

Supplemental Material for: Consequences of Preoperative Cardiac Stress Testing—A Cohort Study

Matthew A Pappas, MD, MPH, Andrew D Auerbach, MD, MPH, Michael W Kattan, PhD, Eugene Blackstone, MD, Michael B Rothberg, MD, MPH, and Daniel I Sessler, MD

1. Additional detail on model of preoperative stress testing, including multiple imputation	3
2. Additional detail on imputation of CPT codes using vectorization	6
3. Log-log plot demonstrating failure of proportional hazards assumption (Schoenfeld residuals)	7
4. Heterogeneity of treatment effect by considered surgery	8
5. Long-term mortality plot	10

1. Additional detail on model of preoperative stress testing, including multiple imputation

Missingness is inherent to EHR data. To minimize biases from missing data, we used multiple imputation with chained equations (MICE) when estimating the probability of preoperative stress testing. Note that our outcome (completed preoperative stress testing before censoring) differs from previous work on physician variation, and therefore has fewer outcomes. Our modeling approach was very similar.

We “imputed” missing CPT codes (see below) separately from and prior to MICE. After that step, over 99% of visits had a CPT code designating the surgery that was under consideration. We used that CPT code for variables that required categorization, including the Gupta-MICA estimated risk.

Of the 159,795 visits in this analysis, 84.7% were missing data in one or more observations. These were most commonly variables that were inconsistently documented in the clinic note. Variables that are documented by exception (such as CHF) were considered complete. The full list of variables in our treatment model is as follows, with the proportion of observations missing each variable listed in the original dataset:

Variable:	Percent of observations missing:
Estimated METs	47.5%
RCRI	42.3%
Functional class	41.2%
ADI (national)	21.1%
BMI	1.5%
Diastolic blood pressure	0.2%
Aortic surgery	0
Peripheral vascular surgery	0
Urologic surgery	0
Gupta-MICA greater than 1%	0
Dx of Congestive heart failure	0
Dx of ischemic heart disease	0
Estimated risk of CAD	0
Current tobacco use	0
Former tobacco use	0
Date	0

We imputed RCRI and functional class as ordinal variables, estimated METs and Area Deprivation Index (ADI) as truncated linear regressions, and BMI and DBP as linear regressions. We included the outcome variable (preoperative stress testing) in all imputation equations.

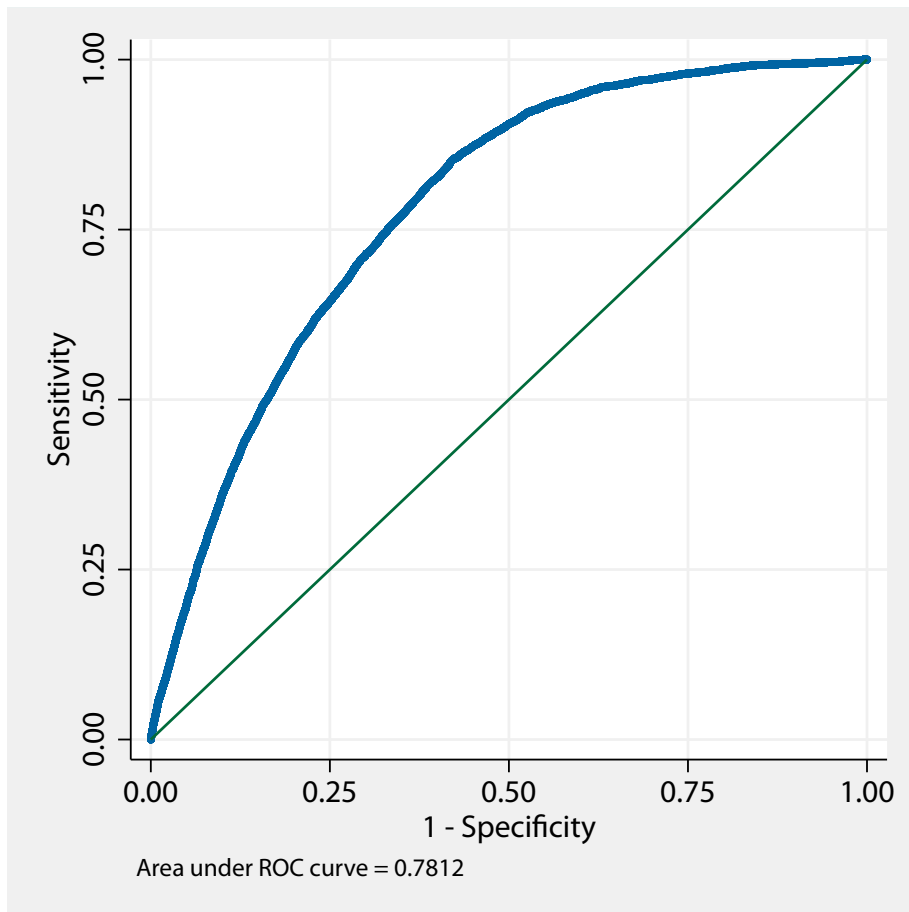
We converted ADI and date to restricted cubic splines (before imputation) to allow nonlinear effects. We also interacted “ischemic heart disease” with “estimated risk of CAD” (calculated from the model published by Genders and colleagues) to allow the estimated pretest probability of CAD to have different weight when a patient did or did not have an existing diagnosis of CAD. Missing ADI was almost entirely a result of uncertain addresses, including patients without a fixed address, patients with more than one available confirmed address, patients whose listed address could not be matched to a census block group (including workplaces and P.O. Boxes), and patients who had a confirmed address in our study period that differed from the last address confirmed before the clinic visit (such that the address at the time of the preoperative clinic visit could have been either the last address confirmed before the clinic visit or the address confirmed during

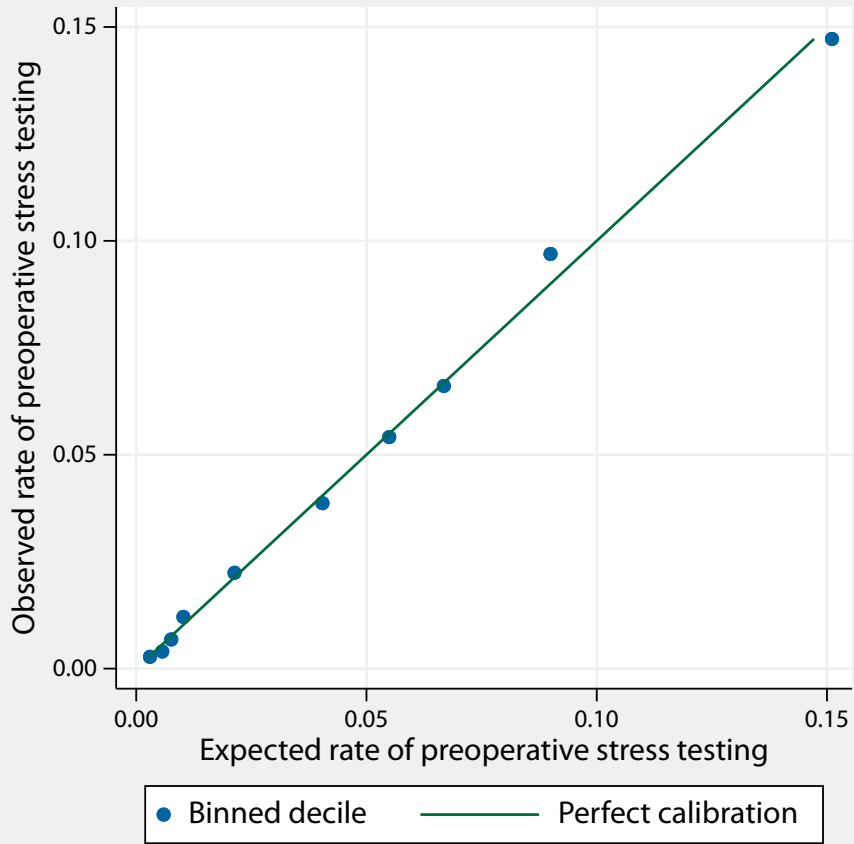
our study period, depending on when the patient moved). Most of those phenomena were more likely for patients living in neighborhoods whose ADI was higher (less advantaged neighborhoods). That is, these data are missing not at random, which leads to different numbers of missing values in each spline.

We used 100 imputations and verified that MICE did not introduce Monte Carlo error based on the White, Royston, and Wood guidelines. The fraction of missing information (FMI, which is the proportion of total sampling variance that is due to missing data) for imputed variables included in our prediction model were as follows:

Variable:	FMI:
METs	45.0%
RCRI	51.2%
Functional class	40.6%
ADI_spline_1	1.8%
ADI_spline_2	6.7%
ADI_spline_3	13.8%
ADI_spline_4	29.6%
BMI	5.7%
DBP	5.7%

The discrimination of this model (area under the receiver-operator curve, AUROC) was 0.78, and the Brier score was 0.04. Visual inspection of a calibration plot did not suggest areas of poor calibration. A receiver-operator curve and a calibration plot are shown below.





2. Additional detail on imputation of CPT codes using vectorization

Vectorization is a technique to represent a word in numerical space, in order to allow comparisons of similarity between words and syntactical relationships. The original algorithm, `word2vec`, has been extended to create a numeric representation of a document of arbitrary length. Using this algorithm (`doc2vec`), we compared the brief surgical descriptions from visits without a designated CPT code against the descriptions of visits with a designated CPT code. For example, consider the following simplified dataset:

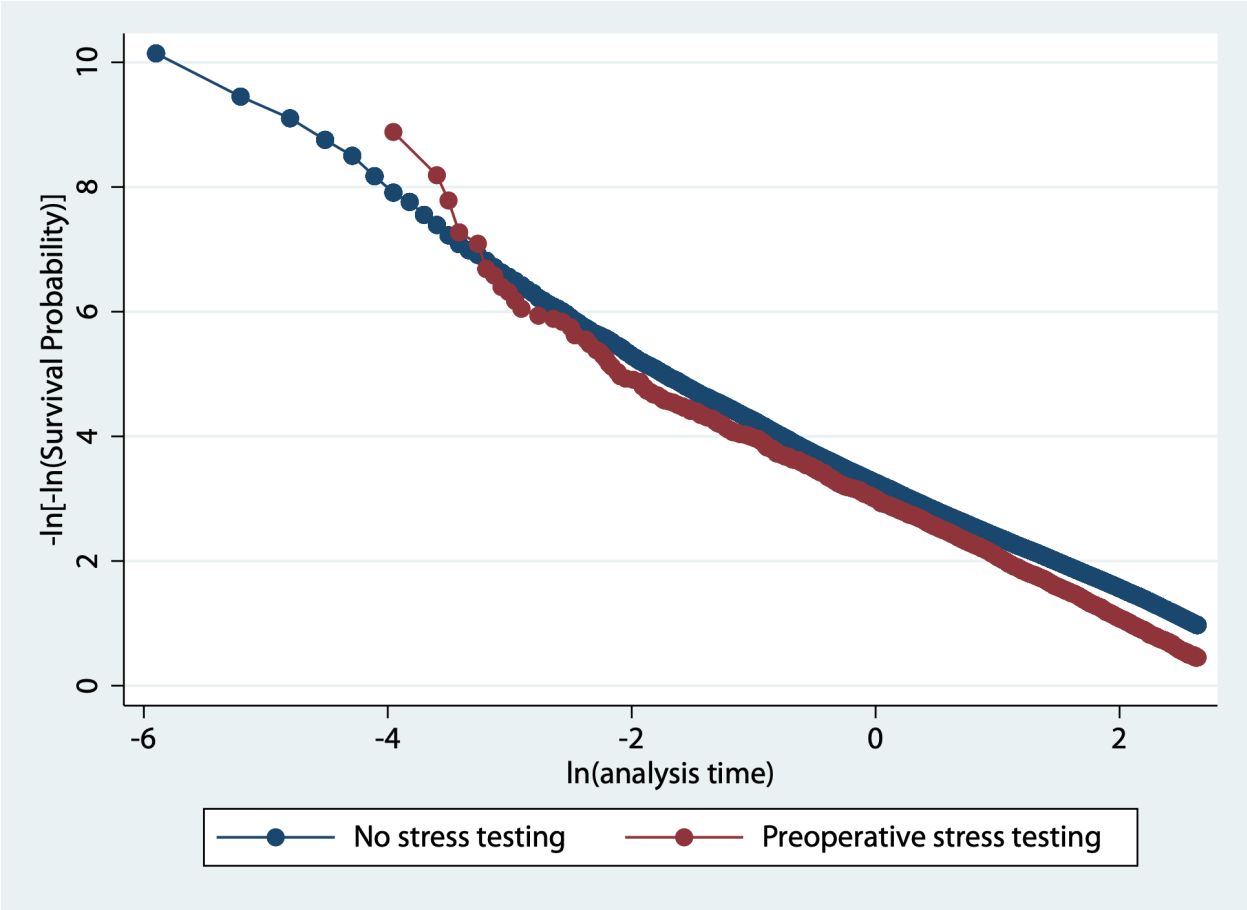
Visit	Description	CPT_code
1	LAPAROSCOPIC TOTAL COLEC	.
2	LAPAROSCOPY HYSTERECTOM	.
3	LAPAROSCOPIC TOTAL COLECTOMY W/ ILEOSTOMY	44210
4	LAPAROSCOPIC HYSTERECTOMY	58570

Here, the text descriptions from visit 1 (“LAPAROSCOPIC TOTAL COLEC”) would be compared against the text descriptions from visits 3 and 4. The similarity between the descriptions from visit 1 and visit 3 (0.999999) is greater than the similarity between visits 1 and 4 (0.167359). Similarly, visit 2’s description is more similar to visit 4’s than to visit 3’s.

After comparing the descriptions of considered surgeries with missing CPT codes against the description of each surgery with a supplied CPT code, we selected the best match as the CPT code from the missing visit. In the example above, the CPT code for visit 1 would be matched as 44210, while the best match for visit 2’s CPT code would be 58570.

Using this process, we compared the surgical descriptions from 27,750 visits with missing CPT codes against the descriptions from 130,607 visits with CPT codes. We “imputed” 27,691 visits using this process; a small number of visits failed when no reasonable match could be algorithmically identified.

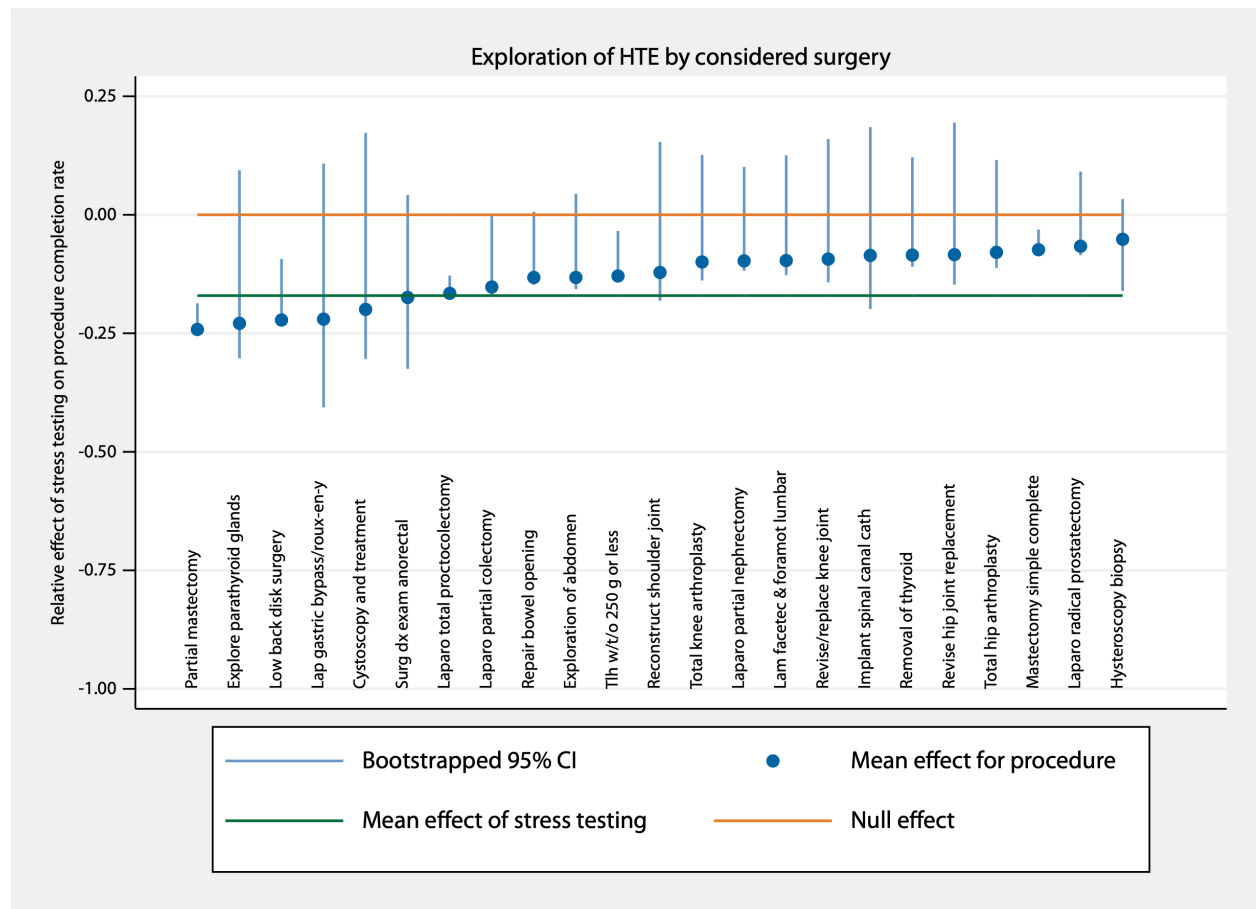
3. Log-log plot demonstrating failure of proportional hazards assumption (Schoenfeld residuals)



4. Heterogeneity of treatment effect by considered surgery

Some considered surgical procedures may be more or less likely to be cancelled following referral for stress testing than other surgical procedures. Below we plot the overall average treatment effect (referral for stress testing leads to a reduction in completion of the considered surgery) along with the point estimate and confidence interval for each CPT code that was considered at 1,000 or more visits. Some procedures (e.g., partial mastectomy, CPT 19301) are more likely than others to be cancelled after referral for stress testing. Other procedures (e.g., Laparoscopic radical prostatectomy) appear equally likely to be completed regardless of stress testing referral.

CPT codes and descriptions are copyrighted by the American Medical Association; the text descriptions are from the equivalent HCPCS code (available from CMS) rather than the CPT code. We extracted CPT codes from our EHR and used these indicators in our analysis. The HCPCS codes associated with each code (as downloaded from CMS) are displayed below.



CPT code (extracted from EHR)	HCPCS description (displayed in above graph for easier interpretation)
19301	Partial mastectomy
19303	Mastectomy simple complete
23472	Reconstruct shoulder joint
27130	Total hip arthroplasty
27134	Revise hip joint replacement
27447	Total knee arthroplasty
27487	Revise/replace knee joint
43644	Lap gastric bypass/roux-en-y
44204	Laparo partial colectomy
44210	Laparo total proctocolectomy
44620	Repair bowel opening
45990	Surg dx exam anorectal
49000	Exploration of abdomen
50543	Laparo partial nephrectomy
52234	Cystoscopy and treatment
55866	Laparo radical prostatectomy
58558	Hysteroscopy biopsy
58571	Tlh w/t/o 250 g or less
60240	Removal of thyroid
60500	Explore parathyroid glands
62350	Implant spinal canal cath
63030	Low back disk surgery
63047	Lam facetec & foramot lumbar

5. Long-term mortality plot

Our cohort includes a median of 7.1 years at risk of mortality. Here we present longer-term mortality by preoperative stress testing.

