SUPPLEMENTAL MATERIAL

Interfacility Transfer of Medicare Beneficiaries with Acute Type A Aortic Dissection and Regionalization of Care in the United States

Interfacility Transfer of Medicare Beneficiaries with Acute Type A Aortic Dissection and Regionalization of Care in the United States

Supplemental Material

Andrew B. Goldstone, MD, PhD* Peter Chiu, MD, MS* Michael Baiocchi, PhD,* Bharathi Lingala, PhD, Justin Lee, MPH, Joseph Rigdon, PhD, Michael P. Fischbein, MD, PhD, Y. Joseph Woo, MD

Supplemental Methods

Funding Sources

This work was conducted with support from a TL1 Clinical Research Training Program and KL2 Mentored Career Development Award of the Stanford Clinical and Translational Science Award to Spectrum (ABG - NIH TL1 TR 001084; PC - NIH KL2 TR 001083), as well as a grant from the Agency for Healthcare Research and Quality (MB – KHS022192A).

Statistical Analysis

In this supplemental statistical analysis section, we will: 1) provide estimates of the effect of receiving care at high-volume versus low-volume centers (third hypothesis); 2) further substantiate our use of instrumental variable methods; 3) evaluate the effect of preferential routing of patients with chest pain to hospitals capable of performing percutaneous coronary intervention; 4) validate our definition of type A aortic dissection; 5) describe our *a priori* power calculation; 6) provide specific details regarding our matching algorithms; 7) explain the intuition behind using restricted mean survival time for effect estimates on overall survival; 8) elaborate on the gamma sensitivity estimate of susceptibility to unmeasured bias; 9) explore the consequences of abandoning the instrumental variable – that is, moving away from using the pseudo-randomization and relying instead upon the assumptions common to propensity score matching (i.e., ignorable treatment assignment); and 10) list the statistical packages used during the analysis.

Using an instrumental variable approach to address confounding from unmeasured covariates

Observational studies represent an alternative to randomized trials to study the comparative effectiveness of different treatments. Unfortunately, a lack of randomization often introduces selection bias and confounding which may obscure the estimation of the true treatment effect. Analytical methods, such as regression, propensity scores, and matching, mitigate measured confounding. However, such methods ignore unmeasured confounders. The instrumental variable method was designed to account for unmeasured confounders; briefly, it seeks to find a randomized experiment embedded within an observational study.

Although age and chronic comorbidities contribute to the risk of death after aortic surgery, the strongest predictor of mortality in patients with acute type A aortic dissection is the gravity of the acute illness, such as circulatory collapse or end-organ malperfusion.^{1, 2} Such variables are not captured within administrative databases, and thus would render an analysis that employs regression, propensity scores, or matching susceptible to unmeasured confounding absent randomization. For this reason, we employed an instrumental variable method as our primary mode of analysis.

An instrumental variable method exploits a variable that influences which treatment subjects receive, and only affects outcome through its influence on treatment.³ For this project, we explored the possibility of using institution practice patterns as a preference-based instrumental variable. We believed a preference-based instrumental variable would be particularly well-suited for this observational study because: 1) the specific treatment (e.g. transferred vs. stayed, high-volume vs. low-volume, rerouted to high-volume vs. not rerouted to high-volume) a patient receives is often determined by the preferences and capabilities of the institution; and 2) for severe illnesses of acute-onset, patients are typically brought by ambulance or present to the nearest hospital that can provide appropriate care for the initial evaluation. For example, a patient with chest pain – the most common chief complaint in patients with acute type A aortic dissection – is much more likely to have an acute coronary syndrome than an aortic dissection, and these patients are often initially brought to hospitals capable of performing a cardiac catheterization. Therefore, the presenting hospital is primarily dictated by a combination of proximity and hospital resources. Because aortic dissection is rare and advanced imaging (i.e. transesophageal echocardiography or computed tomography) is needed to confirm the diagnosis, the initial hospital where patients present is independent of the need for expert thoracic aortic surgery. Therefore, with respect to thoracic aortic

surgery experience, the population is initially pseudo-randomized to present at high-volume, low-volume, or no-volume (obligate transfer) hospitals. Similarly, with respect to patient destination for treatment, a patient that was transferred to a high-volume center might have remained at a low-volume hospital without transfer had they been in a different location when symptoms began. Examination of diseasespecific, hospital-level transfer patterns observed over the 15-year study confirmed institution preferences to be a credibly strong instrumental variable; the majority of hospitals always transferred or never transferred patients with aortic dissection, and among hospitals that always transferred, the destination typically did not change (i.e. most hospitals always transferred to a high-volume institution or always transferred to a low-volume institution, few transferred to a mix of hospitals) (Supplemental Figure 1). The unselected assignment to different levels of aortic surgery experience coupled with the routinized treatment patterns is analogous to an unbalanced, randomized 2x2 factorial design with random assignment to interfacility transfer or no transfer, and random assignment to treatment at a high-volume or low-volume hospital, irrespective of illness severity (Figure 1).

To further understand how Emergency Medical Services (EMS) selects the hospital where each patient is initially brought from the field (and thus assess the validity of the instrumental variable), we drew from two sources of information: (i) the National Emergency Medical Services Information System, or NEMSIS data set and (ii) the county-level EMS transport protocols from the state of California. We reviewed these two sources of information to deepen our understanding of the factors considered in determining the presenting hospital for the patients in our study. The NEMSIS data set is the national database that is used to store EMS data from U.S. States and Territories. The 2015 data set contains more than 30 million EMS activations submitted by 10,137 EMS agencies and representing 49 states and territories during the 2015 calendar year. The NEMSIS data set is a sample, not a census, of all EMS activations nationwide. Simple summary statistics are informative for our consideration. Element E_20 in the NEMSIS asks EMS providers to report "Reason for Choosing Destination." In 2015, 33% of patient care reports indicated that the receiving facility was chosen because it was the closest facility. 52% of patient care reports indicated that either the patient, a family member, or a physician chose the receiving facility. It is likely that a significant proportion of the patient/family decisions were heavily influenced by location. More importantly, as a diagnosis of acute type A aortic dissection is unlikely in the field, the patient and family preference is less likely to be sensitive to, and modified by, the severity of the condition or knowledge of the need for thoracic aortic surgery. Another important insight from the NEMSIS is that choice determination varies by region. In rural and suburban regions, "closest facility" was the primary reason for choosing the destination, but in urban regions, patients, physicians, and family members had more influence in routing. Thus the instrument appears to vary in strength based on population density.

Additionally, we examined EMS protocols from 47 of the 62 California counties. In summary, all of the protocols stated that the top priority in transport decisions should be patient preference. After patient preference, location was the second most commonly referenced reason for choosing a hospital. Nearly all of the protocols (89%) explicitly state that paramedics should consider location while making a transport decision, and the remaining (11%) suggest routing to "the most appropriate" facility. Protocols also often indicate that patients must be transported only to facilities within the geographic constraints of that EMS agency. In San Mateo County, for example, "all patients…will be transported to the hospital of their preference; as long as it is a San Mateo County receiving hospital."

We believe our instrument minimized sorting of patients based on disease complexity and illness severity. We restricted our analysis only to patients who presented to "always" or "never" hospitals with respect to the decision to transfer and the transfer destination (high-volume vs. low-volume). Although measured baseline characteristics were balanced prior to matching, we further improved balance (among measured, and theoretically unmeasured variables) by matching patients on their predicted probability of the instrument. This is important because the patient groups were well balanced across chronic comorbidities and demographics prior to matching on the instrument (which are likely not the most important confounders of outcome and treatment selection in this disease), and highlights a potential

weakness of limiting the analysis to a propensity score matching design. See below for further discussion of the consequences of using a propensity score design in this setting.

Implications of preferential routing to hospitals capable of percutaneous coronary intervention

Over 90% of patients with aortic dissection present with chest pain as the chief complaint. However, an acute coronary syndrome is 100 times more likely than an aortic dissection. Whenever possible, EMS will often bring patients with suspected acute coronary syndromes to hospitals that have cardiac catheterization laboratories. We used the American Hospital Association data set to evaluate whether initially presenting to a percutaneous coronary intervention (PCI) capable hospital may confound our primary endpoint. Effectively every hospital that performed aortic surgery was also capable of performing PCI. However, many hospitals that did not perform aortic surgery (no-volume hospitals) lacked cardiac catheterization laboratories. We examined only patients who presented to no-volume hospitals who were subsequently transferred to high-volume hospitals for surgery (PCI-capable, n=902) patients; not PCI-capable, n=1167 patients). We propensity score-matched patients on whether the initial hospital was capable of PCI (Supplemental Table 7) to yield similar groups. In fact, initially presenting to a PCI-capable hospital did not affect operative mortality after surgery for type A aortic dissection (PCIcapable 22.2%, not PCI-capable, 20.4%, risk difference -1.8%, 95% CI -5.4% to 1.9%). We believe this data helps support the validity of our instrument in that sorting of patients to PCI-capable hospitals does not confound the outcome.

Validation of codes used to define type A aortic dissection

We validated our definition of acute type A aortic dissection within the population treated at Stanford University. Between 2009 and 2015, 232 patients with acute type A aortic dissection were treated at Stanford. Among these patients, 224 (97%) patients had a International Classification of Diseases (ICD9-CM) diagnosis code of 441.0 and Current Procedural Terminology (CPT) code(s) for surgery of the aortic root and/or ascending aorta within the same hospitalization. This high sensitivity for type A aortic dissection is accompanied by a theoretically high specificity: there should be very few false positives among all patients treated at Stanford University – for any reason – over the same time period.

Power calculation

We performed an *a priori* power calculation to ensure the study was adequately powered. We designed the study to achieve at least 90% power to detect a small risk difference in operative mortality (1300 matched sets, 92% power; 2500 matched sets, 99.8% power). In the power calculation, we used an alpha level of 0.05 and assumed a Cohen's h of 0.15. We used a Bonferroni correction to control for the familywise error on three hypotheses.

Details of matching algorithm

Logistic regression was used to predict each patient's probability of the dichotomous instrument given baseline covariates (Supplemental Figure 2). The model included the following variables as fixed effects: age at surgery, age <65 years, year of surgery, male sex, race, geographic region, prior myocardial infarction, prior Alzheimer's dementia, prior atrial fibrillation, prior cataracts, prior chronic kidney disease, prior chronic obstructive pulmonary disease, prior congestive heart failure, prior diabetes mellitus, prior glaucoma, prior hip fracture, prior ischemic heart disease, prior depression, prior osteoporosis, prior rheumatoid or osteoarthritis, prior stroke, prior cancer (breast, colorectal, prostate, lung, or endometrial), prior anemia, prior asthma, prior hyperlipidemia, prior hyperparathyroidism, prior hypothyroidism, prior thoracoabdominal aortic surgery, prior open thoracic aortic surgery, prior endovascular aortic repair, prior thoracic endovascular aortic repair, prior abdominal aortic surgery, prior coronary bypass surgery, prior aortic valve surgery, prior mitral valve surgery, prior tricuspid valve surgery, prior mechanical assist device implantation, and other prior cardiac surgeries.

For each comparison – transferred vs. stayed, high-volume vs. low-volume, rerouted to highvolume vs. not rerouted to high-volume – restricted 1:*k* (transfer and rerouting) or *k*:1 (volume) matching

was performed with the optmatch package;⁴ patients <65 years old were not allowed to be matched to those ≥65 years old. The maximum controls (in 1:*k*) and treated (*k*:1) individuals allowed in each subclass were determined by doubling the ratio of the total number of control:treated patients (for 1:*k*) or treated:control patients (for *k*:1). ⁵ The omitted fraction was set to equal the proportion of patients outside the region of common support. To estimate the effect of transfer, transferred patients were only matched to those who remained at a hospital of equivalent volume class. To estimate the effect of hospital volume, transferred patients were only matched to patients that originated at a hospital of the same volume category; and patients that remained at high-volume centers were only matched to those that remained at low-volume centers. To estimate the effect of rerouting, patients who were transferred to a high-volume hospital (rerouted) were matched to patients that were transferred to a low-volume hospital or remained at a low-volume hospital (not rerouted). Patients that were excluded due to the instrumental variable were similar to those who were included in the analysis with respect to baseline comorbidities and demographics (Supplemental Table 8).

We estimated the average treatment effect on the treated for the effects of transfer and hospital volume. Because we were specifically interested in identifying the number of additional patients who would have benefitted from rerouting to high-volume hospitals if such a policy were implemented, we chose to estimate the average treatment effect on the controls for the effect of rerouting.

Restricted mean survival time

Time-to-event analyses typically employ Cox proportional hazards models. When comparing two groups in this fashion, an assumption is made that the ratio of the two hazard functions for the groups remains constant over time (e.g. proportional hazards). The Cox model is then used to estimate the unknown constant hazard ratio parameter. When there is concern or evidence of non-proportional hazards (e.g. the hazard ratio changes over time), the estimated hazard ratio may not be a meaningful measure of the between-group survival difference.⁶ Yet, the CONSORT guidelines and Cochrane handbook instruct the reporting of hazard ratios to quantify between-group differences in survival.⁷ For this reason, we report hazard ratios in our manuscript, but also chose to report "model-free" statistics based on the restricted mean survival time $(RMST)^{8,9}$ A model-free statistic, the RMST does not rely on assumptions about the underlying hazard functions for each group. The RMST is the population average of the amount of event-free survival time experienced during a set follow-up period – in our study, we chose 15 years because there were still enough patients remaining in the risk set at that time, and it was not beyond the last observed event. The RMST equates to the area under the survival curve. Once the area under the survival curve is calculated, the difference in the RMST between groups and the RMST ratio provide useful estimates of the treatment effect on overall survival. The 15-year RMST difference is interpreted as the average number of additional days gained by the treatment arm over 15 years. The RMST ratio quantifies the percent of life gained by the treatment arm compared to the control arm. Similarly, the area above the survival curve corresponds to the number of days of life lost until the designated time point (e.g. 15 years), and is known as the restricted mean time lost.

Sensitivity analyses

In addition to the sensitivity analyses presented in the main manuscript that quantify the number of unobserved patients required to change our conclusions, we also conducted a sensitivity analysis to explore the extent to which our results were robust to unmeasured bias. Through matching, we assume that patients who appear comparable are in fact comparable. However, consistent with our aforementioned discussion of unmeasured confounding, patients who appear comparable are not always comparable. The gamma sensitivity parameter describes the fold increase in likelihood of receiving treatment between matched individuals.^{10, 11} In the main manuscript, we present the minimum value of gamma necessary to change our conclusion that the null hypothesis should be rejected.

Effects of relying on a propensity score-based design

A major design decision in the primary analysis was to use an instrumental variable approach. This imposed two restrictions: (i) we restricted our analysis to hospitals that observed >1 aortic dissection during the study period in order to establish a treatment pattern and (ii) we further restricted to hospitals that had "always" patterns – either always transferring/staying, and if the hospital transferred it always transferred to the same type of hospital (i.e., either always high-volume or always low-volume). That decision was made *a priori* because it was believed it would increase internal validity by reducing biased estimation due to confounding arising from unobserved covariates. This decision came with a tradeoff, forcing the exclusion of approximately a third of hospitals and 40% of patients. It is important to consider whether we have restricted our analysis such that it is not informative of the wider population (i.e., generalizability of results). Inspection of the SMD columns in Supplemental Table 11 shows that the analyzed populations do not have considerably different baseline covariate distributions from the overall population – with the notable exception of region of country. This observation of baseline patient-level covariates does not completely remove the question of generalizability – perhaps the composition of hospitals significantly changes. To address this concern we performed a follow-on study using propensity score matching, without the exclusion criteria imposed by the instrumental variable analysis. The analysis was nearly identical to our main analysis but in this follow-on study we assumed that there were no unobserved covariates causing selection into the type of care the patient received. For example, if patient A and B both presented at low-level hospital 1 with nearly identical records in our data set, yet patient A was transferred to high-level hospital 2 and patient B was kept at low-level hospital 1, then we assume that these differentiated decisions were not informative of their outcomes except through the care received (i.e., ignorable treatment assignment). Under this assumption we can follow a similar design as in the main analysis, which approximately doubles the number of observations analyzed. In this propensity score analysis the effect of rerouting patients to high-volume hospitals will reduce operative mortality from 30.1% to 21.4% (risk difference -8.7%, 95% CI -10.6% to -6.7%, P<0.001; Supplemental Table 6). This is quite a bit larger than what was found in our main analysis (-7.2% vs. -8.7%). Before interpreting that result, we move to the sub-analysis of the isolated effect of interfacility transfer. It is unlikely that the effect of transferring a patient is beneficial. In fact, it is almost certain there should be a slight increase in probability of mortality due to transfer. Said another way: often the fundamental question in regionalization is if the risk of transfer outweighs the benefit of the care received at the subsequent facility. In this propensity score analysis, the effect of transfer is estimated to have a risk difference of - 1.7% (95% CI -3.2% to -0.001%, P=0.04). Taken literally, this would mean that merely transferring a patient is beneficial. If this were the case, then before initiating treatment, a hospital could improve outcomes by stabilizing a patient and moving the patient around and returning the patient to the same hospital for treatment. That is counterintuitive. Instead, the more plausible interpretation of this result is that there is bias in the estimation, likely due to selection into transferring or staying and/or into where a hospital sends a patient high-volume or low-volume. Comparing the propensity score analysis to the instrumental variable analysis, we see that the instrumental variable analysis had a point estimate of - 0.7% (95% CI, -2.7% to 1.35%, P=0.35) which is indistinguishable from a null effect. The inclusion of more hospitals and patients into our analysis seems, at least in part, to be biasing our estimate toward showing a benefit to regionalization.

Statistical software and packages used to perform the analysis

Data was prepared for analysis in SAS (SAS Institute Inc., Cary, NC) and statistical analyses were performed in R version 3.3.1 (R Foundation, Vienna, Austria). The survey package¹² was used to compare baseline characteristics and operative mortality between groups. The Cochrane Armitage test within the coin package¹³ was used to evaluate trends over time. Distances were calculated from centroid to centroid of each hospital's zip code using the Haversine method implemented by the fields package.¹⁴ Survival analyses were performed with the survival package.¹⁵ RMST differences, RMST ratios, and RMTL ratios were calculated and compared between groups with the survRM2 package.¹⁶ Gamma parameters were calculated with the sensitivitymv package.¹⁷ Mapping and trends over time were plotted with Tableau version 10.0 (Tableau Software, Seattle, WA).

Supplemental Table 1. Definitions of Index Cardiovascular Surgery Procedures

Supplemental Table 1. Definitions of Index Cardiovascular Surgery Procedures

(Continued)

CPT, current procedural terminology

Supplemental Table 2. Definitions of Prior Cardiovascular Surgery Procedures

Supplemental Table 2. Definitions of Prior Cardiovascular Surgery Procedures

(Continued)

Supplemental Table 2. Definitions of Prior Cardiovascular Surgery Procedures

(Continued)

CPT, current procedural terminology; ECMO, extracorporeal membrane oxygenation; ICD9-CM, International Classification of Diseases, 9th Revision, Clinical Modification

Supplemental Table 3. Baseline and operative characteristics of study population before and after propensity score matching – transfer comparison*

Supplemental Table 3. Baseline and operative characteristics of study population before and after propensity score matching – transfer comparison (Continued)*

* Plus-minus values are means +/- standard deviation. Variables excluded from table though well-balanced (SMD <0.1): cataracts, glaucoma, osteoporosis, asthma, hyperparathyroidism, and depression. CABG, coronary artery bypass grafting; COPD, chronic obstructive pulmonary disease; ECMO, extracorporeal membrane oxygenation; EVAR, endovascular aortic repair; SMD, standardized mean difference; TEVAR, thoracic endovascular aortic repair; VAD, ventricular assist device

Supplemental Table 4. Baseline and operative characteristics of study population before and after propensity score matching – volume comparison*

Supplemental Table 4. Baseline and operative characteristics of study population before and after propensity score matching – volume comparison (Continued)*

* Plus-minus values are means +/- standard deviation. Variables excluded from table though well-balanced (SMD <0.1): cataracts, glaucoma, osteoporosis, asthma, hyperparathyroidism, and depression. CABG, coronary artery bypass grafting; COPD, chronic obstructive pulmonary disease; ECMO, extracorporeal membrane oxygenation; EVAR, endovascular aortic repair; SMD, standardized mean difference; TEVAR, thoracic endovascular aortic repair; VAD, ventricular assist device

Supplemental Table 5. Baseline and operative characteristics of study population before and after propensity score matching – regionalization comparison*

Supplemental Table 5. Baseline and operative characteristics of study population before and after propensity score matching – regionalization comparison (Continued)*

* Plus-minus values are means +/- standard deviation. Variables excluded from table though well-balanced (SMD <0.1): cataracts, glaucoma, osteoporosis, asthma, hyperparathyroidism, and depression. CABG, coronary artery bypass grafting; COPD, chronic obstructive pulmonary disease; ECMO, extracorporeal membrane oxygenation; EVAR, endovascular aortic repair; SMD, standardized mean difference; TEVAR, thoracic endovascular aortic repair; VAD, ventricular assist device

 \dagger n=4,313 patients that presented at high-level hospitals so are not part of this analysis

Supplemental Table 6. Between-group differences in operative mortality and long-term mortality for comparisons of transfer, volume and regionalization in propensity scorematched cohort

***** Gamma = 1.32 for this comparison.

† Gamma = 1.41 for this comparison.

The gamma parameter estimates the amount of unmeasured bias necessary to render the finding null. For interpretation, a Gamma = 1.32, implies unobserved covariates would need to increase the odds of low-volume hospital 2.00-fold, and increase the odds of operative mortality 1.62-fold, to render the presented finding null. A Gamma = 1.41 implies unobserved covariates would need to increase the odds of receiving care at a non-regionalized hospital 2.00-fold, and increase the odds of operative mortality 1.82-fold, to render the presented finding null. CI, confidence interval

Supplemental Table 7. Baseline and operative characteristics of patients who presented to PCI-capable vs. not PCI-capable hospitals*

Supplemental Table 7. Baseline and operative characteristics of patients who presented to PCI-capable vs. not PCI-capable hospitals (Continued)*

* Plus-minus values are means +/- standard deviation. Variables excluded from table though well-balanced (SMD <0.1): cataracts, glaucoma, osteoporosis, asthma, hyperparathyroidism, and depression. CABG, coronary artery bypass grafting; COPD, chronic obstructive pulmonary disease; ECMO, extracorporeal membrane oxygenation; EVAR, endovascular aortic repair; SMD, standardized mean difference; TEVAR, thoracic endovascular aortic repair; VAD, ventricular assist device

† Numbers are not necessarily integers due to matching with variable controls and weighting for the average treatment effect

Supplemental Table 8. Baseline and operative characteristics of patients who were included vs. excluded by the instrumental variable design for each comparison (Continued)*

* Plus-minus values are means +/- standard deviation. Variables excluded from table though well-balanced (SMD <0.1): cataracts, glaucoma, osteoporosis, asthma, hyperparathyroidism, and depression. CABG, coronary artery bypass grafting; COPD, chronic obstructive pulmonary

disease; ECMO, extracorporeal membrane oxygenation; EVAR, endovascular aortic repair; SMD, standardized mean difference; TEVAR, thoracic endovascular aortic repair; VAD, ventricular assist device

Supplemental Table 9. Baseline and operative characteristics of transferred vs. not transferred population*

Supplemental Table 9. Baseline and operative characteristics of transferred vs. not transferred population (Continued)*

* Plus-minus values are means +/- standard deviation. Variables excluded from table though well-balanced (SMD <0.1): cataracts, glaucoma, osteoporosis, asthma, hyperparathyroidism, and depression. CABG, coronary artery bypass grafting; COPD, chronic obstructive pulmonary disease; ECMO, extracorporeal membrane oxygenation; EVAR, endovascular aortic repair; SMD, standardized mean difference; TEVAR, thoracic endovascular aortic repair; VAD, ventricular assist device

Supplemental Table 10. Baseline and operative characteristics of patients receiving surgery at low-volume vs. high-volume hospitals*

Supplemental Table 10. Baseline and operative characteristics of patients receiving surgery at low-volume vs. high-volume hospitals (Continued)*

* Plus-minus values are means +/- standard deviation. Variables excluded from table though well-balanced (SMD <0.1): cataracts, glaucoma, osteoporosis, asthma, hyperparathyroidism, and depression. CABG, coronary artery bypass grafting; COPD, chronic obstructive pulmonary disease; ECMO, extracorporeal membrane oxygenation; EVAR, endovascular aortic repair; SMD, standardized mean difference; TEVAR, thoracic endovascular aortic repair; VAD, ventricular assist device

Supplemental Figure 1. Distributions of transfer rates at the hospital- and patient-level

C Distribution of rates of hospital transfer to high-volume centers among hospitals that always transferred patients with acute type A aortic dissection

D Distribution of patients presenting to hospitals with varying transfer rates to highvolume centers among hospitals that always transferred patients with acute type A aortic dissection

Supplemental Figure 1. Distributions of transfer rates at the (*A, C*) hospital- and (*B, D*) patient-level. The hospital transfer patterns demonstrate that most hospitals transfer all of their patients or none of their patients (*A-B*). Among hospitals that always transfer, patients are almost always transferred to a highvolume hospital, or always to a low-volume hospital (*C-D*).

Supplemental Figure 2. Distribution of predicted probabilities of the instrumental variable for each comparison

A Transferred vs. stayed

C High-volume vs. low-volume

E Rerouted to high-volume vs. not rerouted to high-volume

Supplemental Figure 2. Distribution of predicted probabilities of the instrumental variable for each comparison. Patients demonstrated excellent regions of common support for predicted probabilities of the instrumental variable in comparisons of (*A-B*) transferring versus staying at the presenting hospital, (*C-D*) having surgery at high-volume versus low-volume hospitals, and (*E-F*) being rerouted to a high-volume versus not being rerouted to a high-volume hospital.

Study Population: Patients with Acute Type A Aortic Dissection who Presented to Hospitals that "Always" Transfer to High-Volume Centers, "Always" Transfer to Low-Volume Centers, or "Never" Transfer (n=8,956)

Supplemental Figure 3. Patient selection flow diagram. The figure demonstrates the reasons for excluding patients in order to arrive at a population of patients with acute type A aortic dissection who presented to hospitals that always transfer to high-volume centers, always transfer to low-volume centers, or never transfer patients with that diagnosis.

Supplemental Figure 4. Proportion of patients presenting to each category of hospital volume per year. The proportion of patients with an acute type A aortic dissection presenting to high-volume, low-volume, or no-volume hospitals did not change during the study period.

Supplemental Figure 5. Proportion of patients undergoing interfacility transfer per year

Supplemental Figure 5. Proportion of patients undergoing interfacility transfer per year in the United States. The proportion of patients with an acute type A aortic dissection who were transferred to another hospital or stayed at their initial hospital for surgery did not change during the study period.

Supplemental Figure 6. Proportion of patients undergoing surgery at high- vs. low-volume hospitals per year

Supplemental Figure 6. Proportion of patients undergoing surgery at high- vs. low-volume hospitals per year. The proportion of patients with an acute type A aortic dissection who have surgery at high-volume versus low-volume hospitals in the United States did not change during the study period.

Supplemental Figure 7. Operative mortality for repair of type A dissection at high and low-volume hospitals per year

Supplemental Figure 7. Operative mortality for repair of type A dissection at high and low-volume hospitals per year. The operative mortality for repair of acute type A aortic dissection among Medicare beneficiaries steadily declined during the study period at both high- and low-volume hospitals, but was consistently lower at high-volume hospitals in the United States.

Supplemental Figure 8. Age dependent hazard of operative mortality after regionalization for repair of acute type A aortic dissection

Supplemental Figure 8. Age dependent hazard of operative mortality after regionalization for repair of acute type A aortic dissection. The hazard ratio of death for patients rerouted to a high-volume hospital is plotted against age as a continuous variable. The dashed lines represent 95% confidence intervals obtained from bootstrap resampling. The horizontal black line at 1 denotes no difference between rerouting patients to have surgery at high-volume hospitals and not rerouting to high-volume hospitals. The relatively straight line between ages 45 to 80 suggests that the beneficial effect of regionalization may be independent of age. CI, confidence interval; HR, hazard ratio

SUPPLEMENTAL REFERENCES

- 1. Conzelmann LO, Weigang E, Mehlhorn U, Abugameh A, Hoffmann I, Blettner M, Etz CD, Czerny M, Vahl CF Investigators G. Mortality in patients with acute aortic dissection type A: analysis of pre- and intraoperative risk factors from the German Registry for Acute Aortic Dissection Type A (GERAADA). *Eur J Cardiothorac Surg*. 2016;49:e44-52.
- 2. Czerny M, Schoenhoff F, Etz C, Englberger L, Khaladj N, Zierer A, Weigang E, Hoffmann I, Blettner M Carrel TP. The Impact of Pre-Operative Malperfusion on Outcome in Acute Type A Aortic Dissection: Results From the GERAADA Registry. *J Am Coll Cardiol*. 2015;65:2628- 35.
- 3. Baiocchi M, Cheng J Small DS. Instrumental variable methods for causal inference. *Stat Med*. 2014;33:2297-340.
- 4. Hansen BB Klopfer SO. Optimal full matching and related designs via network flows. *J Comp Graph Stat*. 2006;15:609-627.
- 5. Hansen BB. Full matching in an observational study of coaching for the SAT. *JASA*. 2004;99:609-618.
- 6. Uno H, Claggett B, Tian L, Inoue E, Gallo P, Miyata T, Schrag D, Takeuchi M, Uyama Y, Zhao L, Skali H, Solomon S, Jacobus S, Hughes M, Packer M Wei LJ. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. *J Clin Oncol*. 2014;32:2380-5.
- 7. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, Elbourne D, Egger M Altman DG. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.
- 8. Zhao L, Tian L, Uno H, Solomon SD, Pfeffer MA, Schindler JS Wei LJ. Utilizing the integrated difference of two survival functions to quantify the treatment contrast for designing, monitoring, and analyzing a comparative clinical study. *Clin Trials*. 2012;9:570-7.

39

- 9. Royston P Parmar MK. The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. *Stat Med*. 2011;30:2409-21.
- 10. Rosenbaum PR. Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies. *Biometrics*. 2007;63:456-464.
- 11. Rosenbaum PR. *Observational studies.* 2nd ed. New York: Springer; 2002.
- 12. Lumley T. Analysis of complex survey samples. *Journal of Statistical Software*. 2004;9:1-19.
- 13. Horthorn T, Hornik K, van de Wiel MA Zeileis A. Implementing a class of permutation tests: the coin package. *J Stat Softw*. 2008;28:1-23.
- 14. Nychka D, Furrer R, Paige J Sain S. Fields: tools for spatial data. *R package version 84-1*. 2015.
- 15. Therneau TM Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. New York: Springer; 2000.
- 16. Uno H, Tian L, Cronin A Battioui C. survRM2: Comparing Restricted mean Survival Time. R package version 1.0-1. 2015.
- 17. Rosenbaum PR. Sensitivitymv: Sensitivity analysis in observational studies. R package version 1.3. 2015.