Journal of Clinical Epidemiology 181 (2025) 111714

Journal of Clinical Epidemiology

ORIGINAL RESEARCH

Trial emulation to assess the effect of surgery on survival when there are competing risks, with application to patients with thoracic aortic aneurysms

James Murray[a], Caroline Chesang[a], Steve Large[b], Colin Bicknell[c], Carol Freeman[d], Ruth H. Keogh[a], Linda D. Sharples[a,*]

[a]*Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK*
[b]*Department of Cardiac Surgery, Royal Papworth Hospital NHS Foundation Trust, Cambridge, UK*
[c]*Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK*
[d]*Papworth Trials Unit Collaboration, Royal Papworth Hospital NHS Foundation Trust, Cambridge, UK*

Accepted 1 February 2025; Published online 8 February 2025

Abstract

Objectives: This study extends methods to estimate average causal effect of aneurysm repair surgery on (i) overall survival and (ii) aneurysm-related mortality, accounting for competing risks using data from the Effective Treatment for Thoracic Aortic Aneurysm (ETTAA) cohort.

Study Design and Setting: ETTAA, a prospective cohort study, recruited 886 patients between 2014 and 2018. Patients were linked to UK national hospital and mortality databases by National Health Service digital and followed-up for later surgeries and deaths. We compared a strategy of open or endovascular surgery (whichever appropriate) within 12 months of enrollment to ETTAA with no surgery within 12 months using the trial emulation framework and cloning-censoring-weighting (CCW) analysis. Key confounders at baseline were controlled for using inverse probability weighting methods.

Results: In complete case analysis, if everyone received surgery within a 12-month grace period, an estimated 7-year survival probability was 57.4% (95% CI: 47.3%, 67.4%) vs 49.9% (44.0%, 55.0%) if no one received surgery. This benefit was primarily attributable to reduction in aneurysm-related deaths (difference −8.7%, 95% CI: −14.0%, −3.9%), with no significant effect on deaths from other causes. The findings were consistent under sensitivity analyses, including multiple imputation of missing confounders. Our CCW approach addressed selection-for-treatment, allowed for surgery to be received within a grace period, and used appropriate methods to separate aneurysm-related mortality from competing risks.

Conclusion: The study demonstrates the utility of trial emulation and counterfactual methods in estimation of causal effects on competing risks using observational data. The findings suggest a benefit for aneurysm-related survival up to 7 years after enrollment.

Keywords: Aortic aneurysms; Surgery; Survival; Competing risks; Trial emulation

Funding: This work was supported by the National Institute for Health Research [Health Technology Assessment, 11/147/03—Effective Treatments for Thoracic Aortic Aneurysms (ETTAA study): a prospective cohort study]. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research, or the Department of Health. RHK was funded by UK Research and Innovation (Future Leaders Fellowship MR/X,015,017/1). CC was supported by a PhD studentship cofunded by the Economic and Social Research Council and AstraZeneca.

Ethics statement: The authors confirm that the ETTAA study complies with the Declaration of Helsinki; the West Midlands—South Birmingham Research Ethical Committee approved the research protocol, and informed consent was obtained from all participants.

* Corresponding author. London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK.
E-mail address: linda.sharples@lshtm.ac.uk (L.D. Sharples).

https://doi.org/10.1016/j.jclinepi.2025.111714

Plain Language Summary

This study shows how to estimate effects of surgery on different causes of death, when we cannot do a clinical trial, and illustrates this using an example from heart surgery. The aorta is the main artery that carries oxygen-rich blood from the heart to the body. In some people, a part of the vessel wall becomes weak and loses its elastic properties, so it doesn't return to its normal shape after the blood has passed through. This can lead to swelling or bulging in the aorta, called an aneurysm. A thoracic aortic aneurysm, or TAA for short, is an aneurysm in the section of the aorta in the chest (https://www.bhf.org.uk/informationsupport/conditions/thoracic-aortic-aneurysm).

We have used data from the Effective Treatment for Thoracic Aortic Aneurysm (ETTAA) study, which investigated aneurysm growth rates, patient outcomes, quality of life, and costs, in 886 patients diagnosed with TAA. ETTAA compared two surgical treatments, *Open Heart Surgery*, where the section of the aorta that contains the aneurysm is removed and replaced by a new aorta made from a synthetic material, and *Stent Grafting*, where tubes are inserted into arteries to allow blood to flow freely using less invasive "keyhole" surgery. ETTAA reviewed existing research evidence but data comparing the effectiveness of these two approaches to each other and to outcomes without surgery were of sparse or limited quality and outdated. The results of ETTAA up to 2020 have been published in a monograph. (https://pubmed.ncbi.nlm.nih.gov/35094747/).

Two findings from ETTAA motivated this study. First, there were no clinical trials comparing surgery with no surgery and no studies that mimic clinical trials. Second, we had not considered whether surgery overall prevents deaths due to aneurysm or deaths from other causes. We call these two types of death, *competing risks*.

It is unlikely that a clinical trial comparing surgery with no surgery will ever be completed because the number of people who are diagnosed with TAA is small. Also, TAA can become a serious problem if left untreated. On the other hand, surgery for TAA is difficult and can result in serious complications, including death. Therefore, it is important to know how much surgery improves survival related to the aneurysm and whether it improves survival overall.

Recent developments in statistics provided methods for investigating survival in a way which increases confidence in the *cause-effect relationship* between surgery and outcomes. In this study, we show how these statistical methods can be used to estimate the proportion of patients who die from the *competing risks*, if all patients had surgery within 12 months compared with if no patients had surgery within 12 months. We take into account the different times between diagnosis of TAA and surgery and adjust for the main differences between surgery and no surgery patients. Using these methods, we estimate that surgery reduces deaths due to aneurysms at 7 years by 8.7%, with no effect on deaths from other causes. The benefit of surgery was significant by 3 years after diagnosis. We also provide discussion about using routine medical records to repeat this type of study.

1. Introduction

Randomized controlled trials (RCTs) are the gold standard for assessing cause-effect relationships. Based on drug development stages, the idea, development, exploration, assessment and long-term follow-up framework guides the development of new surgical procedures [1]. In addition, well-conducted surgical RCTs exist in cardiovascular disease [2], cancer [3] and orthopedics [4].

However, surgical RCTs remain challenging due to their complexity [2]. Patient and/or surgeon preferences result in difficulty in recruiting to trials and confer selection-for-treatment bias in observational studies. Waiting lists for operating theaters and the need to assemble surgical teams result in delayed treatment, and patients assigned to surgery may die before surgery can take place. When comparing between surgery and no-surgery groups, delay in time to surgery results in immortal time bias if ignored (ie, patients must be alive to receive surgery). One option is to evaluate a strategy of surgery within a set grace period after diagnosis [5]. For example, we might compare a strategy of surgery within a grace period of 12 months, with no surgery within 12 months.

Developments in causal inference tools and thinking, alongside statistical methods, such as inverse probability of treatment weighting, adjustment, and standardization, have increased confidence in estimating causal relationships between treatment and outcomes using observational data [5]. Given challenges in conducting large RCTs [2], these methods can provide insights into potential benefits of surgical interventions; provided key assumptions are valid.

In this context, there are advantages to applying the trial emulation framework. First, it allows for rigorous definition of a target trial that would address the question of interest, including inclusion criteria, treatment strategies, follow-up period, and outcomes. Second, it promotes description of an emulated trial that is as close to the target trial as available data allow, including design and analysis strategies addressing concerns of immortal time bias and treatment selection bias.

What is new?

Key findings:

- Causal analysis for competing risks estimated an 8.7% lower probability of aneurysm-related death rate at 7 years after thoracic aneurysm surgery within 12 months compared to no surgery within 12 months.
- Trial emulation methods, including cloning−censoring−weighting, improved causal inference by addressing selection bias and allowing for surgery to be received within a grace period.
- The prospective cohort design in ETTAA enabled identification of non-RCT patients, accurate classifications of aneurysm-related deaths, and direct recording of aneurysm size.

What this adds to what is known:

- Our approach strengthens confidence in the causal effect of surgery on aneurysm-related mortality by reducing bias in nonrandomized comparisons.
- We highlight the feasibility of emulating a trial using prospective cohort data while ensuring accurate classification of deaths and surgical eligibility.

What is the implication and what should we change now:

- Surgery to repair thoracic aortic aneurysms is successful at decreasing aneurysm-related mortality, with little impact on other cause deaths.
- Provided key confounders are measured and assumptions are justified, observational data can support causal inferences about treatment effects.
- Future studies using routine health records should incorporate additional clinical data or leverage text mining to extract key surgical eligibility criteria.

The ETTAA study was an observational cohort of 886 people with TAAs, which was referred to cardiovascular hospitals for assessment for surgery [6]. Surgery is recommended to prevent life-threatening growth and rupture of aneurysms $>$5-6 cm in diameter, depending on aneurysm location within the aorta and patient factors, although evidence for its effectiveness is sparse [7−9]. TAA diagnosis is uncommon; untreated cases are life-threatening [10−14], so surgery has never been tested in RCTs. Conversely, surgery is difficult and can result in complications, including death [6]. Our primary aim is to assess the potential of trial emulation and causal analysis for estimating the impact of surgery on (i) all-cause deaths and (ii) aneurysm-related deaths, considering death from other causes as a competing risk. The cloning-censoring-weighting (CCW) method is used to address selection and immortal time biases [5]. We estimate the causal effects of surgery within 12 months of enrollment to ETTAA on aneurysm-related deaths and mortality from other causes as a competing risk.

1.1. Objectives were to

- Illustrate trial emulation and causal analysis methods to address selection bias and allow for treatment delay, extending previous work to accommodate competing risks and variance estimation.
- Apply these methods to aneurysm-related deaths in ETTAA, estimating the average (causal) treatment effect (ATE) on survival probability after 7 years of enrollment, if all patients had received aneurysm repair surgery within 12 months vs if no patients had received surgery within 12 months.
- Assess sensitivity of results to length of the grace period within which surgery is received and missing data assumptions.

2. Materials and methods

The strengthening the reporting of observational studies in epidemiology checklist guided reporting.

2.1. ETTAA

The original ETTAA study had prospective design and data collection [6,15−17]. During ETTAA, patients with TAA were referred to a multidisciplinary team (MDT) at 1 of 30 English cardiovascular centers and assessed for either open surgical repair or endovascular stent grafting, as considered appropriate by the MDT. Patients were recruited during 2014-2018 and followed until June 2019. Thereafter, patients were linked to UK National Hospital and mortality databases by National Health Service Digital and followed until March 2023 for new surgeries and deaths.

2.2. Trial emulation

Trial emulation involves describing protocols for (i) the 'target' RCT comparing surgery with no surgery that we would complete if feasible and (ii) an emulated trial that is as close to this as possible within the limitations of the extended ETTAA data. Table 1 summarizes both protocols, which include standard RCT components: definitions of eligibility, surgery, control, assignment procedures and timing, follow-up, and outcomes. In the target trial, patients would be recruited at an MDT meeting and randomized to surgery or no-surgery within a time period. In the observational study, 371 patients underwent surgery at varying

times after enrollment. Approximately, three-quarters of surgery patients ($n = 281$) had the operation within 1 year of enrollment, so we adopted 12 months as the grace period for surgery. The causal question of interest was: What would the difference in overall and cause-specific survival probability at 7 years have been if all patients had undergone surgery within 12 months of enrollment to ETTAA, compared to if no patients had received surgery within 12 months (but could have received it later)?

2.3. Cloning-censoring-weighting (CCW) method

Because there is a delay between the MDT meeting and operation, surgery is only possible for patients surviving the interval. If not accommodated in the methods, benefit of surgery is generally overestimated, which is described as immortal time bias. CCW addresses this issue. Each patient's outcome is observed either under surgery within the grace period or under the control condition but not both. The population ATE can be identified from the observed data conditional on identifying assumptions (see Hernan & Robins [5], below, and supplementary data).

2.3.1. Cloning and defining treatment

Maringe [18] provides a tutorial for the CCW method, which we extend to incorporate competing risks. Briefly, for each patient, two 'clones' are created: one labeled surgery and one labeled control. Patient experience is modified to reflect the clone's treatment label (Fig 1). Surgical clones are censored at the grace period if the patient does not undergo surgery within 12 months; otherwise, they are consistent with surgery. Control clones are censored at surgery if it occurs within the grace period; otherwise, they are consistent with control at 12 months. If death occurs before surgery within the grace period, it contributes equally to both arms, ie, patient clones are censored when their treatment deviates from their labeled treatment, indicating a departure from the target trial protocol, enabling estimation of per-protocol effects. This contrasts with intention-to-treat (ie, the target trial) estimand, which estimates the effect of *assignment to surgery*, including patients irrespective of whether surgery was performed within the grace period. The per-protocol estimand implemented via CCW instead provides a causal estimate of adherence to the strategy of surgery within the grace period vs no surgery in the grace period, censoring crossover between arms. Artificial censoring conferred by definitions of treatment introduces informative censoring, which must be addressed.

2.3.2. Addressing informative censoring due to treatment

Two main methods are proposed for addressing treatment allocation bias, inverse probability of (treatment) censoring weighting and standardization (eg, g-formula), and have been extended to accommodate more flexible modeling assumptions [19,20]. Inverse probability of (treatment) censoring weighting aligns with the CCW approach when treatment strategy incorporates a grace period. In ETTAA, probability of surgery for surgical clones was estimated, conditional on confounders, using logistic regression. The time-dependent probability of remaining surgery-free for control clones was estimated using Cox regression because clones can deviate from control status at any time during the grace period (ie, whenever the patient has surgery). Further details are presented in [18] and the supplementary data. Each clone's contribution to the estimate of the treatment contrast is weighted by the inverse probability of being uncensored. Note that weights change over time as the risk sets change. Clones have large weights if their observed confounders are uncommon for their treatment label. To avoid clones having large influence on the results, we stabilized weights by multiplying them by the marginal (unconditional) probabilities of the treatment label (see supplementary data). To safeguard against positivity assumption violations, we applied a maximum weight based on age, New York Heart Association functional classification, and aneurysm size in all analyses (see supplementary data).

2.3.3. Identifying confounders

Confounders were identified from a literature review, a Delphi study of clinician-stated surgical practice, and empirical analysis of predictors of surgery and survival completed as part of ETTAA [6]. We included confounders and other variables related to outcomes and avoided variables related to surgery but not survival and mediators on the pathway between surgery and survival (Table 2) [21]. Nine baseline variables with few missing measurements were used in complete case analysis. They included maximum aneurysm diameter, recommended by the international guidelines as the main determinant of timing of surgery [22–24]. Creatinine, a marker of impaired renal function related to cardiovascular disease and death, was not collected routinely in ETTAA and was missing for 459 (51.8%) patients (Table 2) and excluded from the base case analysis.

2.3.4. Composite survival and competing risks

For overall mortality (composite of aneurysm-related and other cause deaths) and aneurysm-related mortality, cumulative incidence function curves were estimated using weighted Kaplan-Meier or Aalen-Johansen estimators. Variable times to surgery introduce immortal time bias, distinct from bias due to competing risks, where deaths from other causes preclude aneurysm-related deaths and alter the risk set dynamically. Combined weights, calculated as the product of treatment-censoring and survival-censoring probabilities, were applied to estimate the ATE, ensuring adjustment for differential follow-up and treatment assignment [5]. The competing risks estimand captured both the direct effect of surgery on aneurysm-related death and indirect effects mediated through changes in risks of other causes of death [25]. Administrative censoring (due to end of follow-up or timing of downloads from electronic health records [EHRs]) was assumed uninformative [5]. During

Table 1. Protocol summaries for target and emulated trials to estimate the effect of aneurysm repair surgery on survival for patients with chronic thoracic aortic aneurysms

Target trial	Trial emulation
Key eligibility criteria	
Aged >17, chronic TAA ≥ 4 cm in aortic arch or descending thoracic aorta, suitable for ESG or OSR when discussed at MDT. No previous surgery for this aneurysm, but may have had surgery on ascending or abdominal aorta.	Aged >17, chronic TAA ≥ 4 cm in aortic arch or descending thoracic aorta, suitable for ESG or OSR when discussed at MDT. No previous surgery for this aneurysm, but may have had surgery on ascending or abdominal aorta.
Treatment strategies	
Intervention: surgery. Control: no surgery	Intervention: surgery within the grace period (12 months) after MDT when enrolled into ETTAA. Control: no surgery within the grace period, but may have surgery after 12 months.
Assignment procedures	
Randomized allocation at time zero to either surgery or no surgery. Participants and clinicians aware of whether surgical intervention is assigned.	Participants assigned at time zero to either surgery or no surgery within the grace period. Participants and clinicians aware of whether surgical intervention is assigned.
Time zero	
Randomization occurs at the time of the MDT meeting where the patients are initially assessed.	Participants assigned at the time of the MDT meeting where the patients are initially assessed.
Follow-up period	
Intervention arm: From date of randomization to death, loss to follow-up, withdrawal, or censored at end of the study if alive. Control arm: From randomization to death, loss to follow-up, withdrawal, or censored at end of the study if alive.	Intervention arm: From consent for ETTAA to death, loss to follow-up, withdrawal from ETTAA, or censored at 12 months if no surgery occurs in the grace period. Control arm: From consent for ETTAA to death, loss to follow-up, withdrawal from ETTAA, or censored at the date of surgery if it occurs in the grace period.
Outcomes	
Composite of aneurysm-related and other cause death. Aneurysm-related death with other causes of death a competing risk.	Composite of aneurysm-related and other cause death. Aneurysm-related death with other causes of death a competing risk.
Causal contrast	
For composite, the difference in probability of death at time horizon (7 years). For competing risks, difference in cumulative incidence of aneurysm-related death at the time horizon (total effect).	For composite, the difference in probability of death at time horizon (7 years). For competing risks, the difference in cumulative incidence of aneurysm-related death at the time horizon (total effect). Both contrasts have a per-protocol interpretation.
Analysis plan	
Intention to treat effect estimation using marginal model such as Kaplan-Meier method.	Per-protocol effect estimation using the CCW, with inverse probability of censoring weighting due to clone definition using propensity scores including all confounders.

CCW, cloning-censoring-weighting; ESG, endovascular stent grafting; MDT, multidisciplinary team; OSR, open surgical repair; TAA, thoracic aortic aneurysm.

the original ETTAA study, 106 of 205 (51.7%) deaths were classified as aneurysm-related by participating centers, compared to 63 of 241 (26.1%) using International Classification of Diseases-10 codes in the extended follow-up.

2.3.5. *Identifiability assumptions*

Key assumptions for identifying the causal estimand are conditional exchangeability, positivity, consistency, and correct specification of weighting and outcome models (Figs S3-4). Exchangeability requires that outcome and treatment are independent, conditional on all confounders, and is impossible to fully justify. However, the literature review, Delphi study, and empirical analysis that completed during ETTAA provide reassurance. Positivity requires that both treatment strategies have nonzero probability for all individuals. During ETTAA, 106 patients declined or were not expected to have surgery, despite being eligible. We included this subgroup since they were eligible and may have undergone surgery if referred to a different MDT. Because ETTAA was observational and did not interfere with patient management, the consistency assumption likely holds.

2.3.6. *Implementation*

We implemented these methods using long-form data created through the cloning and censoring approach, splitting follow-up for each clone at times of surgery, death, or censoring. This increased the number of data rows and computation times. Because there were two clones per patient, and

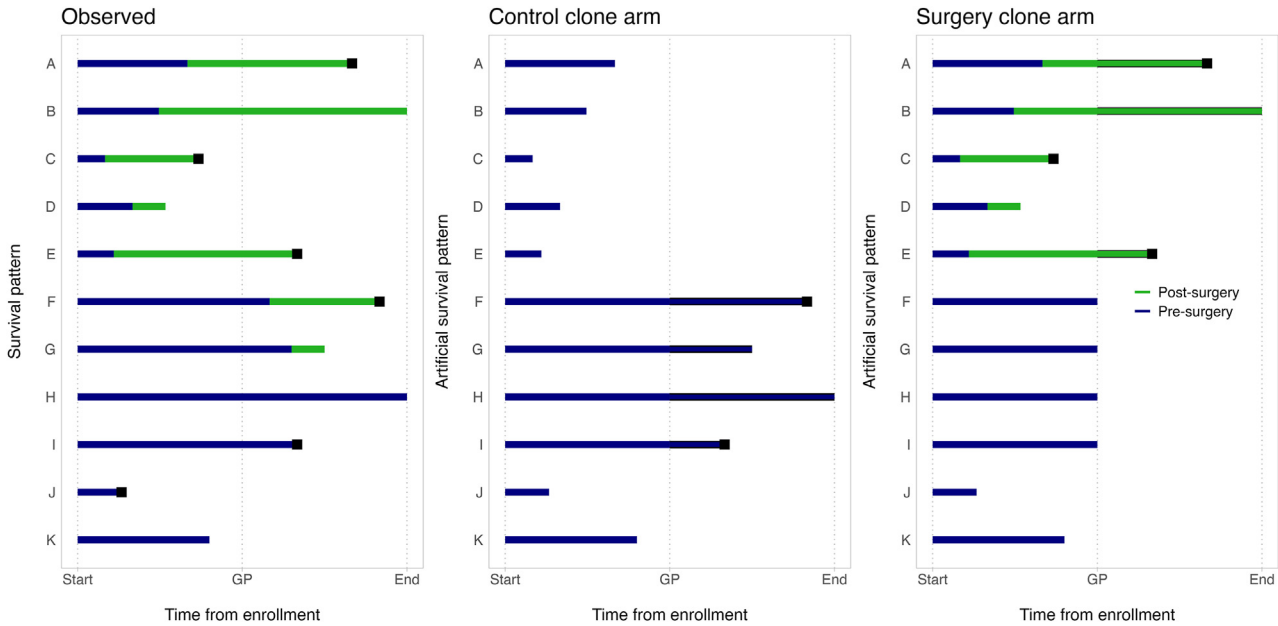

Figure 1. Illustration of the cloning process. Column 1 shows observed data for 11 example patients. Columns 2 and 3 show modifications to data for control and surgery clones that are required to estimate the probability of censoring when patients deviate from the clone treatment label. Green and navy lines represent time before and after surgery up to censoring or death. Green or navy lines with a black background indicate follow-up after the grace period (GP) as used in the analysis model. Black squares represent time of death.

because the analysis uses estimated weights, usual standard errors for risk estimates are invalid. Nonparametric bootstrapping was used to estimate percentile-based 95% CIs, with 250 samples balancing accuracy and efficiency.

2.3.7. *Sensitivity analysis*

The base case analysis included all eligible patients with complete data for nine confounders with few missing values. In sensitivity analysis, we re-estimated treatment contrasts after (i) multiple imputation for nine confounders using the multiple imputation using fully conditional specification [26], with 250 bootstraps, each with one imputation, and reported percentile-based CIs [27], (ii) creatinine added to the multiple imputation of confounders, (iii) excluding patients who refused surgery or otherwise were not actively monitored, (iv) defining grace periods of 3, 6, 9, and 18 months, and (v) removing the cap on weights. A negative control outcome (malignant neoplasm deaths) was tested to evaluate the confounding structure [28].

3. Results

3.1. *ETTAA patients and outcomes*

During 2014-2018, 886 patients were enrolled in ETTAA by 30 MDTs. The mean age was 70.8 years (SD 10.9); 36.2% were women (Table 2). Mean maximum aneurysm diameter was larger for those having surgery. Over 80% of patients had the maximum aneurysm diameter in the descending or thoracoabdominal thoracic aorta. To March 2023, 198 had open surgery and 183 endovascular surgery (total 381). Comorbidities including chronic obstructive pulmonary disease (18.5%), coronary heart disease (19.3%), and diabetes (9.4%) were common. Complete measurements for the nine key confounders were available for 832 (93.9%) patients. Mean serum creatinine was lower among surgery recipients. There were 169 (19.1%) aneurysm-related deaths, 277 (31.3%) deaths from other causes, and 440 (49.7%) patients were alive at last follow-up (Table 2).

3.2. *Base case*

Figures 2 and S1 show estimated cumulative incidence functions for composite survival and aneurysm-related and other-cause deaths. Estimated 7-year survival probability was 57.3% (95% CI: 47.3%, 67.4%) under the strategy of surgery (within 12 months) and 49.9% (95% CI: 44.0%, 55.0%) for the control strategy, a benefit of −7.4% (95% CI: −4.1%, 19.4%) survival at 7 years. Cause-specific incidence curves show that the benefit results from reduced aneurysm-related deaths (difference −8.7%, 95% CI: −14.0%, −3.9%), with no effect on deaths from other causes. Aneurysm-related death appears significantly lower under the surgery strategy from 3 years after enrollment in ETTAA (difference -5.2%, 95% CI: −9.4%, −1.2%). Surgery results in an average of 104.7 days (95% CI: −30.2, 263.8) additional lifetime over 7 years (Figs 4 and S2.1).

Weighted standardized mean differences were well balanced for confounders between surgery and no-surgery clones during the grace period (Fig 3). Postbaseline, maximum aneurysm diameter and age at enrollment, and

Table 2. Summaries of patient characteristics and final outcomes overall and for those having and not having surgery during the study.

Variable	Surgery (n = 381)	No surgery (n = 505)	Overall (n = 886)
Key confounders—required recording in ETTAA			
Age at enrollment; years			
Mean (SD)	68.7 (10.7)	72.4 (10.7)	70.8 (10.9)
Sex n (%)			
Women	133 (34.9%)	188 (37.2%)	321 (36.2%)
Men	248 (65.1%)	317 (62.8%)	565 (63.8%)
Maximum aneurysm diameter			
Mean (SD); cm	6.02 (1.11)	5.42 (1.09)	5.68 (1.14)
Aneurysm site n (%)			
Descending/thoraco-abdominal	318 (83.5%)	416 (82.4%)	734 (82.8%)
Aortic arch	63 (16.5%)	89 (17.6%)	152 (17.2%)
Chronic obstructive pulmonary disease n (%)			
Yes	63 (16.6%)	100 (20.0%)	163 (18.5%)
No	317 (83.4%)	401 (80.0%)	718 (81.5%)
Missing n	1	4	5
New York Heart Association class n (%)			
I	170 (47.2%)	180 (36.7%)	359 (42.2%)
II	131 (36.4%)	184 (37.6%)	315 (37.1%)
III	50 (13.9%)	100 (20.4%)	150 (17.6%)
IV	9 (2.5%)	17 (3.5%)	26 (4.2%)
Missing n	21	15	36
Coronary artery disease n (%)			
Yes	58 (15.5%)	110 (22.2%)	160 (19.3%)
No	317 (84.5%)	384 (77.7%)	701 (80.7%)
Missing n	6	11	17
Diabetes n (%)			
Type I/II	27 (7.1%)	56 (11.2%)	83 (9.4%)
None	354 (92.9%)	446 (88.8%)	800 (90.6%)
Missing n	0	3	3
Connective tissue disorder n (%)			
Yes	29 (7.6%)	26 (5.1%)	55 (6.2%)
No	352 (92.4%)	479 (94.6%)	831 (93.8%)
Confounder—not mandated variable in ETTAA			
Serum creatinine μmol/l			
Mean (SD)	89.7 (30.2)	99.8 (35.2)	94.2 (32.9)
Missing n	144	315	459
Survival status at the end of extended follow-up			
Survival status at the end of study n (%)			
Alive	199 (52.2%)	241 (47.7%)	440 (49.7%)
Aneurysm-related death	76 (19.9%)	93 (18.4%)	169 (19.1%)
Other death	106 (27.8%)	171 (33.9%)	277 (31.3%)

key criteria for surgery timing differed substantially between groups. Figure 3 shows that inverse probability of (treatment) censoring weighting was successful in reducing imbalance in maximum diameter.

3.3. Sensitivity analysis

One or more of nine key confounders was missing for 54 (6.1%) patients. Multiple imputations produced almost identical estimates to the complete case analysis (Table S1). Imputing serum creatinine increased the estimate of benefit of surgery for aneurysm-related survival but also increased the variance around these estimates, with similar conclusions. Estimates were also robust across grace periods between 3 and 18 months and removal of the ad hoc maximum weight, with the surgery strategy consistently having lower incidence of aneurysm-related mortality (Table S1, S2). Excluding patients without planned surgery did not change the effect of the surgery strategy on aneurysm-related death, but this benefit was overturned by greater risk of other deaths for

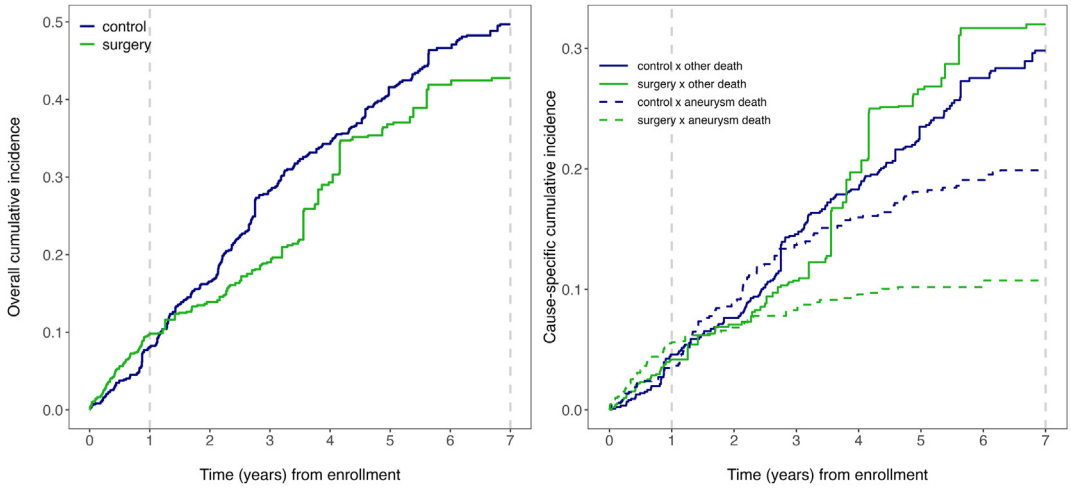

Figure 2. Cumulative incidence curves for composite deaths (aneurysm-related and other cause death) and cause-specific cumulative incidence of aneurysm-related deaths and deaths from other causes for the base case. Vertical dashed lines indicate the grace period (12 months) and the time horizon for estimating the risk difference (7 years).

the surgery strategy, resulting in no net benefit for composite survival. The risk difference for the negative control outcome (cancer death) was consistent with no effect throughout follow-up (95% CI: −0.4% to 0.3% at 1 year; −6.2% to 12.4% at 7 years).

4. Discussion

This study applied trial emulation and causal analysis methods to estimate the ATE of ETTAA patients having surgery within 12 months on mortality, compared with no patients having surgery within 12 months. The CCW method addressed treatment selection and immortal time biases and demonstrated the potential of these methods for estimating causal effects when competing risks exist, using prospective observational data. In contrast to ITT in the target trial, the per-protocol design of the trial emulation estimated actual treatment received within the grace period. The results showed survival benefit for surgery due to a reduction in aneurysm-related deaths. The findings were robust to grace period alterations and missing data. This is the first study of surgery for TAA using CCW to estimate the incidence of cause-specific deaths.

Surgery is a single intervention, without preoperative conditions (eg, weight loss), although some aneurysm repairs may be completed in stages [29,30]. In this study, actual treatment was labeled "surgery" if the first stage began within the grace period, avoiding the immortal time bias that would occur if analysis was based on eventual surgery status, enabling more accurate comparison of survival outcomes. Interventions with a treatment duration

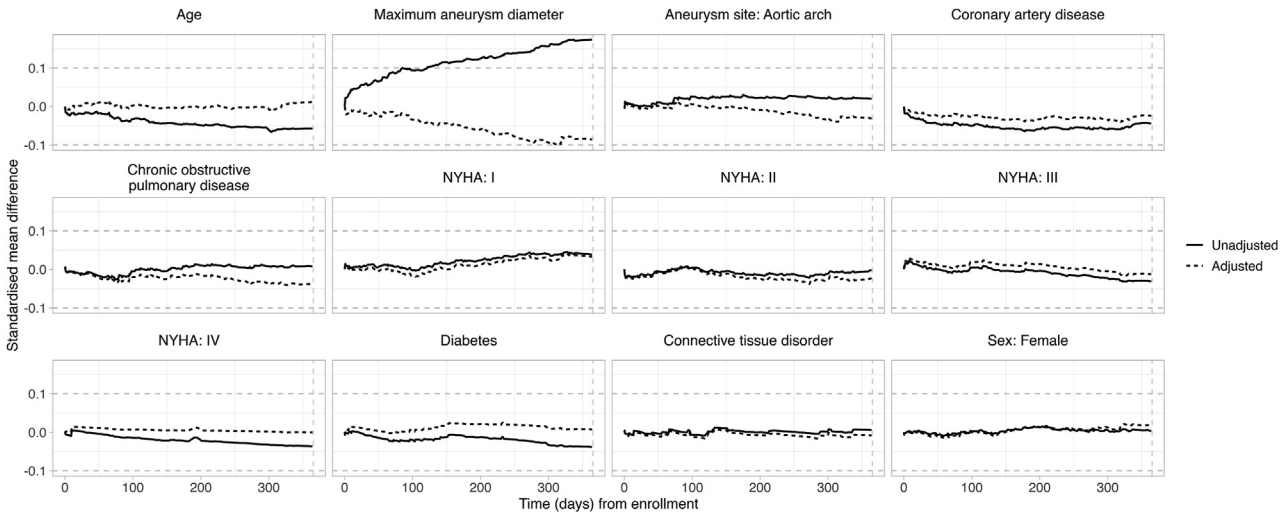

Figure 3. Daily estimates of the standardized mean differences for key confounders between the two groups, without and with adjustment for inverse probability of treatment weighting. Values within ± 0.1 are considered well balanced.

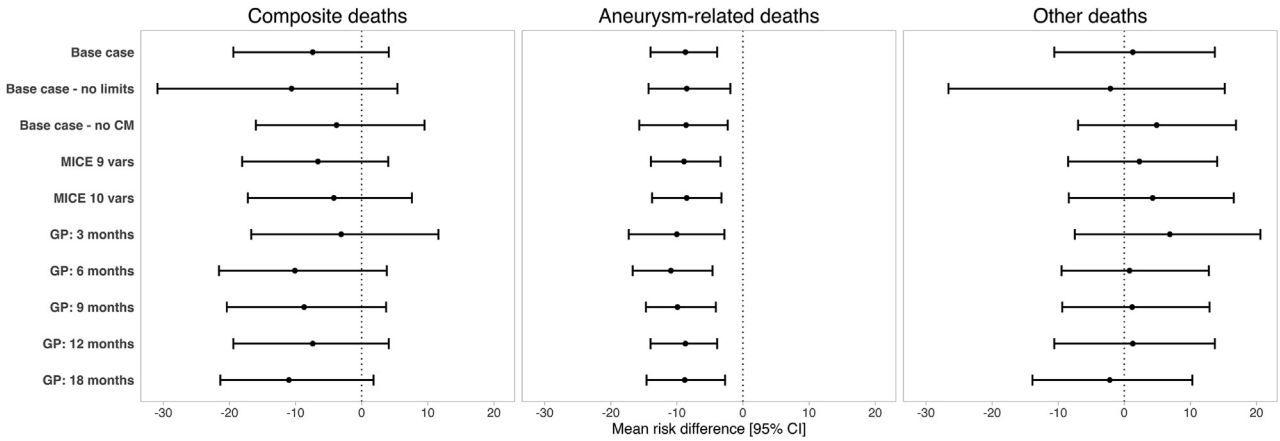

Figure 4. Estimated difference in cumulative incidence at 7 years (95% CI) for overall survival (composite), aneurysm-related death, and death from other causes for the base case and sensitivity to model assumptions (no limits on weights, exclusion of conservative management patients, imputation of 9 or 10 key variables) and grace period (GP) of 3, 6, 9, 12, and 18 months.

(eg, medication), intermittent prescriptions, or specific preoperative preparation require more careful definition of the estimand and related grace periods [31−34].

Patients without planned surgery would be unlikely to enter a target trial, and we could have excluded them from analysis under the positivity assumption. They were retained in our base case because they were eligible and may have been accepted for surgery at a different hospital. Excluding them resulted in lower nonaneurysm mortality for no surgery clones, suggesting that these patients were at high risk of death from comorbidities.

This study highlights benefits of using cohort data in trial emulation. This prospective observational study allowed the identification of eligible patients who declined or had no planned surgery. Capturing treatment decisions and their rationale is challenging in EHR because these details are unrecorded. During ETTAA, sites were more likely to classify deaths as aneurysm-related (about half) compared to EHR follow-up (a quarter). Inconsistencies in cause-of-death reporting between observational studies and EHR may lead to biased estimates of aneurysm-related mortality. Finally, aneurysm size, a critical determinant of intervention timing, is typically found in hospital radiology reports but absent in EHR. Data extraction would require text mining or manual reassessment. Without this data, addressing differences in aneurysm size between surgery and control groups from EHR is not possible, potentially invalidating comparative analyses.

Though prospective studies like ETTAA provide detailed data, their sample size is small relative to EHR, especially for aneurysm-related deaths, resulting in imprecise estimates. Combining ETTAA with EHR using methods such as constrained nonparametric maximum likelihood or empirical Bayes may provide deeper insight into the potential effects of thoracic aneurysm repair strategies [35].

For valid inference, it is important to fully adjust for confounders. Our adjustment for baseline variables was robust, based on literature, formal and informal expert elicitation, and empirical analysis. We did not have access to time-varying confounders beyond ETTAA, in particular, aneurysm diameter. Should these measures become available, they can be accommodated by expanding the CCW method or by jointly modeling growth and mortality [36].

5. Conclusion

For the first time, we demonstrated the utility of trial emulation and the CCW approach to estimate effects of aneurysm repair surgery on cause-specific deaths. These tools provide robust estimation of causal effects in observational studies where the analogous randomized trial(s) are infeasible.

We found benefit of giving all patients surgery within 12 months in reducing aneurysm-related deaths while treating deaths from other causes as a competing risk. Additional data and/or extended follow-ups are necessary for confirmation. Observational studies using EHR face challenges in reproducing these analyses, but synthesizing cohort studies with EHR remains a promising research area.

CRediT authorship contribution statement

James Murray: Writing − review & editing, Writing − original draft, Visualization, Software, Methodology, Formal analysis, Data curation. **Caroline Chesang:** Writing − review & editing, Visualization, Software, Methodology. **Steve Large:** Writing − review & editing, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization. **Colin Bicknell:** Writing − review & editing, Validation, Resources, Project administration, Investigation. **Carol Freeman:** Writing − review & editing, Resources, Project administration, Investigation, Funding acquisition, Data curation. **Ruth H. Keogh:** Writing − review & editing, Visualization, Supervision, Software, Methodology,

Conceptualization. **Linda D. Sharples:** Writing − review & editing, Writing − original draft, Visualization, Validation, Supervision, Software, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

C.B. reports personal fees and nonfinancial support from Medtronic, grants, personal fees, and nonfinancial support from Gore, all outside the submitted work. There are no competing interests for any other author.

Acknowledgments

ETTAA was supported by patient representatives from the Marfan Association UK and Liverpool Aneurysm Support group during study design and public communications, and one patient representative attended all trial steering committee meetings and provided advice throughout ETTAA.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2025.111714.

Data availability

The data used in this article cannot be shared publicly to maintain privacy of the individuals who participated, under the signed consent conditions. The data will be shared on reasonable request to the Chief Investigator of ETTAA, SRL (s.large@nhs.net). The R code used in all analysis is available online at https://github.com/jamesmurray7/ETTAA-CCW.

References

[1] Cook JA, McCulloch P, Blazeby JM, Beard DJ, Marinac-Dabic D, Sedrakyan A, et al. IDEAL framework for surgical innovation 3: randomised controlled trials in the assessment stage and evaluations in the long term study stage. BMJ 2013;346:f2820.

[2] Gaudino M, Kappetein AP, Di Franco A, Bagiella E, Bhatt DL, Boening A, et al. Randomized trials in cardiac surgery: JACC review topic of the week. J Am Coll Cardiol 2020;75:1593−604.

[3] Wong LY, Li Y, Elliott IA, Backhus LM, Berry MF, Shrager JB, et al. Randomized controlled trials in lung cancer surgery: how are we doing? JTCVS Open 2024;18:234−52.

[4] Gusho C, Hoskins W, Ghanem E. A comparison of surgical approaches for hip hemiarthroplasty performed for the treatment of femoral neck fracture: a systematic review and network meta-analysis of randomized controlled trials. JBJS Rev 2024;12:1−3.

[5] Hernán M, Robins JM. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC; 2020.

[6] Sharples L, Sastry P, Freeman C, Gray J, McCarthy A, Chiu YD, et al. Endovascular stent grafting and open surgical replacement for chronic thoracic aortic aneurysms: a systematic review and prospective cohort study. Health Technol Assess 2022;26:1−166.

[7] Salem O, El Beyrouti H, Mulorz J, Schelzig H, Ibrahim A, Oberhuber A, et al. Predictors for reintervention and survival during long-term follow-up after thoracic endovascular aortic repair for descending thoracic aortic aneurysm. J Vasc Surg 2024;80:1408−14017.

[8] Haji-Zeinali AM, Mansouri P, Raeis Hosseini N, Abbasi K, Shirzad M, Jameie M, et al. Five-year survival and complications of thoracic endovascular aortic repair (TEVAR): a single tertiary center registry for all-comers patients. Cardiovasc Revasc Med 2023;51:23−30.

[9] Clouse WD, Hallett JW Jr, Schaff HV, Gayari MM, Ilstrup DM, Melton LJ 3rd. Improved prognosis of thoracic aortic aneurysms: a population-based study. JAMA 1998;280:1926−9.

[10] Davies RR, Goldstein LJ, Coady MA, Tittle SL, Rizzo JA, Kopf GS, et al. Yearly rupture or dissection rates for thoracic aortic aneurysms: simple prediction based on size. Ann Thorac Surg 2002;73:17−27.

[11] Oladokun D, Patterson BO, Sobocinski J, Karthikesalingam A, Loftus I, Thompson MM, et al. Systematic review of the growth rates and influencing factors in thoracic aortic aneurysms. Eur J Vasc Endovasc Surg 2016;51:674−81.

[12] Zafar MA, Chen JF, Wu J, Li Y, Papanikolaou D, Abdelbaky M, et al. Natural history of descending thoracic and thoracoabdominal aortic aneurysms. J Thorac Cardiovasc Surg 2021;161:498−511.

[13] Huang Y, Schaff HV, Dearani JA, Oderich GS, Bower TC, Kalra M, et al. A population-based study of the incidence and natural history of degenerative thoracic aortic aneurysms. Mayo Clin Proc 2021;96:2628−38.

[14] Huang Y, Schaff HV, Bagameri G, Pochettino A, DeMartino RR, Todd A, et al. Differential expansion and outcomes of ascending and descending degenerative thoracic aortic aneurysms. J Thorac Cardiovasc Surg 2024;167:918−26.

[15] Sharples L, Sastry P, Freeman C, Bicknell C, Chiu YD, Vallabhaneni SR, et al. Aneurysm growth, survival, and quality of life in untreated thoracic aortic aneurysms: the effective treatments for thoracic aortic aneurysms study. Eur Heart J 2022;43:2356−69.

[16] Gray J, McCarthy A, Samarakoon D, McMeekin P, Sharples L, Sastry P, et al. Costs of endovascular and open repair of thoracic aortic aneurysms. Br J Surg 2024;111:znad378.

[17] McCarthy A, Gray J, Sastry P, Sharples L, Vale L, Cook A, et al. Systematic review of endovascular stent grafting versus open surgical repair for the elective treatment of arch/descending thoracic aortic aneurysms. BMJ Open 2021;11:e043323.

[18] Maringe C, Benitez Majano S, Exarchakou A, Smith M, Rachet B, Belot A, et al. Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. Int J Epidemiol 2020;49:1719−29.

[19] Kurz CF. Augmented inverse probability weighting and the double robustness property. Med Decis Mak 2022;42:156−67.

[20] Young JG, Hernan MA, Robins JM. Identification, estimation and approximation of risk under interventions that depend on the natural value of treatment using observational data. Epidemiol Methods 2014;3:1−19.

[21] Shiba K, Kawahara T. Using propensity scores for causal inference: pitfalls and tips. J Epidemiol 2021;31:457−63.

[22] Upchurch GR Jr, Escobar GA, Azizzadeh A, Beck AW, Conrad MF, Matsumura JS, et al. Society for Vascular Surgery clinical practice guidelines of thoracic endovascular aortic repair for descending thoracic aortic aneurysms. J Vasc Surg 2021;73:55S−83S.

[23] Riambau V, Bockler D, Brunkwall J, Cao P, Chiesa R, Coppi G, et al. Editor's choice - management of descending thoracic aorta diseases: clinical practice guidelines of the European society for vascular surgery (ESVS). Eur J Vasc Endovasc Surg 2017;53:4−52.

[24] Hiratzka LF, Bakris GL, Beckman JA, Bersin RM, Carr VF, Casey DE Jr, et al. 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM

guidelines for the diagnosis and management of patients with thoracic aortic disease: a report of the American college of cardiology foundation/American heart association task force on practice guidelines, American association for thoracic surgery, American college of radiology, American stroke association, society of cardiovascular anesthesiologists, society for cardiovascular angiography and interventions, society of interventional radiology, society of thoracic surgeons, and society for vascular medicine. Circulation 2010;121:e266−369.

[25] Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernan MA. A causal framework for classical statistical estimands in failure-time settings with competing events. Stat Med 2020;39:1199−236.

[26] van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations inR. J Stat Software 2011;45.

[27] Bartlett JW, Hughes RA. Bootstrap inference for multiple imputation under uncongeniality and misspecification. Stat Methods Med Res 2020;29:3533−46.

[28] Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology 2010;21:383−8.

[29] National Institute for Health and Care Excellence (NICE). Pre- and postoperative interventions to optimise outcomes after abdominal aortic aneurysm repair. London: National Institute for Health and Care Excellence (NICE); 2020.

[30] National Institute for Health and Care Excellence (NICE). Cardiovascular disease: risk assessment and reduction, including lipid modification. 2023. Available at: https://www.nice.org.uk/guidance/ng238/resources/cardiovascular-disease-risk-assessment-and-reduction-including-lipid-modification-pdf-66143902851781. Accessed February 26, 2025.

[31] Wanis KN, Sarvet AL, Wen L, Block JP, Rifas-Shiman SL, Robins JM, et al. Grace periods in comparative effectiveness studies of sustained treatments. J R Stat Soc Ser A Stat Soc 2024;187:796−810.

[32] Madenci AL, Kurgansky KE, Dickerman BA, Gerlovin H, Wanis KN, Smith AD, et al. Estimating the effect of bariatric surgery on cardiovascular events using observational data? Epidemiology 2024;35:721−9.

[33] Katsoulis M, Leyrat C, Hingorani A, Gomes M. Bariatric surgery and cardiovascular disease: the target trial emulation framework provides transparency in articulating the limits of observational studies. Epidemiology 2024;35:730−3.

[34] NHS England. Appendix 9 Guidance for Clinical Commissioning Groups (CCGs): Service Specification Guidance for Obesity Surgery. 2016. Available at: https://www.england.nhs.uk/wp-content/uploads/2016/05/appndx-9-serv-spec-ccg-guid.pdf. Accessed February 26, 2025.

[35] Gu T, Taylor JMG, Cheng W, Mukherjee B. Synthetic data method to incorporate external information into a current study. Can J Stat 2019;47:580−603.

[36] Rizopoulos D, Taylor JM, Papageorgiou G, Morgan TM. Using joint models for longitudinal and time-to-event data to investigate the causal effect of salvage therapy after prostatectomy. Stat Methods Med Res 2024;33:894−908.