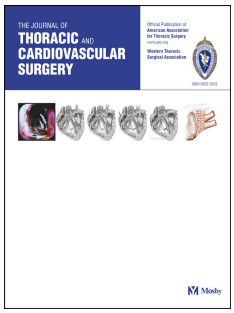
Article:
Nair, SK, Sudarshan, CD, Thorpe, BS et al. (2018) Mini-Stern Trial: A randomised trial comparing mini-sternotomy to full median sternotomy for aortic valve replacement. Journal of Thoracic and Cardiovascular Surgery, 156 (6). pp. 2124-2132. ISSN: 0022-5223

https://doi.org/10.1016/j.jtcvs.2018.05.057

Accepted Manuscript



Mini-Stern Trial: A randomised trial comparing mini-sternotomy to full median sternotomy for aortic valve replacement

Sukumaran K. Nair, FRCS(CTh), Catherine D. Sudarshan, FRCS(CTh), Benjamin S. Thorpe, PhD, Jeshika Singh, PhD, Thasee Pillay, FRCS(CTh), Pedro Catarino, FRCS(CTh), Kamen Valchanov, MD, Massimiliano Codispoti, FRCS(CTh), John Dunning, FRCS(CTh), Yasir Abu-Omar, FRCS(CTh), Narain Moorjani, FRCS(CTh), Claire Matthews, BSc, Carol J. Freeman, MPhil, Julia A. Fox-Rushby, PhD, Linda D. Sharples, PhD.

PII: S0022-5223(18)31482-X

DOI: 10.1016/j.jtcvs.2018.05.057

Reference: YMTC 13056

To appear in: *The Journal of Thoracic and Cardiovascular Surgery*

Received Date: 17 December 2017

Revised Date: 23 April 2018

Accepted Date: 15 May 2018

Please cite this article as: Nair SK, Sudarshan CD, Thorpe BS, Singh J, Pillay T, Catarino P, Valchanov K, Codispoti M, Dunning J, Abu-Omar Y, Moorjani N, Matthews C, Freeman CJ, Fox-Rushby JA, Sharples LD, Mini-Stern Trial: A randomised trial comparing mini-sternotomy to full median sternotomy for aortic valve replacement, *The Journal of Thoracic and Cardiovascular Surgery* (2018), doi: 10.1016/j.jtcvs.2018.05.057.

Mini-Stern Trial: A randomised trial comparing mini-sternotomy to full median sternotomy for aortic valve replacement

Authors:

Sukumaran K Nair FRCS(CTh),[1,4] Catherine D Sudarshan FRCS(CTh),[1] Benjamin S Thorpe PhD,[2] Jeshika Singh PhD,[3] Thasee Pillay FRCS(CTh),[4] Pedro Catarino FRCS(CTh),[1] Kamen Valchanov MD,[1] Massimiliano Codispoti FRCS(CTh),[1] John Dunning FRCS(CTh),[1] Yasir Abu-Omar FRCS(CTh), [1] Narain Moorjani FRCS(CTh),[1] Claire Matthews BSc,[1] Carol J Freeman MPhil,[1] Julia A Fox-Rushby PhD,[3] Linda D Sharples PhD.[5]

Institutions:

1. Papworth Hospital, Cambridge, UK
2. Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, UK
3. Health Economics Research Group (HERG), Brunel University London, London, UK
4. Freeman Hospital, Newcastle upon Tyne, Newcastle upon Tyne, UK
5. London School of Hygiene and Tropical Medicine, Keppel Street, London, UK

Conflict of interest statement and source of funding:

This work was supported by the National Institute for Health Research (NIHR) under its Research for Patient Benefit (RfPB) Programme (grant number PB-PG-0408-16296). The views expressed are those of the author(s) and not necessarily those of the NHS, NIHR or Department of Health.

Corresponding author and contact information:

Mr Sukumaran Nair, Consultant Cardiac Surgeon, Golden Jubilee National Hospital, Agamemnon Street, Glasgow, G81 4DY, United Kingdom

Telephone: 0044 141 951 5280.

Email: Sukumaran.Nair@gjnh.scot.nhs.uk

National Research Ethics Service Approval: 09/H0301/58

Trial Registration: ISRCTN, number 58128724

Article word count: 3500 (including main text and figure legends)

Glossary of Abbreviations

AVR	aortic valve replacement
mAVR	minimal access aortic valve replacement
BMI	body mass index
CI	95% confidence interval
COPD	chronic obstructive pulmonary disease
CPB	cardiopulmonary bypass
FEV_1	forced expiratory volume in one second
FS	full median sternotomy
HR	hazard ratio
HRQoL	health-related quality of life
ICER	incremental cost-effectiveness ratio
LVEF	left ventricular ejection fraction
MS	mini-sternotomy
NHS	National Health Service
OR	odds ratio
QALY	quality-adjusted life year
RCT	randomised control trial
SAE	serious adverse event
SD	standard deviation
TLCO	transfer factor of the lung for carbon monoxide
TOE	transoesophageal echocardiogram
UK	United Kingdom

Central Message

In the UK NHS, compared to conventional median sternotomy approach for surgical AVR, mini-sternotomy did not hasten recovery or hospital discharge, and was not cost-effective.

58 **Perspective Statement**

59 Minimal access surgery is appealing for its perceived advantages including better patient
60 recovery, satisfaction and cost-effectiveness. This RCT conducted within the UK NHS
61 setting did not demonstrate quicker patient recovery or cost-effectiveness associated with
62 mini-sternotomy compared to full median sternotomy approach. These findings are relevant
63 to physicians, patients and health care funders.

64

Structured Abstract

Objective: Aortic valve replacement (AVR) can be performed either through full median sternotomy (FS) or upper mini-sternotomy (MS). The Mini-Stern trial aimed to establish whether MS leads to quicker postoperative recovery and shorter hospital stay after first-time isolated AVR.

Methods: This pragmatic, open-label, parallel RCT compared MS with FS for first-time isolated AVR in two UK NHS hospitals. Primary endpoints were duration of postoperative hospital stay and the time to fitness for discharge from hospital after AVR, analysed in the intent-to-treat population.

Results: In this RCT, 222 patients were recruited and randomised (118 MS, 104 FS). Compared to FS patients, MS patients had longer hospital stay (mean 9.5 vs. 8.6 days) and took longer to achieve fitness for discharge home (mean 8.5 vs. 7.5 days). Adjusting for valve type, sex and surgeon, hazard ratios (HR) from Cox models did not show a statistically significant effect of MS (relative to FS) on either hospital stay (HR 0.874, 95% CI 0.668-1.143, p-value 0.3246) or time to fitness for discharge (HR 0.907, 95% CI 0.688-1.197, p-value 0.4914). During mean follow up of 760 days (MS:745 and FS:777 days), 12 (10%) MS and 7 (7%) FS patients died (HR 1.871, 95% CI 0.723-4.844, p-value 0.1966). Average extra cost for MS was £1,714, during the first 12 months after AVR.

Conclusions: Compared to FS for AVR, MS did not result in shorter hospital stay, faster recovery or improved survival and was not cost-effective. MS approach is not superior to FS for performing AVR.

Word count for Abstract: 248

Introduction

Aortic valve replacement (AVR) is the second commonest cardiac surgery in the UK [1] with an increasing proportion of older patients [1, 2]. Minimal access AVR (mAVR) might shorten hospital stay and postoperative recovery period and could be beneficial if offered safely and cost-effectively.

Currently, most AVRs are performed safely through full median sternotomy (FS) [2-6]. However, mAVR may be associated with less postoperative pain, blood loss, pulmonary and wound complications and shorter hospital stay [2]. The most commonly practised mAVR involves mini-sternotomy (MS), which could potentially hasten postoperative recovery, shorten hospital stay and improve patient satisfaction [2-10].

Most studies comparing MS and FS for AVR are non-randomised. Although systematic reviews with meta-analyses [11, 12] have been conducted, inadequate statistical power and heterogeneity of studies calls for prospective, randomised control trials (RCTs) to assess benefits and risks of mAVR. Published evidence on cost-effectiveness comparing MS to FS is sparse and weak. A recent review comparing cost-effectiveness of FS and MS called for a well-designed RCT to evaluate cost-effectiveness of mAVR up to at least a year after surgery [13]. Recently, a propensity-matched study from the UK national data concluded that mAVR is safe and was associated with shorter postoperative hospital stay [14]. The authors concluded that although general clinical equipoise exists between FS and MS, it is essential to have a well-constructed and adequately powered RCT before widespread adoption of MS. This retrospective study did not analyse cost-effectiveness of either surgical approach.

The Mini-Stern trial assessed whether MS is superior to FS in shortening postoperative recovery time and improving patient outcomes without compromising patient safety. It also assessed cost-effectiveness of MS from the perspective of the UK NHS as a health care provider.

Materials and Methods

Mini-Stern was a two-centre, pragmatic, open-label RCT conducted in the UK. Patients were randomised (1:1) to AVR either by MS or FS.

Sample Size

Considering four published RCTs [5, 6, 9, 10] and two cohort studies [7, 8], a 20% reduction in hospital stay from 11.7 to 9.36 days was considered clinically significant. Based on an internal audit of 252 first-time elective AVRs performed at Papworth Hospital in 2007/08 (mean hospital stay 11.7 days, SD 6.2), to detect this change with 80% power and 2-sided significance of 5%, 110 patients per group were required. As randomisation was performed on the day of surgery after induction of anaesthesia and introduction of the transoesophageal echocardiogram (TOE) probe, no subjects dropped out between randomisation and surgery thereby making the total trial recruitment target, 220 patients.

Recruitment

Adult patients undergoing first-time isolated AVR were included. Exclusion criteria included emergency AVR, LVEF$\leq$ 30%, chest wall deformities, severe COPD (FEV_1 or TLCO < 40% predicted), BMI > 35kg/m^2, concomitant cardiac surgery, redo-surgery and inability to perform TOE. Details of patient enrolment are given in the online protocol.

Randomisation

Randomisation (1:1) used random permuted blocks of variable lengths (6 or 8), stratified by surgeon and valve prosthesis (bio-prosthetic or mechanical). Random allocations were pre-generated, held in secure files by Papworth Trials Unit. During early days of the trial, TOE probe could not be passed in four patients due to technical reasons. These patients underwent the allocated procedure and were included in the trial. Later the Trial Steering Committee decided that under such circumstances, MS would be unsafe and patients should be excluded from the trial to FS. Since eligibility for MS required TOE, in order to avoid post-randomisation drop-out, group allocation for the study subjects was retrieved via telephone by theatre staff soon after anaesthesia and introduction of the TOE probe. Due to the nature of interventions, this trial could not be blinded.

Outcomes

Primary endpoints: Two closely related primary endpoints were measured. Firstly, length of postoperative hospital stay (days between surgery and actual hospital discharge) which is easily measured, a surrogate for early postoperative events and sensitive to outcomes that affect health-related quality of life (HRQoL). Secondly, the interval in days between surgery and the patient being medically fit for discharge. To reduce investigator bias, standard discharge criteria were followed to decide the day of fitness for discharge. This endpoint was chosen to address exogenous effects (social factors, lack of transport, non-availability of space in nursing homes etc.) that commonly delay hospital discharge in the UK.

Clinical secondary endpoints: duration of surgery, total theatre time, aortic cross-clamp and cardiopulmonary bypass (CPB) times, blood loss in the first 12 hours after surgery, transfusion of blood and clotting products in the first 48 hours (blood transfusion trigger was

haemoglobin level < 80g/L), frequency of re-intubation, time to initial extubation, mediastinal drain removal and first independent mobilisation, daily pain scores at rest and on deep breath (over the first ten days or until hospital discharge) on a scale of 0 to 10, LVEF and severity of para-prosthetic regurgitation at hospital discharge and at 6 months, and time to all-cause death. Definitions of adverse events and details of their reporting are in the online protocol. To exclude bias, clinical outcome data were collected by research team who were not involved in routine care of subjects, following standardised protocols.

Non-clinical secondary endpoints: Health-related Quality of Life and Healthcare resource use.

HRQoL: Patients completed EQ-5D-3L [15] and SF-36 [16, 17] questionnaires at baseline, 6 weeks, 6 months and 12 months following surgery. EQ-5D-3L was repeated on fourth postoperative day and at discharge.

Healthcare resource use: Patient-specific resource use collected from hospital records and patient interviews during the primary admission included phases of care including operative surgery, critical care, post-surgical ward care and medications. Post-discharge resource use included attending wound clinics, community nurse visits, physiotherapy sessions, occupational therapy services, medical tests, cost of analgesics and other drugs and further hospitalisation within the first year after AVR.

Surgical details

All participating surgeons were consultants experienced in performing AVR by both FS and MS. They followed the operative surgical protocol as described below.

MS approach: With the patient anaesthetised as per standard protocol, skin was incised from half-way between the suprasternal notch and the sternal angle to the level of the fourth

intercostal space, measuring approximately 8cm. The manubrium was divided in the midline from the suprasternal notch inferiorly and then into the right 4th intercostal space. Thymus was divided and pericardium opened exposing the ascending aorta, aortic root and right atrial appendage. A loading dose of unfractionated heparin 300U/kg followed by boluses of 5000U was administered to achieve activated clotting time above 450 seconds. Aorta was cannulated using a wired flexible aortic cannula. Right atrial appendage was cannulated using a flat venous cannula and CPB commenced. The ascending aorta was cross-clamped and intermittent, antegrade, cold blood cardioplegia administered. The aorta was then incised open in an oblique or transverse fashion, the diseased valve excised and annulus decalcified. A suitably sized aortic valve prosthesis was inserted using either horizontal mattress, 2-0 Ethibond sutures or semi-continuous, 2-0 Prolene sutures. Surgeons adopted either of these suture techniques and adhered to the same technique irrespective of the type of valve prosthesis or the surgical approach. Aortotomy was then closed, heart de-aired, right atrial and ventricular epicardial pacing wires inserted and patient weaned off CPB. After confirming satisfactory functioning of the aortic valve prosthesis by TOE, heparin was reversed with protamine (1mg/100U of heparin). Chest drains were inserted into the anterior mediastinum, posterior pericardial space and pleural space if necessary. Sternal wires were inserted and incision closed in layers. Conversion to FS was performed to ensure patient safety if access was difficult or if intraoperative complications occurred.

FS approach: Anaesthesia and positioning of patients was the same as for MS approach. The skin incision was made between the suprasternal notch and the xiphoid process and sternum divided in the midline from the suprasternal notch to the xiphoid process. A two-stage venous cannula was used for atrial cannulation. Remaining steps were similar to MS approach.

Statistical analysis

Analyses of primary and secondary endpoints used intention-to-treat and included all randomised patients. Unless stated otherwise, statistical models included treatment (MS vs. FS), valve (mechanical vs. bio-prosthetic) and sex as fixed effects, and surgeons as random effects. Hypothesis testing was two-sided at the 5% significance level, with no adjustments for multiple testing. All confidence intervals (CI) were estimated at the 95% confidence level.

Distributions of time-to-event endpoints were compared between study groups using Kaplan-Meier curves and log-rank tests (stratified by sex, valve and surgeon). Hazard ratios (HR) for MS relative to FS were estimated from a Cox model. The null hypothesis of no treatment effect (HR = 1) was tested. Patients who were lost to follow-up, withdrew or died before the event were censored at the latest time they were known to be event-free. Models were checked by plotting Schoenfeld and deviance residuals. For primary endpoints, Cox models were re-fitted using the per-protocol population and in sensitivity analyses (Appendix A. Table A4).

Need for reintubation and other dichotomous endpoints were compared between groups by estimating a MS/FS odds ratio (OR) via logistic regression. EQ-5D, SF-36 and pain scores were modelled using repeated measures linear regression. Where possible, random intercepts and random time coefficients for patients were included. For EQ-5D and SF-36, fixed effects for baseline scores were included. Models were fitted using complete cases, then re-fitted with multiple imputation of missing scores via chained equations.

Serious adverse events (SAEs) were analysed in the safety population according to intervention received. Patients randomised to MS who crossed over to FS prior to surgery were considered to have received FS; those who crossed over after MS had commenced were

considered to have received MS. Rates of SAEs were explored using Poisson regression with a random patient effect.

CONSORT guidelines [18] were followed. Analyses were performed in SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). No interim analyses were undertaken but reports were presented annually to the Data Monitoring and Ethics Committee.

Economic analysis

Unit costs were obtained from nationally published sources in the UK [19, 20, 21, 22] or from the Finance department, Papworth Hospital when the former did not provide the required information. Total cost per patient was calculated by summing resource use items multiplied by unit costs across the in-patient stay and the 12-month postoperative follow-up period (Appendix B. Table B7). Health state utilities from the EQ-5D-3L and SF-36, based on UK value sets [15, 23] were used to generate quality-adjusted life years (QALYs) using the area under the curve method and assigning a value of zero from date of death. Missing values were imputed using chained predictive mean matching, stratified by treatment and conditional on age, sex and baseline EQ-5D-3L.

Differences in mean costs and QALYs were estimated using seemingly unrelated regression, controlling for age, sex, valve, baseline EQ-5D-3L and treatment, to accommodate skewness [24]. Uncertainty in cost-effectiveness was estimated by drawing 1000 bootstrapped samples and conducting probabilistic sensitivity analysis. Results are presented as incremental net monetary benefit at various thresholds of willingness to pay per QALY, cost-effectiveness planes and cost-effectiveness acceptability curves. Deterministic sensitivity analyses explored effects of using complete cases only, SF6D-based QALY estimates, the procedure inpatient

admission only, excluding patients who died and excluding additional equipment costs (Appendix B. Table B11).

Results

Overall 1024 patients were screened between 28 January 2010 and 13 April 2015, of whom 222 were recruited and randomised to MS (118) or FS (104). One-year follow-up was completed on 23 May 2016.

Study groups were similar at baseline except for a non-significant sex imbalance (Table 1). In this trial, MS was not completed in 14 (12%) of 118 patients randomised to MS. Of these patients, 6 (5%) had conversion from MS to FS due to reasons listed in Figure 1. The remaining 8 patients underwent FS after randomisation to MS but without initial MS incision as MS was considered unsafe/impractical. The true rate of intraoperative conversion of MS to FS was therefore 5%. Four patients (2%, Table 2) were censored before discharge: one withdrawal before surgery (FS) and three deaths (all randomised to and received MS). A further thirteen (6%) were censored before fitness for discharge: six discharged to acute hospital (three MS, three FS), seven to long-term care or rehabilitation (three FS, four MS).

Mean time to hospital discharge was longer for MS than FS (9.5 vs. 8.6 days), as was mean time to fitness to discharge (8.5 vs. 7.5 days). However, distributions of these endpoints were similar in both groups (Figure 2, Table 2). The difference was not statistically significant in either primary analyses using Cox models (Figure 3), log-rank tests (Table 2) or sensitivity analyses (Appendix A. Table A4). The gamma-distributed frailty term in the Cox models was estimated to have variance 0.006675 for time to fitness and 0.000100 for time to discharge, suggesting that surgeon heterogeneity was negligible.

Time to drain removal (including drains inserted/retained to treat complications) was longer for MS, but times to extubation and independent mobilisation did not differ significantly between groups (Table 2, Figure 3), nor did numbers of patients re-intubated (six MS vs. five FS, OR 1.039, CI 0.306-3.531, p=0.9512). Statistically significant HRs indicated longer surgery, CPB, cross-clamp and theatre times for MS (Figure 3). No significant differences were seen in blood loss (Appendix A. Table A3), or in numbers of patients requiring transfusion of blood (50 MS vs. 51 FS, OR 0.797, CI 0.453-1.402, p=0.4310) or clotting products (11 MS vs 4 FS, OR 2.616, CI 0.801-8.541, p=0.1112).

Regression models for pain at rest, EQ-5D utilities and SF-36 domain scores (Appendix A. Tables A6, A7, A8) estimated greater rate of improvement over time in MS patients for three SF-36 domains (social functioning, vitality and role physical). After multiple imputation, the difference was only significant for the role physical domain (Appendix A. Table A9). Pain on deep breath was not analysed as only less than half the data were collected due to poor patient compliance.

Nine (4%) patients died within a year of surgery: seven (6%) MS, two (2%) FS. Five deaths were possibly related to treatment (four MS, one FS), none were probably or definitely related (Appendix A. Table A15). Overall, twelve (10%) MS and seven (7%) FS patients died during follow-up (mean follow-up 760 days: 745 MS, 777 FS). Time to all-cause death, adjusted for age, showed a moderately large but statistically non-significant HR (MS/FS) of 1.871 (CI 0.723-4.844, p=0.1966).

Safety analyses excluded one patient who was withdrawn before surgery. There were significantly more SAEs in MS recipients (rate ratio 1.615, CI 1.070-2.437, p=0.0225) (Appendix A. Table A11). The numbers of patients experiencing SAEs were not significantly different (OR 1.559, CI 0.895-2.715, p=0.1161). Incidence of para-prosthetic

regurgitation did not differ significantly between groups (Appendix A. Table A13). Seven patients developed pericardial collection (three MS vs four FS, OR 0.680, CI 0.146-3.178, p=0.6229). Wound infections (including superficial and deep infections) were more common in FS recipients (thirteen FS vs four MS, OR 0.312, CI 0.097-1.005, p=0.0511). Deep sternal wound infection developed in one MS and one FS recipient, neither of whom required plastic surgical repair.

Economic analyses are summarised in Table 4. There was additional cost for MS relative to FS (£1,714 per patient, p=0.0765) in the first year following surgery. MS patients had (non-significant) better EQ-5D-based QALYs (0.03 per patient, p=0.1509). The incremental cost per QALY gained was £61,379, but after adjusting for baseline characteristics, MS had higher costs and lower QALYs (i.e. was dominated). In deterministic and probabilistic sensitivity analyses, MS was either dominated or had a very large cost per QALY, except for the complete case analysis (Appendix B. Tables B11, B12).

Discussion

The UK NHS is a free for patient at point-of-delivery healthcare system. Apart from good recovery, hospital discharge of a significant proportion of elderly patients depends on the timely availability of social care services in the community. The Mini-Stern trial is the first RCT comparing FS and MS for isolated AVR when performed for UK NHS patients.

In this prospective, pragmatic, open-label RCT, MS did not reduce the total duration of hospital stay after AVR. As hospital discharge is sometimes delayed due to social factors, we included time until fit for discharge as a second primary endpoint. This was also not reduced by MS. These endpoints were recorded by physiotherapists based on a common discharge

protocol with specific clinical milestones to achieve, thereby excluding physician-induced bias.

In this study operation, total theatre, aortic cross-clamp and CPB times were significantly prolonged with MS. This was expected as in general, minimal access valve operations take longer [5, 9]. This is justifiable if MS resulted in either faster recovery, shorter postoperative stay, reduced cost of treatment or more importantly a significant reduction in adverse events and therefore superior patient safety. In this RCT, MS did not achieve these benefits and hence we feel that the prolonged operation time, total theatre, cross-clamp and CPB times are not justifiable for performing AVR through MS.

Previously, two meta-analyses [11, 12] concluded that mAVR approaches are superior in certain aspects of postoperative recovery. However, both included studies on mini-thoracotomy approach for AVR, and therefore inferences drawn cannot be extrapolated to MS. A retrospective propensity-matched analysis of data from a UK national database concluded that MS is safe and comparable to conventional AVR [14]. The authors found that MS resulted in a shorter postoperative hospital stay, which disagrees with our findings. However, a propensity-matched study can suffer from selection bias if its matching algorithm produces treatment groups that are unbalanced in some unobserved characteristics. Recently, a retrospective study demonstrated safety of right thoracotomy minimally invasive isolated and concomitant AVR in patients of all age groups [25]. As randomisation balances study groups in known and unknown characteristics, results of the Mini-Stern trial should be more reliable than non-randomised studies.

Previous studies investigating cost-effectiveness provided unclear answers. A report analysing registry data from patients who underwent isolated primary AVR [26] reported lower hospital cost when AVR was performed through right anterior thoracotomy compared to sternotomy-based approaches with no significant differences in outcome. The main reasons attributed to lower costs were earlier hospital discharge and reduced use of blood products. Ghanta et al [27] noted that exclusion of rehabilitation costs could alter this finding. A review by Glauber et al [13], based on uncontrolled studies, noted that higher cost of instruments and devices in mAVR could be offset by economic advantage gained by shorter hospital stay and lower complication rates. The Mini-Stern trial assessed cost-effectiveness using a range of sensitivity analyses, but only the complete case analysis showed MS to be cost-effective, suggesting lower costs but slightly worse outcomes with MS. However, this analysis used a potentially unrepresentative sample of just 90 patients. Our analysis was restricted to the first year following operation without long-term analysis beyond 1 year.

This RCT is robust with many merits including on-table randomisation, comprehensive and independent outcome assessment without physician-bias, longer-term clinical assessment, HRQoL analysis and economic analysis. However there were some limitations. Although we report on secondary endpoints, this trial was powered only to address the primary endpoint. A total of 14 patients (12%) allocated to MS received FS, which could be another limitation. However, only 6 patients (5%) had true conversion after an attempted MS, while 8 patients (6.7%) went on to FS for safety reasons. Although this RCT took place in only two centres, thereby limiting generalisability, recruitment by eight surgeons improves generalisability. A total of 1024 patients were screened to recruit 222 (21.7%) patients. Although this potentially suggests selection bias, only 125 eligible patients (12.2%) failed recruitment while the remaining 667 patients (65.1%) did not meet inclusion criteria. Blinding was not

practical as sternotomy dressings were usually changed 48 hours after surgery and patients became aware of the approach. This could have caused bias in self-reported outcomes. Missing 'pain at rest' data were unlikely to be missing at random, and therefore imputation might not have addressed all potential biases. Despite having two primary outcomes, we did not adjust for multiple testing. However, as neither showed a significant difference between groups, this would not have affected our conclusions.

In conclusion, MS for AVR did not result in quicker recovery or earlier hospital discharge. MS resulted in longer operations, increased costs, and resulted in more SAEs than FS. Overall, this pragmatic RCT did not provide evidence that MS results in better clinical or quality of life outcomes, or that MS is cost-effective compared to FS in the first year after AVR.

Acknowledgement:

Prof Linda Sharples was part-funded by the MRC, UK. We thank Dr Matt Glover, Ms Jacinta Nalpon and Mr Chelliah Paramasivam for their contributions.

401 **Legends**

402 **Central Picture Legend:** Duration of hospital stay after AVR: FS versus MS.

403 **Video Legend:** MS approach for AVR.

404 **Figure 1.** Trial flow diagram.

405 **Figure 2**. Kaplan-Meier curves for primary endpoints. Points indicate censoring and dashed
406 lines represent 95% confidence intervals.

407 **Figure 3**. Forest plot of HRs and 95% confidence intervals from Cox models.

408 **Figure 4**. Cost-effectiveness planes. Proportion of points below each threshold gives the
409 probability that MS is more cost-effective than FS. This probability is 3.7% for willingness to
410 pay £20,000/QALY and 5.1% for willingness to pay £30,000/QALY.

411

412 **Table 1. Baseline characteristics**

	MS (n = 118)	**FS (n = 104)**
Age (years) - Mean (SD)	71.3 (12.3)	72.1 (10.9)
BMI (kg/m^2) – Mean (SD)	26.6 (3.2)	27.7 (3.7)
Sex - frequency (%)		
Female	53 (45%)	57 (55%)
Male	65 (55%)	47 (45%)
Valve type - frequency (%)		
Mechanical	15 (13%)	14 (13%)
Tissue	103 (87%)	90 (87%)
EuroSCORE (%) - Mean (SD)	5.9 (2.1) *	6.1 (2.1)

413 * EuroSCORE was missing for one MS patient.

414

415

416

417 **Table 2. Kaplan-Meier medians (quartiles) for time-to-event endpoints**

	MS (n = 118)	FS (n = 104)	p-value*
Time to discharge (days)	7 (6, 10)	7 (6, 10)	0.6924
Censored	3	1	
Time until fit for discharge (days)	6 (5, 10)	6 (5, 9)	0.5597
Censored	10	7	
Time to independent mobilisation (days)	4 (3, 7)	4 (3, 6)	0.5819
Censored	8	7	
Time to mediastinal drain removal (hours)	26.1 (20.6, 53.3)	22.5 (19.4, 37.8)	0.0157
Censored	2	2	
Time to extubation (hours)	9.2 (7.8, 12.1)	8.3 (6.8, 11.7)	0.5488
Censored	1	1	
Theatre time (minutes)	191 (172, 225)	176 (152, 203)	< 0.0001
Censored	0	0	
CPB time (minutes)	80 (70, 95)	66 (52, 85)	< 0.0001
Censored	0	0	
Cross-clamp time (minutes)	65 (53, 76)	49 (39, 64)	< 0.0001
Censored	0	0	
Surgery duration (minutes)	163 (139, 190)	149 (114, 167)	< 0.0001
Censored	3	4	

418 **Log-rank test. Seven surgery durations were not recorded and censored at 1 minute.*

419

420 **Table 3. Costs, QALYs and Cost-effectiveness**

	Cost and QALYs (with imputation)	**FS (n = 118)**		**MS (n = 104)**	
		Mean Cost per patient	**SD**	**Mean Cost per patient**	**SD**
Primary Admission Costs	Theatre use	£3,824	£1,243	£4,422	£2,053
	Additional surgical items	£16.52	£0.0	£52.0	£0.0
	Critical care (ITU)	£1,834	£3,023	£2,934	£5,030
	Cardiac ward	£2,744	£1,664	£2,676	£1,500
	Physio- and Occupational Therapy	£77	£55	£78	£68
	Rehabilitation	£384	£1,878	£263	£1,621
	Acute hospital	£347	£1,919	£298	£1,971
	Sub-total cost	*£9,226*	*£6,511*	*£10,724*	*£8,850*
Post primary admission costs to 12 months	Hospital Re-admission	£418	£1,475	£575	£1,863
	Follow up tests	£224	£258	£282	£279
	Follow up healthcare visits	£373	£359	£311	£263
	Sub-total cost	*£1,015*	*£1,778*	*£1,168*	*£2,079*
	Drugs	£379	£548	£441	£977
	Total cost over 12 months	*£10,620*	*£7,624*	*£12,333*	*£9,864*
Incremental cost-effectiveness* (probabilistic analysis with baseline	Incremental cost at 12 months (MS-FS)			£2,154.0 (*SE* £36)	
	Incremental EQ-5D-3L QALYs (MS-FS)			-0.0122 (SE 0.0008)	
	ICER			MS dominated by FS	
	NMB (at WTP £20,000/QALY)			-£2,397	
	NMB (at WTP £30,000/QALY)			-£2,519	

adjustment)
SD: standard deviation, SE: standard error, WTP: willingness to pay, NMB: net monetary benefit, ICER: incremental cost-effectiveness ratio. * Incremental costs and effects estimated using SUR, adjusting for baseline differences.

References

[1] The Society for Cardiothoracic Surgery in Great Britain & Ireland. http://bluebook.scts.org/#ActivityRates

[2] Rosengart TK, Feldman T, Borger MA, Vassiliades TA Jr, Gillinov AM, Hoercher KJ, et al. Percutaneous and minimally invasive valve procedures: a scientific statement from the American Heart Association Council on Cardiovascular Surgery and Anesthesia, Council on Clinical Cardiology, Functional Genomics and Translational Biology Interdisciplinary Working Group, and Quality of Care and Outcomes Research Interdisciplinary Working Group. Circulation. 2008;117:1750-67.

[3] Merk DR, Lehmann S, Holzhey DM, Dohmen P, Candolfi P, Misfeld M, et al. Minimal invasive aortic valve replacement surgery is associated with improved survival: a propensity-matched comparison. Eur J Cardiothorac Surg. 2015;47:11-7.

[4] Furukawa N, Kuss O, Aboud A, Schönbrodt M, Renner A, Hakim MK, et al. Ministernotomy versus conventional sternotomy for aortic valve replacement: matched propensity score analysis of 808 patients. Eur J Cardiothorac Surg. 2014;46:221-6.

[5] Bonacchi M, Prifti E, Giunti G, Frati G, Sani G. Does ministernotomy improve postoperative outcome in aortic valve operation? A prospective randomized study. Ann Thorac Surg. 2002;73:460-5.

[6] Moustafa MA, Abdelsamad AA, Zakaria G, Omarah MM. Minimal vs median sternotomy for aortic valve replacement. Asian Cardiovasc Thorac Ann. 2007;15:472-5.

[7] Sharony R, Grossi EA, Saunders PC, Schwartz CF, Ribakove GH, Culliford AT, et al. Minimally invasive aortic valve surgery in the elderly: a case-control study. Circulation. 2003;108 Suppl 1:II43-7.

[8] Bakir I, Casselman FP, Wellens F, Jeanmart H, De Geest R, Degrieck I, et al. Minimally invasive versus standard approach aortic valve replacement: a study in 506 patients. Ann Thorac Surg. 2006;81:1599-604.

[9] Aris A, Camara ML, Montiel J, Delgado LJ, Galan J, Litvan H. Ministernotomy versus median sternotomy for aortic valve replacement: a prospective, randomized study. Ann Thorac Surg. 1999;67:1583-7.

[10] Dogan S, Dzemali O, Wimmer-Greinecker G, Derra P, Doss M, Khan MF, et al. Minimally invasive versus conventional aortic valve replacement: a prospective randomized trial. J Heart Valve Dis. 2003;12:76-80.

[11] Lim JY, Deo SV, Altarabsheh SE, Jung SH, Erwin PJ, Markowitz AH, et al. Conventional versus minimally invasive aortic valve replacement: pooled analysis of propensity-matched data. J Card Surg. 2015;30:125-34.

[12] Phan K, Xie A, Di EM, Yan TD. A meta-analysis of minimally invasive versus conventional sternotomy for aortic valve replacement. Ann Thorac Surg. 2014;98:1499-511.

[13] Glauber M, Ferrarini M, Miceli A. Minimally invasive aortic valve surgery: state of the art and future directions. Ann Cardiothorac Surg. 2015;4:26-32.

[14] Attia RQ, Hickey GL, Grant SW, Bridgewater B, Roxburgh JC, Kumar P, et al. Minimally invasive versus conventional aortic valve replacement. A propensity-matched study from the UK National Data. Innovations. 2016:11:15-23.

[15] Dolan P, Gudex C, Kind P. A social tariff for EuroQoL: results from a UK general population survey (1995). Discussion Paper, no 138, University of York Centre for Health Economics. https://www.york.ac.uk/che/pdf/DP138.pdf

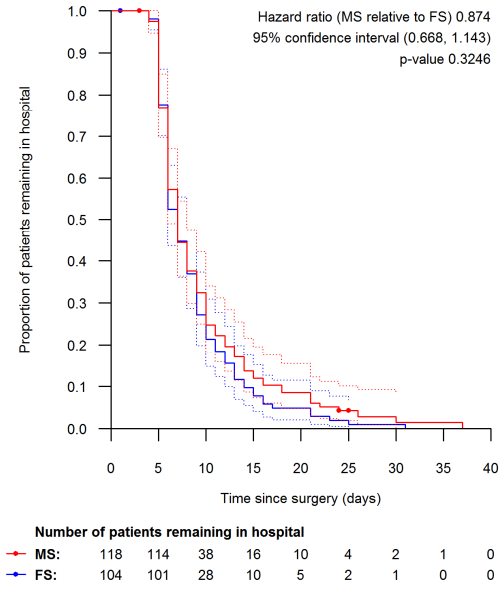
[16] Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T, et al. Validating The SF-36 Health Survey Questionnaire: New Outcome Measure For Primary Care. BMJ. 1992;305:160-4.

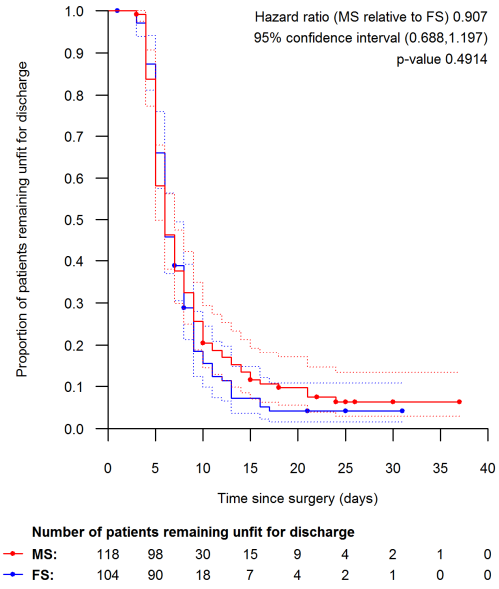
[17] Ware JE, Kosinski M, Gandek B. SF-36 Health Survey: Manual and Interpretation Guide. Lincoln RI: Quality Metric Incorporated; 1993.

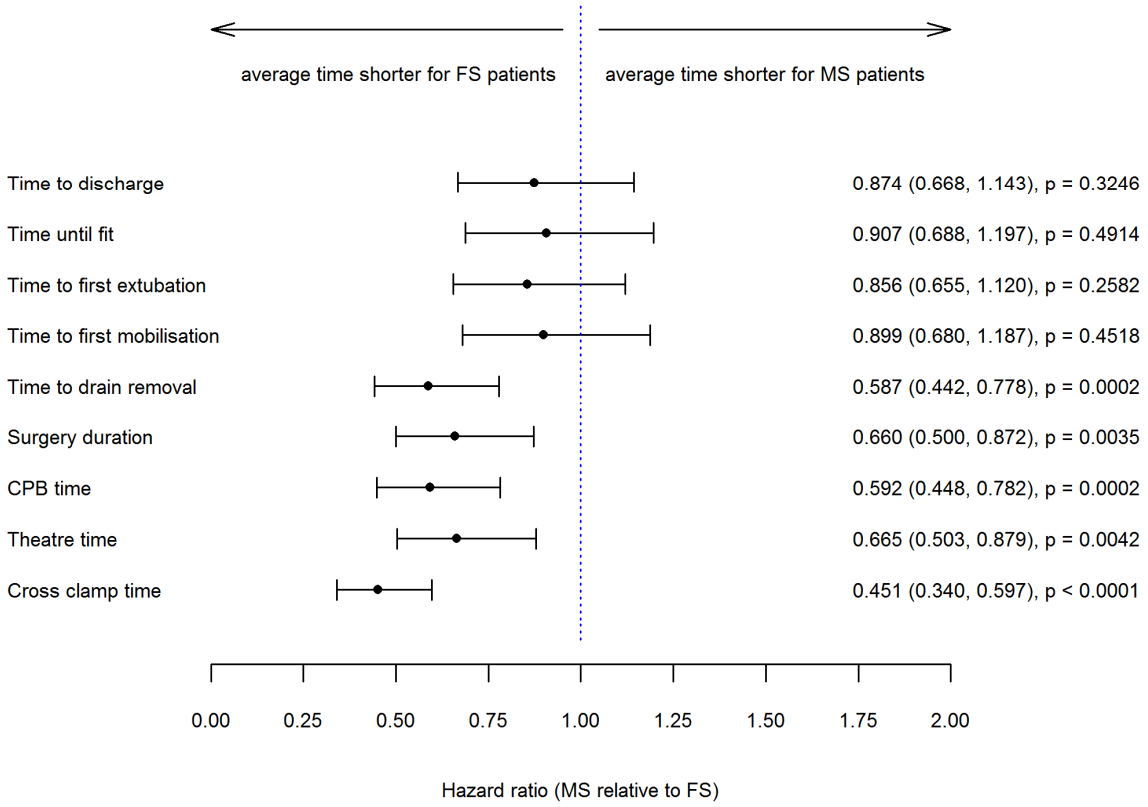
514 [18] Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for
515 reporting parallel group randomised trials. BMJ. 2010;340:c332
516 [19] Joint Formulary Committee. British National Formulary (BNF).
517 https://www.evidence.nhs.uk/formulary/bnf/current (July 2016)
518 [20] Department of Health. NHS reference costs 2014 to 2015.
519 https://www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015 (July 2016)
520 [21] NHS Prescription Services Electronic Drug Tariff. http://www.drugtariff.nhsbsa.nhs.uk/
521 (July 2016)
522 [22] Curtis L, Burns A. Unit Costs of Health and Social Care 2015. Canterbury: Personal
523 Social Services Research Unit, University of Kent. http://www.pssru.ac.uk (July 2016)
524 [23] Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health
525 from the SF-36. J Health Econ. 2002; 21:271-92.
526 [24] Faria, R, Gomes, M., Epstein, D, White, IR. A guide to handling missing data in cost-
527 effectiveness analysis conducted within randomised controlled trials. Pharmacoeconomics.
528 2014;32:1157–1170.
529 [25] Lamelas J, Mawad M, Williams R, Weiss UK, Zhang Q, LaPietra A. Isolated and
530 concomitant minimally invasive minithoracotomy aortic valve surgery. J Thorac Cardiovasc
531 Surg. 2018;155:926-36.
532 [26] Rodriguez E, Malaisrie SC, Mehall JR, Moore M, Salemi A, Ailawadi G, et al.
533 Economic Workgroup on Valvular Surgery, Right anterior thoracotomy aortic valve
534 replacement is associated with less cost than sternotomy-based approaches: a multi-institution
535 analysis of 'real world' data. J Med Econ. 2014;17:846-52.
536 [27] Ghanta RK, Lapar DJ, Kern JA, Kron IL, Speir AM, Fonner E, et al. Minimally invasive
537 aortic valve replacement provides equivalent outcomes at reduced cost compared with

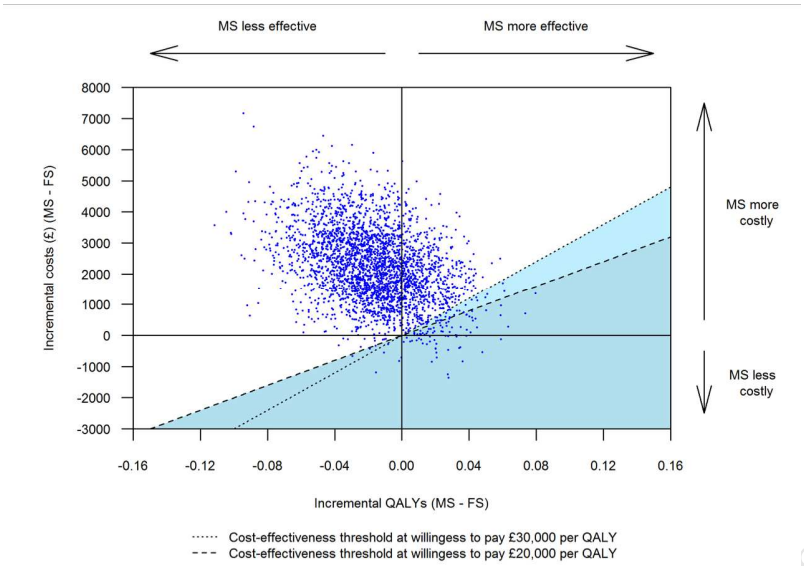
538 conventional aortic valve replacement: A real-world multi-institutional analysis. J Thorac
539 Cardiovasc Surg. 2015;149:1060-5.

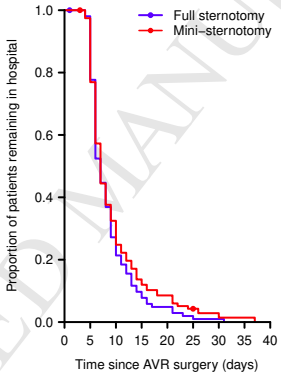
540

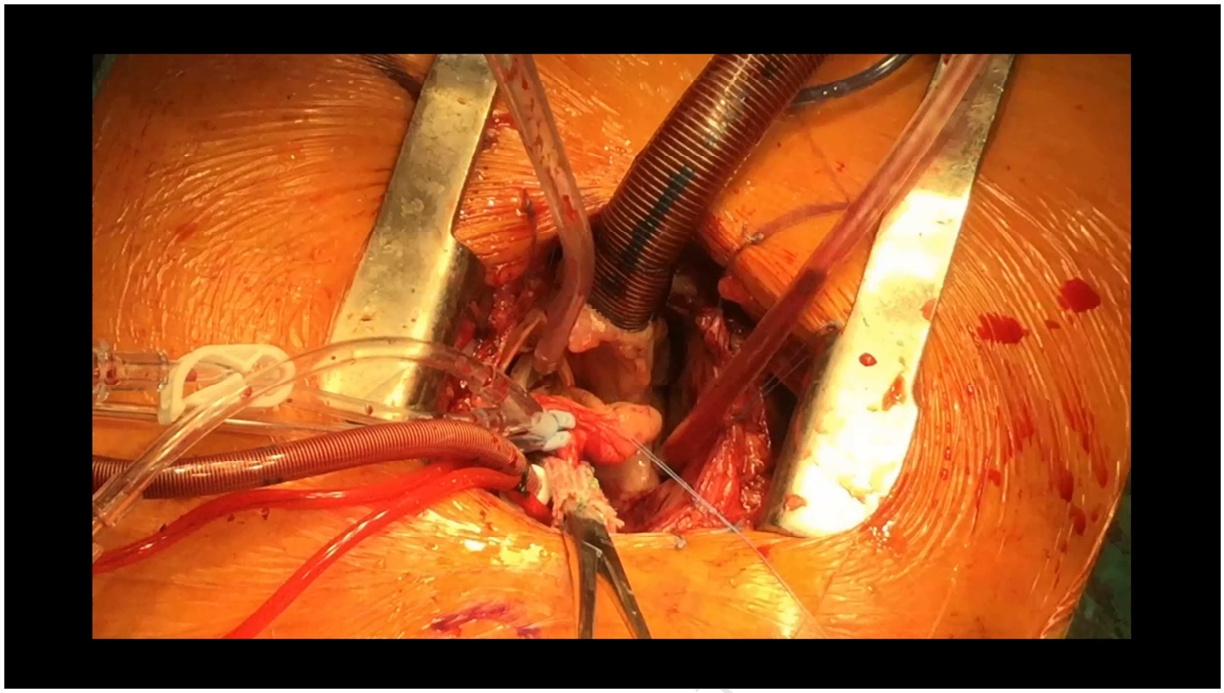




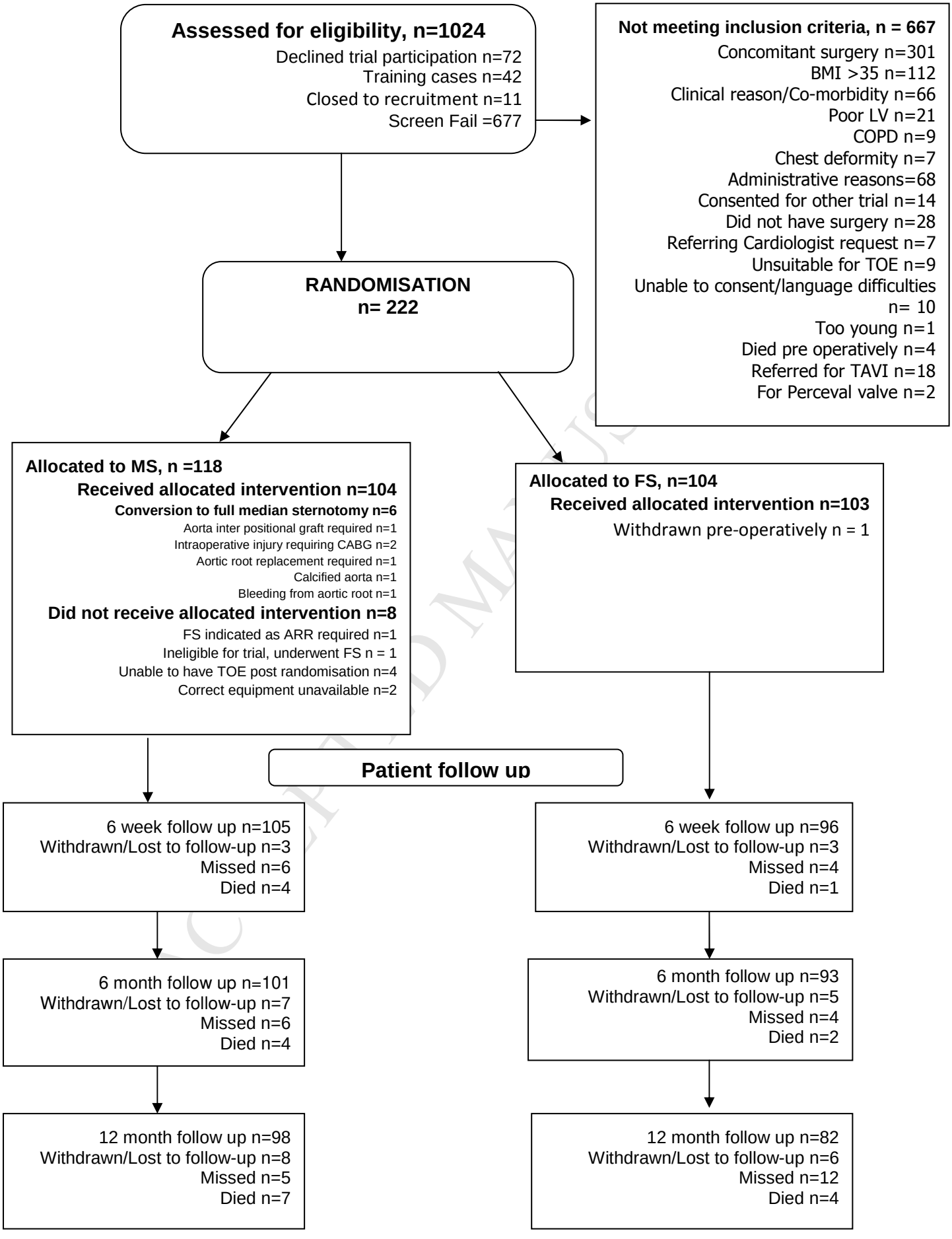








MiniStern Trial. CONSORT Flow Diagram



Appendices

Tables and figures:

Appendix A: Statistical Analysis

Table A1. All patients who underwent a redo sternotomy, or crossed over from mini- to full sternotomy, or were randomised but found to be ineligible

	Allocated treatment	Description	Per-protocol population	Safety population
Redos	FS	Return to theatre for ventricular septal defect closure and redo-AVR.	As FS	As FS
	FS	Return to theatre for tamponade and cardiac arrest. Redo sternotomy for tamponade.	As FS	As FS
	MS	Return to theatre for tamponade MA bleed. Conversion to FS.	As MS	As MS
	MS	Return to theatre for bleeding. Redo FS.	As MS	As MS
	MS	Return to theatre for tamponade. Evacuation of clot/pericardial effusion. Conversion to FS.	As MS	As MS
	MS	Return to theatre for cardiac arrest and tamponade. Emergency re-sternotomy (FS), tamponade and aortotomy repair.	As MS	As MS
	MS	Return to theatre for pericardial collection and early tamponade. PEA arrest. Re-exploration on bypass. Completion FS.	As MS	As MS
	MS	Second return to theatre. Attempted weaning of ECMO and placement of RVAD. Removal of blood clot. Redo sternotomy.	As MS	As MS
Crossovers	MS	Aortic root replacement required, FS indicated.	Excluded	As FS
	MS	FS indicated as unable to perform TOE.	Excluded	As FS
	MS	Aorta interposition graft required.	Excluded	As MS
	MS	FS indicated as unable to have TOE.	Excluded	As FS
	MS	Needed CABG due to intraoperative injury.	Excluded	As MS
	MS	Needed CABG due to intraoperative injury.	Excluded	As MS
	MS	FS indicated as unable to perform TOE.	Excluded	As FS
	MS	Required aortic root replacement, conversion to FS.	Excluded	As MS
	MS	Patient randomised too early - unable to insert TOE probe.	Excluded	As FS
	MS	Did not have correct equipment in theatre.	Excluded	As FS
	MS	Mini-sternotomy equipment not available.	Excluded	As FS
	MS	Bleeding.	Excluded	As MS
	MS	Patient had calcified aorta. Nowhere to cannulate safely.	Excluded	As MS
Ineligible	FS	Withdrawn from trial by surgeon pre-operatively (but post-randomisation) as required AVR and myectomy.	Excluded	Excluded
	FS	Poor quality baseline echocardiogram, with no assessment of LV function.	Excluded	As FS
	MS	Surgeon had not checked echo report until after randomisation. Underwent FS.	Excluded	As FS

Table A2. Additional summaries of in-hospital endpoints

	Mini-sternotomy (n = 118)	Full sternotomy (n = 104)
Time to discharge (days)		
Mean (standard error)	9.5 (0.6)	8.6 (0.5)
Median (95% confidence interval)	7 (6, 8)	7 (6, 8)
Time until fit for discharge (days)		
Mean (standard error)	8.5 (0.5)*	7.5 (0.3)*
Median (95% confidence interval)	6 (5, 7)	6 (6, 7)
Time to first mobilisation (days)		
Mean (standard error)	5.7 (0.5)*	4.9 (0.3)*
Median (95% confidence interval)	4 (3, 4)	4 (-, -)
Time to mediastinal drain removal (hours)		
Mean (standard error)	48.1 (4.8)*	30.0 (1.7)
Median (95% confidence interval)	26.1 (22.8, 42.6)	22.5 (22.0, 22.9)
Time to first extubation (hours)		
Mean (standard error)	13.1 (1.7)*	10.5 (0.7)
Median (95% confidence interval)	9.2 (8.7, 9.9)	8.3 (8.0, 9.2)

Table A2 shows Kaplan-Meier estimates of in-hospital endpoints. Censoring of longest time to event for some endpoints led to underestimation of means and standard errors (highlighted with asterisks). A confidence interval for median time to mobilisation could not be estimated.

Table A3. Additional summaries of operative endpoints

	Mini-sternotomy (n = 118)	Full sternotomy (n = 104)
Theatre time (minutes)		
Mean (standard error)	201.2 (3.9)	181.0 (4.6)
Median (95% confidence interval)	191 (187, 205)	176 (170, 180)
CPB time (minutes)		
Mean (standard error)	82.0 (1.9)	69.5 (2.3)
Median (95% confidence interval)	80 (77, 86)	66 (59, 74)
Cross clamp time (minutes)		
Mean (standard error)	65.5 (1.5)	52.4 (1.6)
Median (95% confidence interval)	65 (61, 69)	49 (45, 53)
Surgery duration (minutes)		
Mean (standard error)	165.5 (3.4)	145.7 (4.3)
Median (95% confidence interval)	163(155, 172)	148.5 (134, 153)
Total theatre time, including repeats/readmissions (minutes)		
Mean (standard error)	221.1 (9.5)	191.2 (6.1)
Median (95% confidence interval)	196 (189, 210)	178.5 (171, 188)
Total CPB time, including repeats/readmissions (minutes)		
Mean (standard error)	85.1 (2.6)	71.1 (2.8)
Median (95% confidence interval)	82 (77, 87)	66 (59, 74)
Total cross clamp time, including repeats/readmissions (minutes)		
Mean (standard error)	66.1 (1.6)	53.5 (2.0)
Median (95% confidence interval)	66 (61, 70)	49 (45, 53)
Volume of blood lost in the first 12 postoperative hours (ml)		
Mean (SD)	310.4 (342.5)	323.2 (267.8)
Median (quartiles)	225 (150, 325)	250 (175, 375)
Transfusion of packed red cells in the first 48 postoperative hours (ml)		
Number of transfused patients (%)	50 (42%)	51 (49%)
Mean (SD) in transfused patients	625.3 (513.2)	442.4 (265.3)
Median (quartiles) in transfused patients	500 (300, 644)	303 (284, 569)
Transfusion of clotting products in the first 48 postoperative hours (ml)		
Number of transfused patients (%)	11 (9%)	4 (4%)
Mean (SD) in transfused patients	920.5 (1438.4)	753.0 (672.5)
Median (quartiles) in transfused patients	332 (183, 1050)	625 (209, 1297)

All estimates for time-to-event endpoints in Table A3 are Kaplan-Meier estimates. Time data were complete, except for seven surgery durations (3 MS, 4 FS) that were not recorded and were therefore censored at 1 minute. Blood data were only missing for one patient (FS group, withdrawn before surgery). Blood transfusion and clotting products data for seven patients at the Freeman hospital were recorded in units and converted to ml (1 unit PRC = 300ml, 1 unit platelets = 245ml, 1 unit FFP = 280ml). Transfusion data were explored using logistic regression models, including fixed effects for treatment, valve and sex, and a random surgeon effect. These analyses did not show a statistically significant difference between MS and FS patients in either need for blood transfusion (MS/FS odds ratio 0.797, confidence interval 0.453 to 1.402, p-value 0.4310) or the need for transfusion of clotting products (MS/FS odds ratio 2.616, confidence interval 0.801 to 8.541, p-value 0.1112).

Table A4. Results from Cox models and log-rank tests for primary and secondary endpoints

	MS/FS hazard ratio (95% confidence interval)	p-value for null hypothesis HR = 1	Log-rank test statistic	p-value from log-rank test
Primary analyses				
Time to discharge	0.874 (0.668,1.143)	0.3246	0.157	0.6924
Time until fit	0.907 (0.688,1.197)	0.4914	0.340	0.5597
Per protocol analyses of primary endpoints				
Time to discharge	0.868 (0.656,1.147)	0.3194	0.200	0.6544
Time until fit	0.915 (0.688,1.218)	0.5443	0.217	0.6415
Sensitivity analyses: age included as an effect in the Cox models				
Time to discharge	0.866 (0.661,1.135)	0.2985	0.157	0.6924
Time until fit	0.902 (0.683,1.192)	0.4685	0.340	0.5597
Sensitivity analyses: EuroSCORE included as an effect in the Cox models				
Time to discharge	0.885 (0.676,1.159)	0.3753	0.157	0.6924
Time until fit	0.936 (0.709,1.236)	0.6400	0.340	0.5597
Sensitivity analyses: censoring times taken as event times:				
Time to discharge	0.884 (0.677,1.153)	0.3625	0.189	0.6639
Time until fit	0.888 (0.680,1.160)	0.3844	0.765	0.3819
Sensitivity analysis: patients assumed to be fit at discharge				
Time until fit	0.879 (0.671, 1.151)	0.3480	0.703	0.4018
Secondary endpoint analyses				
Time until first mobilisation	0.899 (0.680,1.187)	0.4518	0.303	0.5819
Time until drain removal	0.587 (0.442,0.778)	0.0002	5.838	0.0157
Time until first extubation	0.856 (0.655,1.120)	0.2582	0.359	0.5488
Exploratory analyses				
Surgery duration	0.660 (0.500,0.872)	0.0035	17.892	< 0.0001
CPB time	0.592 (0.448,0.782)	0.0002	24.871	< 0.0001
Cross clamp time	0.451 (0.340,0.597)	< 0.0001	42.539	< 0.0001
Theatre time	0.665 (0.503,0.879)	0.0042	16.806	< 0.0001
Total CPB time including repeats/readmissions	0.547 (0.414,0.723)	< 0.0001	20.176	< 0.0001
Total cross clamp time including repeats/readmissions	0.458 (0.346,0.608)	< 0.0001	34.352	< 0.0001
Total theatre time including repeats/readmissions	0.698 (0.531,0.918)	0.0102	5.657	0.0174
Time to death by any cause	1.871 (0.723, 4.844)	0.1966	0.7309	0.3926

Table A4 shows the results of all analyses performed for the primary and secondary time-to-event endpoints, including unplanned, exploratory analyses of secondary endpoints. All secondary endpoint analyses, sensitivity analyses and exploratory analyses were performed using the intent to treat population. All log-rank tests were stratified by valve, sex and surgeon. All Cox models included valve, sex and treatment as fixed effects, and surgeon as a random effect. Exploratory analysis of time to all-cause death included age as a fixed effect in the Cox model. Mean imputation was used for missing EuroSCORE data at baseline (1 MS).

Table A5. Summaries of pain at rest scores in the first ten days following surgery

		Mini-sternotomy (n = 118)	Full sternotomy (n = 104)
Day 1	Mean (SD)	3.5 (2.5)	3.7 (2.4)
	n	100 (85%)	82 (80%)
Day 2	Mean (SD)	3 (2.3)	3.1 (2.5)
	n	89 (75%)	81 (79%)
Day 3	Mean (SD)	2.7 (2.3)	2.4 (2.3)
	n	91 (77%)	83 (81%)
Day 4	Mean (SD)	2.4 (2.1)	2.4 (2.4)
	n	94 (80%)	84 (82%)
Day 5	Mean (SD)	2 (1.9)	2.1 (2)
	n	90 (79%)	80 (79%)
Day 6	Mean (SD)	1.8 (1.7)	2.1 (2)
	n	69 (77%)	61 (76%)
Day 7	Mean (SD)	1.5 (1.8)	1.8 (2)
	n	46 (69%)	42 (78%)
Day 8	Mean (SD)	1.2 (1.4)	1.7 (1.6)
	n	40 (77%)	35 (76%)
Day 9	Mean (SD)	1 (1.8)	0.8 (1.5)
	n	25 (57%)	18 (47%)
Day 10	Mean (SD)	0.7 (1)	1.3 (2)
	n	18 (47%)	12 (43%)

Table A5 shows the number of pain scores taken for each of the 10 days following surgery. The denominator used for each percentage is the number of patients known to be alive and in hospital on the given day.

Table A6. Summaries of EQ-5D utility scores up to the 12 month follow-up

		Mini-sternotomy (n = 118)	Full sternotomy (n = 104)
Baseline	Mean (SD)	0.77 (0.19)	0.70 (0.24)
	n	105 (89%)	95 (91%)
Day 4	Mean (SD)	0.47 (0.29)	0.39 (0.28)
	n	92 (78%)	89 (86%)
Discharge	Mean (SD)	0.60 (0.24)	0.58 (0.24)
	n	103 (87%)	88 (85%)
Six weeks	Mean (SD)	0.74 (0.23)	0.71 (0.21)
	n	106 (90%)	88 (85%)
Six months	Mean (SD)	0.83 (0.25)	0.83 (0.23)
	n	105 (89%)	95 (91%)
Twelve months	Mean (SD)	0.83 (0.29)	0.78 (0.28)
	n	103 (87%)	84 (81%)

For patients who died, EQ-5D scores were taken to be zero following death. Percentages presented in Table A6 were calculated as the number of scores recorded (including the zeros) divided by the number of patients randomised to the group. The difference in mean baseline score was potentially due to the imbalance in gender (the FS group has a greater proportion of females, who reported lower quality of life on average).

Table A7. Summaries of SF-36 domain scores up to the 12 month follow up

			Mini-sternotomy (n = 118)	Full sternotomy (n = 104)
Bodily pain	Baseline	Mean (SD)	70 (25)	64 (28)
		n	104 (88%)	96 (92%)
	Six weeks	Mean (SD)	61 (24)	60 (23)
		n	105 (89%)	90 (87%)
	Six months	Mean (SD)	79 (27)	74 (28)
		n	104 (88%)	94 (90%)
	Twelve months	Mean (SD)	76 (31)	72 (32)
		n	99 (84%)	86 (83%)
General health	Baseline	Mean (SD)	62 (20)	58 (22)
		n	104 (88%)	94 (90%)
	Six weeks	Mean (SD)	70 (20)	66 (20)
		n	104 (88%)	91 (88%)
	Six months	Mean (SD)	71 (24)	66 (24)
		n	103 (87%)	94 (90%)
	Twelve months	Mean (SD)	68 (26)	62 (26)
		n	100 (85%)	86 (83%)
Mental health	Baseline	Mean (SD)	74 (18)	67 (21)
		n	104 (88%)	95 (91%)
	Six weeks	Mean (SD)	72 (22)	73 (19)
		n	104 (88%)	91 (88%)
	Six months	Mean (SD)	80 (21)	74 (22)
		n	103 (87%)	94 (90%)
	Twelve months	Mean (SD)	76 (26)	73 (23)
		n	100 (85%)	86 (83%)
Physical functioning	Baseline	Mean (SD)	54 (26)	47 (28)
		n	105 (89%)	96 (92%)
	Six weeks	Mean (SD)	63 (22)	56 (23)
		n	105 (89%)	91 (88%)
	Six months	Mean (SD)	78 (27)	70 (28)
		n	104 (88%)	94 (90%)
	Twelve months	Mean (SD)	74 (30)	67 (31)
		n	100 (85%)	86 (83%)
Role emotional	Baseline	Mean (SD)	67 (40)	55 (46)
		n	104 (88%)	94 (90%)
	Six weeks	Mean (SD)	60 (44)	63 (43)
		n	104 (88%)	90 (87%)
	Six months	Mean (SD)	81 (35)	72 (42)
		n	104 (88%)	94 (90%)
	Twelve months	Mean (SD)	76 (39)	71 (42)
		n	98 (83%)	85 (82%)
Role physical	Baseline	Mean (SD)	33 (41)	23 (38)
		n	103 (87%)	96 (92%)

	Six weeks	Mean (SD)	19 (32)	20 (33)
		n	103 (87%)	90 (87%)
	Six months	Mean (SD)	65 (42)	59 (44)
		n	103 (87%)	94 (90%)
	Twelve months	Mean (SD)	64 (44)	52 (46)
		n	98 (83%)	85 (82%)
Social functioning	Baseline	Mean (SD)	66 (30)	61 (29)
		n	104 (88%)	94 (90%)
	Six weeks	Mean (SD)	66 (29)	68 (27)
		n	104 (88%)	91 (88%)
	Six months	Mean (SD)	85 (26)	78 (28)
		n	102 (86%)	93 (89%)
	Twelve months	Mean (SD)	81 (30)	78 (30)
		n	98 (83%)	85 (82%)
Vitality	Baseline	Mean (SD)	46 (25)	40 (23)
		n	104 (88%)	95 (91%)
	Six weeks	Mean (SD)	50 (22)	48 (22)
		n	104 (88%)	90 (87%)
	Six months	Mean (SD)	64 (23)	57 (23)
		n	103 (87%)	94 (90%)
	Twelve months	Mean (SD)	60 (26)	54 (26)
		n	100 (85%)	86 (83%)

An in-house implementation of the standard scoring algorithm for the developmental version of SF-36 was used. For patients who died, SF-36 scores were taken to be zero following death. Percentages presented in Table A7 were calculated as the number of scores recorded (including the zeros) divided by the number of patients randomised to the group. The differences in mean baseline scores were potentially due to the imbalance in gender (the FS group has a greater proportion of females, who reported lower quality of life on average).

Table A8. Estimated treatment effects (MS - FS) and treatment-time interactions for SF-36 domain scores up to 12 months, EQ-5D utility scores up to 12 months and pain scores up to discharge

	Effect (MS – FS)	95% confidence interval	p-value
Pain at rest (n = 219)			
Treatment effect	0.0	(-0.7, 0.6)	0.9766
Treatment-time (days) interaction	0.0	(-0.1, 0.1)	0.8190
EQ-5D utility scores (n = 197)			
Treatment effect	0.02	(-0.03, 0.07)	0.5148
Treatment-time (months) interaction	0.00	(-0.01, 0.01)	0.9731
SF-36 physical functioning (n = 192)			
Treatment effect	1.2	(-6.2, 8.7)	0.7414
Treatment-time (months) interaction	0.3	(-0.2, 0.9)	0.2387
SF-36 role physical (n = 190)			
Treatment effect	-8.3	(-21.1, 4.5)	0.2025
Treatment-time (months) interaction	1.7	(0.3, 3.1)	0.0169
SF-36 bodily pain (n = 191)			
Treatment effect	-0.7	(-9.1, 7.8)	0.8792
Treatment-time (months) interaction	0.3	(-0.5, 1.1)	0.4331
SF-36 general health (n = 189)			
Treatment effect	-1.0	(-7.5, 5.5)	0.7710
Treatment-time (months) interaction	0.3	(-0.2, 0.8)	0.2224
SF-36 vitality (n = 190)			
Treatment effect	-2.1	(-8.8, 4.5)	0.5273
Treatment-time (months) interaction	0.6	(0.1, 1.2)	0.0293
SF-36 social functioning (n = 189)			
Treatment effect	-5.5	(-14.1, 3.1)	0.2093
Treatment-time (months) interaction	1.0	(0.2, 1.7)	0.0183
SF-36 role emotional (n = 189)			
Treatment effect	-6.2	(-18.6, 6.2)	0.3255
Treatment-time (months) interaction	1.1	(-0.1, 2.3)	0.0699
SF-36 mental health (n = 190)			
Treatment effect	-3.2	(-9.7, 3.4)	0.3431
Treatment-time (months) interaction	0.5	(-0.0, 1.0)	0.0702

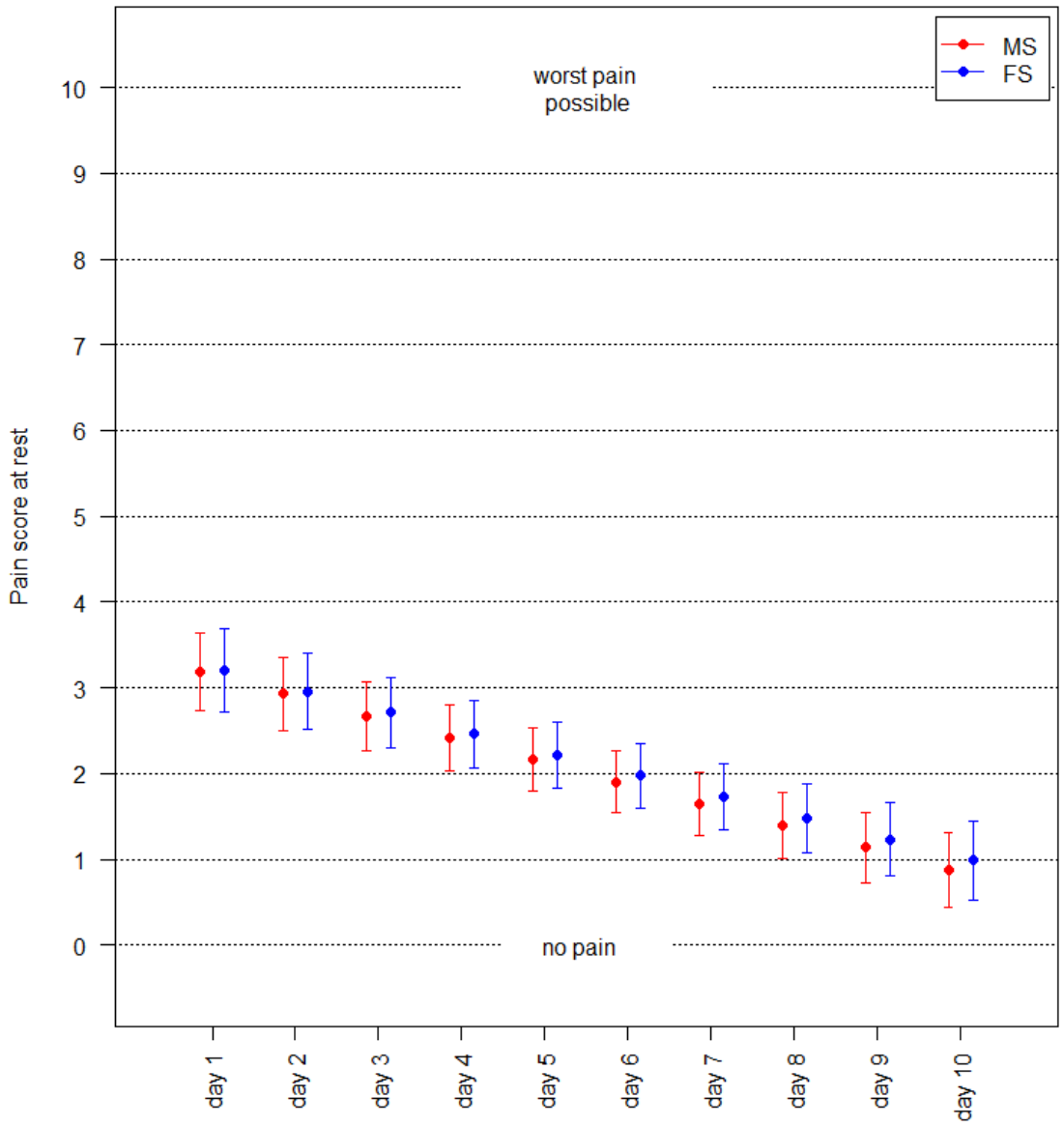
Table A8 shows results of complete case analyses of questionnaire data, under a missing completely at random assumption, including only patients with at least one analysable follow-up questionnaire. For each analysis, the n in parentheses is number of patients used to fit the model. For pain and SF-36 scores, some random effects were estimated to have a variance of 0 and were excluded from the models (surgeon effect for pain, and both the surgeon effect and random slope for SF-36). The slope (time coefficient) was estimated to be negative for pain and positive for all EQ-5D and SF-36 scores. This suggests improvement over time in each score. Evidence of greater rate of improvement over time for MS patients (statistically significant, positive interaction term) was seen for three SF-36 domains (role physical, vitality, and social functioning), but no others.

Table A9. Estimated treatment effects (MS-FS) and treatment-time interactions for SF-36 domain scores up to 12 months, EQ-5D utility scores up to 12 months and pain scores up to discharge, after multiple imputation of missing scores

	Effect (MS – FS)	95% confidence interval	p-value
Pain at rest			
Treatment effect	0.0	(-0.7, 0.6)	0.9059
Treatment-time (days) interaction	0.0	(-0.1, 0.1)	0.9685
EQ-5D utility scores			
Treatment effect	0.01	(-0.04, 0.06)	0.8203
Treatment-time (months) interaction	0.00	(-0.01, 0.01)	0.9094
SF-36 physical functioning			
Treatment effect	2.0	(-4.9, 8.9)	0.5744
Treatment-time (months) interaction	0.2	(-0.3, 0.8)	0.3996
SF-36 role physical			
Treatment effect	-6.6	(-18.7, 5.4)	0.2808
Treatment-time (months) interaction	1.5	(0.1, 2.8)	0.0310
SF-36 bodily pain			
Treatment effect	-0.1	(-9.0, 7. 7)	0.9748
Treatment-time (months) interaction	0.3	(-0.4, 1.1)	0.4091
SF-36 general health			
Treatment effect	1.1	(-5.0, 7.3)	0.7175
Treatment-time (months) interaction	0.2	(-0.3, 0.7)	0.3373
SF-36 vitality			
Treatment effect	-0.5	(-6.9, 5.9)	0.8798
Treatment-time (months) interaction	0.4	(-0.2, 1.0)	0.1733
SF-36 social functioning			
Treatment effect	-4.4	(-12.4, 3.5)	0.2756
Treatment-time (months) interaction	0.7	(0.0, 1.5)	0.0589
SF-36 role emotional			
Treatment effect	-4.6	(-16.4, 7.2)	0.4415
Treatment-time (months) interaction	0.8	(-0.4, 2.0)	0.1790
SF-36 mental health			
Treatment effect	-2.5	(-8.6, 3.5)	0.4113
Treatment-time (months) interaction	0.4	(-0.1, 0.9)	0.1195

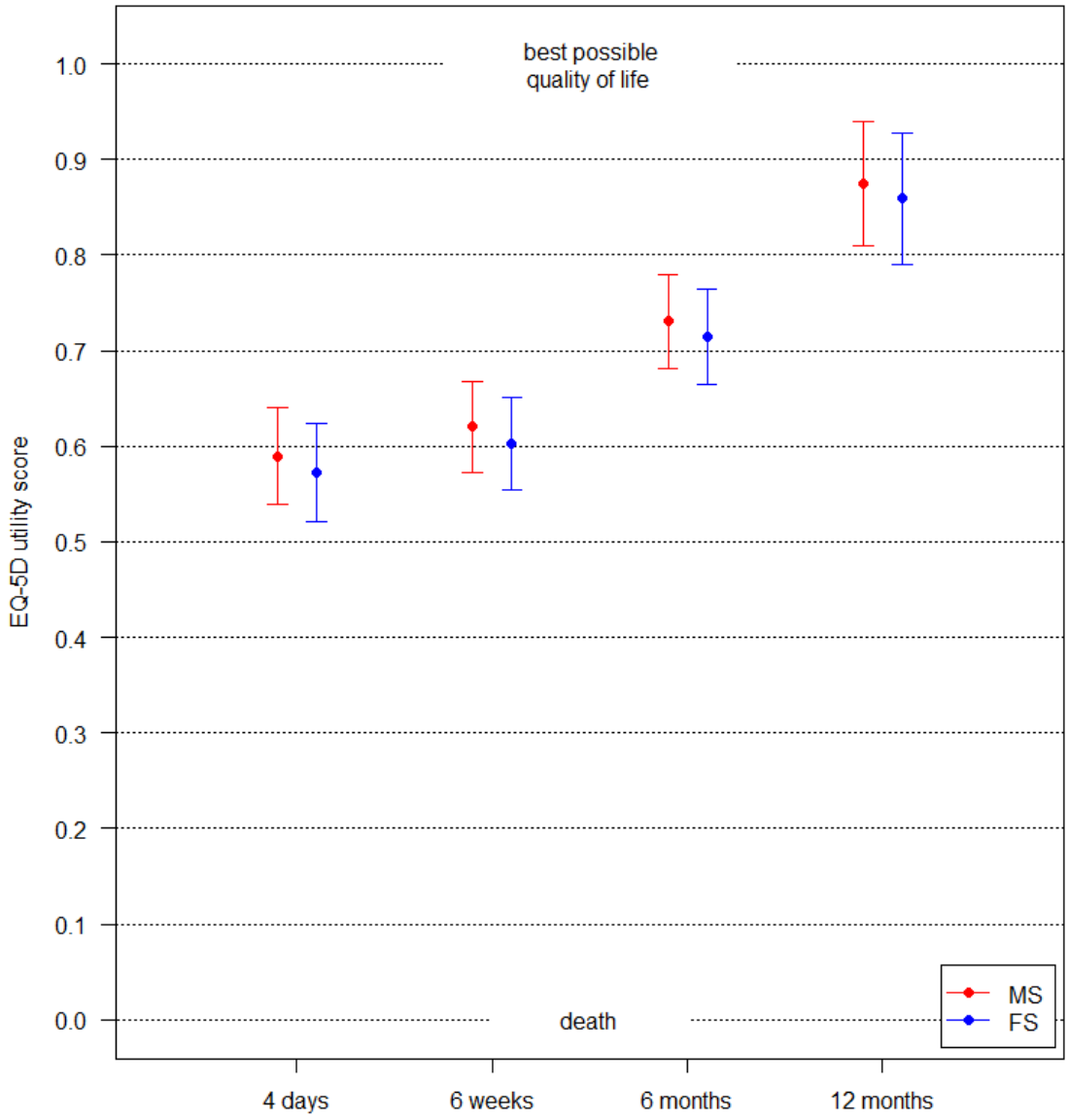
Table A9 shows the results from analysing the questionnaire data using multiple imputation to handle missing observations, under a missing at random assumption. For each analysis, missing data were imputed from models that included all other variables used in the analysis, along with CCS grading and NYHA grading as auxiliary variables. The method used was multiple imputation by chained equations with predictive mean matching. Estimates from 100 imputed data sets were combined using Rubin's rules. Pain was only imputed for patients known to be alive and in hospital, not for patients who had died or had already been discharged. Evidence of greater rate of improvement over time for MS patients (statistically significant, positive interaction term) was seen only for one SF-36 domain.

Figure A1. Forest plots of mean pain scores for the first 10 days following surgery, with 95% confidence intervals



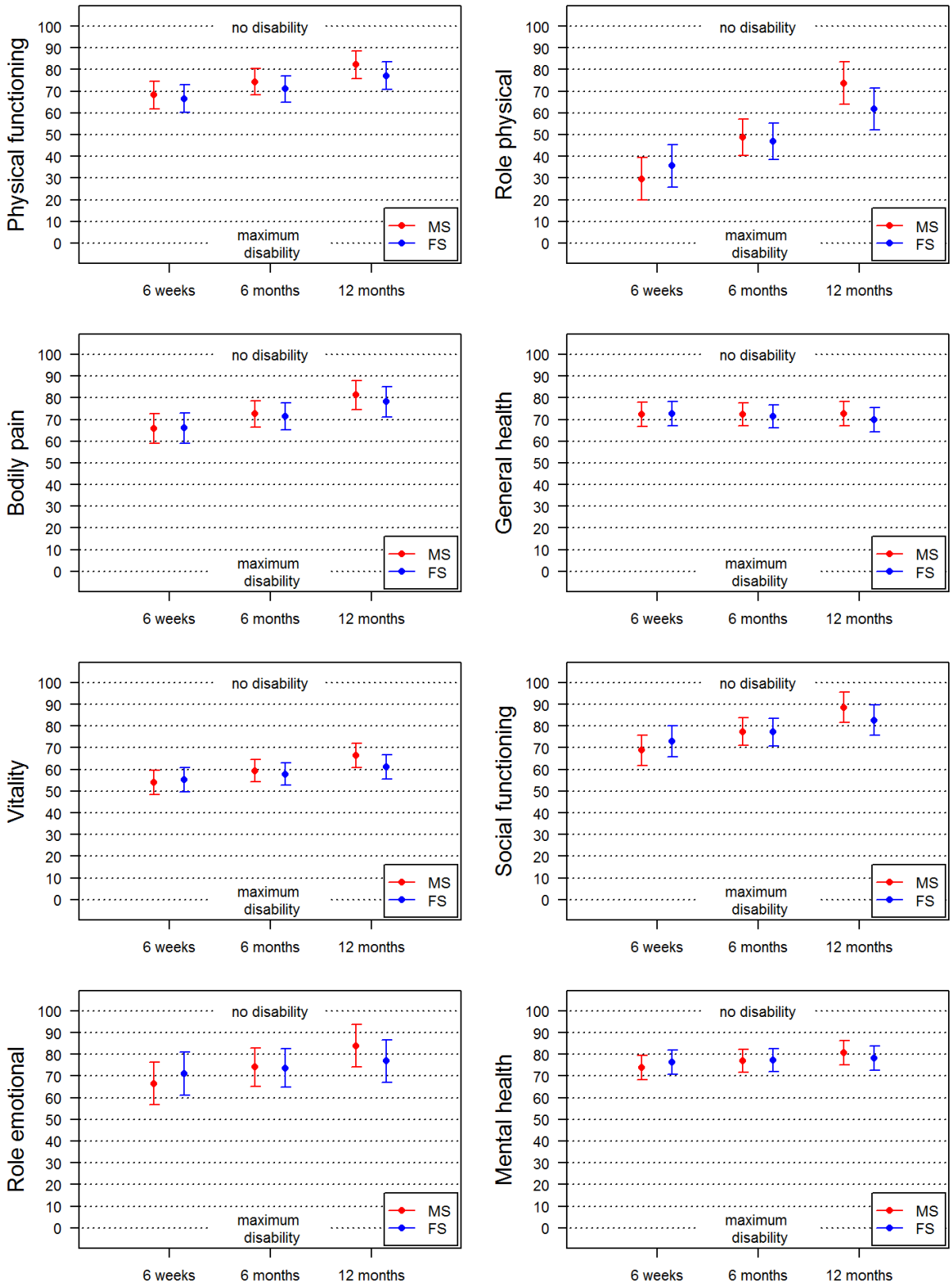
In Figure A1, means on each day were adjusted for sex and valve type, and were estimated from the complete case analysis.

Figure A2. Forest plot of mean EQ-5D scores at each follow-up time, with 95% confidence intervals



In Figure A2, means at each follow-up time were adjusted for baseline EQ-5D, sex and valve type, and were estimated from the complete case analysis.

Figure A3. Forest plot of mean SF36 domain scores at each follow-up time, with 95% confidence intervals



In Figure A3, means at each follow-up time were adjusted for baseline domain score, sex and valve type, and were estimated from the complete case analysis. A score of 100 represents no disability, and a score of 0 represents maximum disability.

Table A10. Summaries heart function (LVEF) and respiratory function (FEV_1)

	Mini-sternotomy (n = 118)	Full sternotomy (n = 104)
FEV_1 (litres):		
Baseline visit		
Mean (SD)	2.3 (0.7)	2.3 (0.8)
Median (quartiles)	2.2 (1.8, 2.7)	2.2 (1.7, 2.6)
n	115	101
Discharge		
Mean (SD)	1.6 (0.6)	1.6 (0.6)
Median (quartiles)	1.5 (1.2, 1.8)	1.5 (1.2, 1.9)
n	82	69
6 week visit		
Mean (SD)	2.1 (0.8)	2.1 (0.7)
Median (quartiles)	2 (1.5, 2.5)	1.9 (1.6, 2.5)
n	92	84
6 month visit		
Mean (SD)	2.2 (0.7)	2.1 (0.7)
Median (quartiles)	2.1 (1.7, 2.6)	1.9 (1.4, 2.4)
n	91	82
LVEF (%):		
Baseline visit		
Mean (SD)	61.9 (9.1)	62.4 (8.6)
Median (quartiles)	62.5 (57.5, 67.5)	63 (57.5, 67.0)
n	117	101
Discharge		
Mean (SD)	59.9 (9.7)	59 (10.2)
Median (quartiles)	62 (55.0, 65.0)	58 (55.0, 64.5)
n	106	96
6 month visit		
Mean (SD)	61.2 (8.1)	61.8 (9.7)
Median (quartiles)	61 (56.0, 67.5)	62.5 (56.3, 68.0)
n	97	88

FEV_1 is forced expiratory volume in one second, measured by hand-held spirometry. LVEF is left ventricular ejection fraction, measured by echocardiography. No analyses were planned for these endpoints.

Table A11. Frequency of non-fatal SAEs (number of patients) within one year of surgery, by treatment received

	Mini-sternotomy (n = 110)	Full sternotomy (n = 111)	Total (n = 221)
Cardiac (including atrial fibrillation, conduction problems, need for permanent pacemaker)	43 (29)	27 (21)	70 (50)
Respiratory	20 (14)	9 (8)	29 (22)
Injury/procedural	19 (11)	7 (6)	26 (17)
Non-cardiorespiratory infection (including wound)	7 (7)	12 (9)	19 (16)
Urinary	11 (10)	8 (6)	19 (16)
Surgical and medical procedures	9 (6)	7 (7)	16 (13)
Nervous system	8 (8)	7 (7)	15 (15)
Cardiorespiratory infection (including endocarditis, device-related infections, chest infection)	9 (9)	6 (5)	15 (14)
Vascular	9 (9)	1 (1)	10 (10)
Psychiatric	5 (5)	5 (5)	10 (10)
Gastro-intestinal – diarrhoea	7 (6)	3 (3)	10 (9)
Gastro-intestinal – other	7 (7)	1 (1)	8 (8)
General disorders	4 (4)	3 (2)	7 (6)
Metabolic	2 (2)	3 (2)	5 (4)
Blood/lymph	4 (3)	1 (1)	5 (4)
Neoplasms	1 (1)	1 (1)	2 (2)
Hepatitis/cholecystitis	1 (1)	1 (1)	2 (2)
Musculoskeletal	2 (2)	0 (0)	2 (2)
Skin/tissue	0 (0)	1 (1)	1 (1)
Eye	0 (0)	1 (1)	1 (1)
Immune	0 (0)	1 (1)	1 (1)
Total	168 (56)	105 (46)	273 (102)

Among the nervous system SAEs recorded in Table A11, strokes were suffered by 3 FS recipients and 2 MS recipients. No patient suffered more than one stroke.

Table A12. Frequencies of non-death SAEs (and number of patients experiencing them), within a year of surgery, at each level of severity, expectedness and relatedness, by treatment received

	Mini-sternotomy (n = 110)	Full sternotomy (n = 111)	Total (n = 221)
Cardiorespiratory:			
Severity			
Severe	26 (14)	14 (11)	40 (25)
Moderate	34 (24)	24 (18)	58 (42)
Mild	12 (11)	4 (4)	16 (15)
Expectedness			
Expected	69 (38)	42 (30)	111 (68)
Unexpected	3 (2)	0 (0)	3 (2)
Relatedness			
Probably related	4 (4)	2 (2)	6 (6)
Possibly related	50 (30)	32 (25)	82 (55)
Unrelated	18 (13)	8 (6)	26 (19)
Total	72 (38)	42 (30)	114 (68)
Non-cardiorespiratory:			
Severity			
Severe	40 (21)	24 (15)	64 (36)
Moderate	43 (29)	31 (21)	74 (50)
Mild	13 (11)	8 (5)	21 (16)
Expectedness			
Expected	68 (34)	45 (27)	113 (61)
Unexpected	28 (15)	18 (15)	46 (30)
Relatedness			
Probably related	9 (5)	5 (5)	14 (10)
Possibly related	37 (22)	30 (20)	67 (42)
Unrelated	50 (27)	28 (20)	78 (47)
Total	96 (41)	63 (34)	159 (75)

The only unexpected events in the MS group were a bilateral pleural effusion in one patient, and bronchial aspiration and peri-arrest event in another. Both patients completely recovered. Exploratory analysis in the safety population, using logistic regression (with fixed treatment, valve and sex effects, and a random surgeon effect), did not show a statistically significant difference between MS and FS recipients in the odds of suffering a non-death SAE within the first year (MS/FS odds ratio 1.559, confidence interval 0.895 to 2.715 and p-value 0.1161). An exploratory Poisson regression (with a fixed effect for treatment and a random patient effect) did show a greater rate of such SAEs for MS recipients (MS/FS rate ratio 1.615, confidence interval 1.070 to 2.437, p-value 0.0225). There were 7 pericardial tamponades in total (4 for FS recipients, 3 for MS recipients, only one per patient), but logistic regression (without the random surgeon effect) did not produce a statistically significant result (MS/FS odds ratio 0.680, confidence interval 0.146 to 3.178, p-value 0.6229).

Table A13. Frequency of paraprosthetic regurgitation, by treatment received

	Mini-sternotomy (n = 110)	Full sternotomy (n = 111)	Total (n = 221)
Discharge			
No regurgitation	84	85	169
Mild regurgitation	19	16	35
Moderate regurgitation	0	0	0
Severe regurgitation	0	0	0
n	101	103	204
6 month visit			
No regurgitation	77	82	159
Mild regurgitation	18	10	28
Moderate regurgitation	0	0	0
Severe regurgitation	0	0	0
n	95	92	187

Paraposthetic regurgitation was explored using logistic regressions at each time point. These were performed as complete case analyses, in the safety population. Logistic regression models included fixed treatment, valve and sex effects, and a random surgeon effect. They did not show a statistically significant difference between MS recipients and FS recipients in the odds of regurgitation, either at discharge (MS/FS odds ratio 1.163, confidence interval 0.553 to 2.445, p-value 0.6883) or at 6 months (MS/FS odds ratio 1.880, confidence interval 0.798 to 4.430, p-value 0.1480).

Table A14. All wound infections within the first year after surgery, by treatment received

Treatment received	Relationship	Description
FS	Possibly related	Superficial sternal wound infection.
FS	Possibly related	Sternal wound infection. Returned to theatre for debridement and 2x wires removed.
FS	Possibly related	Sternal wound infection.
FS	Possibly related	Sternal wound breakdown. Debridement and excision of sinuses. PICC line inserted for 6 weeks IV antibiotics.
FS	Possibly related	Drain site wound infection.
FS	Possibly related	Wound infection - small area at lower end of sternum.
FS	Possibly related	Small sternal wound infection.
FS	Probably related	Sternal wound infection.
FS	Probably related	Sternal wound infection.
FS	Possibly related	Sternal wound infection.
FS	Probably related	Sternal wound infection.
FS	Possibly related	Sternal wound infection. Antibiotics commenced.
FS	Possibly related	Sternal wound infection - requiring hospital admission. Treated with antibiotics.
FS	Possibly related	Wound Infection. Commenced on antibiotics and daily dressings.
MS	Possibly related	Readmission, wound infection, iv/oral flucloxacillin.
MS	Possibly related	MRSA sternal wound infection.
MS	Probably related	Sternal wound infection. Admitted to NGTH with fever, chest pain, SOB and discharging sternal wound. Commenced IV flucloxacillin. Swab taken, VAC dressing applied.
MS	Possibly related	Wound infection at base of sternotomy. Wound swab taken, grown K.pneumoniae. Commenced antibiotics - amoxycillin.

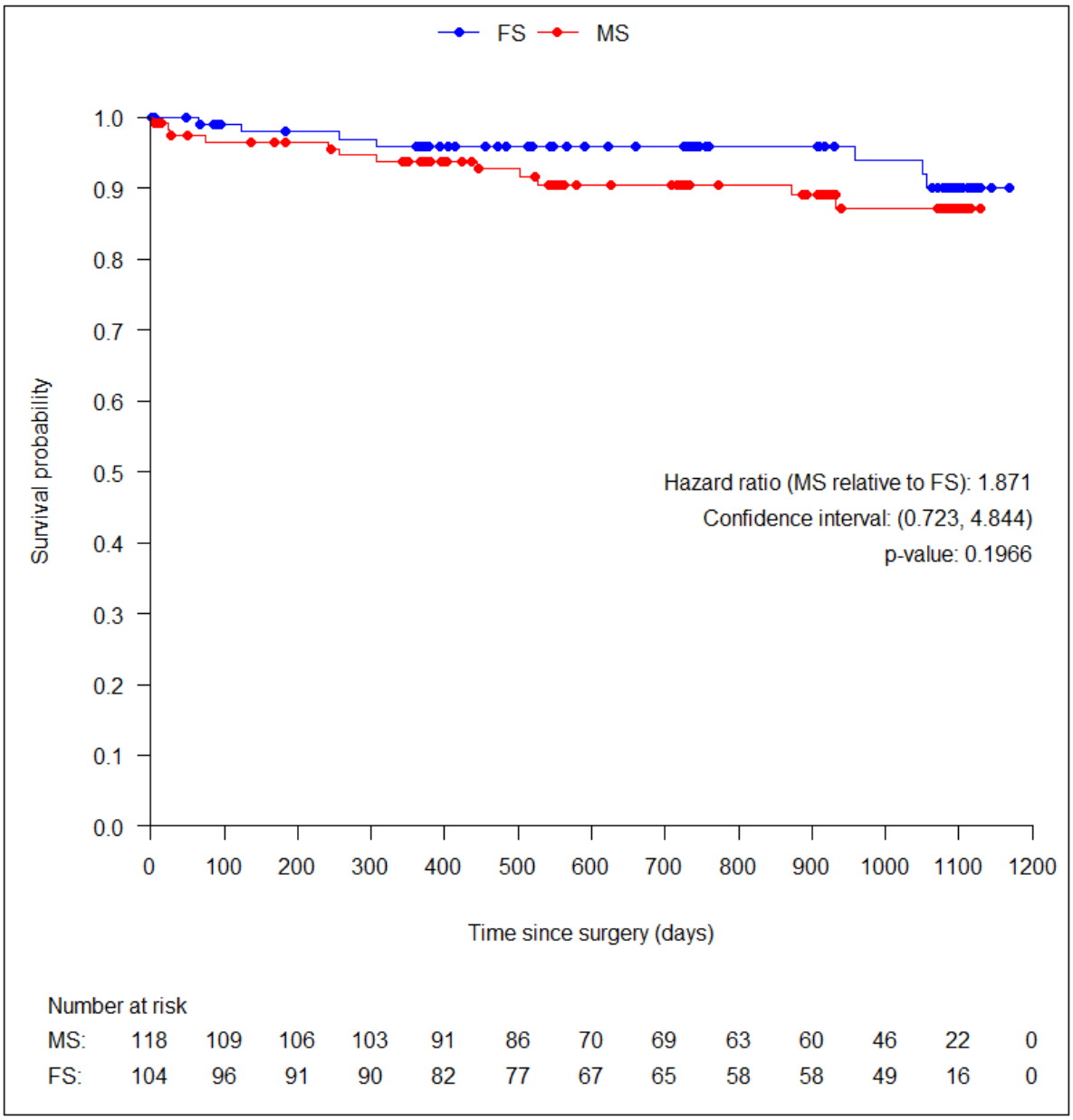
In total, 4 MS recipients and 13 FS recipients suffered wound infections within a year of surgery (one FS recipient suffered two infections). No patients who received a mechanical valve suffered a wound infection. Odds of wound infection were explored via logistic regression (complete case analysis in the safety population, with fixed treatment and sex effects, and with a random surgeon effect). The odds of suffering at least one wound infection were estimated to be lower for MS recipients than for FS recipients (MS/FS odds ratio 0.312, confidence interval 0.097 to 1.005, p-value 0.0511). Only two infections were categorised as deep (1 MS, 1FS).

Table A15. All deaths

	Treatment received	Treatment allocated	Cause	Relationship to treatment	Days from surgery to death
Cardiorespiratory	FS	FS	Endocarditis and sepsis.	Possibly related	124
	FS	FS	Lung infection.	Unrelated	1050
	FS	FS	Respiratory failure, pneumonia, chronic lymphocytic leukaemia.	Unrelated	1057
	MS	MS	Cardiac arrest and pericardial tamponade 2 days after surgery. Heart failure and left anterior pneumothorax 3 days after surgery.	Possibly related	3
	MS	MS	Type 2 respiratory failure and shock, multi-organ failure.	Possibly related	24
	MS	MS	Post-op arrest on HDU on day of surgery. Heart failure 26 days after surgery.	Possibly related	26
	MS	MS	Lower respiratory tract infection. Type 2 respiratory failure. NSTEMI during hospital admission.	Unrelated	75
	MS	MS	Endocarditis, infected valve. Refused all treatment including antibiotics. Palliation only.	Possibly related	241
	MS	MS	Exacerbation of COPD.	Unrelated	307
	MS	MS	Ischaemic heart disease.	Unrelated	502
	MS	MS	Myocardial infarction.	Unrelated	933
Non-cardiorespiratory	FS	FS	Sepsis.	Unrelated	66
	FS	FS	Metastatic prostate cancer.	Unrelated	256
	FS	FS	B cell lymphoma.	Unrelated	308
	FS	FS	Embolus of left common femoral artery, advanced colorectal cancer, AS, CHF.	Unrelated	958
	MS	MS	Metastatic bladder cancer.	Unrelated	257
	MS	MS	Death due to malignant tumour of oesophagus	Unrelated	445
	MS	MS	Diffuse large B cell lymphoma.	Unrelated	527
	MS	MS	Spontaneous subdural haemorrhage.	Unrelated	873

Table A15 shows that none of the patients who died were considered to be crossovers from MS to FS. However, there were three deaths among patients who were allocated and received MS but who were returned to theatre for redo FS. These were the deaths, all categorised as cardiorespiratory in Table A15, which occurred at 3, 26 and 933 days after surgery.

Figure A4. Kaplan-Meier curves for time to death by any cause



Patients are grouped by the treatment allocated to them. Patients who had no fatal events recorded were censored at the last time they were known to be alive. Times of censoring are indicated by points on the curves.

Appendix B: Economic Evaluation

Introduction

This trial collected data on resource and health service use for each patient during their in-patient stay through to the end of follow-up at 1 year. The economic analysis compared the costs and quality of life impacts of full and mini-sternotomy and assessed the cost-effectiveness of mini-sternotomy as an alternative to full median sternotomy.

The methods section first presents the unit costs, resource use data and the methods used to aggregate resource use and utility data at a patient level. The methods used to document and impute missing data follow. The last part describes the construction of incremental cost-effectiveness ratios and representation of uncertainty.

Results are presented first for raw data (with and without imputation) for costs and QALYs separately, followed by estimations of costs and QALYs that account for baseline differences. The final section provides results of probabilistic and deterministic sensitivity analyses.

Methods

Unit costs

All resource use data collected formed part of the patient-specific case-report form. Trained research nurses extracted data for inpatient stays from individual patient records. Face-to-face interviews with patients, by research nurses, provided data for quality of life as well as health service use during follow-up.

Multiplying the unit costs by each unit of resource use and summing these resource costs across each patient's 12 month follow-up from date of operation enabled aggregation of total cost per patient. Table B1 provides the unit costs used, with source of data. Where possible, national estimates of unit prices were used (e.g. PSSRU 2015 [1], NHS Ref 2014-15 [2]) to increase generalisability.

All resources were used once by patients (e.g. a GP visit or specific test), with the exception of two capital items used during surgery; the horizontal saw and defibrillator handles, both acquired for mini-sternotomy. These costs were apportioned, using clinical opinion, to each patient assuming a lifespan of 20 years and that surgeons undertake a total of 255 mini-sternotomies over five years.

Table B1. Unit costs

Item	Source	Consultation time/Codes	Mean 2014/15	SD
GP Visits	PSSRU 2015. 10.8b	Per patient contact lasting 17.2 minutes	£65.00	£13.00
GP Home Visits	PSSRU 2015. 10.8b	Per patient contact lasting 11.7 minutes	£45.00	£9.00
Nurse (GP Practice) Visits	PSSRU 2015. 10.6	Per patient contact 15.5 minutes	£14.47	£2.89
Nurse (Specialist Community) Home Visits	PSSRU 2015. 10.4	Per patient contact 15.5 minutes	£19.38	£3.88
Physiotherapy (outpatient)	NHS Ref 2014-15	Code: WF01A	£16.13	£3.23
Occupational Therapy (outpatient)	NHS Ref 2014-15	Code: WF01A	£16.67	£3.33
Physiotherapy (inpatient)	PSSRU 2015. 13.1	Per patient contact lasting 20 minutes	£12.67	£2.53
Occupational Therapy (inpatient)	PSSRU 2015. 13.2	Per patient contact lasting 20 minutes	£12.67	£2.53
Physiotherapy (home)	PSSRU 2015. 8.4.1	Per patient contact lasting 20 minutes	£27.00	£5.40
Theatre use	Papworth estimate		£20.00	£4.00
Horizontal surgical saw	Papworth estimate	20 year life span and are used in 255 surgeries in every 5 years	£3,138.22	£3.1
Paediatric internal cardioversion paddles			£161.71	£0.2

Internal paddle handle			£670.00	£0.7
*Reprocessing cost of defibrillator paddles for each surgery**		Per patient	£2.40	£2.40
Single use saw blade for mini-sternotomy		Per patient	£15.80	£15.80
Single use saw blade for full sternotomy		Per patient	£48.00	£48.00
Adult Critical Care	NHS Ref 2014-15	Total/weighted average	£1,274.92	£583.33
Specialised Ward	NHS Ref 2014-15	Code: SD01A	£387.96	£77.59
General Ward	NHS Ref 2014-15	Code: SD03A	£103.01	£20.60
Rehabilitation	PSSRU (1.3) 2015		£158.57	£31.71
24 hour Blood Pressure Monitoring	Lovibond et al. 2011, [3]		£61.47	£12.29
Radiography (chest)	Auguste et al. 2011, [4]		£3.46	£0.69
Echo TTE	NHS Ref 2014-15	Simple Echocardiogram	£83.94	£16.79
Echo TOE	NHS Ref 2014-15	Complex Echocardiogram	£128.49	£25.70
Echo Stress	NHS Ref 2014-15	Complex Echocardiogram	£128.49	£25.70
24 hour ECG	NHS Ref 2014-15	Electrocardiogram Monitoring	£140.69	£28.14
12 hour ECG	NHS Ref 2014-15	Electrocardiogram Monitoring	£140.69	£28.14
Exercise Tolerance Test	NHS Ref 2014-15	Electrocardiogram Monitoring	£140.69	£28.14
MRI scan	NHS Ref 2014-15	Total/weighted average	£146.15	£56.64
Full Pulmonary Function Testing	NHS Ref 2014-15	Code: DZ52Z	£55.32	£11.06
Cardiac Rehabilitation	NHS Ref 2014-15	Code: VC38Z	£97.84	£19.57
Cardio Clinic	NHS Ref 2014-15	Code: WF01A	£123.02	£24.60
Pacemaker	NHS Ref 2014-15	Code: EY08E	£76.32	£15.26
Blood tests	NHS Ref 2014-15	Code: DAPS08	£3.46	£0.69
Arrhythmia clinic	NHS Ref 2014-15	Total/weighted average	£131.14	£26.23
Wound clinic	NHS Ref 2014-15	Code: N25AF/AN	£54.93	£10.99
A&E visit	NHS Ref 2014-15	Total/weighted average	£140.59	£141.05
Computerised Tomography Scan	NHS Ref 2014-15	Total/weighted average	£122.31	£48.86

***** The lead clinician confirmed that: defibrillator is not routinely used and that the cost of paddles should apply to 30% of patients; and the cost of external defibrillator plates should be excluded for mini-sternotomy as the plate is only used when it is not possible to insert the paddles.

Patient-level aggregation of cost

This section describes the aggregation of costs, by patient, for the inpatient stay, post-discharge follow-up to 12 months and drug use.

Hospital stay: The time in the hospital from randomisation to discharge was disaggregated into theatre time, critical care unit (CCU) stay and cardiac ward stay as shown in Table B2. The total length of stay comprised time spent in surgery (measured in minutes), CCU (measured in hours) and cardiac ward (measured in days). Theatre time included duration of re-operations where applicable (a few patients had up to two returns to theatre) and corresponding CCU stays were added to the CCU hours. The total stay in the hospital, calculated using theatre time, critical care and ward stay, was compared with direct calculation of duration using date of operation and date of discharge to validate the breakdown of patient stay. After discharge from hospital, the majority of patients were discharged home but some were referred on to acute hospitals or rehabilitation centres (short or long term) for more care, and the costs of this additional stay were included.

Post-discharge: Resource use after discharge and up to twelve months post randomisation was collected at 6 week, 6 month and 12 month follow-up visits, with resource use divided into three categories: hospital admissions, tests and healthcare visits. A total of 28 different healthcare resources were used and aggregated over the follow-up period. For example, if a patient reported 1 blood test in discharge to 6 week follow-up period, 2 blood tests between 6 week to 6 month period and none after that, resource use was costed as £10.38 (3*£3.46) post discharge .

Table B2. Summary of resource use (without imputation)

		Full Sternotomy			Mini-sternotomy		
Primary Admission Costs	**Unit of measurement**	**Obs**	**Mean resource use/patient**	**SD**	**Obs**	**Mean resource use/patient**	**SD**
Theatre	Minutes	104	191.19	62.15	118	221.11	102.65
Critical care (ITU)	Hours	103	34.67	57.17	118	55.24	94.69
Cardiac ward	Days	103	7.09	4.31	118	6.90	3.87
Rehabilitation*	Days	103	2.45	11.90	117	1.68	10.27
Acute hospital*	Days	103	0.90	4.97	117	0.74	5.09
Physiotherapy (inpatient)	Days	103	5.90	4.21	117	5.90	5.16
Occupational therapy (inpatient)	Days	103	0.17	0.58	118	0.24	0.69
Follow-up (post discharge)							
ITU	Days	81	0.00	0.00	94	0.03	0.31
General ward	Days	92	2.87	14.37	101	0.86	3.43
Cardiac ward	Days	92	0.40	1.49	100	1.15	4.32
24 hour BP Monitoring	No. of tests	80	0.16	0.56	94	0.19	1.26
Radiography (chest)	No. of tests	80	0.49	0.89	94	0.64	0.90
Computerised Tomography Scan	No. of tests	80	0.14	0.52	94	0.15	0.51
Echo TTE	No. of tests	80	0.41	0.69	94	0.55	0.84
Echo TOE	No. of tests	80	0.03	0.22	92	0.03	0.18
Echo Stress	No. of tests	80	0.01	0.11	93	0.01	0.10
24 hour ECG	No. of tests	80	0.11	0.39	94	0.15	0.46
12 hour ECG	No. of tests	80	0.69	0.91	94	0.90	1.18
Exercise Tolerance Test	No. of tests	80	0.08	0.27	93	0.06	0.25
MRI scan	No. of tests	79	0.03	0.16	94	0.05	0.23
Full Pulmonary Function Testing	No. of tests	80	0.05	0.22	94	0.03	0.18
Blood test	No. of tests	81	0.05	0.22	94	0.06	0.35
A&E visit	No. of visits	80	0.09	0.28	94	0.22	0.51
Arrhythmia clinic	No. of visits	80	0.03	0.16	94	0.00	0.00
Cardiac Rehabilitation	No. of visits	79	0.84	2.76	93	0.32	1.43
Cardio Clinic	No. of visits	79	0.48	0.68	94	0.49	0.73
GP Home Visits	No. of visits	79	0.23	0.64	94	0.30	0.75
GP Visits	No. of visits	80	2.00	2.34	94	2.20	2.31
Nurse (Specialist Community) Home Visits	No. of visits	80	0.31	1.12	94	0.39	1.18
Nurse (GP Practice) Visits	No. of visits	80	2.10	10.02	92	0.75	1.46
Occupational therapy (outpatient)	No. of visits	80	0.11	0.71	94	0.06	0.62
Pacemaker	No. of visits	79	0.08	0.68	93	0.06	0.38
Physiotherapy (home)	No. of visits	80	0.05	0.35	94	0.00	0.00
Physiotherapy (outpatient)	No. of visits	80	0.04	0.19	94	0.01	0.10
Wound clinic	No. of visits	80	0.06	0.29	94	0.02	0.15
*discharged to convalescence/long term care/acute hospital instead of home							

Drugs: Drug use was matched to a corresponding unit cost using the NHS Electronic Drug tariff [5] and BNF [6] to sum costs across drug type for each patient.

Information on drugs administered during the primary admission was complete, with total amount of each drug per patient checked against patient prescriptions. However drug use post-discharge was self-reported and it was not possible to verify or retrieve any further data on this over the follow-up period.

Health State Utilities: This data was collected using EQ-5D-3L and SF-36 questionnaires. EQ-5D-3L responses were converted to utility values using Dolan et al (1995) [7] and to quality-adjusted life years (QALYs) for the trial period using the area under the curve method. SF-36 data was mapped to SF-6D utility values based on the ScHARR (School of Health and Related Research, University of Sheffield) algorithm and were converted to QALY scores (Brazier et al 2002 [8]). A value of 0 was assigned from date of death.

Missing data

The patterns of missing data for resource use and utilities were tested using Pearson Chi square goodness of fit and Wilcoxon rank sum tests for being missing at random and completely at random using the following variables: age, sex, treatment and health status at baseline (EQ-5D). The baseline characteristics assessed were not statistically significantly different between the two groups and multiple imputations were used for economic analysis. Patients were assigned zero cost and zero utility value from point of death.

Hospital stay: For primary admission, there were a few item non-responses for resource use data but no censored data. Complete information was available on all respondents barring one participant who withdrew from the trial after operation.

Post-discharge: The frequency of missing data for resource use after discharge is provided in Table B3 for the two groups. Imputation models did not converge at month twelve and resource use was aggregated over time, i.e. imputation was carried out for the aggregate value for each item rather than at each time period. The proportion of missing values in the aggregated utility data ranged from 11% to 25% in resource use post discharge (Table B3).

Table B3. Missing follow-up resource use

Follow up Resource Use	Full Sternotomy	Mini-sternotomy	Total
6 weeks			
Missing	3	4	7
Lost to follow up	4	6	10
Dead	1	4	5
Observations	96	104	200
6 months			
Missing	2	5	7
Lost to follow up	8	9	17
Dead	2	6	8
Observations	92	98	190
12 months			
Missing	9	4	13
Lost to follow up	11	13	24
Dead	4	7	11
Observations	80	94	174
Total	104	118	222

Drugs: Only drugs taken from randomisation to 12 month follow up period were accounted for (covering 3,078 drug uses of 118 different drugs). A number of assumptions (about quantity/dose and length of administration) were used to minimise the degree of missing information on drugs used. For example, when dosage or

frequency of dose per day was missing, the mode usage among trial participants was used or, if not available, the BNF dosage was used. Duration of medicinal use was calculated using start and stop dates for drugs used in primary admission and follow-up. However, when start/stop dates were missing, replies to a "yes/no" question on use of drugs at follow-up time points informed duration. For example if a drug was taken during inpatient stay, 6 week, 6 month and 12 month follow up, the drug was said to be used for entire 12 month trial period. However further assumptions about duration of medication were used when data was less forthcoming; for example drugs which were being taken only at 12 month follow up, without start date or stop date specified, were assumed to have been taken according to prescription every day for an average of three months (based on expert consultation). 58 records had insufficient information on usage for such personalised manual imputation, requiring predictive mean matching (conditioned on patient ID and name of drug).

Health State Utilities: EQ-5D-3L and SF-6D utility data were imputed at each follow-up as presented in Table B4, and percent of missing value ranged from 9% to 23%. Further breakdown of missing data for resource use and HRQoL questionnaires, and imputation required for each variable is provided in Table B4.

Table B4. Incomplete data and imputation

	Full Sternotomy				Mini-sternotomy			
Resource Use	**Complete**	**Incomplete**	**Imputed**	**Total**	**Complete**	**Incomplete**	**Imputed**	**Total**
Primary admission								
Theatre time (minutes)	104	0	0	104	118	0	0	118
Critical care stay (hours)	103	1	1	104	118	0	0	118
Cardiac ward stay (days)	103	1	1	104	118	0	0	118
Rehabilitation days*	103	1	1	104	117	1	1	118
Acute hospital days*	103	1	1	104	117	1	1	118
Physiotherapy visits	103	1	1	104	117	1	1	118
Occupational therapy visits	103	1	1	104	118	0	0	118
Follow-up (post discharge)								
Post discharge ITU days	81	23	23	104	94	24	24	118
Post discharge general ward stay	92	12	12	104	101	17	17	118
Post discharge cardiac ward stay	92	12	12	104	100	18	18	118
24 hour BP Monitoring	80	24	24	104	94	24	24	118
Radiography (chest)	80	24	24	104	94	24	24	118
Computerised Tomography Scan	80	24	24	104	94	24	24	118
Echo TTE	80	24	24	104	94	24	24	118
Echo TOE	80	24	24	104	92	26	26	118
Echo Stress	80	24	24	104	93	25	25	118
24 hour ECG	80	24	24	104	94	24	24	118
12 hour ECG	80	24	24	104	94	24	24	118
Exercise Tolerance Test	80	24	24	104	93	25	25	118
MRI scan	79	25	25	104	94	24	24	118
Pulmonary Function Testing	80	24	24	104	94	24	24	118
Blood test	81	23	23	104	94	24	24	118
A&E visit	80	24	24	104	94	24	24	118
Arrhythmia clinic	80	24	24	104	94	24	24	118
Cardiac Rehabilitation	79	25	25	104	93	25	25	118
Cardio Clinic	79	25	25	104	94	24	24	118
GP Home Visits	79	25	25	104	94	24	24	118

GP Visits	80	24	24	104	94	24	24	118
Nurse (Specialist Community) Home Visits	80	24	24	104	94	24	24	118
Nurse (GP Practice) Visits	80	24	24	104	92	26	26	118
Occupational therapy	80	24	24	104	94	24	24	118
Pacemaker	79	25	25	104	93	25	25	118
Physiotherapy (home)	80	24	24	104	94	24	24	118
Physiotherapy	80	24	24	104	94	24	24	118
Wound clinic	80	24	24	104	94	24	24	118
EQ-5D Score								
Baseline	95	9	9	104	105	13	13	118
4 Days Post Operation	89	15	15	104	92	26	26	118
Discharge	88	16	16	104	103	15	15	118
6 weeks follow-up	88	16	16	104	106	12	12	118
6 months follow-up	95	9	9	104	105	13	13	118
12 months follow-up	84	20	20	104	103	15	15	118
SF-6D Score								
Baseline	89	15	15	104	101	17	17	118
6 weeks follow-up	88	16	16	104	102	16	16	118
6 months follow-up	90	14	14	104	102	16	16	118
12 months follow-up	82	22	22	104	91	27	27	118

Imputation

Missing values were imputed conditional on sex, age, type of replacement valve used, risk classification measured using New York Heart Association (NYHA) Functional Classification and Canadian Cardiovascular Society (CCS) grading of angina. To avoid loss in efficiency, missing values for resource use and utility values at different time points were replaced using multiple imputations by chained equations.

Chained predictive mean matching was used to replace missing data for resource use and quality of life variables, and a total of 20 imputed datasets were created, stratified by treatment group. The imputed resource use is summarised in Table B5. However while conducting probabilistic analysis using bootstrap method; multiple imputation was carried out only once for each iteration with a total of 1000 iterations to adequately retain between imputation variance. The distribution of imputed values was visually checked for comparability with the observed data.

Table B5. Summary of resource use

		Full Sternotomy			Mini-sternotomy		
Primary Admission Costs	**Unit of measurement**	**Obs**	**Mean resource use/ patient**	**SD**	**Obs**	**Mean resource use/ patient**	**SD**
Theatre	Minutes	104	191.19	62.15	118	221.11	102.65
Critical care (ITU)	Hours	104	34.52	56.91	118	55.24	94.69
Cardiac ward	Days	104	7.07	4.29	118	6.90	3.87
Rehabilitation*	Days	104	2.42	11.84	118	1.66	10.22
Acute hospital*	Days	104	0.89	4.95	118	0.77	5.08
Physiotherapy (inpatient)	Days	104	5.88	4.20	118	5.94	5.15
Occupational therapy (inpatient)	Days	104	0.17	0.58	118	0.24	0.69

Follow-up (post discharge)							
ITU	Days	104	0.00	0.00	118	0.03	0.28
General ward	Days	104	2.61	13.55	118	0.77	3.20
Cardiac ward	Days	104	0.38	1.43	118	1.19	4.14
24 hour BP Monitoring	No. tests	104	0.18	0.52	118	0.17	1.13
Radiography (chest)	No. tests	104	0.55	0.87	118	0.61	0.83
CT Scan	No. tests	104	0.16	0.48	118	0.16	0.49
Echo TTE	No. tests	104	0.42	0.66	118	0.56	0.79
Echo TOE	No. tests	104	0.02	0.20	118	0.05	0.19
Echo Stress	No. tests	104	0.01	0.10	118	0.01	0.09
24 hour ECG	No. tests	104	0.13	0.41	118	0.16	0.44
12 hour ECG	No. tests	104	0.72	0.85	118	0.94	1.17
Exercise Tolerance Test	No. tests	104	0.07	0.24	118	0.06	0.23
MRI scan	No. tests	104	0.02	0.15	118	0.06	0.22
Full Pulmonary Function Testing	No. tests	104	0.06	0.22	118	0.03	0.16
Blood test	No. tests	104	0.06	0.21	118	0.07	0.33
A&E visit	No. visits	104	0.13	0.31	118	0.24	0.50
Arrhythmia clinic	No. visits	104	0.02	0.14	118	0.00	0.00
Cardiac Rehabilitation	No. visits	104	1.07	2.78	118	0.34	1.36
Cardio Clinic	No. visits	104	0.47	0.62	118	0.52	0.72
GP Home Visits	No. visits	104	0.27	0.64	118	0.25	0.68
GP Visits	No. visits	104	2.00	2.16	118	2.17	2.18
Nurse (Specialist Community) Home Visits	No. visits	104	0.38	1.06	118	0.47	1.22
Nurse (GP Practice) Visits	No. visits	104	1.93	8.83	118	0.71	1.32
Occupational therapy	No. visits	104	0.15	0.70	118	0.05	0.55
Pacemaker	No. visits	104	0.06	0.59	118	0.08	0.39
Physiotherapy (home)	No. visits	104	0.05	0.32	118	0.00	0.00
Physiotherapy	No. visits	104	0.05	0.20	118	0.02	0.11
Wound clinic	No. visits	104	0.06	0.28	118	0.03	0.15
*discharged to convalescence/long term care/acute hospital instead of home							

Adjustment method

To account for differences in baseline utility values, as well as skewness, censoring and confounding in cost data, linear regression models were used to provide adjusted estimates of mean values. Control variables used were age, sex, valve, EQ-5D-3L baseline value and treatment arm. The type of valve used for replacement was also controlled for, because it was used as a stratification factor in the randomisation.

Incremental cost effectiveness analysis and sensitivity analyses

Differences in estimated costs and EQ-5D QALYs between trial arms, using raw data with imputation, were tested using two-sample t-test with equal variances.

Incremental cost-effectiveness ratios were also constructed using adjusted mean estimates of costs and QALYs using 'seemingly unrelated regression', to account for correlation between costs and effects at the patient-level. This regression technique relies on the multivariate normality of the group-specific mean costs and QALYs, and is valid where the individual costs and QALYs are skewed (Faria et al 2014, [9]).

Probabilistic Sensitivity Analysis (PSA) was used to characterise the uncertainty of input parameters and a bootstrap approach (with 1000 bootstrapped samples) was applied to estimate the precision of results. The probability that mini-sternotomy is cost-effective when compared to full sternotomy is presented, at varying willingness to pay (WTP) threshold values, using a Cost Effectiveness Acceptability Curve (CEAC) and incremental net monetary benefit.

Deterministic sensitivity analyses and scenario analysis were used to explore the robustness of cost-effectiveness results that adopted different methodological approaches or assumptions (see Table B6). Baseline characteristics were assessed using Chi square and rank sum test, to assess whether patients included in the complete case analysis were different from those outside the complete case analysis.

Table B6. Summary of deterministic sensitivity and scenario analyses undertaken

Sensitivity analyses	Rationale
1. Complete case analysis	Only including respondents with no missing values across all variables and across follow-up; to check results in sample requiring no missing value imputation
2. Excluding patients who died during primary admission	Patients who died during primary admission were the main cost driver and required substantial surgical time and cardiac care; to assess whether excluding these patients would change recommendations.
3. Excluding additional equipment cost required	Assuming the additional equipment required for the surgeries already exists in the trusts;
4. Excluding follow-up resource use	To test the assumption that the cost difference between the two arms were accrued during primary admission, to allow comparison with literature that missed these costs, but still retain benefits as captured in other studies.
5. Excluding follow-up resource use and utility data	Data up to discharge had few missing values; also to assess impact of having a shorter cut-off time point for trial (as wider literature had) but provide a less biased analysis that measures benefits but not costs.
6. Use SF-6D utility values	SF-6D values used as an alternative construction for QALYs

Results

The comparison of mean costs per patient up to one year (see Table B7), using raw data with imputation, shows that mini-sternotomy was £1,714 more than median sternotomy although this was not statistically significant. The higher costs resulted from longer surgery time, additional equipment and longer time in critical care. EQ-5D QALYs were very slightly higher in the mini-sternotomy arm compared with full sternotomy (difference 0.0279), but this was not statistically significantly so (see Table B8), and there was no statistically significant difference in SF6D QALYs either. Figures B1 and B2 illustrate the distribution of total costs and QALYs across the patients in the trial.

Table B9 summarises the comparison of costs and QALYs. The additional cost of gaining an additional QALY using mini-sternotomy rather than median sternotomy when imputed using PMM method is £61,379 and the net monetary loss at a willingness to pay (WTP) of £20,000 is £1,155.

Seemingly unrelated regression analysis of costs and QALYs, adjusted for baseline characteristics showed that, in terms of QALYS, mini-sternotomy was not statistically significantly different from full sternotomy. Table B10 also shows that the coefficient for cost was positive, indicating mini-sternotomy was more costly than full sternotomy and that this difference was statistically significant. Mini-sternotomy is therefore dominated by median sternotomy. The cost effectiveness plane for the analysis is illustrated in Figure B3.

The probabilistic sensitivity analysis shows (see Figure B4) that, at a WTP per QALY of £20,000, there is a 3.7% likelihood that mini-sternotomy is cost-effective compared with median sternotomy and that this likelihood rises to 5.1% at a WTP of £30,000/QALY. The net monetary benefit of mini-sternotomy is negative across all WTP threshold values (Figure B5).

Deterministic sensitivity analyses (see Table B11) showed that mini-sternotomy was either dominated or had a huge ICER. The one exception to this was the complete case analysis (CCA-cost-effectiveness), which found mini-sternotomy to be cost-effective. The intervention cost less but also had slightly worse outcomes in this sample size, which was limited to only 90 cases. The result indicates a saving of £10,000 for a loss of one QALY. The sample is not representative of those with missing data and consisted a larger proportion of females than the sample outside the CCA-cost-effectiveness sample. The sensitivity analyses conducted using PSA (Table B12) consistently found full sternotomy to be a superior intervention to mini-sternotomy. The cost effectiveness planes for the sensitivity analyses are illustrated in Figure B6.

Table B7: Comparison of mean costs (SD) per patient up to 12 months post-randomisation (with imputation) (UK pounds, 2015)

		Full Sternotomy			Mini-sternotomy		
	Mean Unit cost	**Obs**	**Mean cost/ patient**	**SD**	**Obs**	**Mean cost/ patient**	**SD**
Primary Admission Costs							
Additional surgical items							
Horizontal surgical saw	£3,138.2	104	£0.0	£0.0	118	£3.1	£0.0
Single use saw blade for mini-sternotomy	£48.0	104	£0.0	£0.0	118	£48.0	£0.0
Single use saw blade for full sternotomy	£15.8	104	£15.8	£0.0	118	£0.0	£0.0
Paediatric internal cardioversion paddles	£161.7	104	£0.0	£0.0	118	£0.2	£0.0
Reprocessing cost of defibrillator paddles for each surgery	£2.4	104	£2.4	£0.0	118	£2.4	£0.0
Internal paddle handle	£670.0	104	£0.0	£0.0	118	£0.7	£0.0
*Cost of additional surgical items***		104	£16.52	£0.0	118	£52.0	£0.0
Theatre	£20.0	104	£3,823.8	£1,243.0	118	£4,422.2	£2,053.0
Critical care (ITU)	£1,274.9	104	£1,833.8	£3,023.2	118	£2,934.2	£5,029.9
Cardiac ward	£388.0	104	£2,743.7	£1,664.0	118	£2,676.3	£1,499.9
Rehabilitation*	£158.6	104	£384.2	£1,877.6	118	£263.4	£1,621.3
Acute hospital*	£388.0	104	£346.9	£1,918.9	118	£297.5	£1,971.3
Physiotherapy (inpatient)	£12.7	104	£74.5	£53.2	118	£75.2	£65.3
Occupational therapy (inpatient)	£12.7	104	£2.1	£7.3	118	£3.0	£8.7
Subtotal (primary admission)	**-**	**104**	**£9225.7**	**£6510.8**	**118**	**£10723.9**	**£8850.2**
Post Primary Admission Costs							
Hospital Admission							
ITU	£1,274.9	104	£0.0	£0.0	118	£32.4	£352.1
General ward	£103.0	104	£268.4	£1,395.4	118	£79.4	£329.5
Cardiac ward	£388.0	104	£149.2	£554.8	118	£463.6	£1,606.4
Tests							
24 hour Blood Pressure Monitoring	£61.5	104	£10.9	£32.0	118	£10.2	£69.5
Radiography (chest)	£3.5	104	£19.4	£30.9	118	£21.6	£29.5
Computerised Tomography Scan	£122.3	104	£19.4	£58.6	118	£19.7	£59.8
Echo TTE	£83.9	104	£35.1	£55.2	118	£46.9	£66.6
Echo TOE	£128.5	104	£2.5	£25.2	118	£6.5	£24.3
Echo Stress	£128.5	104	£1.2	£12.6	118	£1.1	£11.8
24 hour ECG	£140.7	104	£18.3	£57.2	118	£22.7	£62.3
12 hour ECG	£140.7	104	£101.5	£119.6	118	£132.9	£165.0
Exercise Tolerance Test	£140.7	104	£9.5	£34.0	118	£8.9	£32.6

MRI scan	£146.2	104	£3.5	£21.3	118	£9.3	£32.5
Full Pulmonary Function Testing	£55.3	104	£3.2	£12.4	118	£1.6	£9.1
Blood test	£3.5	104	£0.0	£0.1	118	£0.0	£0.1
Healthcare visits							
A&E visit	£140.6	104	£18.9	£43.0	118	£33.4	£70.4
Arrhythmia clinic	£131.1	104	£2.5	£18.1	118	£0.0	£0.0
Cardiac Rehabilitation	£97.8	104	£104.4	£271.9	118	£33.6	£133.4
Cardio Clinic	£123.0	104	£57.4	£76.3	118	£63.6	£88.1
GP Home Visits	£45.0	104	£12.1	£28.9	118	£11.3	£30.4
GP Visits	£65.0	104	£129.7	£140.6	118	£141.3	£141.8
Nurse (Specialist Community) Home Visits	£19.4	104	£7.3	£20.6	118	£9.0	£23.6
Nurse (GP Practice) Visits	£14.5	104	£28.0	£127.7	118	£10.3	£19.2
Occupational therapy (outpatient)	£16.7	104	£2.5	£11.7	118	£0.8	£9.2
Pacemaker	£76.3	104	£4.4	£44.9	118	£6.1	£29.5
Physiotherapy (home)	£27.0	104	£1.4	£8.6	118	£0.0	£0.0
Physiotherapy (outpatient)	£16.1	104	£0.8	£3.4	118	£0.3	£1.9
Wound clinic	£54.9	104	£3.4	£15.2	118	£1.6	£8.3
Subtotal (post-primary admission)	**-**	**104**	**£1014.9**	**£1777.5**	**118**	**£1168.2**	**£2077.9**
Drugs (total)	**-**	**104**	**£379.4**	**£548.2**	**118**	**£441.4**	**£976.7**
Total cost		**104**	**£10,620.0**	**£7,623.8**	**118**	**£12,333.5**	**£9,864.2**

*discharged to convalescence/long term care/acute hospital instead of home

**mean cost per patient estimated by assuming that the saw, paddle and handle have a twenty year life span and are used in 255 surgeries in every 5 years; NB: defib (paddle, handle and sterilisation cost) applicable in only 30% of cases

Table B8. Summary of utility values and QALYs

	Full Sternotomy			**Mini-sternotomy**		
EQ-5D	**Obs**	**Mean Utility**	**SD**	**Obs**	**Mean Utility**	**SD**
Baseline	104	0.6988	0.24	118	0.7793	0.18
4 Days Post Operation	104	0.3721	0.29	118	0.4430	0.28
Discharge	104	0.5815	0.23	118	0.5940	0.25
6 weeks follow-up	104	0.6930	0.21	118	0.7195	0.24
6 months follow-up	104	0.8272	0.22	118	0.8322	0.24
12 months follow-up	104	0.7584	0.29	118	0.8253	0.29
EQ-5D QALYs	104	0.7699	0.19	118	0.7978	0.21
	Full Sternotomy			**Mini-sternotomy**		
SF-6D	**Obs**	**Mean Utility**	**SD**	**Obs**	**Mean Utility**	**SD**
Baseline	104	0.6418	0.11	118	0.6802	0.12
6 weeks follow-up	104	0.6327	0.10	118	0.6356	0.14
6 months follow-up	104	0.7184	0.16	118	0.7332	0.19
12 months follow-up	104	0.6868	0.19	118	0.7058	0.23
SF-6D QALYs	104	0.6847	0.12	118	0.6989	0.16

Figure B1. Distribution of total cost

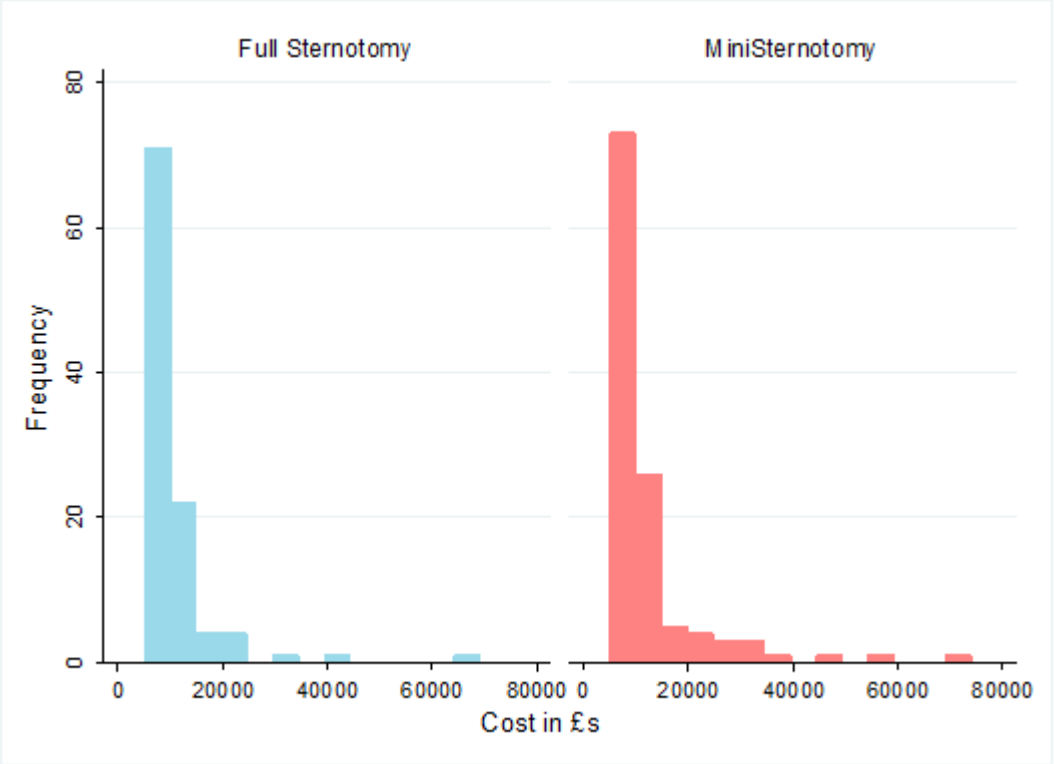


Figure B2. Distribution of QALYs

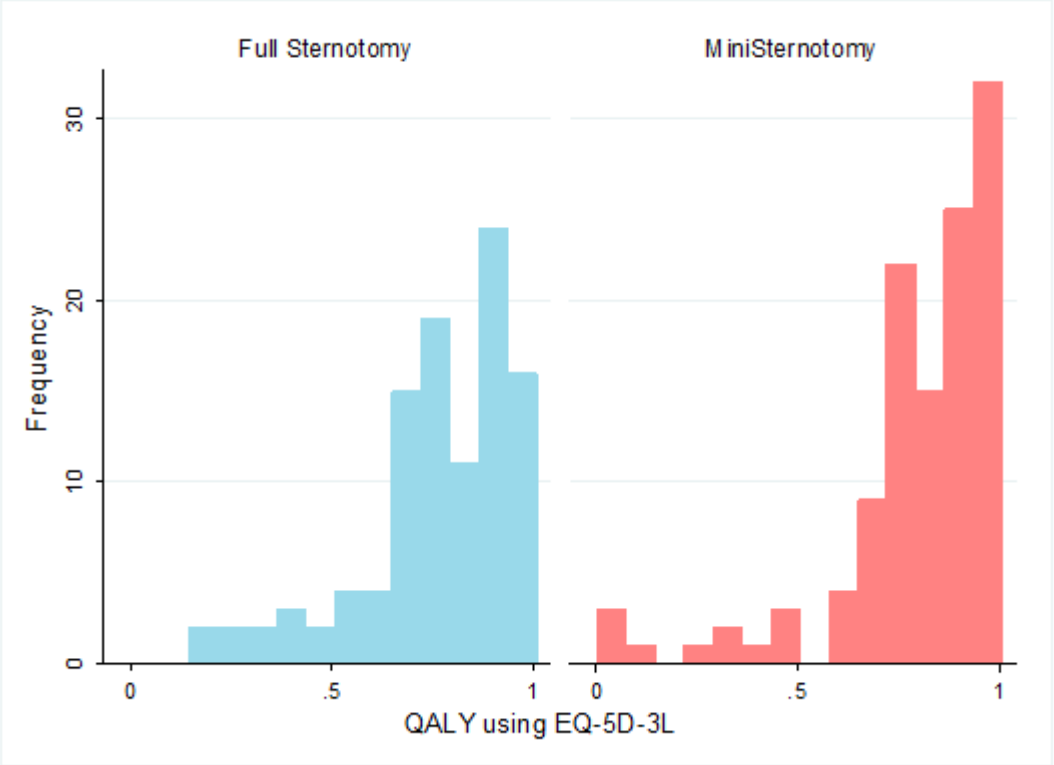


Table B9. Comparison of costs and QALYS (raw data, with imputation)

	Full Sternotomy (n=104)		Mini-sternotomy (n=114)	
	Mean	**SD**	**Mean**	**SD**
Total costs over 12 months	£10,620	£7,624	£12,334	£9,864
Incremental cost at 12 months (MS-FS)		-	£1,714	
Total EQ5D3L QALYs	0.7699	0.19	0.7978	0.21
Incremental EQ5D3L QALYs (MS-FS)		-	0.0279	
ICER		-	£61,379	
INMB at WTP of £20,000/QALY		-	-£1,155	
INMB at WTP of £30,000/QALY		-	-£876	

Table B10. Regression estimates of costs and QALYs

Dependant variable: EQ5D QALYs					
	Coefficient	**Std. Err.**	**P value**	**[95% Conf. Interval]**	
Mini-sternotomy	-0.0040	0.0245	0.87	-0.0520	0.0440
Male	0.0250	0.0246	0.31	-0.0231	0.0732
Age	-0.0051	0.0014	0.00	-0.0078	-0.0024
Baseline EQ-5D score	0.3037	0.0590	0.00	0.1880	0.4194
Tissue valve	0.0794	0.0459	0.08	-0.0107	0.1694
Constant	0.7391	0.1093	0.00	0.5249	0.9533
Dependant variable: Total Cost (£)					
	Coefficient	**Std. Err.**	**P value**	**[95% Conf. Interval]**	
Mini-sternotomy	2010.22	1201.57	0.09	-344.82	4365.25
Male	-1275.52	1205.23	0.29	-3637.73	1086.70
Age	98.32	67.58	0.15	-34.13	230.77
Baseline EQ-5D score	-983.50	2896.40	0.73	-6660.34	4693.33
Tissue valve	-853.43	2254.14	0.71	-5271.45	3564.60
Constant	5704.71	5362.01	0.29	-4804.64	16214.06

Table B11. Deterministic sensitivity analysis (using difference MS - FS, adjusted for baseline)

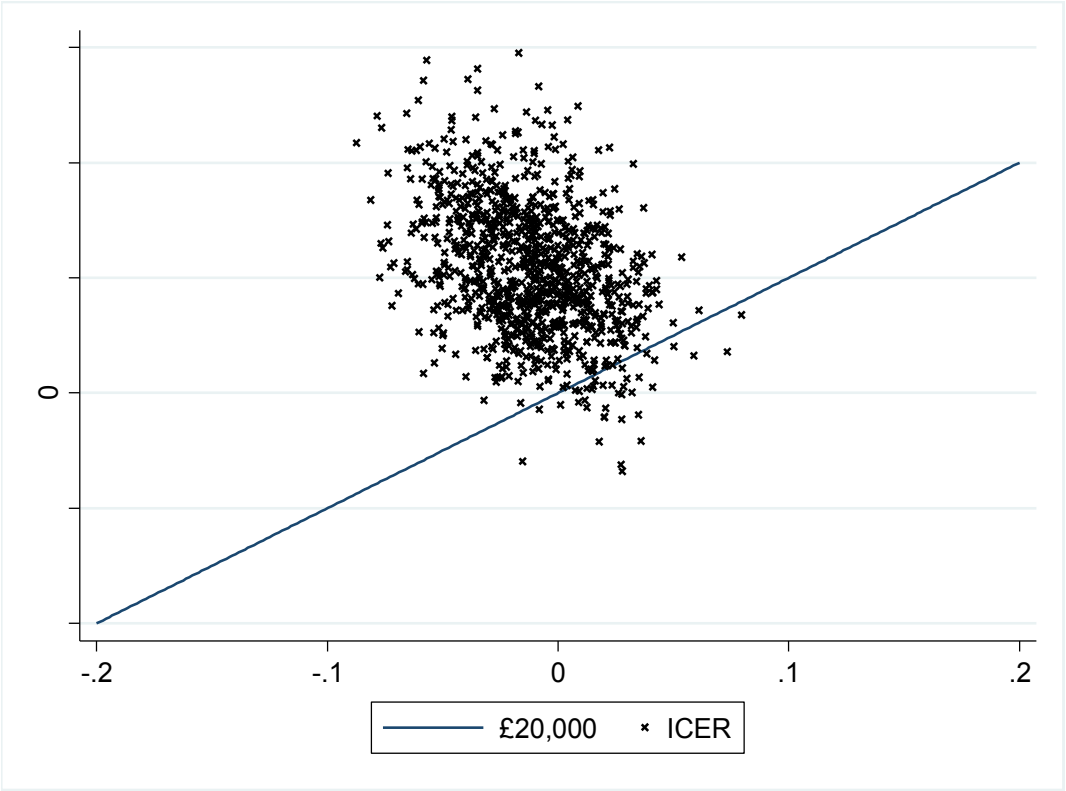
	Obs	Incremental cost over 12 months (MS-FS)		Incremental QALYs over 12 months (MS-FS)		ICER	INMB at £20,000 per QALY	INMB at £30,000 per QALY
		Mean	Std Error	Mean	Std Error			
Missing values imputed by PMM	222	£2,010	£1,202	-0.0040	0.0245	Dominated	-£2,089.26	-£2,128.78
Using SF6D QALYs	222	£2,010	£1,202	-0.0017	0.0178	Dominated	-£2,044.44	-£2,061.55
Assuming there is no additional equipment required for the two procedures	222	£1,975	£1,202	-0.0040	0.0245	Dominated	-£2,053.73	-£2,093.26
Excluding follow-up resource use	222	£1,664	£1,060	-0.0040	0.0245	Dominated	-£1,742.98	-£1,782.50
Complete case analysis	90	-£150	£661	-0.0145	0.0334	£10,333.62	-£139.89	-£284.60
Excluding patients who died during primary admission	219	£1,408	£1,128	0.0172	0.0216	£81,905.62	-£1,064.40	-£892.46
Including costs and QALY data only up to discharge	222	£1,664	£1,060	0.0013	0.0009	£1,316,409.02	-£1,638.66	-£1,626.02

Table B12. Probabilistic sensitivity analysis (using difference MS - FS, adjusted for baseline)

	Obs	Incremental cost over 12 months (MS-FS)		Incremental QALYs over 12 months (MS-FS)		ICER	INMB at £20000	INMB at £30000
		Mean	Std Error	Mean	Std Error			
Missing values imputed by PMM and adjusted	1000	£2,154	£36	-0.0122	0.0008	Dominated	-£2,396.99	-£2,518.59
Using SF6D QALYs	1000	£2,154	£36	-0.0075	0.0006	Dominated	-£2,303.03	-£2,377.66
Assuming there is no additional equipment required for the two procedures	1000	£2,245	£40	-0.0096	0.0008	Dominated	-£2,437.25	-£2,533.50
Excluding follow-up resource use	1000	£1,835	£35	-0.0131	0.0008	Dominated	-£2,096.58	-£2,227.15
Complete case analysis	1000	-£111	£22	-0.0121	0.0011	£9,170.78	-£130.56	-£251.12
Excluding patients who died during primary admission	1000	£1,433	£32	0.0147	0.0007	£97,425.25	-£1,138.55	-£991.50
Including costs and QALY data only up to discharge	1000	£1,835	£35	0.0008	0.0000	£2,415,384.92	-£1,820.25	-£1,812.65

Figure B3. Cost effectiveness plane (using difference MS-FS, adjusted for baseline)

B3.1 Using EQ-5D to estimate QALY



B3.2 Using SF-36 to estimate QALY

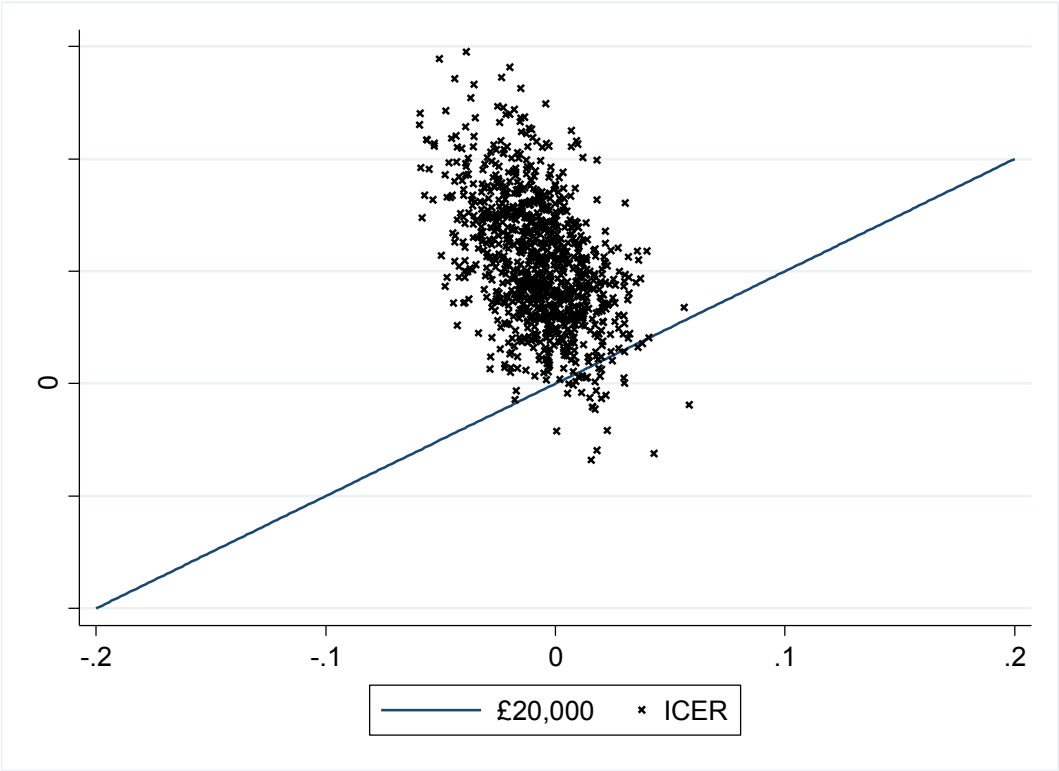


Figure B4. Cost-effectiveness acceptability curve (EQ-5D)

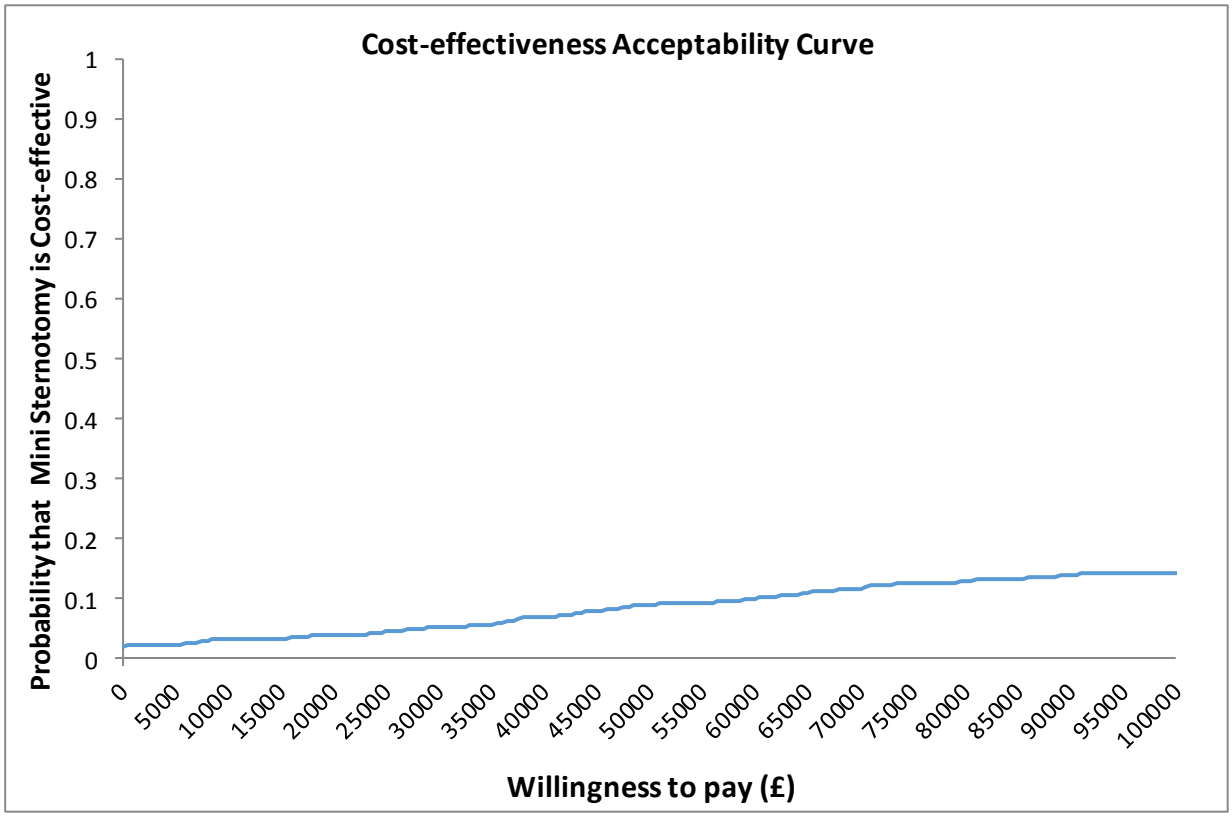


Figure B5. Net monetary benefit (controlling for baseline characteristics and missing data)

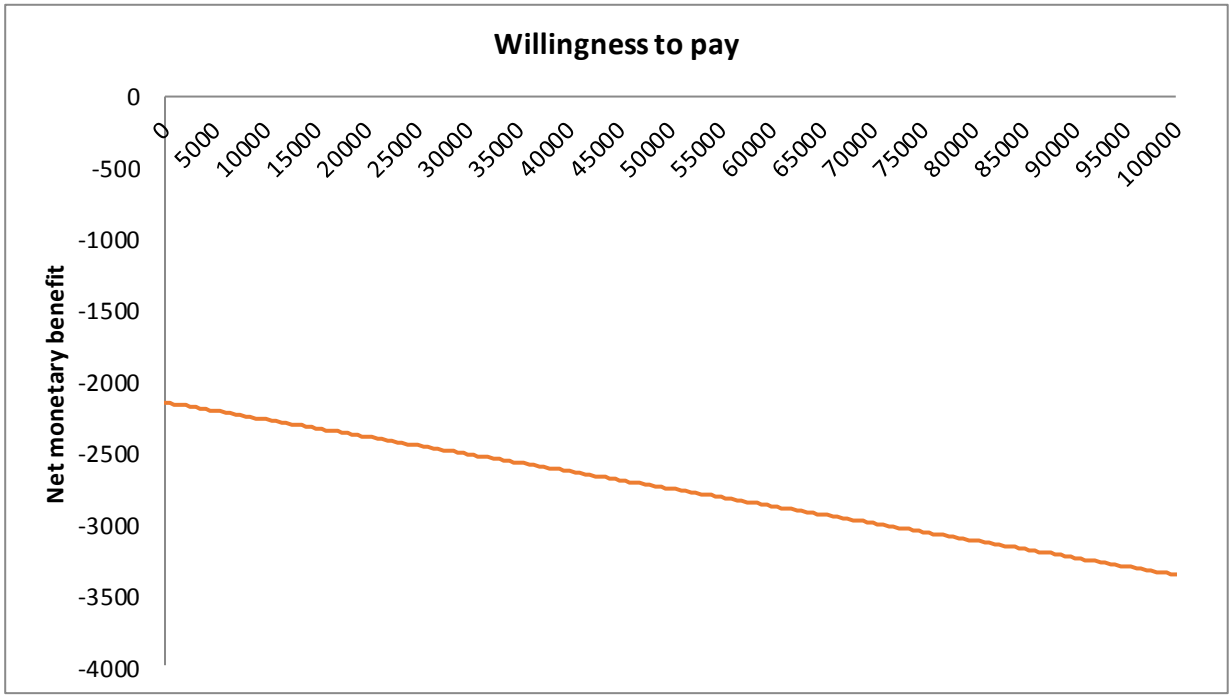
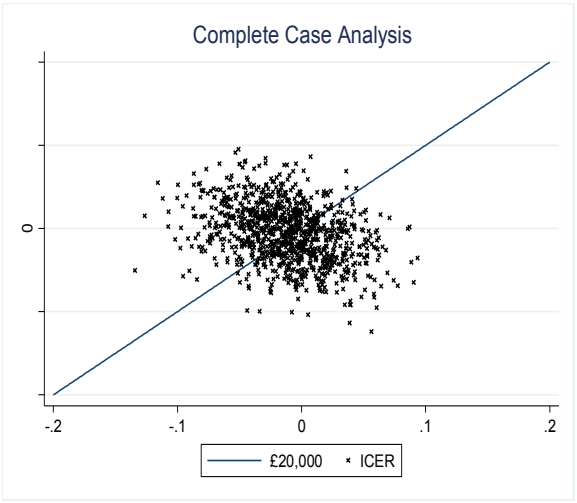
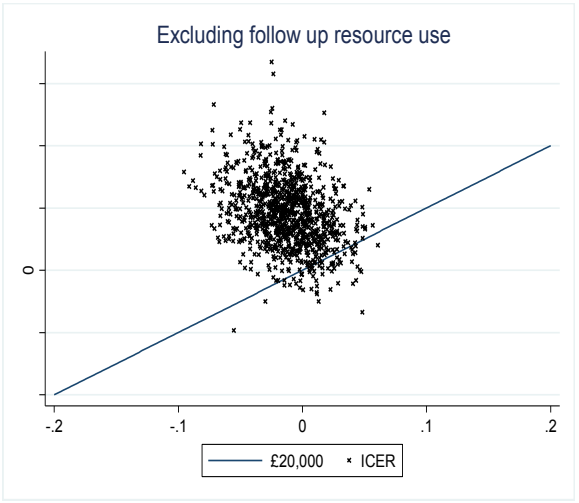
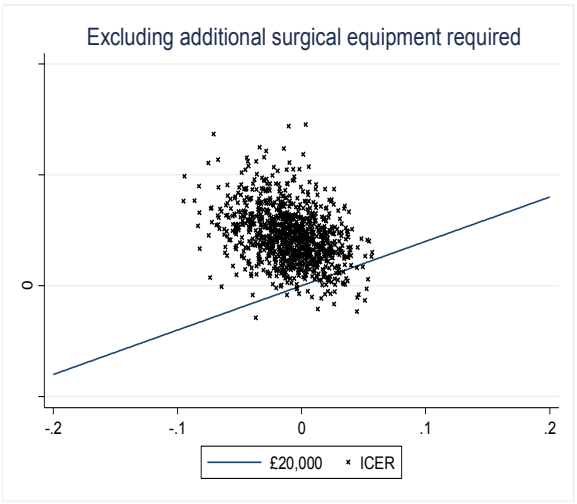
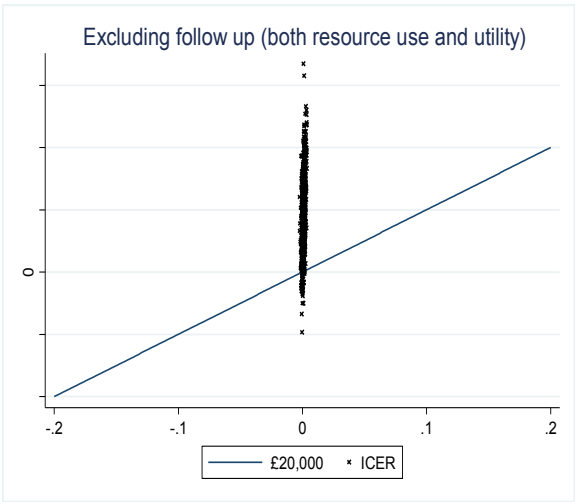


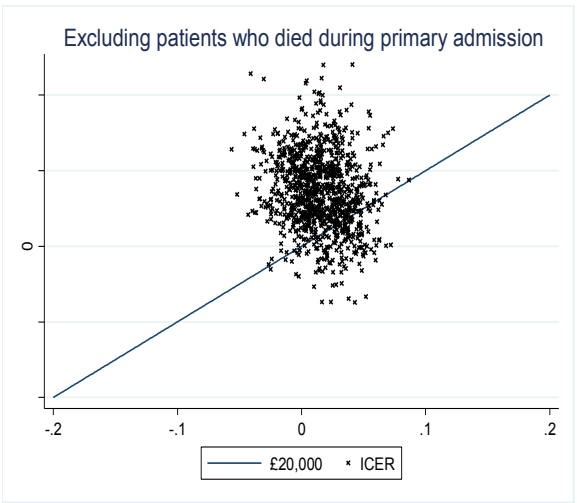
Figure B6. Sensitivity analyses using difference (MS - FS), adjusted for baseline











References

[1] Personal Social Services Research Unit (PSSRU). Unit costs of Health and Social Care (2015): Accessed July 2016. http://www.pssru.ac.uk

[2] Department of Health. NHS Reference costs (2014/15): Published November 2015: Accessed July 2016. https://www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015

[3] Lovibond K, Jowett S, Barton P, Caulfield M, Heneghan C, Hobbs FD, Hodgkinson J, Mant J, Martin U, Williams B, Wonderling D, McManus RJ. Cost-effectiveness of options for the diagnosis of high blood pressure in primary care: a modelling study. Lancet. 2011; 378:1219–1230

[4] Auguste P, Barton P, Hyde C, et al. An economic evaluation of positron emission tomography (PET) and positron emission tomography/computed tomography (PET/CT) for the diagnosis of breast cancer recurrence. Health Technol Assess 15: iii-iv, 1-54, 2011

[5] NHS Electronic Drug Tariff (2016): Accessed July 2016. http://www.nhsbsa.nhs.uk/PrescriptionServices/4940.aspx

[6] BNF. British National Formulary (2016): Accessed July 2016. https://www.nice.org.uk/about/what-we-do/evidence-services/british-national-formulary

[7] Dolan P, Gudex C, Kind P, Williams A. A social tariff for EuroQol: results from a UK general population survey University of York Center for Health Economics 1995:1-24

[8] Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21:271-92

[9] Faria, R., Gomes, M., Epstein, D., & White, I. R. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. Pharmacoeconomics. 2014 Dec; 32(12):1157-70.