay > 7 days and adverse events within 30 days (wound, pulmonary, cardiac, renal and infectious complications)
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.4. Were cohorts from the same time period? Low risk – cohorts were drawn from the same time period. 1.5. Were cohorts from the same place? Low risk – all participants were selected from the same national surgical registry 1.6. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included, which is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – cohort controls for age and sex. 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – a broad range of relevant comorbidities were controlled for. 2.3. Does study control appropriately for AAA characteristics? High risk – AAA characteristics were not controlled for. 2.4. Could any adjustment variables have been affected by the intervention? High risk –analysis controlled for duration of operation.

Full citation	**Raval MV, and Eskandari MK. (2012) Outcomes of elective abdominal aortic aneurysm repair among the elderly: endovascular versus open repair. Surgery. 151(2):245-60.**
	Data collection 3.1. Is method of data collection likely to have identified suitable participants accurately? Moderate risk – participants were identified using detailed surgical registry with diagnosis and procedure codes specified. 3.2. Is method of data collection likely to record perioperative outcomes accurately? Low risk – detailed surgical registries were used to assess outcomes. 3.3. Is method of data collection likely to record long-term outcomes accurately? Low risk – no long-term outcomes were assessed. **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? Low risk – model specification/fit was assessed using the C-statistic and the Hosmer–Lemeshow test 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? Low risk - variables with considerable amounts of missing data were accounted for in the analyses. 4.3. Have different methods been compared within the study? High risk – no different methods were compared. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? High risk – number of events is <10 times greater than number of variables considered. 6.2 Were interactions between treatment and other covariates considered? High risk – no interactions were considered **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Salata K, Hussain MA, de Mestral C, Greco E, Aljabri BA, Mamdani M, Forbes TL, Bhatt DL, Verma S, Al-Omran M. Comparison of Outcomes in Elective Endovascular Aortic Repair vs Open Surgical Repair of Abdominal Aortic Aneurysms. JAMA network open. 2019 Jul 3;2(7):e196578.**
Study details	Study design: retrospective propensity-matched cohort study Location(s): Canada Study period: April 2003 to March 2016

Full citation	**Salata K, Hussain MA, de Mestral C, Greco E, Aljabri BA, Mamdani M, Forbes TL, Bhatt DL, Verma S, Al-Omran M. Comparison of Outcomes in Elective Endovascular Aortic Repair vs Open Surgical Repair of Abdominal Aortic Aneurysms. JAMA network open. 2019 Jul 3;2(7):e196578.**
	Aim of the study: to assess the differences between EVAR and OSR for elective AAA repair in long-term survival, major adverse cardiovascular event (MACE)–free survival, reintervention, and secondary rupture
Participants	Sample size: EVAR group, n=4,010; OSR group, n= 4,010 Inclusion criteria: all elective EVARs and OSRs of AAA performed in Ontario, Canada, in patients 40 years and older Exclusion criteria: Patients with multiple AAA repair procedures listed on their index admission Baseline characteristics Mean age: EVAR group, 73.1 years; OSR group, 72.8 years Gender: EVAR group, 82.5% male; OSR group, 81.6% male Mean aneurysm diameter: not reported Coronary artery disease: EVAR group, 16.8%; OSR group, 16.7% MI: EVAR group, 5.6%; OSR group, 5.5% Congestive heart failure: EVAR group, 13.1%; OSR group, 12.4% Hypertension: EVAR group, 80.9%; OSR group, 80.5% Diabetes: EVAR group, 29.2%; OSR group, 28.6% COPD: EVAR group, 40.2%; OSR group, 39.3% Chronic kidney disease: EVAR group, 2.2%; OSR group, 1.8%
Methods	Data collection: Linked data across 11 administrative health databases Analysis: The propensity score for repair approach was calculated using a logistic regression model incorporating all covariates as potential confounders. Patients who received EVAR or OSR were matched 1:1 using the greedy nearest-neighbour method with a calliper width of 0.2 SD units. Balance of covariates was assessed using standardized differences, with differences less than 0.1 indicating good balance. Residual confounding was assessed using the distribution of tracer variables not used to specify the propensity score.
Intervention	EVAR
Comparator	OSR
Outcomes	30-day mortality, 30-day cardiovascular events, long-term survival, reinterventions
Study Appraisal using NICE's bespoke risk of bias	**Selection** 1.1. Were cohorts from the same time period? Low risk – the recruitment was >5 years. However, the year of recruitment was controlled for 1.2. Were cohorts from the same place? Low risk – same country 1.3. Is the definition of AAA the same across cohorts? High risk – authors note that the algorithm used to identify cases is validated to find infrarenal EVAR and infrarenal, pararenal, and juxtarenal OSR.

Full citation	**Salata K, Hussain MA, de Mestral C, Greco E, Aljabri BA, Mamdani M, Forbes TL, Bhatt DL, Verma S, Al-Omran M. Comparison of Outcomes in Elective Endovascular Aortic Repair vs Open Surgical Repair of Abdominal Aortic Aneurysms. JAMA network open. 2019 Jul 3;2(7):e196578.**
assessment tool	**Confounding** 2.1. Does study control appropriately for demographics? Low risk 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – broad range of individual comorbidities and Charlson score 2.3. Does study control appropriately for AAA characteristics? High risk – none 2.4. Could any adjustment variables have been affected by the intervention? Low risk **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – administrative registry with high-level diagnosis and procedure codes 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk –administrative registry with high-level diagnosis and procedure codes 3.3. Is method of data collection likely to record long-term outcomes accurately? Low risk – direct use of reliable routine data registry **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – none reported 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? Low risk 4.3. Have different methods been compared within the study? High risk **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? Low risk 5.2. Was overlap / common support appropriately assessed? High risk – no checks reported 5.3. Has balancing of the covariates been demonstrated? Low risk – standardised differences (all ≤ 0.05) **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? N/A 6.2 Were interactions between treatment and other covariates considered? N/A **Overall risk of bias:** Moderate risk **Directness:** directly applicable

Full citation	**Schermerhorn ML, Buck DB, O'Malley. et al. (2015) Long-term outcomes of abdominal aortic aneurysm in the Medicare population. NEJM. 373:328-338** **Schermerhorn ML, Giles KA, Sachs T, et al. (2011) Defining perioperative mortality after open and endovascular Aortic Aneurysm Repair in the US Medicare Population. J Am Coll Surg. 212(3):349-355** **Schermerhorn ML, O'Malley AJ, Jhaveri A, et al. (2008) Endovascular vs. open repair of abdominal aortic aneurysms in the Medicare population. NEJM. 359:464-474**
Study details	Study design: retrospective propensity-matched cohort study Location(s): USA Study period: January 2001 to December 2008 Aim of the study: to compare endovascular repair with open repair with respect to the long-term outcomes of each procedure
Participants	Sample size: EVAR group, n=39,966; OSR group, n= 39,966 Inclusion criteria: patients were included in the study if they had been continuously enrolled in traditional Medicare Parts A and B for at least 2 years before the repair, had received a discharge diagnosis of abdominal aortic aneurysm, and had undergone open repair or endovascular repair. Exclusion criteria: ruptured abdominal aortic aneurysms, thoracic aneurysms, thoracoabdominal aortic aneurysms, or aortic dissections. In addition, they excluded those who had undergone visceral bypass or renal bypass Baseline characteristics Mean age: EVAR group, 75.7 years; OSR group, 75.5 years Gender: EVAR group, 77.7% male; OSR group, 77.6% male Mean aneurysm diameter: not reported MI in previous 6 months: EVAR group, 1.6%; OSR group, 1.6% MI in previous 7-24 months: EVAR group, 6.8%; OSR group, 6.7% Valvular heart disease: EVAR group, 8.7%; OSR group, 8.6% Congestive heart failure: EVAR group, 11.7%; OSR group, 11.6% Peripheral vascular disease: EVAR group, 19.8%; OSR group, 19.4% Neurovascular disease: EVAR group, 13.9%; OSR group, 13.9% Hypertension: EVAR group, 63.2%; OSR group, 62.9% Diabetes: EVAR group, 16.1%; OSR group, 15.9% COPD: EVAR group, 27.8%; OSR group, 27.8% Renal failure: EVAR group, 5.6%; OSR group, 5.5% End-stage renal disease: EVAR group, 0.4%; OSR group, 0.4% Obesity: EVAR group, 2.0%; OSR group, 2.0%

Full citation	**Schermerhorn ML, Buck DB, O'Malley. et al. (2015) Long-term outcomes of abdominal aortic aneurysm in the Medicare population. NEJM. 373:328-338** **Schermerhorn ML, Giles KA, Sachs T, et al. (2011) Defining perioperative mortality after open and endovascular** **Aortic Aneurysm Repair in the US Medicare Population. J Am Coll Surg. 212(3):349-355** **Schermerhorn ML, O'Malley AJ, Jhaveri A, et al. (2008) Endovascular vs. open repair of abdominal aortic aneurysms in the Medicare population. NEJM. 359:464-474**
Methods	Data collection: Medicare medical records Analysis: Propensity score matching
Intervention	EVAR
Comparator	OSR
Outcomes	30-day mortality, in-hospital mortality, long-term survival/mortality, length of stay, discharge to home, adverse events (including readmission)
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.4. Were cohorts from the same time period? Low risk – the recruitment was >5 years. However, the year of recruitment was controlled for 1.5. Were cohorts from the same place? Low risk – same country 1.6. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included, which is likely to have led to implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk 2.3. Does study control appropriately for AAA characteristics? High risk – none 2.4. Could any adjustment variables have been affected by the intervention? Low risk **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – administrative registry with high-level diagnosis and procedure codes 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk –administrative registry with high-level diagnosis and procedure codes 3.3. Is method of data collection likely to record long-term outcomes accurately? Low risk – direct use of reliable routine data registry **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – none reported 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? Low risk 4.3. Have different methods been compared within the study? Low risk **Analysis – matching**

Full citation	**Schermerhorn ML, Buck DB, O'Malley. et al. (2015) Long-term outcomes of abdominal aortic aneurysm in the Medicare population. NEJM. 373:328-338** **Schermerhorn ML, Giles KA, Sachs T, et al. (2011) Defining perioperative mortality after open and endovascular** **Aortic Aneurysm Repair in the US Medicare Population. J Am Coll Surg. 212(3):349-355** **Schermerhorn ML, O'Malley AJ, Jhaveri A, et al. (2008) Endovascular vs. open repair of abdominal aortic aneurysms in the Medicare population. NEJM. 359:464-474**
	5.1. Is the matching algorithm reported and reasonable? Low risk 5.2. Was overlap / common support appropriately assessed? High risk – no checks reported 5.3. Has balancing of the covariates been demonstrated? Moderate risk – conventional hypothesis tests, with no evidence of significant differences (except for age, where difference in mean appears minor – 75.7 -v- 75.5) **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? N/A 6.2 Were interactions between treatment and other covariates considered? N/A **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Schwarze ML, Shen Y, Hemmerich J, et al. (2009) Age-related trends in utilization and outcome of open and endovascular repair for abdominal aortic aneurysm in the United States, 2001-2006. J Vasc Surg. 50(4):722-729.**
Study details	Study design: retrospective cohort study Location(s): USA Study period: 2001 to 2006 Aim of the study: to compare utilisation and age-specific outcomes between EVAR and OSR for the treatment of AAA
Participants	Sample size: EVAR group, n=90,925; OSR group, n=75,222 Inclusion criteria: people who underwent EVAR or OSR of unruptured AAA were included. Note: there is no indication that selection was limited to only people with infrarenal AAA. Exclusion criteria: people <50 years, those with ruptured AAA, aortic dissection, thoracic or thoracoabdominal aortic aneurysms, coarctation of the aorta, Marfan syndrome and other congenital anomalies, gonadal dysgenesis-Turner syndrome, or polyarteritis nodosa were excluded. Baseline characteristics: Mean age: not reported Gender: not reported Mean aneurysm diameter: not reported

Full citation	**Schwarze ML, Shen Y, Hemmerich J, et al. (2009) Age-related trends in utilization and outcome of open and endovascular repair for abdominal aortic aneurysm in the United States, 2001-2006. J Vasc Surg. 50(4):722-729.**
	Comorbidities: not reported
Methods	Data collection: participants were identified using diagnosis and procedure codes (such as ICD9 and OPCS4 codes) to query an administrative database that used high-level diagnosis and procedure codes: the Nationwide Inpatient Sample (NIS). NIS data is an anonymised, stratified sample of 20% of hospitals including specialty, community and public hospitals, and academic medical centres. All data on comorbidities and outcomes were extracted from the NIS database. Analysis: multivariate linear regression was used to examine the risk-adjusted association between the type of procedure and length of stay. Multivariate logistic regression was performed to assess the risk-adjusted effect of the procedure performed on in-hospital mortality, discharge to home, and the occurrence of adverse events.
Intervention	EVAR
Comparator	OSR
Outcomes	In-hospital mortality, discharge to home, adverse events, and length of stay
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Moderate risk – ≥5-yr recruitment with no adjustment for year of operation. 1.2. Were cohorts from the same place? Low risk – all participants were selected from the same national administrative database 1.3. Is the definition of AAA the same across cohorts? High risk – authors highlight that 'only anatomically suitable infrarenal AAAs were treated by EVAR. **Confounding** 2.1. Does study control appropriately for demographics? Moderate risk – only gender was controlled for in the analyses. 2.2. Does study control appropriately for comorbidity and/or fitness? Moderate – authors stated that the models adjusted for comorbidities but no further details were provided. 2.3. Does study control appropriately for AAA characteristics? High risk – no AAA characteristics were controlled for. 2.4. Could any adjustment variables have been affected by the intervention? Low risk - no post-intervention variables which could mediate the treatment effect were controlled for. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – administrative registries with high-level diagnosis and procedure codes were used. 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – administrative registries with high-level diagnosis and procedure codes were used. 3.3. Is method of data collection likely to record long-term outcomes accurately? Low risk – no long-term outcomes assessed. **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – no checks were performed.

Full citation	**Schwarze ML, Shen Y, Hemmerich J, et al. (2009) Age-related trends in utilization and outcome of open and endovascular repair for abdominal aortic aneurysm in the United States, 2001-2006. J Vasc Surg. 50(4):722-729.**
	4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – the impact of missing data was not factored into the analyses. 4.3. Have different methods been compared within the study? High risk – different methods were not compared. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? High risk – authors do not explicitly state how many covariates (namely comorbidities) were considered in the regression models. 6.2 Were interactions between treatment and other covariates considered? High risk – no interactions were considered **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Symonides B, Śliwczyński A, Gałązka Z, et al. (2018) Short- and long-term survival after open versus endovascular repair of abdominal aortic aneurysm-Polish population analysis. PLoS One. 2018 Jun 14;13(6):e0198966.**
Study details	Study design: retrospective cohort study Location(s): Poland Study period: January 2011 to March 2016 Aim of the study: to compare short and long-term mortality and readmissions in patients with unruptured AAA treated by EVAR or OSR
Participants	Sample size of unmatched cohort: EVAR group, n=5,469; OSR group, n=2,336 Sample size of unmatched cohort: EVAR group, n=2,336; OSR group, n=2,336 Inclusion criteria: people with unruptured AAA who underwent elective EVAR or OSR were included. Note: there is no indication that selection was limited to only people with infrarenal AAA. Exclusion criteria: ruptured and people with thoracoabdominal aneurysms were excluded Baseline characteristics of the matched cohorts: Mean age (SD): EVAR group, 68.7 (8.0) years; OSR group, 68.5 (7.7) years Gender: EVAR group, 85.3% male; OSR group, 84.8% male Mean aneurysm diameter: not reported

Full citation	**Symonides B, Śliwczyński A, Gałązka Z, et al. (2018) Short- and long-term survival after open versus endovascular repair of abdominal aortic aneurysm-Polish population analysis. PLoS One. 2018 Jun 14;13(6):e0198966.**
	Hypertension: EVAR group, 48.7%; OSR group, 48.8% Chronic renal failure: EVAR group, 2.7%; OSR group, 2.7% Diabetes: EVAR group, 10.3%; OSR group, 11.3% Coronary heart disease: EVAR group, 9.1%; OSR group, 9.5% Stroke: EVAR group, 1.6%; OSR group, 1.9%
Methods	Data collection: patients were identified and outcome data were collected using ICD9 codes to query reimbursement data in a database managed by the only public and obligatory health insurer in Poland (the National Health Fund). Authors highlighted that the database tracks all patient admissions, main diagnoses, concomitant diseases and medical procedures longitudinally throughout the entire country. Additionally, the database contains information on birth and death dates. Analysis: propensity matching was performed, and supplemented with Cox proportional hazard regression. Propensity score analysis was performed by matching patients while controlling for age, gender and concomitant diseases. Similarly, Cox proportional-hazards analyses were performed controlling for age, gender, concomitant diseases and readmissions (expressed as total number of events).
Intervention	EVAR
Comparator	OSR
Outcomes	30-day mortality, readmissions, and long-term survival
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Moderate risk – ≥5-yr recruitment with no adjustment for year of operation. 1.2. Were cohorts from the same place? Low risk – cohorts were selected from the same national healthcare insurer database. 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included, which is likely to have led to implicitly different inclusion criteria across treatment arms.. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – the study controlled for age and sex. 2.2. Does study control appropriately for comorbidity and/or fitness? Moderate risk – authors stated that comorbidities were controlled for but did not specify which ones. 2.3. Does study control appropriately for AAA characteristics? High risk – AAA characteristics were not controlled for. 2.4. Could any adjustment variables have been affected by the intervention? Low risk – in the propensity score matching analysis, no post-intervention variables which could mediate the treatment effect were controlled for. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – participants were identified using a health insurance provider's administrative database.

Full citation	**Symonides B, Śliwczyński A, Gałązka Z, et al. (2018) Short- and long-term survival after open versus endovascular repair of abdominal aortic aneurysm-Polish population analysis. PLoS One. 2018 Jun 14;13(6):e0198966.**
	3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – outcomes were assessed using a health insurance provider's administrative database. 3.3. Is method of data collection likely to record long-term outcomes accurately? Low risk – direct use of reliable routine data registry **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk- no checks were reported. 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – no details were provided. 4.3. Have different methods been compared within the study? Moderate risk - Different methods were compared but all relied on the same assumption about selection. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? High risk – no algorithm was reported. 5.2. Was overlap / common support appropriately assessed? High risk – no details were provided by authors. 5.3. Has balancing of the covariates been demonstrated? Moderate risk – conventional hypothesis tests were performed, with no evidence of significant differences. **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? High risk – it was not possible to ascertain whether the sample size was adequate relative to number of covariates considered because authors did not provide a list of all covariates considered. 6.2 Were interactions between treatment and other covariates considered? High risk – no interactions were considered. **Overall risk of bias:** High risk **Directness: Directly applicable**

Full citation	**Tarbunou YA, Smith JB, Kruse RL, Vogel TR. Outcomes associated with hyperglycemia after abdominal aortic aneurysm repair. Journal of vascular surgery. 2019 Mar 1;69(3):763-73.**
Study details	Study design: retrospective propensity-matched cohort study Location(s): USA Study period: September 2008 to March 2014 Aim of the study: to evaluate outcomes and complications associated with AAA repair in patients with postoperative hyperglycemia.
Participants	Sample size: EVAR group, n=1,486; OSR group, n= 992 Inclusion criteria: Patients who underwent endovascular or open repair for a nonruptured AAA

Full citation	**Tarbunou YA, Smith JB, Kruse RL, Vogel TR. Outcomes associated with hyperglycemia after abdominal aortic aneurysm repair. Journal of vascular surgery. 2019 Mar 1;69(3):763-73.**
	Exclusion criteria: Patients who were younger than 21 years, had an emergent or urgent admission, or had no postoperative medication or laboratory data in Health Facts and patients whose postoperative blood glucose levels were below 80 mg/dL (hypoglycemic) Baseline characteristics (whole group) Mean age: 69.4 Gender: 71.7% male Mean aneurysm diameter: not reported Coronary artery disease: 8.1% Diabetes: 22.1% Chronic kidney disease: 11.2%
Methods	Data collection: Patients were identified from Cerner Health Facts (Cerner Corporation, North Kansas City, Mo), a database composed of electronic clinical records from hospital systems that have Cerner Corporation's electronic health record (ICD-9-CM) Analysis: Multivariable logistic regression was used to examine the association between postoperative hyperglycemia and infection, mortality, and readmission after adjusting for patient and hospital characteristics. We calculated odds ratios (ORs) and 95% CIs. We assessed model discrimination with the C statistic. Model calibration over the range of risk was assessed with the Hosmer–Lemeshow goodness-of-fit test.
Intervention	EVAR
Comparator	OSR
Outcomes	In-hospital mortality
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Moderate risk – ≥5-yr recruitment with no adjustment for year of operation 1.2. Were cohorts from the same place? Low risk – same country 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included, with no adjustment for anatomy. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – age and sex adjusted for 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – broad range of individual comorbidities 2.3. Does study control appropriately for AAA characteristics? High risk – none 2.4. Could any adjustment variables have been affected by the intervention? High risk – postoperative hyperglycaemia and medications adjusted for **Data collection**

Full citation	**Tarbunou YA, Smith JB, Kruse RL, Vogel TR. Outcomes associated with hyperglycemia after abdominal aortic aneurysm repair. Journal of vascular surgery. 2019 Mar 1;69(3):763-73.**
	3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – administrative registry with high-level diagnosis and procedure codes 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk –administrative registry with high-level diagnosis and procedure codes 3.3. Is method of data collection likely to record long-term outcomes accurately? N/A **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? Low risk – Hosmer–Lemeshow and c-statistic 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – no consideration of missing data 4.3. Have different methods been compared within the study? High risk **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? High risk – 58 events and 18 covariates 6.2 Were interactions between treatment and other covariates considered? High risk – none **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	**Wald R, Waikar SS, Liangos O, et al. (2006) Acute renal failure after endovascular vs open repair of abdominal aortic aneurysm. J Vasc Surg. 43(3):460-466;**
Study details	Study design: retrospective cohort study Location(s): USA Study period: January to December 2002 Aim of the study: to compare outcomes of people who underwent EVAR and OSR for unruptured AAA.
Participants	Sample size: EVAR group, n=2,651; OSR group, n=3,865 Inclusion criteria: people over 18 years with a primary diagnosis of unruptured AAA who underwent EVAR or OSR were included. Note: people with complex AAA or symptomatic AAA may have also been included. Exclusion criteria: patients who underwent an aorto-renal bypass in addition to AAA repair and those receiving dialysis were excluded

Full citation	**Wald R, Waikar SS, Liangos O, et al. (2006) Acute renal failure after endovascular vs open repair of abdominal aortic aneurysm. J Vasc Surg. 43(3):460-466;**
	Baseline characteristics: Mean age (SD): EVAR group, 73.5 (10.4) years; OSR group, 71.6 (10.8) years Gender: EVAR group, 85.3% male; OSR group, 77.0% male Mean aneurysm diameter: not reported Chronic kidney disease: EVAR group, 7.3%; OSR group, 8.4% Congestive heart failure: EVAR group, 7.2%; OSR group, 11.6% Chronic lung disease: EVAR group, 30.8%; OSR group, 36.7% Chronic liver disease: EVAR group, 0.7%; OSR group, 1.0% Diabetes: EVAR group, 13.5%; OSR group, 10.3%
Methods	Data collection: participants were identified using diagnosis and procedure codes (such as ICD9 and OPCS4 codes) to query an administrative database that used high-level diagnosis and procedure codes: the Nationwide Inpatient Sample (NIS). NIS data is an anonymised, stratified sample of 20% of acute care non-federal American hospitals in 33 states. All data on comorbidities and outcomes were extracted from the NIS database. Analysis: Multivariate logistic regression was the primary method of analysis. Authors stated that all demographic, clinical, and hospital-related variables were included in the regression models. Furthermore, investigators selected multiplicative interaction terms (eg, chronic kidney disease*procedure type) to evaluate for effect modification. The Hosmer and Lemeshow test was used to assess model calibration. To adjust for additional unmeasured confounding, investigators fit a logistic regression model using procedure type as the dependent variable to generate propensity scores for receipt of EVAR. Only covariates with p values <0.05 were included in this model. The propensity score was then used to rank individuals according to their likelihood of receiving a given treatment and the association between procedure type and acute renal failure was then evaluated across quintiles of the propensity score.
Intervention	EVAR
Comparator	OSR
Outcomes	
Study Appraisal using NICE's bespoke risk of bias assessment tool	**Selection** 1.1. Were cohorts from the same time period? Low risk – cohorts were drawn from the same time period. 1.2. Were cohorts from the same place? Low risk – all participants were selected from the same national patient database. 1.3. Is the definition of AAA the same across cohorts? Moderate risk – all infrarenal and complex AAA included, which is likely to have led to implicitly different inclusion criteria across treatment arms.. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – the study controlled for age and gender. 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – a broad number of relevant comorbidities were controlled for.

Full citation	Wald R, Waikar SS, Liangos O, et al. (2006) Acute renal failure after endovascular vs open repair of abdominal aortic aneurysm. J Vasc Surg. 43(3):460-466;
	2.3. Does study control appropriately for AAA characteristics? High risk – AAA characteristics were not controlled for. 2.4. Could any adjustment variables have been affected by the intervention? Low risk - no post-intervention variables which could mediate the treatment effect were controlled for. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – an administrative registry with high-level diagnosis and procedure codes was used. 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – an administrative registry with high-level diagnosis and procedure codes was used. 3.3. Is method of data collection likely to record long-term outcomes accurately? N/A – no long term outcomes were assessed **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? Low risk – model fit was assessed using the Hosmer–Lemeshow test. 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – no indication that the effect missing data were considered. 4.3. Have different methods been compared within the study? Low risk – different methods were compared but all relied on the same assumption about selection. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? Low risk – number of events is ≥10 times greater than number of variables considered. 6.2 Were interactions between treatment and other covariates considered? Low risk – investigators selected multiplicative interaction terms (eg, chronic kidney disease*procedure type) to evaluate for effect modification **Overall risk of bias:** High risk **Directness:** directly applicable

Full citation	Williams TK, Schneider EB, Black JH, et al. (2013) Disparities in outcomes for Hispanic patients undergoing endovascular and open abdominal aortic aneurysm repair. Ann Vasc Surg. 27(1):29-37.
Study details	Study design: retrospective cohort study

Full citation	**Williams TK, Schneider EB, Black JH, et al. (2013) Disparities in outcomes for Hispanic patients undergoing endovascular and open abdominal aortic aneurysm repair. Ann Vasc Surg. 27(1):29-37.**
	Location(s): USA Study period: 2005 to 2008 Aim of the study: to examine the influence of race and ethnicity on the outcomes of EVAR and OSR of unruptured AAA and its effect on costs.
Participants	Sample size: EVAR group, n=62,728; OSR group, n=24,253 Inclusion criteria: people who underwent repair of unruptured AAA were included. Patients who possessed ICD-9-CM codes for both EVAR and OSR were included in the EVAR group, as this likely represented patients undergoing open conversion during the same hospitalisation. Analyses were performed using an intention-to-treat approach. Note: some people were operated upon under emergency circumstances (likely to be due to symptomatic AAA). Furthermore, there is no indication that the study was limited to people with infrarenal AAA. Exclusion criteria: people < 18 years, > 99 years, those with ruptured aneurysms, thoracic aneurysms, mycotic aneurysms, syphilitic aneurysms, or traumatic aneurysms were excluded. Baseline characteristics: Mean age: not reported Gender: EVAR group, 82.6% male; OSR group, 75.6% male Mean aneurysm diameter: not reported Cerebrovascular disease: EVAR group, 4.7%; OSR group, 5.1% Renal disease: EVAR group, 8.9%; OSR group, 10.8% Congestive heart failure: EVAR group, 7.9%; OSR group, 10.1% COPD: EVAR group, 31.2%; OSR group, 38.1% Emergency admission: EVAR group, 8.3%; OSR group, 20.7%
Methods	Data collection: participants were identified using diagnosis and procedure codes (such as ICD9 and OPCS4 codes) to query an administrative database that used high-level diagnosis and procedure codes: the Nationwide Inpatient Sample (NIS). NIS data is an anonymised, stratified sample of 20% of acute care non-federal American hospitals in 33 states. All data on comorbidities and outcomes were extracted from the NIS database. Analysis: multivariate regression analysis was performed adjusting for age, gender, race, comorbidities utilising the Charlson comorbidity index, procedure type, insurance type, and hospital characteristics. No further details were provided.
Intervention	EVAR
Comparator	OSR
Outcomes	In-hospital mortality
Study Appraisal	**Selection** 1.1. Were cohorts from the same time period? Low risk – cohorts were drawn from the same time period.

Full citation	Williams TK, Schneider EB, Black JH, et al. (2013) Disparities in outcomes for Hispanic patients undergoing endovascular and open abdominal aortic aneurysm repair. Ann Vasc Surg. 27(1):29-37.
using NICE's bespoke risk of bias assessment tool	1.2. Were cohorts from the same place? Low risk – all participants were selected from the same international database. 1.3. Is the definition of AAA the same across cohorts? High risk – a considerably higher proportion of participants in the EVAR group were treated under emergency admissions, indicating the likelihood of symptomatic AAAs. This highlights implicitly different inclusion criteria across treatment arms. **Confounding** 2.1. Does study control appropriately for demographics? Low risk – study controls for age and sex. 2.2. Does study control appropriately for comorbidity and/or fitness? Low risk – comorbidities (using the Charlson comorbidity index) were controlled for. 2.3. Does study control appropriately for AAA characteristics? High risk – no AAA characteristics were controlled for. 2.4. Could any adjustment variables have been affected by the intervention? Low risk - no post-intervention variables which could mediate the treatment effect were controlled for. **Data collection** 3.1. Is method of data collection likely to have identified suitable participants accurately? High risk – an administrative registry with high-level diagnosis and procedure codes was used. 3.2. Is method of data collection likely to record perioperative outcomes accurately? High risk – an administrative registry with high-level diagnosis and procedure codes was used. 3.3. Is method of data collection likely to record long-term outcomes accurately? Low risk – no long term outcomes were assessed **Analysis – general** 4.1. Were any checks conducted on model specification and/or fit? High risk – no checks were reported. 4.2. Are missing outcome data and covariates reported and, if necessary, adjusted for? High risk – there is no indication that the effect of missing data was considered. 4.3. Have different methods been compared within the study? High risk – different methods were not compared. **Analysis – matching** 5.1. Is the matching algorithm reported and reasonable? N/A 5.2. Was overlap / common support appropriately assessed? N/A 5.3. Has balancing of the covariates been demonstrated? N/A **Analysis – simple multivariable models** 6.1 Is sample size adequate relative to number of covariates considered? Low risk – number of events is ≥10 times greater than number of variables considered. 6.2 Were interactions between treatment and other covariates considered? High risk – no interactions were considered. **Overall risk of bias:** High risk

Full citation	Williams TK, Schneider EB, Black JH, et al. (2013) Disparities in outcomes for Hispanic patients undergoing endovascular and open abdominal aortic aneurysm repair. Ann Vasc Surg. 27(1):29-37.
	Directness: directly applicable