BMJ Open

Variation in preoperative stress testing by patient, physician, and surgical type – a cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-048052
Article Type:	Original research
Date Submitted by the Author:	17-Dec-2020
Complete List of Authors:	Pappas, Matthew; Cleveland Clinic, Department of Hospital Medicine Blackstone, Eugene; Cleveland Clinic, Miller Family Heart and Vascular Institute Sessler, Daniel; Cleveland Clinic, Outcomes Research Auerbach, A; University of California, San Francisco Kattan, Michael; Cleveland Clinic Milinovich, Alex; Cleveland Clinic Rothberg, Michael B.; Cleveland Clinic, Cleveland, USA, Department of Internal Medicine
Keywords:	INTERNAL MEDICINE, Adult surgery < SURGERY, Adult anaesthesia < ANAESTHETICS, Adult cardiology < CARDIOLOGY, HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Variation in preoperative stress testing by patient, physician, and surgical type – a cohort study

Matthew A Pappas, MD, MPH,[1,2,3] Daniel I Sessler, MD,[3,4] Andrew D Auerbach, MD, MPH,[5] Michael W Kattan, PhD,[6] Alex Milinovich, BA,[6] Eugene H Blackstone, MD,[7] and Michael B Rothberg, MD, MPH[1]

1: Center for Value-based Care Research, Cleveland Clinic, Cleveland, OH
2: Department of Hospital Medicine, Cleveland Clinic, Cleveland, OH
3: Outcomes Research Consortium, Cleveland, OH
4: Department of Outcomes Research, Anesthesiology Institute, Cleveland Clinic, Cleveland, OH
5: Department of Hospital Medicine, University of California, San Francisco, San Francisco, CA
6: Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH
7: Miller Family Heart and Vascular Institute, Cleveland Clinic, Cleveland, OH

Correspondence to Matthew A Pappas: pappasm@ccf.org; 216.444.9565; 9500 Euclid Ave., Mail Stop G-10, Cleveland, OH 44195

Short title/running head: Variation in preoperative stress testing
Word count: 3,013
Abstract word count: 264
Number of references: 37
Number of tables: 3
Number of figures: 2
Number of appendices: 1

Financial support: This work was supported by the National Heart, Lung, and Blood Institute (NHLBI 1K08HL141598).
Competing Interests: None

Keywords (MEDLINE): Perioperative care
Preoperative care
Risk assessment
Elective surgical procedures
Cohort studies
Exercise test
Tomography, Emission-Computed, Single-Photon
Echocardiography, Stress

Abstract

Objectives:

To describe variation in and drivers of contemporary preoperative cardiac stress testing

Setting:

A dedicated preoperative risk assessment and optimization clinic at a large integrated medical center from 2008 through 2018

Participants:

A cohort of 118,552 adult patients seen by 104 physicians across 159,795 visits to a preoperative risk assessment and optimization clinic

Main Outcome:

Referral for preoperative stress testing, including nuclear, echocardiographic, or electrocardiographic-only stress testing, following the clinic visit, within 30 days, and before major surgery

Results:

A total of 8,303 visits (5.2%) resulted in referral for preoperative stress testing. Key patient factors associated with preoperative stress testing included predicted surgical risk, patient functional status, a previous diagnosis of ischemic heart disease, tobacco use, and body mass index. Patients living in either the most- or least-deprived census block groups were more likely to be tested. Patients were tested more frequently before aortic, peripheral vascular, or urologic interventions than before other surgical subcategories. Even after fully adjusting for patient and surgical factors, provider effects remained important:

marginal testing rates differed by 3-fold in relative terms and around 2.5% in absolute terms between the 5^{th} and 95^{th} percentile physicians. Rates of stress testing appear to be decreasing over time.

Conclusions:

In this large cohort of patients seen for preoperative risk assessment at a single health system, decisions to refer patients for preoperative stress testing are influenced by a number of factors other than estimated perioperative risk and functional status, the key considerations in current guidelines. Use of preoperative stress testing appears to have decreased over time and remains highly dependent on the provider.

Strengths and limitations of this study:

- We identified a large cohort of patients considering noncardiac surgery, with detailed clinical data from each visit.
- We tested predictor variables across multiple different constructs that could be related to preoperative stress testing.
- We accounted for clustering by physician and patient, testing different structures to ensure that variance was partitioned as accurately as possible.
- Although missing data is inherent to all studies using data from electronic medical records, we used multiple imputation by chained equations to mitigate the biases that missing data could introduce.

Glossary of Abbreviations:

ACC/AHA: American College of Cardiology/American Heart Association
ASA: American Society of Anesthesiologists
CAD: coronary artery disease
METs: Estimated metabolic equivalents
MICA: Myocardial infarction or cardiac arrest calculator
RCRI: Revised cardiac risk index

Introduction

The 2014 American College of Cardiology/American Heart Association (ACC/AHA) guidelines recommend preoperative stress testing for patients whose predicted risk of a major adverse cardiac event exceeds 1% and whose functional status is poor or unknown, when results from stress testing would change clinical management.[1]

However, clinicians use different risk prediction tools, which identify different patients as having elevated risk.[2,3] In addition, multiple methods of assessing functional status are in use, which again can lead to variations in patient populations chosen for stress testing.[2,4–7] Thus, the final decision to proceed with stress testing can become something closer to a provider-level judgment than a guideline-driven protocol.[8] Variation in use of stress testing can have substantial cost implications and potentially prompt subsequent tests and procedures with little clinical benefit.[9]

To understand contemporary use and drivers of preoperative cardiac stress testing, we sought to describe variation and predictors of preoperative stress testing using rich clinical data from a large integrated health system.

Methods

The Internal Medicine Preoperative Assessment, Consultation and Treatment (IMPACT) Center assesses patients prior to noncardiac surgery at the Cleveland Clinic. In the years

from 2008 through 2018, we captured 118,552 patients seen in this clinic by 104 physicians across 159,795 visits. Among this cohort, we identified scheduled and completed preoperative cardiac stress tests, here defined as those within the 30 days following the clinic visit and before noncardiac surgery. This study was approved by the Cleveland Clinic Institutional Review Board. All analyses were performed in Stata (version 14; College Station, TX).

Predictor variables

We theorized that six underlying constructs would be related to stress test ordering: predicted perioperative risk, functional status, social and financial support, medical comorbidities, physician tendencies and experience, and time. We created a random effects logit model for each construct to refine variables included in our final model. Within each submodel, we pruned variables according to Bayesian information criteria (BIC). For continuous variables, we assessed for nonlinear or categorical relationships using visual examination of binned scatter plots. To avoid overfitting, we limited our number of candidate predictors to fewer than 1 predictor variable per 15 preoperative stress tests, including tested interactions, nonlinear effects, and discarded predictors. We estimated that we had approximately 539 degrees of freedom for analysis, with fewer for cluster-level variables depending on model structure (described below).

For measures of perioperative risk, we tested Revised Cardiac Risk Index (RCRI), Myocardial Infarction or Cardiac Arrest (MICA), and MICA's categorization of surgeries using a previously-published crosswalk.[2,10,11] (Although different procedures likely have different intrinsic cardiac risk, we used the MICA categorization of surgeries to avoid overfitting.[11,12]) Upon finding that few surgical categories were associated with different stress testing rates, we replaced that multinomial variable with indicator variables for each category associated with different testing rates in our data (aortic, peripheral vascular, and urologic surgeries). We tested both documented and calculated RCRI, which may differ for a variety of reasons including lab results between the clinic visit and documentation, erroneous diagnoses/chart lore, outside records unavailable in the electronic medical record, and misconceptions about how RCRI is calculated. We treated both estimates of RCRI as continuous to force a monotonic relationship (no theory would support lower testing rates at higher predicted cardiac risk). We tested MICA-predicted probability both as a continuous variable and dichotomized at 1%. Although we used the American Society of Anesthesiologists (ASA) class to calculate MICA, we did not test that separately, as it is usually documented only after the patient's visit to this clinic during subsequent anesthesiologist evaluation.

For measures of functional status, we tested estimated metabolic equivalents (METs), which in this clinic is based on a semi-quantitative questionnaire, and the physician's subjective global assessment of function, which is comparable to the Eastern Cooperative Oncology Group (ECOG) score.[13] For measures of social and financial support, we tested area deprivation index (a measure of socioeconomic disadvantage based on education, employment, housing quality, and poverty measures by census block group), race, ethnicity, marital status, and age (here dichotomized at age 65 to reflect changes in access to care with universal Medicare eligibility).[14] For measures of medical comorbidities and illness, we considered age, vital signs at the clinic visit, diagnoses of coronary artery disease, cerebrovascular disease, or congestive heart failure, diabetes, use of insulin, creatinine, tobacco use, and predicted probability of obstructive coronary artery disease.[15] For measures of physician tendencies and experience, we tested (on the date of each visit) years of post-residency practice (a proxy for overall experience) and the number of previous encounters the physician had completed in our dataset (a proxy for experience in preoperative risk assessment more specifically). For measures of time, we theorized that patients who had previous cardiac stress tests would be less likely to be referred for preoperative stress testing, and that physicians would give greater weight to more recent tests (i.e., the relationship would be time-dependent). We used the date of the visit as a continuous variable to test for changing stress test rates over time. To assess for changes related to publication of the current ACC/AHA guideline, we created a dichotomous variable for whether the visit occurred before or after said guideline's December 9, 2014 publication, and tested for interactions between that term and other predictors.[1]

Model structure

Conceptually, visits could be thought of as clustered by patient (with physician-level variables at the level of the visit) or by physician (with patient-level variables at the level of the visit). We tested different structures using empty models and calculated intraclass correlation coefficients to estimate what proportion of variance was explained by unmeasured patient-level or physician-level factors. With visits clustered by physician, approximately 0.4% of variance in stress test ordering was at the level of the physician. When we clustered visits by patient, approximately 4.9% of variance in stress test ordering was at the level of the patient. We therefore developed our model using physician- and visit-level variables clustered by patient, including physician ID as a visit-level indicator ("dummy") variable. This approach drops some low volume providers for whom outcomes are overfitted but should capture more unmeasured variance at both the patient and provider levels.

Multivariable modeling

Finally, we added the remaining predictor variables from each submodel into a multivariable logistic regression model. We again pruned predictors based on BIC and examined for nonlinear or categorical relationships. We revisited our model structure using the final predictor variables, comparing models clustered by patient or physician based on BIC.

Because results of a multilevel logistic regression with interaction terms have limited intuitive meaning, we calculated marginal effects for reporting. Holding all other variables at their medians, we estimated the effect of changing one predictor variable at a time.

Data extraction and missingness

Our methods for extracting data from the electronic medical record have been described previously.[16] We considered patients as having each considered diagnosis if it had been documented at any time before or at the analyzed visit. For creatinine and other lab testing, we used the most recent measurement up to and including the day of the clinic visit. We used multiple imputation by chained equations to address missing data and previously described standards to ensure multiple imputation did not introduce Monte Carlo error.[17] We imputed predictor variables other than those with negligible missing data: age, sex, vital signs, and binary variables indicating previous diagnoses.

Patient and Public Involvement

Patients were not involved in the design of this study.

Results

Overall, 5.2% of visits to the preoperative clinic led to a cancelled or completed preoperative stress test (8,303/159,795), with 5.1% (8,085; 97.4% of those referred) completing the test. Patient demographics, selected risk factors, and proportions of missing data are shown in Tables 1 and 2. Unadjusted physician referral rates are shown in

Figure 1.

Marginal testing rates across each predictor variable, with other variables at their respective medians, are shown in Table 3. In general, patients were more likely to be referred for preoperative stress testing as estimated perioperative risk increased and for specific categories of surgeries (aortic, peripheral vascular, and urologic). Of those, patients undergoing aortic surgery were most likely to be referred for stress testing, though our dataset included relatively few aortic surgeries and confidence intervals for that predictor were wide. Even after adjusting for all other factors, different providers were more likely to refer for preoperative stress testing than others: a visit to the 95^{th} percentile physician in this clinic would result in stress testing 3.8% of the time, while a visit to the 5^{th} percentile physician in this clinic would result in testing around 1.2% of the time.

Other important patient variables included the physician's subjective assessment of global patient function, METs, socioeconomic advantage or disadvantage compared to the median, BMI, diastolic blood pressure, existing diagnoses of ischemic heart disease or congestive heart failure, estimated probability of obstructive coronary artery disease, and tobacco use. Visits later in our dataset were less likely to result in a preoperative stress test compared with earlier visits. Each of these variables, while significant, appeared to exert less influence than surgical categories, estimated surgical risk, and provider. Fully adjusted provider marginal rates are shown in Figure 2.

Results were very similar for models clustered by patient or physician. Information criteria would slightly favor a model clustered by physician (BIC: 51040) compared to a model clustered by patient with a physician indicator variable (BIC: 52009). Meanwhile, the model clustered by physician had a slightly lower R^2 (0.1896) compared with a model clustered by patient with physician as an indicator variable (0.1907).

Discussion

In this cohort of patients seen for preoperative risk assessment at a single health system, we have identified key drivers of preoperative stress testing, which include type of surgery, estimated surgical risk, and patient functional status. Our results demonstrate use of preoperative stress testing in a real-world cohort.

Current guidelines recommend preoperative stress testing for patients whose predicted perioperative adverse cardiac event risk exceeds 1% and whose functional capacity is poor or unknown, when such testing would change clinical management. Although we cannot determine from these data whether physicians thought testing would change management, and predicted surgical risk scores have poor concordance across the 1% threshold, patients able to perform four or more metabolic equivalents of activity made up nearly one-third of all stress test referrals.[2,10,18] Our data suggest that substantial numbers of preoperative stress tests are against current guidance.

Still, predicted surgical risk is a key driver of preoperative stress testing. Testing rates increased with increasing RCRI, without a clear dichotomization at any particular value of RCRI. And interestingly, although MICA was essentially never documented, a MICA-predicted surgical risk of greater than 1% appears to be a better single predictor variable than RCRI. Physicians could be trying to incorporate the guideline-recommended threshold into their decision-making while relying on cohorts with different calibration, or could be deliberately avoiding a stark dichotomization of risk at 1%.[10,18]

As with predicted surgical risk, physicians appeared to consider functional status as something between the dichotomy of current guidance and continuum of risk encountered in clinical practice. Testing rates were higher among less functional patients and lower among patients able to achieve higher METs, but neither were especially important predictors in our model. It seems probable that clinical decision-making is more nuanced than we can discern from our data source. For example, physicians could reasonably be more inclined to test before a pancreaticoduodenectomy than laparoscopic cholecystectomy in view of those procedures' very different metabolic demands, but we lack sufficient power to test individual surgical procedures without overfitting.[12] In any case, these variables explained little variance in testing rates.

Surgical category also offers insights into testing rationale. Patients were tested more frequently before aortic or peripheral vascular interventions, perhaps reflecting persistent beliefs that patients with coronary artery disease should be identified and revascularized before vascular surgery.[19–21] However, patients undergoing vascular surgery are generally evaluated elsewhere at our institution, leaving our sample small and confidence intervals wide. We also note that patients are more likely to be referred for stress testing before urologic surgery, in spite of no higher intrinsic cardiac risk, compared to other common surgical categories. Anecdotally, physicians practicing in this clinic have reported that a number of urologists at our institution are reluctant to operate on high-risk patients unless those patients first undergo preoperative testing. While investigating such a hypothesis would require a different approach than ours, clearly every physician in a preoperative clinic functions within a larger system of care and must build consensus among a team of treating physicians.

Other significant predictor variables include tobacco use, BMI, diastolic blood pressure, ischemic heart disease, and a patient's census block group. Because each of these are correlated with risk of obstructive coronary artery disease, one possibility is that data unavailable to us (such as outside records) led to some portion of the preoperative stress testing we observed. Although that remains possible, multiple observations argue against a simple explanation that these variables are proxies for coronary disease risk. First, higher probability of obstructive coronary artery disease, calculated based on available data, was associated with lower likelihood of preoperative testing. Second, patients residing in the wealthiest and poorest census tracts were approximately as likely to be referred for stress testing, with patients in the middle of the socioeconomic range less likely. Finally, diabetes was not associated with testing. It would seem that either physicians in our dataset did not incorporate patients' pretest probability of obstructive coronary disease in their decision to refer for testing, or that their assessments were poorly calibrated.

Rates of testing have declined over time in our dataset, in contrast with increasing testing rates suggested by other datasets.[22,23] Our data are unable to answer whether this decline is specific to our site over this time period or part of a wider change in practice. Our dataset does not suggest a clear change in testing rates after the release of current guidelines, and therefore argues against a causal relationship between publication of the current guideline and changes in testing rates.

Our model demonstrates physician practice variation: with all other predictors held at their medians, the physician at the 95th percentile was around three times as likely to order preoperative stress testing than the physician at the 5th percentile. But we would caution against using our results, or others, for profiling individual providers, which is generally a low-reliability exercise and prone to gaming.[24] Our dataset is among the largest clinical datasets of preoperative risk assessment, but true outliers are rare and most physicians are not detectably different from the mean after adjustment (see Figure 2). As with other observations of physician practice variation, ours suggests a deeper failure: that our field has yet to fully understand how preoperative stress testing might be used to mitigate perioperative cardiac risk.[25]

Operative intervention carries inherent cardiac risk, and stress testing may reflect physician discomfort with the malpractice or cognitive liabilities that cardiac risk incurs.[26–30] Of course, stress testing does not in itself mitigate operative risk: it can inform diagnosis and prognosis, but without an intervention that is allocated based on test results and reduces cardiac risk, it cannot be therapeutic. Such an intervention has proven elusive: preoperative revascularization did not reduce cardiac risk in the largest randomized trial to date, beta blockers are more likely harmful than helpful, and other interventions (e.g., statins) that may be allocated differently based on stress testing likely have modest effects, if any.[31–34] Intraoperative care or postoperative testing patterns could differ based on whether a preoperative stress test was performed, but what practices in those settings might reduce the risk of major adverse cardiac events remain equally unclear. While estimated perioperative cardiac risk appears to drive stress testing, it remains to be seen how stress testing might reduce perioperative cardiac risk.

Although we have made every effort to ensure the internal validity of our data, analysis, and results, we cannot be sure about the representativeness of our findings. Our data may not adequately represent drivers of or variation in preoperative stress testing before some common types of surgery, including ophthalmologic surgery, who are evaluated elsewhere in our institution.[35] This clinic has made substantial efforts to provide uniform care, which could have reduced physician variation in our dataset, and we cannot analyze variation across health system or region, which can also be substantial.[36] As with any single-center study, we would urge caution when generalizing to other settings. For example, we found higher rates of testing before urologic surgery than would be expected for cardiac risk; other centers may have different surgical categories with testing out of proportion to surgical risk. The need to build a consensus plan of care among a treatment team is true across institutions, but the particulars of our institution's consensus may not be.

Additionally, as with other observational studies, our analytical choices are difficult to separate from our theoretical framework, and may influence our results in various ways.[37] For example, we rejected physician experience as a predictor of testing in favor of a random effect for each physician and the date of each visit due to AIC and BIC. Still, experience differs by physician and necessarily accrues over time. A reasonable investigator with a different theoretical model could assume broadly stable testing rates over time and conclude that testing decreases as physicians gain experience (see Supplemental Appendix). Time in particular is rife with potential confounders of this sort.

But while the limitations of our study reflect the limitations of any single center observational study, the detailed clinical data available to us offers distinct advantages. We have demonstrated real-world use of preoperative stress testing before a wide range of possible surgical interventions, using visit-level data to comprehensively assess variation in and predictors of preoperative cardiac stress testing.

In summary, use of preoperative stress testing varied with estimated surgical risk, patient functional status, socioeconomic status, ischemic heart disease, congestive heart failure, body mass index, diastolic blood pressure, surgical category, and provider. The fraction of patients referred for stress testing appears to be declining over time, but testing remains common and highly dependent on the provider. The value of preoperative stress testing remains to be established.

Funding Sources:

This work was funded by NHLBI 1K08HL141598.

Conflict of Interest Disclosures:

None

Author Contributions:

Dr. Pappas: This author conceived the study, collected the data, performed the analysis, and wrote the initial draft of the manuscript.
Dr. Sessler: This author helped improve the analysis and revised the manuscript.
Dr. Auerbach: This author helped improve the analysis and revised the manuscript.
Dr. Kattan: This author helped improve the analysis and revised the manuscript.
Mr. Milinovich: This author helped obtain and validate data.
Dr. Blackstone: This author helped improve the analysis and revised the manuscript.
Dr. Rothberg: This author helped improve the analysis and revised the manuscript.

References

1. Fleisher LA, Fleischmann KE, Auerbach AD, et al. 2014 ACC/AHA guideline on perioperative cardiovascular evaluation and management of patients undergoing noncardiac surgery: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;130(24):2215-2245. doi:10.1161/CIR.0000000000000105
2. Pappas MA, Sessler DI, Rothberg MB. Anticipated Rates and Costs of Guideline-Concordant Preoperative Stress Testing. *Anesth Analg*. 2019;128(2):241-246. doi:10.1213/ANE.0000000000003754
3. Glance LG, Faden E, Dutton RP, et al. Impact of the choice of risk model for identifying low-risk patients using the 2014 American college of cardiology/American Heart Association perioperative guidelines. *Anesthesiology*. 2018;129(5):889-900. doi:10.1097/ALN.0000000000002341
4. Hlatky MA, Boineau RE, Higginbotham MB, et al. A brief self-administered questionnaire to determine functional capacity (The Duke Activity Status Index). *Am J Cardiol*. 1989;64(10):651-654. doi:10.1016/0002-9149(89)90496-7
5. Reilly DF, McNeely MJ, Doerner D, et al. Self-reported exercise tolerance and the risk of serious perioperative complications. *Arch Intern Med*. 1999;159(18):2185-2192. doi:10.1001/archinte.159.18.2185
6. Melon CC, Eshtiaghi P, Luksun WJ, Wijeysundera DN. Validated questionnaire vs physicians' judgment to estimate preoperative exercise capacity. *JAMA Intern Med*.

2014;174(9):1507-1508. doi:10.1001/jamainternmed.2014.2914

7. Wijeysundera DN, Pearse RM, Shulman MA, et al. Assessment of functional capacity before major non-cardiac surgery: an international, prospective cohort study. *Lancet*. 2018;391(10140):2631-2640. doi:10.1016/S0140-6736(18)31131-0

8. Eddy DM. Variations in physician practice: The role of uncertainty. *Health Aff*. 1984;3(2):74-89. doi:10.1377/hlthaff.3.2.74

9. Mark DB, Federspiel JJ, Cowper PA, et al. Economic Outcomes With Anatomical Versus Functional Diagnostic Testing for Coronary Artery Disease. *Ann Intern Med*. 2016;165(2):94-102. doi:10.7326/M15-2639

10. Lee TH, Marcantonio ER, Mangione CM, et al. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. *Circulation*. 1999;100(10):1043-1049. doi:10.1161/01.CIR.100.10.1043

11. Gupta PK, Gupta H, Sundaram A, et al. Development and validation of a risk calculator for prediction of cardiac risk after surgery. *Circulation*. 2011;124(4):381-387. doi:10.1161/CIRCULATIONAHA.110.015701

12. Liu JB, Liu Y, Cohen ME, Ko CY, Sweitzer BJ. Defining the Intrinsic Cardiac Risks of Operations to Improve Preoperative Cardiac Risk Assessments. *Anesthesiology*. 2018;128(2):283-292. doi:10.1097/ALN.0000000000002024

13. Oken MM, Creech RH, Davis TE. Toxicology and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol*. 1982;5(6):649-655. doi:10.1097/00000421-198212000-00014

14. Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible - The neighborhood atlas. *N Engl J Med*. 2018;378(26):2456-2458. doi:10.1056/NEJMp1802313

15. Genders TSS, Steyerberg EW, Hunink MGM, et al. Prediction model to estimate presence of coronary artery disease: Retrospective pooled analysis of existing cohorts. *BMJ*. 2012;344(7862). doi:10.1136/bmj.e3485

16. Milinovich A, Kattan MW. Extracting and utilizing electronic health data from Epic for research. *Ann Transl Med*. 2018;6(3):42-42. doi:10.21037/atm.2018.01.13

17. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30(4):377-399. doi:10.1002/sim.4067

18. Duceppe E, Parlow J, MacDonald P, et al. Canadian Cardiovascular Society Guidelines on Perioperative Cardiac Risk Assessment and Management for Patients Who Undergo Noncardiac Surgery. *Can J Cardiol*. 2017;33(1):17-32. doi:10.1016/j.cjca.2016.09.008

19. Eagle KA, Singer DE, Brewster DC, Darling RC, Mulley AG, Boucher CA. Dipyridamole-thallium scanning in patients undergoing vascular surgery. Optimizing preoperative evaluation of cardiac risk. *JAMA*. 1987;257(16):2185-2189.

20. Hertzer NR, Beven EG, Young JR, et al. Coronary artery disease in peripheral vascular patients. A classification of 1000 coronary angiograms and results of surgical management. *Ann Surg*. 1984;199(2):223-233. doi:10.1097/00000658-198402000-00016

21. Boucher CA, Brewster DC, Darling RC, Okada RD, Strauss HW, Pohost GM. Determination of Cardiac Risk by Dipyridamole-Thallium Imaging before Peripheral Vascular Surgery. *N Engl J Med*. 1985;312(7):389-394. doi:10.1056/NEJM198502143120701

22. Sheffield KM, McAdams PS, Benarroch-Gampel J, et al. Overuse of preoperative cardiac stress testing in medicare patients undergoing elective noncardiac surgery. *Ann Surg*. 2013;257(1):73-80. doi:10.1097/SLA.0b013e31826bc2f4

23. Sigmund AE, Stevens ER, Blitz JD, Ladapo JA. Use of preoperative testing and physicians' response to professional society guidance. *JAMA Intern Med*. 2015;175(8):1352-1359. doi:10.1001/jamainternmed.2015.2081

24. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *J Am Med Assoc*. 1999;281(22):2098-2105. doi:10.1001/jama.281.22.2098

25. Wennberg JE, Freeman JL, Culp WJ. Are Hospital Services Rationed in New Haven or Over-Utilized in Boston? *Lancet*. 1987;329(8543):1185-1189. doi:10.1016/S0140-6736(87)92152-0

26. Saint S, Vaughn VM, Chopra V, Fowler KE, Kachalia A. Perception of Resources Spent on Defensive Medicine and History of Being Sued Among Hospitalists: Results from a National Survey. *J Hosp Med*. 2018;13(1):26-29. doi:10.12788/jhm.2800

27. Feinstein AR. The 'Chagrin Factor' and Qualitative Decision Analysis. *Arch Intern*

Med. 1985;145(7):1257-1259. doi:10.1001/archinte.1985.00360070137023

28. Rolfe A, Burton C. Reassurance after Diagnostic Testing with a Low Pretest Probability of Serious Disease: Systematic Review and Meta-analysis. *JAMA Intern Med*. 2013;173(6):407-416. doi:10.1001/jamainternmed.2013.2762

29. Van der Weijden T, Van Velsen M, Dinant GJ, Van Hasselt CM, Grol R. Unexplained complaints in general practice: Prevalence, patients' expectations, and professionals' test-ordering behavior. *Med Decis Mak*. 2003;23(3):226-231. doi:10.1177/0272989X03023003004

30. Little P, Dorward M, Warner G, Stephens K, Senior J, Moore M. Importance of patient pressure and perceived pressure and perceived medical need for investigations, referral, and prescribing in primary care: Nested observational study. *Br Med J*. 2004;328(7437):444-446. doi:10.1136/bmj.38013.644086.7c

31. McFalls EO, Ward HB, Moritz TE, et al. Coronary-Artery Revascularization before Elective Major Vascular Surgery. *N Engl J Med*. 2004;351(27):2795-2804. doi:10.1056/NEJMoa041905

32. Blessberger H, Kammler J, Domanovits H, et al. Perioperative beta-blockers for preventing surgery-related mortality and morbidity. *Cochrane Database Syst Rev*. March 2018. doi:10.1002/14651858.CD004476.pub3

33. Berwanger O, Manach Y Le, Suzumura EA, et al. Association between pre-operative statin use and major cardiovascular complications among patients undergoing non-cardiac surgery: The VISION study. *Eur Heart J*. 2016;37(2):177-185.

doi:10.1093/eurheartj/ehv456

34. Berwanger O, de Barros e Silva PGM, Barbosa RR, et al. Atorvastatin for high-risk statin-naïve patients undergoing noncardiac surgery: The Lowering the Risk of Operative Complications Using Atorvastatin Loading Dose (LOAD) randomized trial. *Am Heart J*. 2017;184:88-96. doi:10.1016/j.ahj.2016.11.001

35. Kerr EA, Chen J, Sussman JB, Klamerus ML, Nallamothu BK. Stress Testing Before Low-Risk Surgery: So Many Recommendations, So Little Overuse. *JAMA Intern Med*. 2015;175(4):645. doi:10.1001/jamainternmed.2014.7877

36. Home - Dartmouth Atlas of Health Care. https://www.dartmouthatlas.org/. Accessed November 11, 2019.

37. Palpacuer C, Hammas K, Duprez R, Laviolle B, Ioannidis JPA, Naudet F. Vibration of effects from diverse inclusion/exclusion criteria and analytical choices: 9216 different ways to perform an indirect comparison meta-analysis. *BMC Med*. 2019;17(1):174. doi:10.1186/s12916-019-1409-3

Table 1. Patient and Surgical Characteristics.

		Total	Percent of category	Percent of all visits	Completed preoperative stress test	Percent of all preoperative stress tests
Age		159,795	100.0%	100.0%	8,303	100.0%
Sex	Female	88,738	55.5%	55.5%	4,079	49.1%
	Male	71,055	44.5%	44.5%	4,224	50.9%
Previous diagnosis of ischemic heart disease	No	128,505	80.4%	80.4%	4,831	58.2%
	Yes	31,290	19.6%	19.6%	3,472	41.8%
Previous diagnosis of congestive heart failure	No	146,556	91.7%	91.7%	6,922	83.4%
	Yes	13,239	8.3%	8.3%	1,381	16.6%
Previous diagnosis of cerebrovascular disease	No	141,519	88.6%	88.6%	6,567	79.1%
	Yes	18,276	11.4%	11.4%	1,736	20.9%
Systolic Blood Pressure		159,488	100.0%	99.8%	8,285	99.8%
Diastolic Blood Pressure		159,481	100.0%	99.8%	8,284	99.8%
Body Mass Index		157,473	100.0%	98.5%	8,155	98.2%
Creatinine (RCRI categorization)	≤ 2.0 mg/dL	151,885	97.1%	95.1%	7,695	92.7%
	> 2.0 mg/dL	4,487	2.9%	2.8%	542	6.5%
Creatinine (MICA categorization)	≤ 1.5 mg/dL	144,369	90.3%	90.3%	7,128	85.8%
	> 1.5 mg/dL	12,003	7.5%	7.5%	1,109	13.4%
	Unknown	3,423	2.1%	2.1%	66	0.8%
Prescribed insulin	No	147,610	92.4%	92.4%	7,136	85.9%
	Yes	12,185	7.6%	7.6%	1,167	14.1%
RCRI surgical category	High risk	27,709	23.8%	17.3%	872	10.5%
	Other	88,929	76.2%	55.7%	360	4.3%
Area Deprivation Index		126,076	100.0%	78.9%	7,091	85.4%
RCRI (documented)	0	50,785	75.1%	31.8%	1,548	18.6%
	1	12,642	18.7%	7.9%	988	11.9%
	2	3,321	4.9%	2.1%	402	4.8%
	3	742	1.1%	0.5%	108	1.3%
	4	151	0.2%	0.1%	21	0.3%
	5	15	0.0%	0.0%	2	0.0%
RCRI (calculated)	0	56,879	52.0%	35.6%	355	4.3%

	1	36,020	32.9%	22.5%	393	4.7%
	2	11,143	10.2%	7.0%	245	3.0%
	3	4,006	3.7%	2.5%	98	1.2%
	4	1,172	1.1%	0.7%	33	0.4%
	5	204	0.2%	0.1%	7	0.1%
	6	12	0.0%	0.0%	0	0.0%
MICA risk estimate	≤ 1%	71,448	44.7%	44.7%	608	7.3%
	> 1%	88,347	55.3%	55.3%	7,695	92.7%
Physician subjective assessment of patient global function	1	17,991	19.1%	11.3%	489	5.9%
	1-2	13,386	14.2%	8.4%	634	7.6%
	2	40,829	43.4%	25.6%	1,739	20.9%
	2-3	8,588	9.1%	5.4%	646	7.8%
	3	9,310	9.9%	5.8%	732	8.8%
	3-4	1,922	2.0%	1.2%	253	3.0%
	4	1,999	2.1%	1.3%	175	2.1%
Insurance	Medicare	20,744	52.2%	13.0%	1,192	14.4%
	Medicaid	2,384	6.0%	1.5%	83	1.0%
	Private	14,764	37.1%	9.2%	307	3.7%
	Other listed insurer	1,881	4.7%	1.2%	53	0.6%
ASA Class	1	3,325	3.0%	2.1%	9	0.1%
	2	34,026	30.9%	21.3%	181	2.2%
	3	65,298	59.3%	40.9%	854	10.3%
	4	7,454	6.8%	4.7%	157	1.9%
MICA surgical category	Anorectal	1,213	1.0%	0.8%	3	0.0%
	Aortic	115	0.1%	0.1%	34	0.4%
	Bariatric	703	0.6%	0.4%	6	0.1%
	Brain	4,780	4.1%	3.0%	15	0.2%
	Breast	8,541	7.4%	5.3%	17	0.2%
	Cardiac	237	0.2%	0.1%	10	0.1%
	Ear, nose, throat	3,640	3.1%	2.3%	19	0.2%
	Foregut/hepatopancreatobiliary	3,252	2.8%	2.0%	37	0.4%
	Gallbladder, appendix, adrenals, or spleen	1,974	1.7%	1.2%	17	0.2%
	Gynecologic	7,458	6.4%	4.7%	51	0.6%
	Hernia	2,778	2.4%	1.7%	16	0.2%

	Intestines	16,081	13.9%	10.1%	87	1.0%
	Neck	3,076	2.7%	1.9%	18	0.2%
	Nonesophageal thoracic	290	0.2%	0.2%	4	0.0%
	Orthopedic	26,005	22.4%	16.3%	208	2.5%
	Other abdomen	1,970	1.7%	1.2%	19	0.2%
	Peripheral vascular	1,355	1.2%	0.8%	110	1.3%
	Skin	7,126	6.1%	4.5%	78	0.9%
	Spinal	11,994	10.3%	7.5%	59	0.7%
	Urologic	13,423	11.6%	8.4%	414	5.0%
	Vein	48	0.0%	0.0%	3	0.0%
Before or after release of current guideline	Before	98,465	61.6%	61.6%	5,911	71.2%
	After	61,330	38.4%	38.4%	2,392	28.8%
Able to perform activities of at least 4 METs	No	90,260	56.5%	56.5%	5,713	68.8%
	Yes	69,535	43.5%	43.5%	2,590	31.2%
Tobacco use	Current smoker	18,806	12.4%	11.8%	1,028	12.4%
	Former smoker	62,067	40.9%	38.8%	3,834	46.2%
	Never smoker	70,966	46.7%	44.4%	2,976	35.8%
Probability of obstructive CAD		159,793	100.0%	100.0%	8,303	100.0%

Table 2: Summary of continuous variables.

	Mean ± SD
Age (years)	58.8 ± 15.2
Systolic Blood Pressure	129 ± 19.1
Diastolic Blood Pressure	73 ± 11.3
Body Mass Index	30.1 ± 7.6
Creatinine	1.02 ± 0.85
Area Deprivation Index	54.2 ± 24.6
Estimated METs of activity	5.22 ± 1.35
MICA risk estimate	0.021 ± 0.025
Previous patients seen by physician in clinic	1,587 ± 1,536
Predicted probability of obstructive CAD	0.129 ± 0.136

Table 3. Marginal results for each variable in our final model, with all other variables held at their medians.

For example, with all other variables at their respective medians, a visit on June 30, 2008 would have resulted in preoperative stress testing approximately 3.5% of the time, while a visit on June 30, 2018 would have resulted in preoperative stress testing approximately 1.3% of the time. Provider effects are summarized for space considerations; full marginal results by physician are included in the Supplemental Appendix.

Predictor	Value	Marginal Rate	95% CI	
MICA estimate > 1%	0	0.6%	0.6%	0.7%
	1	7.1%	6.7%	7.5%
RCRI (documented)	0	2.2%	2.0%	2.3%
	1	2.7%	2.5%	2.9%
	2	3.4%	3.0%	3.8%
	3	4.2%	3.5%	5.0%
	4	5.1%	4.0%	6.5%
	5	6.4%	4.7%	8.6%
Subjective assessment of patient function	1	2.1%	1.9%	2.2%
	2	2.4%	2.2%	2.5%
	3	2.8%	2.6%	3.0%
	4	3.2%	2.9%	3.5%
Estimated METs	2	3.3%	2.9%	3.7%
	4	2.7%	2.5%	2.9%
	8	1.8%	1.6%	1.9%
Body mass index	20	2.1%	1.9%	2.3%
	30	2.4%	2.2%	2.5%
	40	2.7%	2.5%	2.9%
Diastolic blood pressure	70	2.3%	2.2%	2.5%
	90	2.5%	2.3%	2.7%
	110	2.7%	2.4%	3.0%
Ischemic heart disease	No	2.1%	2.0%	2.3%
	Yes	3.6%	3.3%	3.9%
Congestive heart	No	2.4%	2.2%	2.5%

failure	Yes	2.1%	1.9%	2.3%
Area deprivation index	10	2.8%	2.5%	3.1%
	50	2.2%	2.0%	2.3%
	90	2.6%	2.4%	2.8%
Predicted probability of obstructive CAD	5%	2.6%	2.4%	2.7%
	10%	2.4%	2.3%	2.6%
	20%	2.2%	2.1%	2.4%
Tobacco use	Current smoker	2.6%	2.3%	2.8%
	Former smoker	2.5%	2.3%	2.7%
	Neither	2.2%	2.1%	2.4%
Date	2008.06.30	3.5%	3.2%	3.8%
	2013.06.30	2.6%	2.4%	2.8%
	2018.06.30	1.3%	1.2%	1.4%
Surgical category	Aortic	23.4%	6.0%	91.1%
	Peripheral vascular	8.7%	6.7%	11.3%
	Urologic	9.2%	8.3%	10.2%
	Other	1.9%	1.7%	2.0%
Physician (summary)	Lowest	1.0%	0.1%	4.4%
	5th percentile	1.2%	0.6%	2.6%
	Median	2.3%	2.1%	2.6%
	95th percentile	3.8%	3.2%	4.5%
	Highest	6.1%	2.7%	13.5%

Figure 1: Unadjusted rates of preoperative stress testing, by physician.

[Attached separately as Figure_01.tif]

Figure 2: Mean marginal rates of preoperative stress testing, by physician.

[Attached separately as Figure_02.tif]

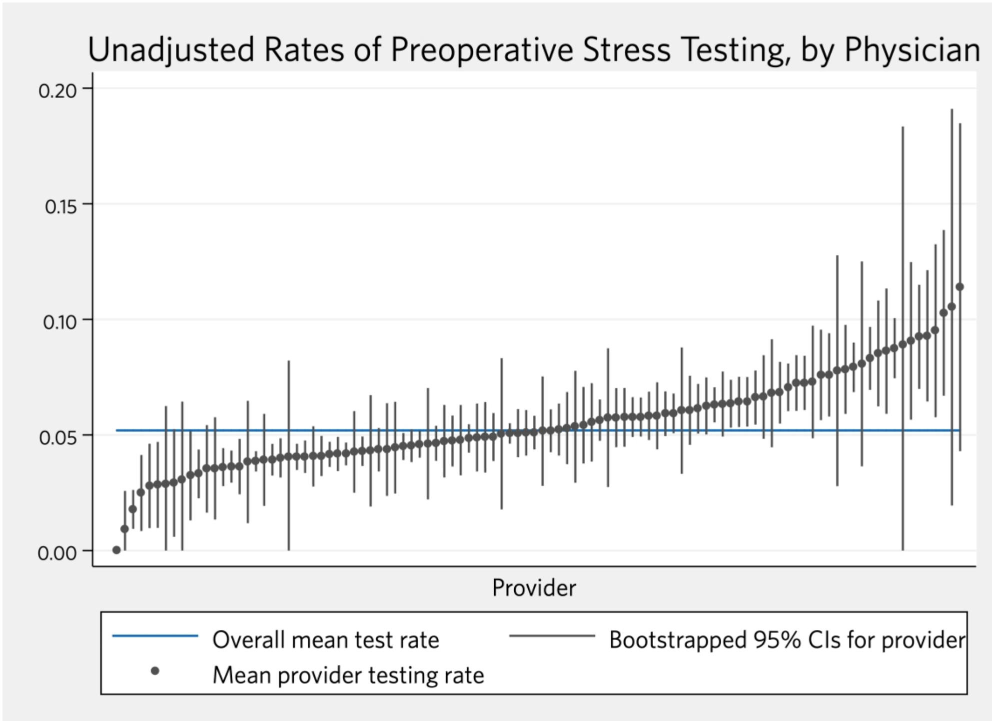


Figure 1: Unadjusted rates of preoperative stress testing, by physician

569x413mm (144 x 144 DPI)

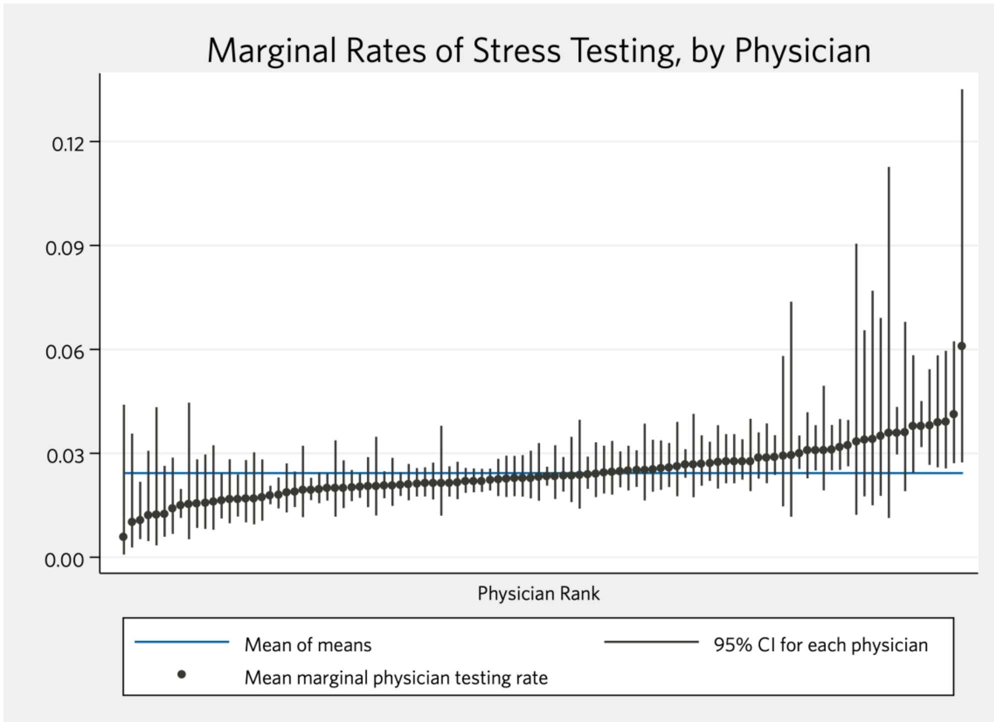


Figure 2: Mean marginal rates of preoperative stress testing, by physician

569x413mm (144 x 144 DPI)

Supplemental Material for: Variation in preoperative stress testing by patient, physician, and surgical type

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

1. Regression results, clustered by physician

Results have been edited/trimmed for clarity and brevity. Interaction terms have been replaced in the table below with exponents (^2) when applicable. All dichotomous variables (those prepended with "1." in the table) use a value of 0 as the referent.

```
Random-effects logistic regression             Number of obs     =    154,171
Group variable: Physician_ID                   Number of groups  =        104
Observations per group:                                     min =          7
                                                            avg =    1,482.4
                                                            max =      8,245
```

Completed or cancelled test	Beta Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
1.aortic_surgery	2.522167	.6078339	4.15	0.002	1.166378	3.877956
1.peripheral_vascular_surgery	1.531945	.1152814	13.29	0.000	1.285334	1.778556
1.urologic_surgery	1.5871	.0657605	24.13	0.000	1.446111	1.728088
documented_RCRI	.2135885	.03032	7.04	0.000	.1506825	.2764944
1.Gupta_greater_than_1pct	2.422218	.0472362	51.28	0.000	2.329492	2.514944
Estimated METs	-.1030793	.0128133	-8.04	0.000	-.1284818	-.0776767
functional_class	.1444786	.0214833	6.73	0.000	.1020898	.1868674
ADI_national	-.0141123	.0025089	-5.62	0.000	-.0190783	-.0091463
ADI_national^2	.0001325	.0000221	5.99	0.000	.0000888	.0001763
BMI	.0120129	.0016333	7.36	0.000	.0088076	.0152182
DBP	.0037949	.0011446	3.32	0.001	.0015493	.0060406
1.ischemic_heart_disease	.5187915	.0390133	13.30	0.000	.4412007	.5963823
predicted_prob_of_CAD	-.9799052	.1105385	-8.86	0.000	-1.198045	-.761765
1.congestive_heart_failure	-.1196815	.0476752	-2.51	0.014	-.2142709	-.0250922
date	.002128	.0004881	4.36	0.000	.0011708	.0030853
date^2	-6.15e-08	1.25e-08	-4.92	0.000	-8.61e-08	-3.70e-08
1.current_smoker	.1436825	.0424903	3.38	0.001	.0603125	.2270526
1.former_smoker	.1288456	.0290712	4.43	0.000	.0718283	.1858628
constant	-23.48573	4.748773	-4.95	0.000	-32.79838	-14.17308
/lnsig2u	-3.592461	.2658187			-4.113895	-3.071026
sigma_u	.1659232	.0220527			.1278436	.2153451
rho	.0082988	.0021877			.0049434	.0138999

2. Physician marginal rates of stress testing

Due to space constraints, only selected physician marginal results are displayed in Table 3. Here we present full marginal results by physician. These are also displayed visually in Figure 2.

Physician Rank	Marginal Rate	95% CI
1	0.58%	0.08% – 4.40%
2	1.00%	0.28% – 3.57%
3	1.07%	0.52% – 2.18%
4	1.20%	0.47% – 3.08%
5	1.22%	0.34% – 4.33%
6	1.25%	0.59% – 2.64%
7	1.39%	0.68% – 2.88%
8	1.50%	1.13% – 1.97%
9	1.52%	0.52% – 4.46%
10	1.55%	0.85% – 2.83%
11	1.56%	0.82% – 2.97%
12	1.60%	0.79% – 3.23%
13	1.64%	1.12% – 2.42%
14	1.67%	0.98% – 2.83%
15	1.68%	1.17% – 2.39%
16	1.68%	1.01% – 2.80%
17	1.70%	0.95% – 3.03%
18	1.73%	1.05% – 2.82%
19	1.78%	1.53% – 2.06%
20	1.80%	1.41% – 2.30%
21	1.87%	1.29% – 2.71%
22	1.89%	1.45% – 2.47%
23	1.93%	1.16% – 3.22%
24	1.95%	1.65% – 2.29%

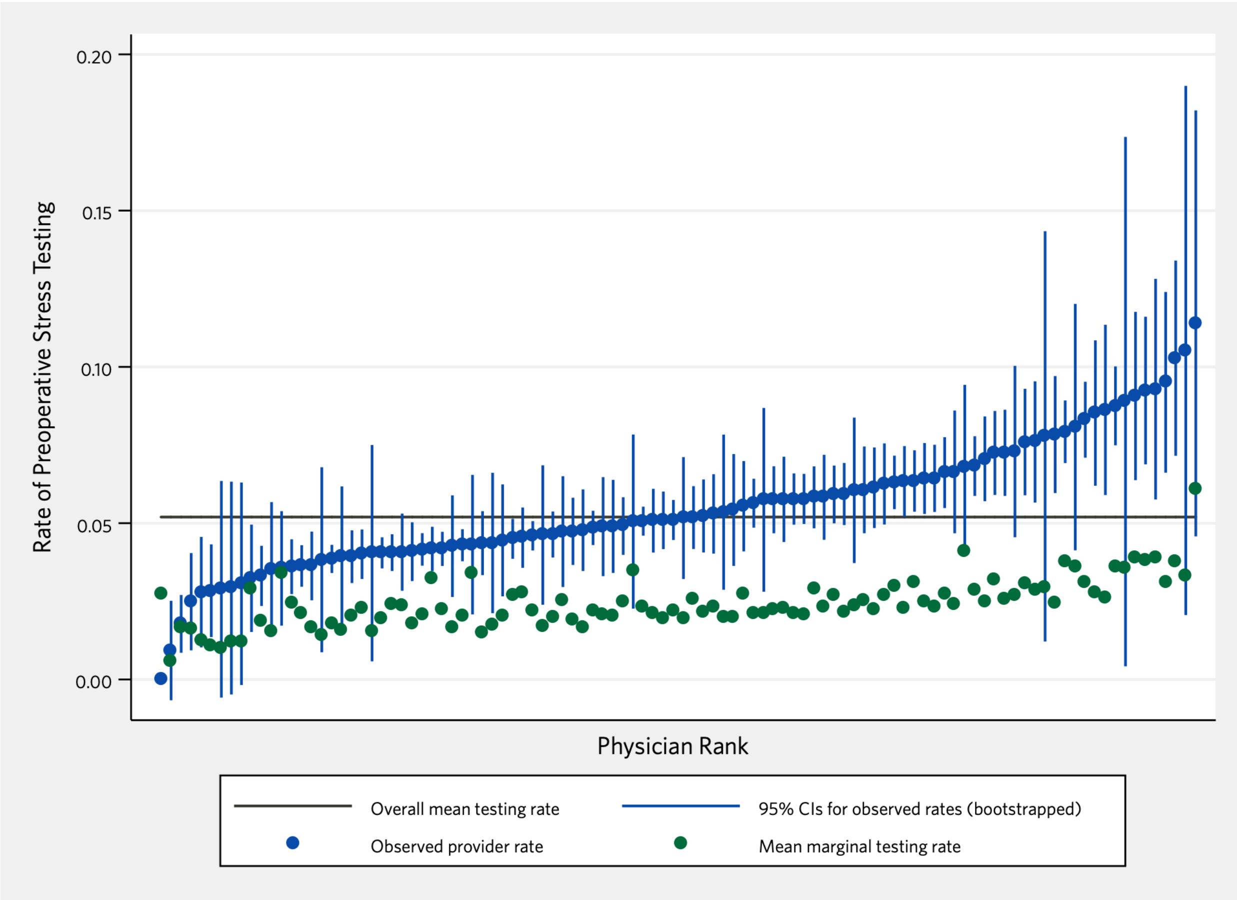
Physician Rank	Marginal Rate	95% CI
25	1.96%	1.56% – 2.45%
26	1.99%	1.63% – 2.42%
27	1.99%	1.17% – 3.38%
28	1.99%	1.42% – 2.80%
29	2.02%	1.62% – 2.52%
30	2.03%	1.71% – 2.42%
31	2.04%	1.45% – 2.89%
32	2.05%	1.20% – 3.48%
33	2.06%	1.70% – 2.48%
34	2.06%	1.48% – 2.87%
35	2.09%	1.77% – 2.46%
36	2.10%	1.64% – 2.69%
37	2.12%	1.75% – 2.57%
38	2.13%	1.75% – 2.60%
39	2.13%	1.67% – 2.73%
40	2.13%	1.20% – 3.80%
41	2.14%	1.75% – 2.62%
42	2.15%	1.68% – 2.76%
43	2.19%	1.86% – 2.58%
44	2.19%	1.87% – 2.56%
45	2.20%	1.89% – 2.55%
46	2.22%	1.93% – 2.56%
47	2.24%	1.76% – 2.84%
48	2.26%	1.74% – 2.93%
49	2.27%	1.76% – 2.93%
50	2.27%	1.75% – 2.95%
51	2.29%	1.70% – 3.08%

Physician Rank	Marginal Rate	95% CI
52	2.32%	1.63% – 3.30%
53	2.32%	2.07% – 2.61%
54	2.33%	1.68% – 3.23%
55	2.34%	1.89% – 2.89%
56	2.35%	1.59% – 3.48%
57	2.36%	1.40% – 3.97%
58	2.39%	1.97% – 2.90%
59	2.40%	1.74% – 3.32%
60	2.43%	1.84% – 3.22%
61	2.46%	1.81% – 3.36%
62	2.48%	2.02% – 3.05%
63	2.49%	1.93% – 3.22%
64	2.50%	2.03% – 3.09%
65	2.51%	1.64% – 3.86%
66	2.53%	1.89% – 3.39%
67	2.56%	1.95% – 3.37%
68	2.58%	2.03% – 3.30%
69	2.62%	1.76% – 3.91%
70	2.68%	2.29% – 3.12%
71	2.68%	1.73% – 4.14%
72	2.70%	2.07% – 3.52%
73	2.71%	2.20% – 3.33%
74	2.75%	1.98% – 3.81%
75	2.75%	2.13% – 3.55%
76	2.75%	2.13% – 3.55%
77	2.76%	2.24% – 3.41%
78	2.76%	1.91% – 4.00%

Physician Rank	Marginal Rate	95% CI
79	2.86%	2.27% – 3.60%
80	2.87%	2.14% – 3.86%
81	2.89%	2.38% – 3.52%
82	2.92%	1.47% – 5.81%
83	2.94%	1.17% – 7.38%
84	2.99%	2.55% – 3.51%
85	3.09%	2.28% – 4.19%
86	3.09%	2.50% – 3.81%
87	3.09%	1.93% – 4.95%
88	3.09%	2.51% – 3.82%
89	3.17%	2.52% – 3.99%
90	3.22%	2.62% – 3.96%
91	3.33%	1.23% – 9.05%
92	3.39%	1.76% – 6.55%
93	3.40%	1.50% – 7.69%
94	3.50%	1.77% – 6.91%
95	3.57%	1.13% – 11.27%
96	3.59%	2.97% – 4.34%
97	3.60%	1.91% – 6.79%
98	3.77%	2.44% – 5.83%
99	3.79%	3.18% – 4.51%
100	3.80%	2.67% – 5.43%
101	3.89%	2.60% – 5.83%
102	3.91%	2.56% – 5.96%
103	4.11%	2.71% – 6.23%
104	6.08%	2.74% – 13.51%

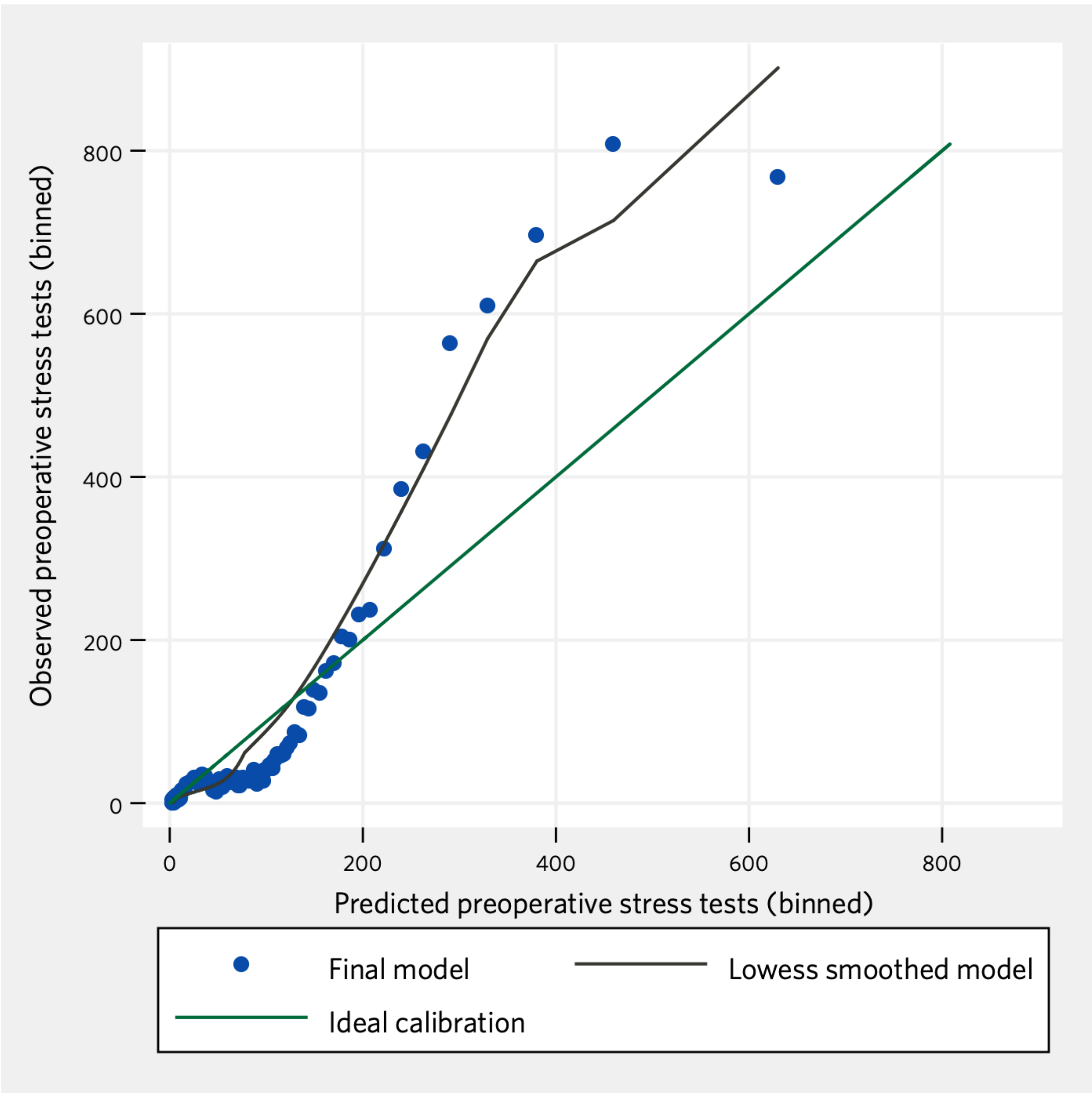
3. Physician rates of stress testing, unadjusted and marginal

Figure 1 demonstrates unadjusted rates of stress testing, and Figure 2 demonstrates marginal predictions for each physician, controlling for all other factors. Here we overlay the marginal results on the unadjusted results to demonstrate the effect of adjustment.



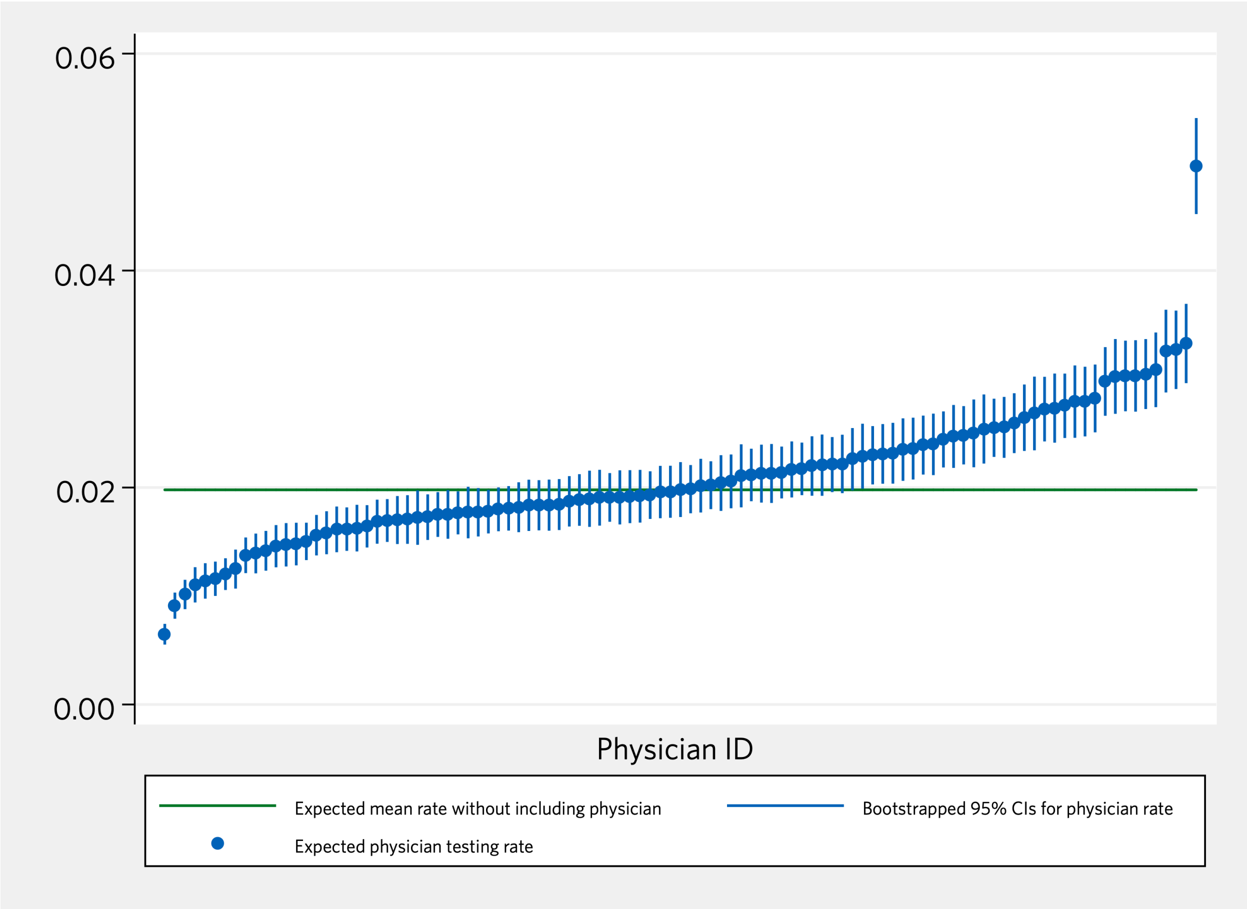
4. Calibration of final model

Based on the plot above, we suspected less-than-ideal calibration of our final model. We do not know of a universally accepted method to assess the calibration of a multilevel model on multiply-imputed data, but in most of our assessments this model fails Hosmer-Lemeshow Goodness-of-Fit testing, and the calibration plot shown here (binned into centiles) indeed suggests poor calibration. We emphasize again that our goal here is to explain variance in testing, not to guide future physicians in who should be referred for stress testing or to enable individual physician profiling.



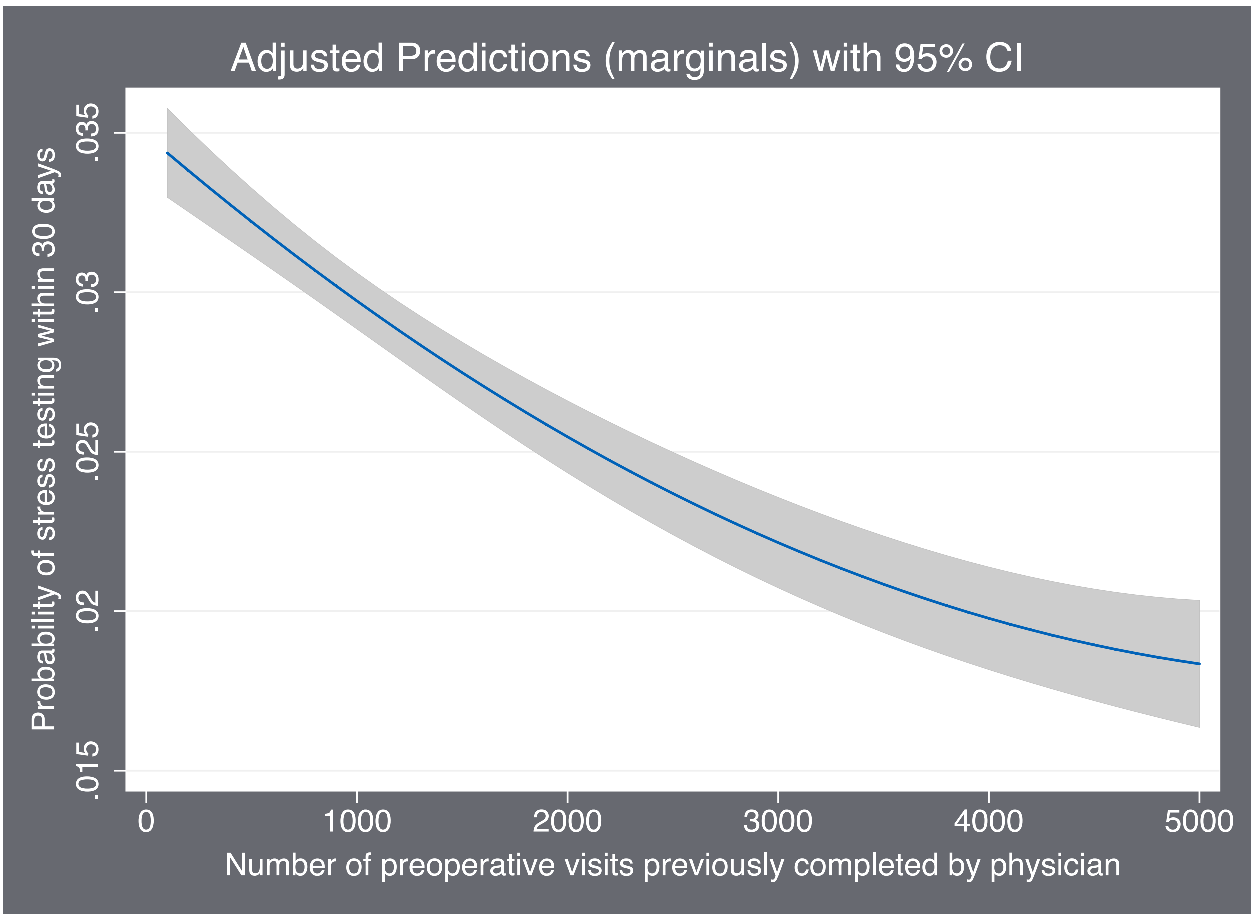
5. Expected testing rates with an identical population

One way to contextualize provider effects is to imagine that each provider sees an identical panel of patients and estimate the consequent differences in outcomes. Here, we sampled 1,000 patients from our original population and estimated rates of stress testing if that same cohort were seen by each physician in our dataset. The overall mean is the expected rate for this small cohort without controlling for physician ID.



6. Marginal testing rate as a function of physician experience

As with all datasets, our conclusions are a product of many decisions. For example, although we rejected physician experience as a predictor of testing rate in favor of date and a physician-specific random effect, reasonable investigators could disagree. To generate the graph below, we replaced date in our model with the number of visits each physician had completed between the beginning of our dataset and the visit in question (a proxy for preoperative experience). We then computed and graphed marginal probabilities as described in our primary results.



STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2-3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4-5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	4-5
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5-8
Bias	9	Describe any efforts to address potential sources of bias	9-10
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5-10
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-10
		(*b*) Describe any methods used to examine subgroups and interactions	5-10
		(*c*) Explain how missing data were addressed	9-10
		(*d*) If applicable, explain how loss to follow-up was addressed	5
		(*e*) Describe any sensitivity analyses	5
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	5,10
		(b) Give reasons for non-participation at each stage	17
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	26-27
		(b) Indicate number of participants with missing data for each variable of interest	26
		(c) Summarise follow-up time (eg, average and total amount)	5
Outcome data	15*	Report numbers of outcome events or summary measures over time	10

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10-11, 26, 28
		(*b*) Report category boundaries when continuous variables were categorized	26
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	28
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11-12
Discussion			
Key results	18	Summarise key results with reference to study objectives	12-15
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	16-18
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	18
Generalisability	21	Discuss the generalisability (external validity) of the study results	17
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	1,19

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

Variation in preoperative stress testing by patient, physician, and surgical type – a cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-048052.R1
Article Type:	Original research
Date Submitted by the Author:	19-Jul-2021
Complete List of Authors:	Pappas, Matthew; Cleveland Clinic, Department of Hospital Medicine Sessler, Daniel; Cleveland Clinic, Outcomes Research Auerbach, A; University of California, San Francisco Kattan, Michael; Cleveland Clinic Milinovich, Alex; Cleveland Clinic Blackstone, Eugene; Cleveland Clinic, Miller Family Heart and Vascular Institute Rothberg, Michael B.; Cleveland Clinic, Cleveland, USA, Department of Internal Medicine
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Cardiovascular medicine, Anaesthesia, Surgery, General practice / Family practice
Keywords:	INTERNAL MEDICINE, Adult surgery < SURGERY, Adult anaesthesia < ANAESTHETICS, Adult cardiology < CARDIOLOGY, HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Variation in preoperative stress testing by patient, physician, and surgical type – a cohort study

Matthew A Pappas, MD, MPH,[1,2,3] Daniel I Sessler, MD,[3,4] Andrew D Auerbach, MD, MPH,[5] Michael W Kattan, PhD,[6] Alex Milinovich, BA,[6] Eugene H Blackstone, MD,[7] and Michael B Rothberg, MD, MPH[1]

1: Center for Value-based Care Research, Cleveland Clinic, Cleveland, OH
2: Department of Hospital Medicine, Cleveland Clinic, Cleveland, OH
3: Department of OUTCOMES RESEARCH, Anesthesiology Institute, Cleveland Clinic, Cleveland, OH
4: OUTCOMES RESEARCH Consortium, Cleveland, OH
5: Department of Hospital Medicine, University of California, San Francisco, San Francisco, CA
6: Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH
7: Miller Family Heart and Vascular Institute, Cleveland Clinic, Cleveland, OH

Correspondence to Matthew A Pappas: pappasm@ccf.org; 216.444.9565; 9500 Euclid Ave., Mail Stop G-10, Cleveland, OH 44195

Short title/running head:	Variation in preoperative stress testing
Word count:	3,644
Abstract word count:	292
Number of references:	39
Number of tables:	3
Number of figures:	2
Number of appendices:	1

Financial support: This work was supported by the National Heart, Lung, and Blood Institute (NHLBI 1K08HL141598).

Competing Interests: None

Keywords (MEDLINE): Perioperative care; Preoperative care; Risk assessment; Elective surgical procedures; Cohort studies; Exercise test; Tomography, Emission-Computed, Single-Photon; Echocardiography, Stress

Abstract

Objectives:

To describe variation in and drivers of contemporary preoperative cardiac stress testing.

Setting:

A dedicated preoperative risk assessment and optimization clinic at a large integrated medical center from 2008 through 2018.

Participants:

A cohort of 118,552 adult patients seen by 104 physicians across 159,795 visits to a preoperative risk assessment and optimization clinic.

Main Outcome:

Referral for stress testing before major surgery, including nuclear, echocardiographic, or electrocardiographic-only stress testing, within 30 days after a clinic visit.

Results:

A total of 8,303 visits (5.2%) resulted in referral for preoperative stress testing. Key patient factors associated with preoperative stress testing included predicted surgical risk, patient functional status, a previous diagnosis of ischemic heart disease, tobacco use, and body mass index. Patients living in either the most- or least-deprived census block groups were more likely to be tested. Patients were tested more frequently before aortic, peripheral vascular, or urologic interventions than before other surgical subcategories. Even after fully adjusting for patient and surgical factors, provider effects remained important: marginal testing rates differed by a factor-of-three in relative terms and around 2.5% in

absolute terms between the 5th and 95th percentile physicians. Stress testing frequency decreased over the time period; controlling for patient and physician predictors, a visit in 2008 would have resulted in stress testing approximately 3.5% of the time, while a visit in 2018 would have resulted in stress testing approximately 1.3% of the time.

Conclusions:

In this large cohort of patients seen for preoperative risk assessment at a single health system, decisions to refer patients for preoperative stress testing are influenced by various factors other than estimated perioperative risk and functional status, the key considerations in current guidelines. The frequency of preoperative stress testing has decreased over time, but remains highly provider-dependent.

Strengths and limitations of this study:

- We identified a large cohort of patients considering noncardiac surgery, with detailed clinical data from each visit.
- We tested predictor variables across various constructs potentially related to preoperative stress testing.
- We accounted for clustering by physician and patient, testing different structures in an effort to optimally partition variance.
- We used multiple imputation by chained equations to mitigate potential biases from missing data.

Glossary of Abbreviations:

ACC/AHA: American College of Cardiology/American Heart Association
ASA: American Society of Anesthesiologists
CAD: coronary artery disease
METs: Estimated metabolic equivalents
MICA: Myocardial infarction or cardiac arrest calculator
RCRI: Revised cardiac risk index

Introduction

The 2014 American College of Cardiology/American Heart Association (ACC/AHA) guidelines recommend preoperative stress testing for patients whose predicted risk of a major adverse cardiac event exceeds 1% and whose functional status is poor or unknown, when results from stress testing would change clinical management.[1]

However, clinicians use various risk prediction tools, which identify different patients as having elevated risk.[2,3] Additionally, multiple methods of assessing functional status are used, which again can lead to variation in patients selected for stress testing.[2,4–7] Thus, the final decision to proceed with stress testing can become something closer to a provider-level judgment than a guideline-driven protocol.[8] Variation in use of stress testing can have substantial cost implications and potentially prompt subsequent tests and procedures with little clinical benefit.[9]

To understand contemporary use and drivers of preoperative cardiac stress testing, we sought to describe variation and predictors of preoperative stress testing using rich clinical data from a large integrated health system.

Methods

The Internal Medicine Preoperative Assessment, Consultation and Treatment (IMPACT) Center assesses patients prior to noncardiac surgery at the Cleveland Clinic. In the years

from 2008 through 2018, 118,552 patients were seen in our clinic by 104 physicians across 159,795 visits. Among this cohort, we identified scheduled and completed preoperative cardiac stress tests, here defined as those within the 30 days after a clinic visit and before noncardiac surgery. This study was approved by the Cleveland Clinic Institutional Review Board (IRB #18-1076).

Natural language processing was performed in python (version 3.7.8) using regular expressions and the spacy.io library (version 2.3.2). All analyses were performed in Stata (version 14; College Station, TX). Data used are from our electronic health record and are not available for outside access.

Predictor variables

We theorized that six underlying constructs would be related to stress test ordering: predicted perioperative risk, functional status, social and financial support, medical comorbidities, physician tendencies and experience, and time. We created a random effects logit model for each construct to refine variables included in our final model. Within each submodel, we pruned variables according to Bayesian information criteria (BIC). For continuous variables, we assessed for nonlinear or categorical relationships using visual examination of binned scatter plots. To avoid overfitting, we limited candidate predictors to fewer than 1 predictor variable per 15 preoperative stress tests, including tested interactions, nonlinear effects, and discarded predictors. We estimated that we had approximately 539 degrees of freedom for analysis, with fewer for cluster-level variables depending on model structure (described below).

For measures of perioperative risk, we tested Revised Cardiac Risk Index (RCRI), Myocardial Infarction or Cardiac Arrest (MICA), and MICA's categorization of surgeries using a previously-published crosswalk.[2,10,11] (Although different procedures likely have different intrinsic cardiac risk, we used the MICA categorization of surgeries to avoid overfitting.[11,12]) Upon finding that few surgical categories were associated with different stress testing rates, we replaced that multinomial variable with indicator variables for each category associated with different testing rates in our data (aortic, peripheral vascular, and urologic surgeries). As a separate sensitivity analysis, we excluded patients seen in advance of cardiac or vascular surgery, who are generally evaluated elsewhere in our institution. We tested both documented and calculated RCRI, which may differ for a variety of reasons including lab results between the clinic visit and documentation, erroneous diagnoses/chart lore, outside records unavailable in the electronic medical record, and misconceptions about how RCRI is calculated. We treated both estimates of RCRI as continuous to force a monotonic relationship (no theory would support lower testing rates at higher predicted cardiac risk). We tested MICA-predicted probability both as a continuous variable and dichotomized at 1%. Although we used the American Society of Anesthesiologists (ASA) physical status to calculate MICA, we did not test that separately because it is assigned at the time of surgery.

For measures of functional status, we tested estimated metabolic equivalents (METs), which in this clinic is based on a semi-quantitative questionnaire, and the physician's subjective global assessment of function, which is comparable to the Eastern Cooperative Oncology Group (ECOG) score.[13] For measures of social and financial support, we tested area deprivation index (a measure of socioeconomic disadvantage based on education, employment, housing quality, and poverty measures by census block group), race, ethnicity, marital status, and age (here dichotomized at age 65 to reflect changes in access to care with universal Medicare eligibility).[14] For measures of medical comorbidities and illness, we considered age, vital signs at the clinic visit, diagnoses of coronary artery disease, cerebrovascular disease, or congestive heart failure, diabetes, use of insulin, creatinine, tobacco use, and predicted probability of obstructive coronary artery disease.[15]

To accurately capture the predicted probability of obstructive coronary artery disease among patients without an existing diagnosis, we applied natural language processing to extract three pain characteristics from the full-text clinic note, when patients were documented to have chest pain: (1) substernal, (2) provoked by exertion, and (3) relieved by either rest or nitroglycerin. Notes with all three chest pain criteria were considered to document "typical" chest pain, notes with two criteria were considered to document "atypical" chest pain, and other documentation of chest pain was considered to represent non-specific chest pain.[15] For visits without documented chest pain, we estimated pretest probability as though patients had non-specific chest pain. We used an interaction with our variable for a previous diagnosis of coronary artery disease, such that this estimated probability was considered a predictor only when patients did not have an existing diagnosis.

For measures of physician tendencies and experience, we tested (on the date of each visit) years of post-residency practice (a proxy for overall experience) and the number of previous encounters the physician had completed in our dataset (a proxy for experience in preoperative risk assessment more specifically). For measures of time, we theorized that patients who had previous cardiac stress tests would be less likely to be referred for preoperative stress testing, and that physicians would give greater weight to more recent tests (i.e., the relationship would be time-dependent). We used the date of the visit as a continuous variable to test for changing stress test rates over time. To assess for changes related to publication of the current ACC/AHA guideline, we created a dichotomous variable for whether the visit occurred before or after said guideline's December 9, 2014 publication, and tested for interactions between that term and other predictors.[1]

Because unstable angina would be a potential indication for cardiac testing regardless of upcoming surgery, we investigated the frequency of angina in a subset of notes between June 2013 and July 2016; an EHR template used during this period included a structured questionnaire of symptoms. One such symptom was "Angina within 30 days".

Model structure

Conceptually, visits could be thought of as clustered by patient (with physician-level variables at the level of the visit) or by physician (with patient-level variables at the level of the visit). We tested different structures using empty models and calculated intraclass correlation coefficients to estimate the proportion of variance explained by unmeasured patient-level or physician-level factors.

With visits clustered by physician, approximately 0.4% of variance in stress test ordering was at the level of the physician. When we clustered visits by patient, approximately 4.9% of variance in stress test ordering was at the level of the patient. We therefore developed our model using physician- and visit-level variables clustered by patient, including physician ID as a visit-level indicator ("dummy") variable. This approach drops some low volume providers for whom outcomes are overfitted but should capture more unmeasured variance at both the patient and provider levels.

Multivariable modeling

Finally, we added the remaining predictor variables from each submodel into a multivariable logistic regression model. We again pruned predictors based on BIC and examined for nonlinear or categorical relationships. We revisited our model structure using the final predictor variables, comparing models clustered by patient or physician based on BIC.

Because results of a multilevel logistic regression with interaction terms have limited intuitive meaning, we calculated marginal effects for reporting. Holding all other variables at their medians, we estimated the effect of changing one predictor variable at a time.

Data extraction and missingness

Our methods for extracting data from the electronic medical record have been described previously.[16] We considered patients as having each considered diagnosis if it had been documented at any time before or at the analyzed visit. For creatinine and other lab testing, we used the most recent measurement up to and including the day of the clinic visit. We used multiple imputation by chained equations to address missing data and previously described standards to ensure multiple imputation did not introduce Monte Carlo error.[17] We imputed predictor variables other than those with negligible missing data: age, sex, vital signs, and binary variables indicating previous diagnoses.

Patient and Public Involvement

Patients were not involved in the design of this study.

Results

Overall, 5.2% of visits to the preoperative clinic led to a cancelled or completed preoperative stress test (8,303/159,795), with 5.1% (8,085; 97.4% of those referred) completing the test. Patient demographics, selected risk factors, and proportions of missing data are shown in Tables 1 and 2. Unadjusted physician referral rates are shown in Figure 1.

Marginal testing rates across each predictor variable, with other variables at their respective medians, are shown in Table 3. In general, patients were more likely to be referred for preoperative stress testing as estimated perioperative risk increased and for specific categories of surgeries (aortic, peripheral vascular, and urologic). Of those, patients undergoing aortic surgery were most likely to be referred for stress testing, though our dataset included relatively few aortic surgeries and confidence intervals for that predictor were wide. Even after adjusting for all other factors, different providers were more likely to refer for preoperative stress testing than others: a visit to the 95th percentile physician in this clinic would result in stress testing 3.8% of the time, while a visit to the 5th percentile physician in this clinic would result in testing around 1.2% of the time.

Other important patient variables included the physician's subjective assessment of global patient function, METs, socioeconomic advantage or disadvantage compared to the median, BMI, diastolic blood pressure, existing diagnoses of ischemic heart disease or congestive heart failure, estimated probability of obstructive coronary artery disease, and tobacco use. Visits later in our dataset were less likely to result in a preoperative stress test compared with earlier visits. Each of these variables, while significant, appeared to exert less influence than surgical categories, estimated surgical risk, and provider. Fully adjusted provider marginal rates are shown in Figure 2.

Results were very similar for models clustered by patient or physician. Information criteria would slightly favor a model clustered by physician (BIC: 51040) compared to a model clustered by patient with a physician indicator variable (BIC: 52009). Meanwhile, the model clustered by physician had a slightly lower R^2 (0.1896) compared with a model clustered by patient with physician as an indicator variable (0.1907).

Between June 2013 and July 2016, 23,034 visits used our EHR template that included structured entry of “Angina within 30 days”. Of those, 48 (0.2%) were marked as “Yes”. Of 107 other visits flagged by natural language processing as potentially including unstable angina, manual chart review of a random sample of 50 visits showed 3 negations missed by natural language processing, 32 descriptions of historical symptoms that had prompted testing or intervention previously, 2 quotations of test reports (coronary catheterization reports that included unstable angina as the indication for testing), and 13 cases of unstable angina.

Discussion

In this cohort of patients seen for preoperative risk assessment at a single health system, we have identified key drivers of preoperative stress testing, which include type of surgery, estimated surgical risk, and patient functional status. Our results demonstrate use of preoperative stress testing in a real-world cohort.

Current guidelines recommend preoperative stress testing for patients whose predicted perioperative adverse cardiac event risk exceeds 1% and whose functional capacity is poor or unknown, when such testing would change clinical management. Although we cannot determine from these data whether physicians thought testing would change management, and predicted surgical risk scores have poor concordance across the 1% threshold, patients able to perform four or more metabolic equivalents of activity made up nearly one-third of all stress test referrals.[2,10,18] Our data suggest that a substantial fraction of preoperative stress tests were inconsistent with current guidance.

Predicted surgical risk was nonetheless a key driver of preoperative stress testing. Testing rates increased with increasing RCRI, without a clear dichotomization at any particular value of RCRI. And interestingly, although MICA was essentially never documented, a MICA-predicted surgical risk of greater than 1% appears to be a better single predictor variable than RCRI. Physicians could be trying to incorporate the guideline-recommended threshold into their decision-making while relying on cohorts with different calibration, or could be deliberately avoiding a stark dichotomization of risk at 1%.[10,18]

As with predicted surgical risk, physicians appeared to consider functional status as something between the dichotomy of current guidance and continuum of risk encountered in clinical practice. Testing rates were higher among less functional patients and lower among patients able to achieve higher METs, but neither were especially important predictors in our model. It seems probable that both clinical skill and clinical decision-making are more nuanced than we can discern from our data source. For example, physicians likely vary in both their ability to elicit anginal equivalents and their interpretation of potentially ambiguous symptoms. And even if they were presented with equivalent information, various physicians might reasonably make different decisions about testing based on factors we are not able to investigate. For example, physicians could reasonably be more inclined to test before a pancreaticoduodenectomy than a laparoscopic cholecystectomy in view of those procedures' very different metabolic demands, but we lack sufficient power to test individual surgical procedures without overfitting.[12] In any case, functional status and patient variables other than predicted perioperative risk explained little variance in testing rates.

Surgical category also offers insights into testing rationale. Patients were tested more frequently before aortic or peripheral vascular interventions, perhaps reflecting persistent beliefs that patients with coronary artery disease should be identified and revascularized before vascular surgery.[19–21] However, patients undergoing vascular surgery are generally evaluated elsewhere at our institution, leaving our sample small and confidence intervals wide. We also note that patients are more likely to be referred for stress testing before urologic surgery, after controlling for patient risk factors and despite the fact that urologic procedures are not associated with higher intrinsic cardiac risk than other common surgical categories. Anecdotally, physicians practicing in this clinic have reported that a number of urologists at our institution are reluctant to operate on high-risk patients unless those patients first undergo preoperative testing. While investigating such a hypothesis would require a different approach than ours, clearly every physician in a preoperative clinic functions within a larger system of care and must build consensus among a team of treating physicians.

Other significant predictor variables include tobacco use, BMI, diastolic blood pressure, ischemic heart disease, and a patient's census block group. Because each of these are correlated with risk of obstructive coronary artery disease, one possibility is that data unavailable to us (such as outside records) led to some portion of the preoperative stress testing we observed. Although that remains possible, multiple observations argue against a simple explanation that these variables are proxies for coronary disease risk. First, higher probability of obstructive coronary artery disease, calculated based on available data, was associated with lower likelihood of preoperative testing. Second, patients residing in the wealthiest and poorest census tracts were approximately as likely to be referred for stress testing, with patients in the middle of the socioeconomic range less likely. Finally, diabetes was not associated with testing. It would seem that either physicians in our dataset did not incorporate patients' pretest probability of obstructive coronary disease in their decision to refer for testing, or that their assessments were poorly calibrated.

Angina or its equivalents do not appear to be a frequent rationale for testing in this cohort. Around 0.2% of notes that used a templated review of pertinent symptoms noted angina within 30 days, parsing of free text notes did not identify unstable angina with appreciable frequency, and many cases identified through natural language processing appeared not to be unstable angina on manual chart review of sampled visits. Although this could represent a failure to document findings that were present during the visit, it would seem more likely that a preoperative visit before elective noncardiac surgery is an inefficient tool to screen for angina.

The frequency of stress testing declined over time in our dataset, in contrast with increasing testing rates suggested in other contexts.[22,23] A recent cross-sectional analysis of claims data from patients who had total hip or knee arthroplasty also identified decreasing testing frequency over a similar period.[24] Our cohort study begins with a visit to a preoperative risk assessment clinic, incorporates detailed clinical data, and is not limited to patients who have completed orthopedic surgery. Our analysis thus extends previous understanding by showing that the reduction is not limited to orthopedic procedures, not a result of selecting patients not referred for stress testing for elective surgery, and not consequent to lower predicted cardiac risk. Taken together, these two analyses with different limitations suggest a shift in practice away from preoperative cardiac stress testing. Neither analysis suggests a clear change in testing frequency after the release of current guidelines. Although there can certainly be time lags between publication and consequent practice change, our findings argue against a causal relationship between publication of the current guideline and near-term changes in testing rates.[25]

Our model demonstrates physician practice variation: with all other predictors held at their medians, the 95th percentile physician was around three times more likely to order preoperative stress testing than the 5th percentile physician. But we caution against using our results, or others, for profiling individual providers, which is generally a low-reliability exercise and prone to gaming.[26] Our dataset is among the largest clinical datasets of preoperative risk assessment, but true outliers are rare and most physicians are not detectably different from the mean after adjustment (see Figure 2). As with other observations of physician practice variation, ours suggests a deeper failure: that we do not yet understand how best to use preoperative stress testing to mitigate perioperative cardiac risk.[27]

Surgery carries inherent cardiac risk, and stress testing may reflect physician discomfort with the malpractice or cognitive liabilities that cardiac risk entails.[28–32] Stress testing can inform diagnosis and prognosis, but outcomes will only improve if testing results in interventions that reduce perioperative risk. Such interventions have proven elusive: preoperative revascularization did not reduce cardiac risk in the largest randomized trial to date, beta blockers are more likely harmful than helpful, and other interventions (e.g., statins) that may be allocated differently based on stress testing likely have modest effects, if any.[33–36] Intraoperative care or postoperative testing patterns could differ based on whether a preoperative stress test was performed, but what practices in those settings might reduce the risk of major adverse cardiac events remain equally unclear. While estimated perioperative cardiac risk appears to drive stress testing, it remains to be seen how stress testing might reduce perioperative cardiac risk.

Although we have made every effort to ensure the internal validity of our data, analysis, and results, our data may not adequately represent drivers of or variation in preoperative stress testing before some common types of surgery, including ophthalmologic surgery, which is evaluated elsewhere in our institution.[37] Our IMPACT clinic has made substantial efforts to provide uniform care, which could have reduced physician variation in our dataset, and we cannot analyze variation across health system or region, which can also be substantial.[38] As with any single-center study, results should be extrapolated to other settings with caution. For example, we observed higher testing frequency before urologic surgery than would be expected for cardiac risk; other centers may have different surgical categories with testing out of proportion to surgical risk. The need to build a consensus plan of care among a treatment team is true across institutions, but the particulars of our institution's consensus may not be.

Additionally, as with other observational studies, our analytical choices are difficult to separate from our theoretical framework, and may influence our results in various ways.[39] For example, we rejected physician experience as a predictor of testing in favor of a random effect for each physician and the date of each visit due to our prespecified analytic criteria (rejecting predictor variables that worsened AIC and BIC). Still, experience differs by physician and necessarily accrues over time. A reasonable investigator with a different theoretical model could assume broadly stable testing rates over time and conclude that testing decreases as physicians gain experience (see Supplemental Appendix). Time in particular is rife with potential confounders of this sort. Recent work using other datasets also identified reduced stress testing over time (described above), offering reassurance that our analytic criteria led to the best conclusion. Still, the effect of experience on appropriate testing could be an avenue for further investigation.

But while the limitations of our study reflect the limitations of any single center observational study, the detailed clinical data available to us offers distinct advantages over earlier work. We have demonstrated real-world use of preoperative stress testing before a wide range of possible surgical interventions, using visit-level data to comprehensively assess variation in and predictors of preoperative cardiac stress testing.

In summary, the frequency of preoperative stress testing varied with estimated surgical risk, patient functional status, socioeconomic status, ischemic heart disease, congestive heart failure, body mass index, diastolic blood pressure, surgical category, and provider. The fraction of patients referred for stress testing appears to be declining over time, but testing remains common and highly dependent on the provider. The value of preoperative stress testing remains to be established.

Funding Sources:

This work was funded by NHLBI 1K08HL141598.

Conflict of Interest Disclosures:

None

Author Contributions:

Dr. Pappas: This author conceived the study, collected the data, performed the analysis, and wrote the initial draft of the manuscript.
Dr. Sessler: This author helped improve the analysis and revised the manuscript.
Dr. Auerbach: This author helped improve the analysis and revised the manuscript.
Dr. Kattan: This author helped improve the analysis and revised the manuscript.
Mr. Milinovich: This author helped obtain and validate data.
Dr. Blackstone: This author helped improve the analysis and revised the manuscript.
Dr. Rothberg: This author helped improve the analysis and revised the manuscript.

Ethical Approval:

This study was approved by the Cleveland Clinic Institutional Review Board (IRB #18-1076).

Data Availability Statement

Data used are from our electronic health record and are not available for outside access.

References

1. Fleisher LA, Fleischmann KE, Auerbach AD, et al. 2014 ACC/AHA guideline on perioperative cardiovascular evaluation and management of patients undergoing noncardiac surgery: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;130(24):2215-2245. doi:10.1161/CIR.0000000000000105

2. Pappas MA, Sessler DI, Rothberg MB. Anticipated Rates and Costs of Guideline-Concordant Preoperative Stress Testing. *Anesth Analg*. 2019;128(2):241-246. doi:10.1213/ANE.0000000000003754

3. Glance LG, Faden E, Dutton RP, et al. Impact of the choice of risk model for identifying low-risk patients using the 2014 American college of cardiology/American Heart Association perioperative guidelines. *Anesthesiology*. 2018;129(5):889-900. doi:10.1097/ALN.0000000000002341

4. Hlatky MA, Boineau RE, Higginbotham MB, et al. A brief self-administered questionnaire to determine functional capacity (The Duke Activity Status Index). *Am J Cardiol*. 1989;64(10):651-654. doi:10.1016/0002-9149(89)90496-7

5. Reilly DF, McNeely MJ, Doerner D, et al. Self-reported exercise tolerance and the risk of serious perioperative complications. *Arch Intern Med*. 1999;159(18):2185-2192. doi:10.1001/archinte.159.18.2185

6. Melon CC, Eshtiaghi P, Luksun WJ, Wijeysundera DN. Validated questionnaire vs physicians' judgment to estimate preoperative exercise capacity. *JAMA Intern Med*.

2014;174(9):1507-1508. doi:10.1001/jamainternmed.2014.2914

7. Wijeysundera DN, Pearse RM, Shulman MA, et al. Assessment of functional capacity before major non-cardiac surgery: an international, prospective cohort study. *Lancet*. 2018;391(10140):2631-2640. doi:10.1016/S0140-6736(18)31131-0

8. Eddy DM. Variations in physician practice: The role of uncertainty. *Health Aff*. 1984;3(2):74-89. doi:10.1377/hlthaff.3.2.74

9. Mark DB, Federspiel JJ, Cowper PA, et al. Economic Outcomes With Anatomical Versus Functional Diagnostic Testing for Coronary Artery Disease. *Ann Intern Med*. 2016;165(2):94-102. doi:10.7326/M15-2639

10. Lee TH, Marcantonio ER, Mangione CM, et al. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. *Circulation*. 1999;100(10):1043-1049. doi:10.1161/01.CIR.100.10.1043

11. Gupta PK, Gupta H, Sundaram A, et al. Development and validation of a risk calculator for prediction of cardiac risk after surgery. *Circulation*. 2011;124(4):381-387. doi:10.1161/CIRCULATIONAHA.110.015701

12. Liu JB, Liu Y, Cohen ME, Ko CY, Sweitzer BJ. Defining the Intrinsic Cardiac Risks of Operations to Improve Preoperative Cardiac Risk Assessments. *Anesthesiology*. 2018;128(2):283-292. doi:10.1097/ALN.0000000000002024

13. Oken MM, Creech RH, Davis TE. Toxicology and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol*. 1982;5(6):649-655. doi:10.1097/00000421-198212000-00014

14. Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible - The neighborhood atlas. *N Engl J Med*. 2018;378(26):2456-2458. doi:10.1056/NEJMp1802313

15. Genders TSS, Steyerberg EW, Hunink MGM, et al. Prediction model to estimate presence of coronary artery disease: Retrospective pooled analysis of existing cohorts. *BMJ*. 2012;344(7862). doi:10.1136/bmj.e3485

16. Milinovich A, Kattan MW. Extracting and utilizing electronic health data from Epic for research. *Ann Transl Med*. 2018;6(3):42-42. doi:10.21037/atm.2018.01.13

17. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30(4):377-399. doi:10.1002/sim.4067

18. Duceppe E, Parlow J, MacDonald P, et al. Canadian Cardiovascular Society Guidelines on Perioperative Cardiac Risk Assessment and Management for Patients Who Undergo Noncardiac Surgery. *Can J Cardiol*. 2017;33(1):17-32. doi:10.1016/j.cjca.2016.09.008

19. Eagle KA, Singer DE, Brewster DC, Darling RC, Mulley AG, Boucher CA. Dipyridamole-thallium scanning in patients undergoing vascular surgery. Optimizing preoperative evaluation of cardiac risk. *JAMA*. 1987;257(16):2185-2189. http://www.ncbi.nlm.nih.gov/pubmed/3560400. Accessed November 11, 2019.

20. Hertzer NR, Beven EG, Young JR, et al. Coronary artery disease in peripheral vascular patients. A classification of 1000 coronary angiograms and results of surgical management. *Ann Surg*. 1984;199(2):223-233. doi:10.1097/00000658-

198402000-00016

21. Boucher CA, Brewster DC, Darling RC, Okada RD, Strauss HW, Pohost GM. Determination of Cardiac Risk by Dipyridamole-Thallium Imaging before Peripheral Vascular Surgery. *N Engl J Med*. 1985;312(7):389-394. doi:10.1056/NEJM198502143120701

22. Sheffield KM, McAdams PS, Benarroch-Gampel J, et al. Overuse of preoperative cardiac stress testing in medicare patients undergoing elective noncardiac surgery. *Ann Surg*. 2013;257(1):73-80. doi:10.1097/SLA.0b013e31826bc2f4

23. Sigmund AE, Stevens ER, Blitz JD, Ladapo JA. Use of preoperative testing and physicians' response to professional society guidance. *JAMA Intern Med*. 2015;175(8):1352-1359. doi:10.1001/jamainternmed.2015.2081

24. Rubin DS, Hughey R, Gerlach RM, Ham SA, Ward RP, Nagele P. Frequency and Outcomes of Preoperative Stress Testing in Total Hip and Knee Arthroplasty from 2004 to 2017. *JAMA Cardiol*. 2021;6(1):13-20. doi:10.1001/jamacardio.2020.4311

25. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: The National Academies Press; 2001. doi:10.17226/10027

26. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *J Am Med Assoc*. 1999;281(22):2098-2105. doi:10.1001/jama.281.22.2098

27. Wennberg JE, Freeman JL, Culp WJ. Are Hospital Services Rationed in New Haven

or Over-Utilized in Boston? *Lancet*. 1987;329(8543):1185-1189. doi:10.1016/S0140-6736(87)92152-0

28. Saint S, Vaughn VM, Chopra V, Fowler KE, Kachalia A. Perception of Resources Spent on Defensive Medicine and History of Being Sued Among Hospitalists: Results from a National Survey. *J Hosp Med*. 2018;13(1):26-29. doi:10.12788/jhm.2800

29. Feinstein AR. The 'Chagrin Factor' and Qualitative Decision Analysis. *Arch Intern Med*. 1985;145(7):1257-1259. doi:10.1001/archinte.1985.00360070137023

30. Rolfe A, Burton C. Reassurance after Diagnostic Testing with a Low Pretest Probability of Serious Disease: Systematic Review and Meta-analysis. *JAMA Intern Med*. 2013;173(6):407-416. doi:10.1001/jamainternmed.2013.2762

31. Van der Weijden T, Van Velsen M, Dinant GJ, Van Hasselt CM, Grol R. Unexplained complaints in general practice: Prevalence, patients' expectations, and professionals' test-ordering behavior. *Med Decis Mak*. 2003;23(3):226-231. doi:10.1177/0272989X03023003004

32. Little P, Dorward M, Warner G, Stephens K, Senior J, Moore M. Importance of patient pressure and perceived pressure and perceived medical need for investigations, referral, and prescribing in primary care: Nested observational study. *Br Med J*. 2004;328(7437):444-446. doi:10.1136/bmj.38013.644086.7c

33. McFalls EO, Ward HB, Moritz TE, et al. Coronary-Artery Revascularization before Elective Major Vascular Surgery. *N Engl J Med*. 2004;351(27):2795-2804. doi:10.1056/NEJMoa041905

34. Blessberger H, Kammler J, Domanovits H, et al. Perioperative beta-blockers for preventing surgery-related mortality and morbidity. *Cochrane Database Syst Rev*. March 2018. doi:10.1002/14651858.CD004476.pub3

35. Berwanger O, Manach Y Le, Suzumura EA, et al. Association between pre-operative statin use and major cardiovascular complications among patients undergoing non-cardiac surgery: The VISION study. *Eur Heart J*. 2016;37(2):177-185. doi:10.1093/eurheartj/ehv456

36. Berwanger O, de Barros e Silva PGM, Barbosa RR, et al. Atorvastatin for high-risk statin-naïve patients undergoing noncardiac surgery: The Lowering the Risk of Operative Complications Using Atorvastatin Loading Dose (LOAD) randomized trial. *Am Heart J*. 2017;184:88-96. doi:10.1016/j.ahj.2016.11.001

37. Kerr EA, Chen J, Sussman JB, Klamerus ML, Nallamothu BK. Stress Testing Before Low-Risk Surgery: So Many Recommendations, So Little Overuse. *JAMA Intern Med*. 2015;175(4):645. doi:10.1001/jamainternmed.2014.7877

38. Home - Dartmouth Atlas of Health Care. https://www.dartmouthatlas.org/. Accessed November 11, 2019.

39. Palpacuer C, Hammas K, Duprez R, Laviolle B, Ioannidis JPA, Naudet F. Vibration of effects from diverse inclusion/exclusion criteria and analytical choices: 9216 different ways to perform an indirect comparison meta-analysis. *BMC Med*. 2019;17(1):174. doi:10.1186/s12916-019-1409-3

29

Table 1. Patient and Surgical Characteristics.

			Total	Percent of category	Percent of all visits	Completed preoperative stress test	Percent of all preoperative stress tests
Age			159,795	100.0%	100.0%	8,303	100.0%
Sex	Female		88,738	55.5%	55.5%	4,079	49.1%
	Male		71,055	44.5%	44.5%	4,224	50.9%
Previous diagnosis of ischemic heart disease	No		128,505	80.4%	80.4%	4,831	58.2%
	Yes		31,290	19.6%	19.6%	3,472	41.8%
Previous diagnosis of congestive heart failure	No		146,556	91.7%	91.7%	6,922	83.4%
	Yes		13,239	8.3%	8.3%	1,381	16.6%
Previous diagnosis of cerebrovascular disease	No		141,519	88.6%	88.6%	6,567	79.1%
	Yes		18,276	11.4%	11.4%	1,736	20.9%
Systolic Blood Pressure			159,488	100.0%	99.8%	8,285	99.8%
Diastolic Blood Pressure			159,481	100.0%	99.8%	8,284	99.8%
Body Mass Index			157,473	100.0%	98.5%	8,155	98.2%
Creatinine (RCRI categorization)	≤ 2.0 mg/dL		151,885	97.1%	95.1%	7,695	92.7%
	> 2.0 mg/dL		4,487	2.9%	2.8%	542	6.5%
Creatinine (MICA categorization)	≤ 1.5 mg/dL		144,369	90.3%	90.3%	7,128	85.8%
	> 1.5 mg/dL		12,003	7.5%	7.5%	1,109	13.4%
	Unknown		3,423	2.1%	2.1%	66	0.8%
Prescribed insulin	No		147,610	92.4%	92.4%	7,136	85.9%
	Yes		12,185	7.6%	7.6%	1,167	14.1%
RCRI surgical category	High risk		27,709	23.8%	17.3%	872	10.5%
	Other		88,929	76.2%	55.7%	360	4.3%
Area Deprivation Index			126,076	100.0%	78.9%	7,091	85.4%
RCRI (documented)		0	50,785	75.1%	31.8%	1,548	18.6%
		1	12,642	18.7%	7.9%	988	11.9%
		2	3,321	4.9%	2.1%	402	4.8%
		3	742	1.1%	0.5%	108	1.3%
		4	151	0.2%	0.1%	21	0.3%
		5	15	0.0%	0.0%	2	0.0%
RCRI (calculated)		0	56,879	52.0%	35.6%	355	4.3%

	1	36,020	32.9%	22.5%	393	4.7%
	2	11,143	10.2%	7.0%	245	3.0%
	3	4,006	3.7%	2.5%	98	1.2%
	4	1,172	1.1%	0.7%	33	0.4%
	5	204	0.2%	0.1%	7	0.1%
	6	12	0.0%	0.0%	0	0.0%
MICA risk estimate	≤ 1%	71,448	44.7%	44.7%	608	7.3%
	> 1%	88,347	55.3%	55.3%	7,695	92.7%
Physician subjective assessment of patient global function	1	17,991	19.1%	11.3%	489	5.9%
	1-2	13,386	14.2%	8.4%	634	7.6%
	2	40,829	43.4%	25.6%	1,739	20.9%
	2-3	8,588	9.1%	5.4%	646	7.8%
	3	9,310	9.9%	5.8%	732	8.8%
	3-4	1,922	2.0%	1.2%	253	3.0%
	4	1,999	2.1%	1.3%	175	2.1%
Insurance	Medicare	20,744	52.2%	13.0%	1,192	14.4%
	Medicaid	2,384	6.0%	1.5%	83	1.0%
	Private	14,764	37.1%	9.2%	307	3.7%
	Other listed insurer	1,881	4.7%	1.2%	53	0.6%
ASA Class	1	3,325	3.0%	2.1%	9	0.1%
	2	34,026	30.9%	21.3%	181	2.2%
	3	65,298	59.3%	40.9%	854	10.3%
	4	7,454	6.8%	4.7%	157	1.9%
MICA surgical category	Anorectal	1,213	1.0%	0.8%	3	0.0%
	Aortic	115	0.1%	0.1%	34	0.4%
	Bariatric	703	0.6%	0.4%	6	0.1%
	Brain	4,780	4.1%	3.0%	15	0.2%
	Breast	8,541	7.4%	5.3%	17	0.2%
	Cardiac	237	0.2%	0.1%	10	0.1%
	Ear, nose, throat	3,640	3.1%	2.3%	19	0.2%
	Foregut/hepatopancreatobiliary	3,252	2.8%	2.0%	37	0.4%
	Gallbladder, appendix, adrenals, or spleen	1,974	1.7%	1.2%	17	0.2%
	Gynecologic	7,458	6.4%	4.7%	51	0.6%
	Hernia	2,778	2.4%	1.7%	16	0.2%

	Intestines	16,081	13.9%	10.1%	87	1.0%
	Neck	3,076	2.7%	1.9%	18	0.2%
	Nonesophageal thoracic	290	0.2%	0.2%	4	0.0%
	Orthopedic	26,005	22.4%	16.3%	208	2.5%
	Other abdomen	1,970	1.7%	1.2%	19	0.2%
	Peripheral vascular	1,355	1.2%	0.8%	110	1.3%
	Skin	7,126	6.1%	4.5%	78	0.9%
	Spinal	11,994	10.3%	7.5%	59	0.7%
	Urologic	13,423	11.6%	8.4%	414	5.0%
	Vein	48	0.0%	0.0%	3	0.0%
Before or after release of current guideline	Before	98,465	61.6%	61.6%	5,911	71.2%
	After	61,330	38.4%	38.4%	2,392	28.8%
Able to perform activities of at least 4 METs	No	90,260	56.5%	56.5%	5,713	68.8%
	Yes	69,535	43.5%	43.5%	2,590	31.2%
Tobacco use	Current smoker	18,806	12.4%	11.8%	1,028	12.4%
	Former smoker	62,067	40.9%	38.8%	3,834	46.2%
	Never smoker	70,966	46.7%	44.4%	2,976	35.8%
Probability of obstructive CAD		159,793	100.0%	100.0%	8,303	100.0%

Table 2: Summary of continuous variables.

	Mean ± SD
Age (years)	58.8 ± 15.2
Systolic Blood Pressure	129 ± 19.1
Diastolic Blood Pressure	73 ± 11.3
Body Mass Index	30.1 ± 7.6
Creatinine	1.02 ± 0.85
Area Deprivation Index	54.2 ± 24.6
Estimated METs of activity	5.22 ± 1.35
MICA risk estimate	0.021 ± 0.025
Previous patients seen by physician in clinic	1,587 ± 1,536
Predicted probability of obstructive CAD	0.129 ± 0.136

Table 3. Marginal results for each variable in our final model, with all other variables held at their medians.

For example, with all other variables at their respective medians, a visit on June 30, 2008 would have resulted in preoperative stress testing approximately 3.5% of the time, while a visit on June 30, 2018 would have resulted in preoperative stress testing approximately 1.3% of the time. Provider effects are summarized for space considerations; full marginal results by physician are included in the Supplemental Appendix.

Predictor	Value	Marginal Rate	95% CI	
MICA estimate > 1%	0	0.6%	0.6%	0.7%
	1	7.1%	6.7%	7.5%
RCRI (documented)	0	2.2%	2.0%	2.3%
	1	2.7%	2.5%	2.9%
	2	3.4%	3.0%	3.8%
	3	4.2%	3.5%	5.0%
	4	5.1%	4.0%	6.5%
	5	6.4%	4.7%	8.6%
Subjective assessment of patient function	1	2.1%	1.9%	2.2%
	2	2.4%	2.2%	2.5%
	3	2.8%	2.6%	3.0%
	4	3.2%	2.9%	3.5%
Estimated METs	2	3.3%	2.9%	3.7%
	4	2.7%	2.5%	2.9%
	8	1.8%	1.6%	1.9%
Body mass index	20	2.1%	1.9%	2.3%
	30	2.4%	2.2%	2.5%
	40	2.7%	2.5%	2.9%
Diastolic blood pressure	70	2.3%	2.2%	2.5%
	90	2.5%	2.3%	2.7%
	110	2.7%	2.4%	3.0%
Ischemic heart disease	No	2.1%	2.0%	2.3%
	Yes	3.6%	3.3%	3.9%
Congestive heart	No	2.4%	2.2%	2.5%

failure	Yes	2.1%	1.9%	2.3%
Area deprivation index	10	2.8%	2.5%	3.1%
	50	2.2%	2.0%	2.3%
	90	2.6%	2.4%	2.8%
Predicted probability of obstructive CAD	5%	2.6%	2.4%	2.7%
	10%	2.4%	2.3%	2.6%
	20%	2.2%	2.1%	2.4%
Tobacco use	Current smoker	2.6%	2.3%	2.8%
	Former smoker	2.5%	2.3%	2.7%
	Neither	2.2%	2.1%	2.4%
Date	2008.06.30	3.5%	3.2%	3.8%
	2013.06.30	2.6%	2.4%	2.8%
	2018.06.30	1.3%	1.2%	1.4%
Surgical category	Aortic	23.4%	6.0%	91.1%
	Peripheral vascular	8.7%	6.7%	11.3%
	Urologic	9.2%	8.3%	10.2%
	Other	1.9%	1.7%	2.0%
Physician (summary)	Lowest	1.0%	0.1%	4.4%
	5th percentile	1.2%	0.6%	2.6%
	Median	2.3%	2.1%	2.6%
	95th percentile	3.8%	3.2%	4.5%
	Highest	6.1%	2.7%	13.5%

Figure 1: Unadjusted rates of preoperative stress testing, by physician.

[Attached separately as Figure_01.tif]

Figure 2: Mean marginal rates of preoperative stress testing, by physician.

[Attached separately as Figure_02.tif]

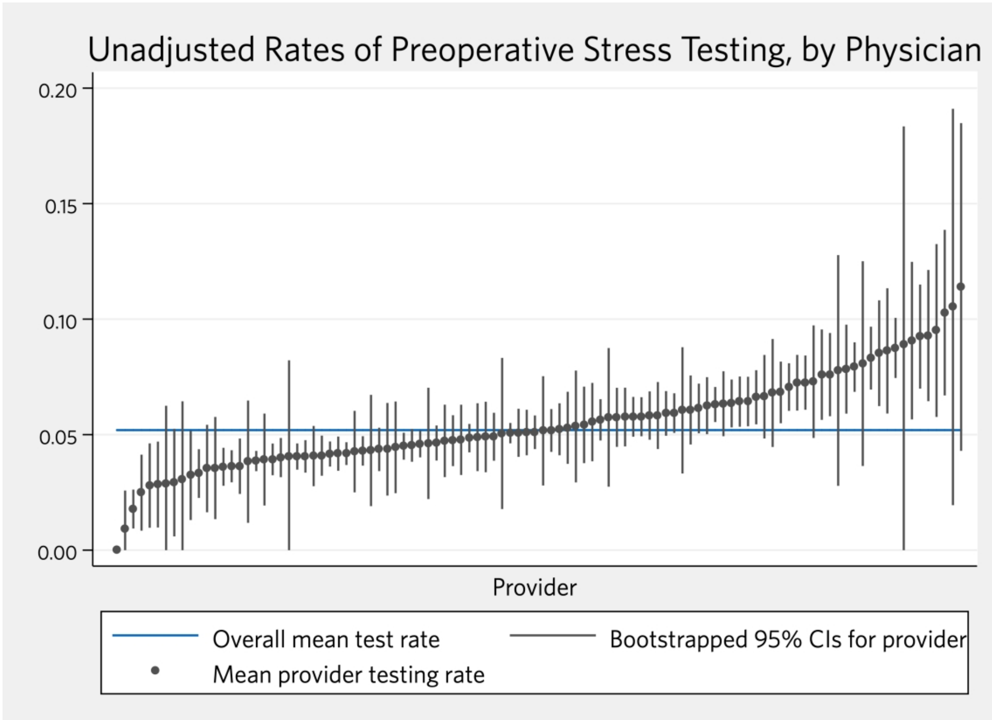


Figure 1: Unadjusted rates of preoperative stress testing, by physician

569x413mm (144 x 144 DPI)

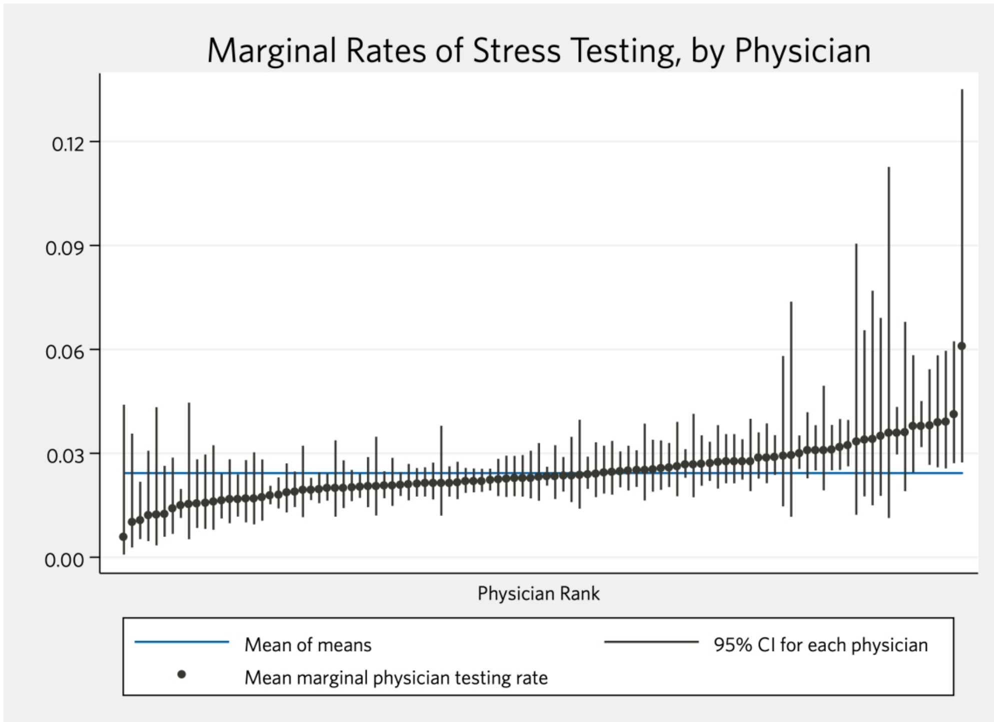


Figure 2: Mean marginal rates of preoperative stress testing, by physician

569x413mm (144 x 144 DPI)

Supplemental Material for: Variation in preoperative stress testing by patient, physician, and surgical type

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

1. Regression results, clustered by physician

Results have been edited/trimmed for clarity and brevity. Interaction terms have been replaced in the table below with exponents (^2) when applicable. All dichotomous variables (those prepended with "1." in the table) use a value of 0 as the referent.

```
Random-effects logistic regression              Number of obs     =    154,171
Group variable: Physician_ID                    Number of groups  =        104
Observations per group:                                       min =          7
                                                              avg =    1,482.4
                                                              max =      8,245
```

Completed or cancelled test	Beta Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
1.aortic_surgery	2.522167	.6078339	4.15	0.002	1.166378	3.877956
1.peripheral_vascular_surgery	1.531945	.1152814	13.29	0.000	1.285334	1.778556
1.urologic_surgery	1.5871	.0657605	24.13	0.000	1.446111	1.728088
documented_RCRI	.2135885	.03032	7.04	0.000	.1506825	.2764944
1.Gupta_greater_than_1pct	2.422218	.0472362	51.28	0.000	2.329492	2.514944
Estimated METs	-.1030793	.0128133	-8.04	0.000	-.1284818	-.0776767
functional_class	.1444786	.0214833	6.73	0.000	.1020898	.1868674
ADI_national	-.0141123	.0025089	-5.62	0.000	-.0190783	-.0091463
ADI_national^2	.0001325	.0000221	5.99	0.000	.0000888	.0001763
BMI	.0120129	.0016333	7.36	0.000	.0088076	.0152182
DBP	.0037949	.0011446	3.32	0.001	.0015493	.0060406
1.ischemic_heart_disease	.5187915	.0390133	13.30	0.000	.4412007	.5963823
predicted_prob_of_CAD	-.9799052	.1105385	-8.86	0.000	-1.198045	-.761765
1.congestive_heart_failure	-.1196815	.0476752	-2.51	0.014	-.2142709	-.0250922
date	.002128	.0004881	4.36	0.000	.0011708	.0030853
date^2	-6.15e-08	1.25e-08	-4.92	0.000	-8.61e-08	-3.70e-08
1.current_smoker	.1436825	.0424903	3.38	0.001	.0603125	.2270526
1.former_smoker	.1288456	.0290712	4.43	0.000	.0718283	.1858628
constant	-23.48573	4.748773	-4.95	0.000	-32.79838	-14.17308
/lnsig2u	-3.592461	.2658187			-4.113895	-3.071026
sigma_u	.1659232	.0220527			.1278436	.2153451
rho	.0082988	.0021877			.0049434	.0138999

2. Physician marginal rates of stress testing

Due to space constraints, only selected physician marginal results are displayed in Table 3. Here we present full marginal results by physician. These are also displayed visually in Figure 2.

Physician Rank	Marginal Rate	95% CI
1	0.58%	0.08% – 4.40%
2	1.00%	0.28% – 3.57%
3	1.07%	0.52% – 2.18%
4	1.20%	0.47% – 3.08%
5	1.22%	0.34% – 4.33%
6	1.25%	0.59% – 2.64%
7	1.39%	0.68% – 2.88%
8	1.50%	1.13% – 1.97%
9	1.52%	0.52% – 4.46%
10	1.55%	0.85% – 2.83%
11	1.56%	0.82% – 2.97%
12	1.60%	0.79% – 3.23%
13	1.64%	1.12% – 2.42%
14	1.67%	0.98% – 2.83%
15	1.68%	1.17% – 2.39%
16	1.68%	1.01% – 2.80%
17	1.70%	0.95% – 3.03%
18	1.73%	1.05% – 2.82%
19	1.78%	1.53% – 2.06%
20	1.80%	1.41% – 2.30%
21	1.87%	1.29% – 2.71%
22	1.89%	1.45% – 2.47%
23	1.93%	1.16% – 3.22%
24	1.95%	1.65% – 2.29%

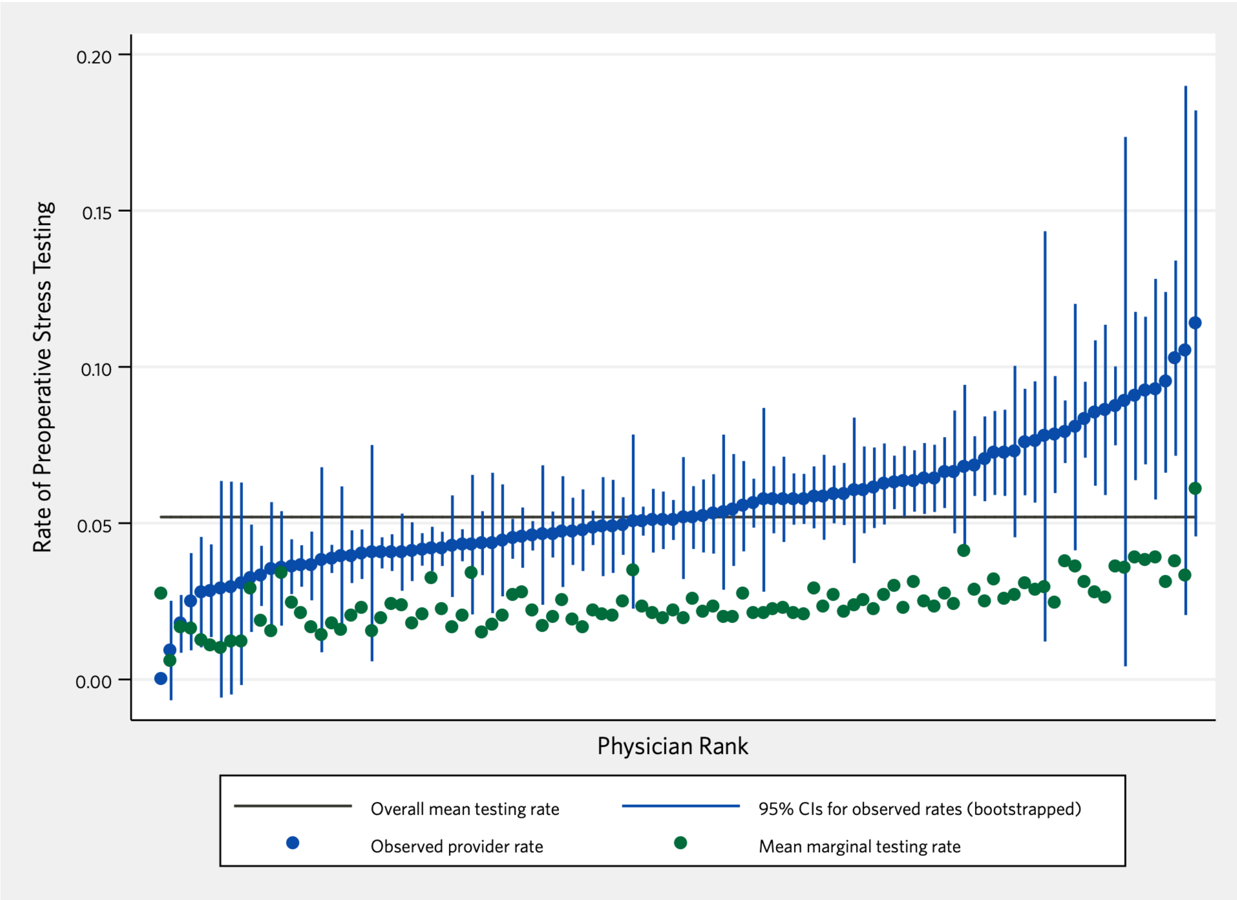
Physician Rank	Marginal Rate	95% CI
25	1.96%	1.56% – 2.45%
26	1.99%	1.63% – 2.42%
27	1.99%	1.17% – 3.38%
28	1.99%	1.42% – 2.80%
29	2.02%	1.62% – 2.52%
30	2.03%	1.71% – 2.42%
31	2.04%	1.45% – 2.89%
32	2.05%	1.20% – 3.48%
33	2.06%	1.70% – 2.48%
34	2.06%	1.48% – 2.87%
35	2.09%	1.77% – 2.46%
36	2.10%	1.64% – 2.69%
37	2.12%	1.75% – 2.57%
38	2.13%	1.75% – 2.60%
39	2.13%	1.67% – 2.73%
40	2.13%	1.20% – 3.80%
41	2.14%	1.75% – 2.62%
42	2.15%	1.68% – 2.76%
43	2.19%	1.86% – 2.58%
44	2.19%	1.87% – 2.56%
45	2.20%	1.89% – 2.55%
46	2.22%	1.93% – 2.56%
47	2.24%	1.76% – 2.84%
48	2.26%	1.74% – 2.93%
49	2.27%	1.76% – 2.93%
50	2.27%	1.75% – 2.95%
51	2.29%	1.70% – 3.08%

Physician Rank	Marginal Rate	95% CI
52	2.32%	1.63% – 3.30%
53	2.32%	2.07% – 2.61%
54	2.33%	1.68% – 3.23%
55	2.34%	1.89% – 2.89%
56	2.35%	1.59% – 3.48%
57	2.36%	1.40% – 3.97%
58	2.39%	1.97% – 2.90%
59	2.40%	1.74% – 3.32%
60	2.43%	1.84% – 3.22%
61	2.46%	1.81% – 3.36%
62	2.48%	2.02% – 3.05%
63	2.49%	1.93% – 3.22%
64	2.50%	2.03% – 3.09%
65	2.51%	1.64% – 3.86%
66	2.53%	1.89% – 3.39%
67	2.56%	1.95% – 3.37%
68	2.58%	2.03% – 3.30%
69	2.62%	1.76% – 3.91%
70	2.68%	2.29% – 3.12%
71	2.68%	1.73% – 4.14%
72	2.70%	2.07% – 3.52%
73	2.71%	2.20% – 3.33%
74	2.75%	1.98% – 3.81%
75	2.75%	2.13% – 3.55%
76	2.75%	2.13% – 3.55%
77	2.76%	2.24% – 3.41%
78	2.76%	1.91% – 4.00%

Physician Rank	Marginal Rate	95% CI		
79	2.86%	2.27%	–	3.60%
80	2.87%	2.14%	–	3.86%
81	2.89%	2.38%	–	3.52%
82	2.92%	1.47%	–	5.81%
83	2.94%	1.17%	–	7.38%
84	2.99%	2.55%	–	3.51%
85	3.09%	2.28%	–	4.19%
86	3.09%	2.50%	–	3.81%
87	3.09%	1.93%	–	4.95%
88	3.09%	2.51%	–	3.82%
89	3.17%	2.52%	–	3.99%
90	3.22%	2.62%	–	3.96%
91	3.33%	1.23%	–	9.05%
92	3.39%	1.76%	–	6.55%
93	3.40%	1.50%	–	7.69%
94	3.50%	1.77%	–	6.91%
95	3.57%	1.13%	–	11.27%
96	3.59%	2.97%	–	4.34%
97	3.60%	1.91%	–	6.79%
98	3.77%	2.44%	–	5.83%
99	3.79%	3.18%	–	4.51%
100	3.80%	2.67%	–	5.43%
101	3.89%	2.60%	–	5.83%
102	3.91%	2.56%	–	5.96%
103	4.11%	2.71%	–	6.23%
104	6.08%	2.74%	–	13.51%

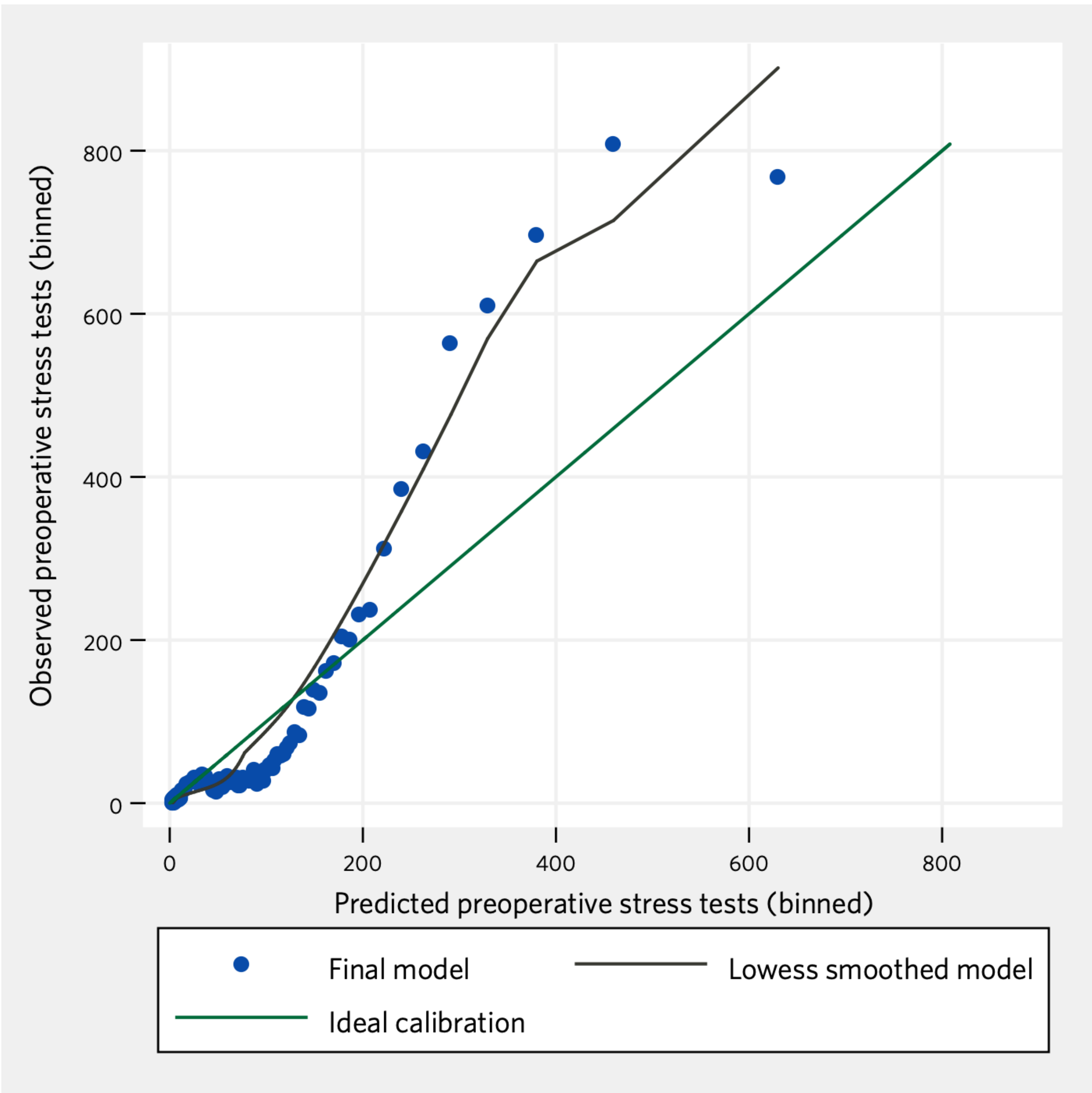
3. Physician rates of stress testing, unadjusted and marginal

Figure 1 demonstrates unadjusted rates of stress testing, and Figure 2 demonstrates marginal predictions for each physician, controlling for all other factors. Here we overlay the marginal results on the unadjusted results to demonstrate the effect of adjustment.



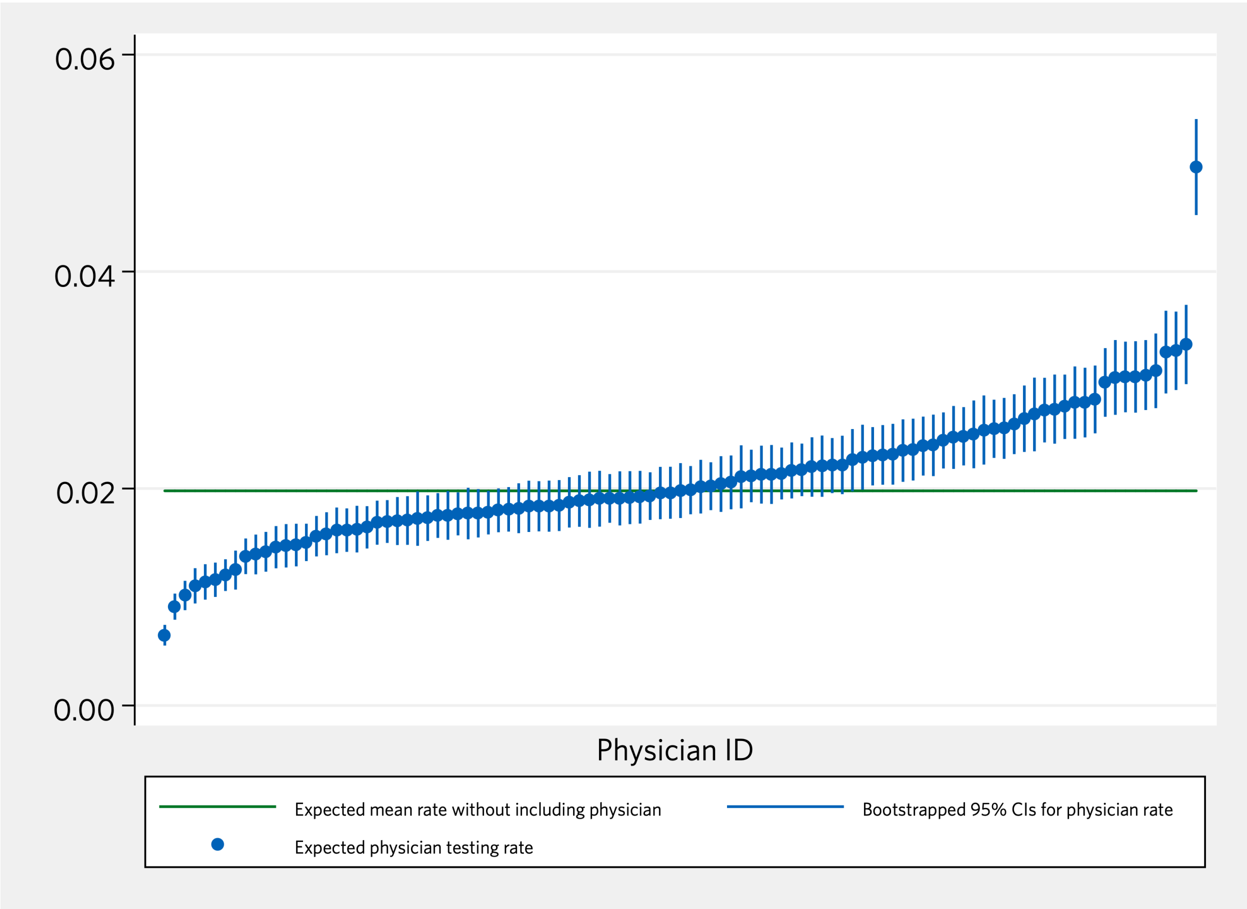
4. Calibration of final model

Based on the plot above, we suspected less-than-ideal calibration of our final model. We do not know of a universally accepted method to assess the calibration of a multilevel model on multiply-imputed data, but in most of our assessments this model fails Hosmer-Lemeshow Goodness-of-Fit testing, and the calibration plot shown here (binned into centiles) indeed suggests poor calibration. We emphasize again that our goal here is to explain variance in testing, not to guide future physicians in who should be referred for stress testing or to enable individual physician profiling.



5. Expected testing rates with an identical population

One way to contextualize provider effects is to imagine that each provider sees an identical panel of patients and estimate the consequent differences in outcomes. Here, we sampled 1,000 patients from our original population and estimated rates of stress testing if that same cohort were seen by each physician in our dataset. The overall mean is the expected rate for this small cohort without controlling for physician ID.



6. Sensitivity analysis excluding patients planned for aortic or vascular surgery

We repeated our analysis while excluding patients who were considered for aortic or vascular surgery. As in all models with dichotomous outcomes, the fixed variance leads to different effect sizes when using a different list of predictors. All effects are in the same direction as in our base-case analysis, as shown below. All changes in effect size are smaller than the smallest effect size in our base-case model (congestive heart failure). We have highlighted results where marginal rates differ from the base case by 0.02% or greater. This is an arbitrary threshold based on the intuition that a difference of less than 1 test per 500 visits is small. Due to rounding, some cells with less than a 0.2% absolute difference are displayed as differences of 0.2% in the cells below.

This analysis includes a total of 151,213 visits.

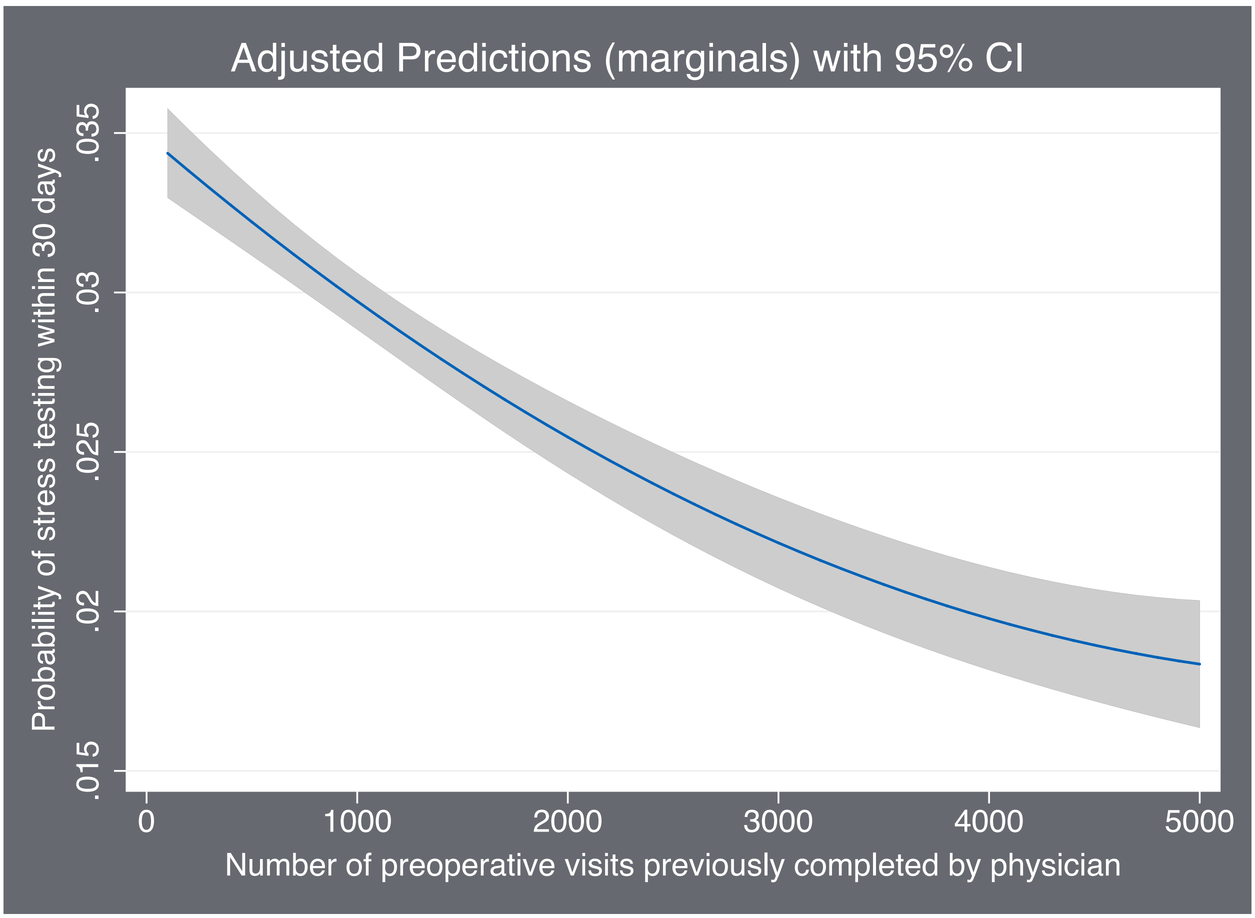
Predictor	Value	Mean marginal rate, base case	95% CI	Mean marginal rate with aortic and vascular surgery patients excluded	95% CI
MICA > 1%	0	0.6%	(0.6% – 0.7%)	0.6%	(0.5% – 0.7%)
	1	7.1%	(6.7% – 7.5%)	6.8%	(6.4% – 7.1%)
Documented RCRI	0	2.2%	(2.0% – 2.3%)	2.1%	(1.9% – 2.2%)
	1	2.7%	(2.5% – 2.9%)	2.6%	(2.4% – 2.8%)
	2	3.4%	(3.0% – 3.8%)	3.2%	(2.8% – 3.6%)
	3	4.2%	(3.5% – 5.0%)	3.9%	(3.3% – 4.7%)
	4	5.1%	(4.0% – 6.5%)	4.9%	(3.8% – 6.2%)
	5	6.4%	(4.7% – 8.6%)	6.0%	(4.4% – 8.2%)
Functional class	1	2.1%	(1.9% – 2.2%)	2.0%	(1.8% – 2.1%)
	2	2.4%	(2.2% – 2.5%)	2.2%	(2.1% – 2.4%)
	3	2.8%	(2.6% – 3.0%)	2.6%	(2.4% – 2.8%)
	4	3.2%	(2.9% – 3.5%)	2.9%	(2.6% – 3.3%)
Estimated metabolic equivalents	2	3.3%	(2.9% – 3.7%)	3.0%	(2.7% – 3.4%)
	4	2.7%	(2.5% – 2.9%)	2.5%	(2.3% – 2.7%)
	8	1.8%	(1.6% – 1.9%)	1.7%	(1.5% – 1.9%)
Body mass index	20	2.1%	(1.9% – 2.3%)	2.0%	(1.8% – 2.1%)
	30	2.4%	(2.2% – 2.5%)	2.2%	(2.1% – 2.4%)
	40	2.7%	(2.5% – 2.9%)	2.5%	(2.3% – 2.7%)

Predictor	Value	Mean marginal rate, base case	95% CI	Mean marginal rate with aortic and vascular surgery patients excluded	95% CI
Diastolic blood pressure	70	2.3%	(2.2% – 2.5%)	2.2%	(2.1% – 2.4%)
	90	2.5%	(2.3% – 2.7%)	2.4%	(2.2% – 2.6%)
	110	2.7%	(2.4% – 3.0%)	2.6%	(2.3% – 2.9%)
Ischemic heart disease	0	2.1%	(2.0% – 2.3%)	2.0%	(1.9% – 2.2%)
	1	3.6%	(3.3% – 3.9%)	3.3%	(3.1% – 3.7%)
Congestive heart failure	0	2.4%	(2.2% – 2.5%)	2.3%	(2.1% – 2.4%)
	1	2.1%	(1.9% – 2.3%)	2.0%	(1.8% – 2.2%)
Area deprivation index	10	2.8%	(2.5% – 3.1%)	2.6%	(2.4% – 2.9%)
	50	2.2%	(2.0% – 2.3%)	2.1%	(1.9% – 2.2%)
	90	2.6%	(2.4% – 2.8%)	2.4%	(2.2% – 2.7%)
Predicted probability of obstructive coronary artery disease	5%	2.6%	(2.4% – 2.7%)	2.4%	(2.3% – 2.6%)
	10%	2.4%	(2.3% – 2.6%)	2.3%	(2.2% – 2.5%)
	20%	2.2%	(2.1% – 2.4%)	2.1%	(1.9% – 2.2%)
Tobacco use	Current smoker	2.6%	(2.3% – 2.8%)	2.4%	(2.2% – 2.7%)
	Former smoker	2.5%	(2.3% – 2.7%)	2.4%	(2.2% – 2.6%)
	Neither	2.2%	(2.1% – 2.4%)	2.1%	(1.9% – 2.2%)
Date	2008.06.30	3.5%	(3.2% – 3.8%)	3.3%	(3.0% – 3.6%)
	2013.06.30	2.6%	(2.4% – 2.8%)	2.5%	(2.3% – 2.7%)
	2018.06.30	1.3%	(1.2% – 1.4%)	1.2%	(1.1% – 1.3%)
Surgical category	Aortic	23.4%	(6.0% –91.1%)	–	
	Peripheral vascular	8.7%	(6.7% –11.3%)	–	
	Urologic	9.2%	(8.3% –10.2%)	8.8%	(7.9% – 9.9%)
	Other	1.9%	(1.7% – 2.0%)	1.8%	(1.7% – 2.0%)
Physician (summary)	Lowest	1.0%	(0.1% – 4.4%)	0.5%	(0.1% – 3.6%)
	5th percentile	1.2%	(0.6% – 2.6%)	0.8%	(0.3% – 1.8%)
	Median	2.3%	(2.1% – 2.6%)	1.6%	(1.2% – 2.0%)

Predictor	**Value**	**Mean marginal rate, base case**	**95% CI**	**Mean marginal rate with aortic and vascular surgery patients excluded**	**95% CI**
	95th percentile	3.8%	(3.2% – 4.5%)	2.7%	(1.8% – 4.1%)
	Highest	6.1%	(2.7% –13.5%)	4.5%	(1.8% –11.3%)

7. Marginal testing rate as a function of physician experience

As with all datasets, our conclusions are a product of many decisions. For example, although we rejected physician experience as a predictor of testing rate in favor of date and a physician-specific random effect, reasonable investigators could disagree. To generate the graph below, we replaced date in our model with the number of visits each physician had completed between the beginning of our dataset and the visit in question (a proxy for preoperative clinic experience). We then computed and graphed marginal probabilities as described in our primary results.



STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2-3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4-5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	4-5
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5-8
Bias	9	Describe any efforts to address potential sources of bias	9-10
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5-10
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-10
		(*b*) Describe any methods used to examine subgroups and interactions	5-10
		(*c*) Explain how missing data were addressed	9-10
		(*d*) If applicable, explain how loss to follow-up was addressed	5
		(*e*) Describe any sensitivity analyses	5
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	5,10
		(b) Give reasons for non-participation at each stage	17
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	26-27
		(b) Indicate number of participants with missing data for each variable of interest	26
		(c) Summarise follow-up time (eg, average and total amount)	5
Outcome data	15*	Report numbers of outcome events or summary measures over time	10

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10-11, 26, 28
		(*b*) Report category boundaries when continuous variables were categorized	26
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	28
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11-12
Discussion			
Key results	18	Summarise key results with reference to study objectives	12-15
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	16-18
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	18
Generalisability	21	Discuss the generalisability (external validity) of the study results	17
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	1,19

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

Variation in preoperative stress testing by patient, physician, and surgical type – a cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-048052.R2
Article Type:	Original research
Date Submitted by the Author:	08-Sep-2021
Complete List of Authors:	Pappas, Matthew; Cleveland Clinic, Department of Hospital Medicine Sessler, Daniel; Cleveland Clinic, Outcomes Research Auerbach, A; University of California, San Francisco Kattan, Michael; Cleveland Clinic Milinovich, Alex; Cleveland Clinic Blackstone, Eugene; Cleveland Clinic, Miller Family Heart and Vascular Institute Rothberg, Michael B.; Cleveland Clinic, Cleveland, USA, Department of Internal Medicine
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Cardiovascular medicine, Anaesthesia, Surgery, General practice / Family practice
Keywords:	INTERNAL MEDICINE, Adult surgery < SURGERY, Adult anaesthesia < ANAESTHETICS, Adult cardiology < CARDIOLOGY, HEALTH SERVICES ADMINISTRATION & MANAGEMENT

Variation in preoperative stress testing by patient, physician, and surgical type – a cohort study

Matthew A Pappas, MD, MPH,[1,2,3] Daniel I Sessler, MD,[3,4] Andrew D Auerbach, MD, MPH,[5] Michael W Kattan, PhD,[6] Alex Milinovich, BA,[6] Eugene H Blackstone, MD,[7] and Michael B Rothberg, MD, MPH[1]

1: Center for Value-based Care Research, Cleveland Clinic, Cleveland, OH
2: Department of Hospital Medicine, Cleveland Clinic, Cleveland, OH
3: Department of OUTCOMES RESEARCH, Anesthesiology Institute, Cleveland Clinic, Cleveland, OH
4: OUTCOMES RESEARCH Consortium, Cleveland, OH
5: Department of Hospital Medicine, University of California, San Francisco, San Francisco, CA
6: Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH
7: Miller Family Heart and Vascular Institute, Cleveland Clinic, Cleveland, OH

Correspondence to Matthew A Pappas: pappasm@ccf.org; 216.444.9565; 9500 Euclid Ave., Mail Stop G-10, Cleveland, OH 44195

Short title/running head:	Variation in preoperative stress testing
Word count:	3,644
Abstract word count:	292
Number of references:	39
Number of tables:	3
Number of figures:	2
Number of appendices:	1

Financial support: This work was supported by the National Heart, Lung, and Blood Institute (NHLBI 1K08HL141598).
Competing Interests: None

Keywords (MEDLINE): Perioperative care; Preoperative care; Risk assessment; Elective surgical procedures; Cohort studies; Exercise test; Tomography, Emission-Computed, Single-Photon; Echocardiography, Stress

Abstract

Objectives:

To describe variation in and drivers of contemporary preoperative cardiac stress testing.

Setting:

A dedicated preoperative risk assessment and optimization clinic at a large integrated medical center from 2008 through 2018.

Participants:

A cohort of 118,552 adult patients seen by 104 physicians across 159,795 visits to a preoperative risk assessment and optimization clinic.

Main Outcome:

Referral for stress testing before major surgery, including nuclear, echocardiographic, or electrocardiographic-only stress testing, within 30 days after a clinic visit.

Results:

A total of 8,303 visits (5.2%) resulted in referral for preoperative stress testing. Key patient factors associated with preoperative stress testing included predicted surgical risk, patient functional status, a previous diagnosis of ischemic heart disease, tobacco use, and body mass index. Patients living in either the most- or least-deprived census block groups were more likely to be tested. Patients were tested more frequently before aortic, peripheral vascular, or urologic interventions than before other surgical subcategories. Even after fully adjusting for patient and surgical factors, provider effects remained important: marginal testing rates differed by a factor-of-three in relative terms and around 2.5% in

absolute terms between the 5th and 95th percentile physicians. Stress testing frequency decreased over the time period; controlling for patient and physician predictors, a visit in 2008 would have resulted in stress testing approximately 3.5% of the time, while a visit in 2018 would have resulted in stress testing approximately 1.3% of the time.

Conclusions:

In this large cohort of patients seen for preoperative risk assessment at a single health system, decisions to refer patients for preoperative stress testing are influenced by various factors other than estimated perioperative risk and functional status, the key considerations in current guidelines. The frequency of preoperative stress testing has decreased over time, but remains highly provider-dependent.

Strengths and limitations of this study:

- We identified a large cohort of patients considering noncardiac surgery, with detailed clinical data from each visit.
- We tested predictor variables across various constructs potentially related to preoperative stress testing.
- We accounted for clustering by physician and patient, testing different structures in an effort to optimally partition variance.
- We used multiple imputation by chained equations to mitigate potential biases from missing data.

Glossary of Abbreviations:

ACC/AHA: American College of Cardiology/American Heart Association
ASA: American Society of Anesthesiologists
CAD: coronary artery disease
METs: Estimated metabolic equivalents
MICA: Myocardial infarction or cardiac arrest calculator
RCRI: Revised cardiac risk index

Introduction

The 2014 American College of Cardiology/American Heart Association (ACC/AHA) guidelines recommend preoperative stress testing for patients whose predicted risk of a major adverse cardiac event exceeds 1% and whose functional status is poor or unknown, when results from stress testing would change clinical management.[1]

However, clinicians use various risk prediction tools, which identify different patients as having elevated risk.[2,3] Additionally, multiple methods of assessing functional status are used, which again can lead to variation in patients selected for stress testing.[2,4–7] Thus, the final decision to proceed with stress testing can become something closer to a provider-level judgment than a guideline-driven protocol.[8] Variation in use of stress testing can have substantial cost implications and potentially prompt subsequent tests and procedures with little clinical benefit.[9]

To understand contemporary use and drivers of preoperative cardiac stress testing, we sought to describe variation and predictors of preoperative stress testing using rich clinical data from a large integrated health system.

Methods

The Internal Medicine Preoperative Assessment, Consultation and Treatment (IMPACT) Center assesses patients prior to noncardiac surgery at the Cleveland Clinic. In the years

from 2008 through 2018, 118,552 patients were seen in our clinic by 104 physicians across 159,795 visits. Among this cohort, we identified scheduled and completed preoperative cardiac stress tests, here defined as those within the 30 days after a clinic visit and before noncardiac surgery. This study was approved by the Cleveland Clinic Institutional Review Board (IRB #18-1076).

Natural language processing was performed in python (version 3.7.8) using regular expressions and the spacy.io library (version 2.3.2). All analyses were performed in Stata (version 14; College Station, TX). Data used are from our electronic health record and are not available for outside access.

Predictor variables

We theorized that six underlying constructs would be related to stress test ordering: predicted perioperative risk, functional status, social and financial support, medical comorbidities, physician tendencies and experience, and time. We created a random effects logit model for each construct to refine variables included in our final model. Within each submodel, we pruned variables according to Bayesian information criteria (BIC). For continuous variables, we assessed for nonlinear or categorical relationships using visual examination of binned scatter plots. To avoid overfitting, we limited candidate predictors to fewer than 1 predictor variable per 15 preoperative stress tests, including tested interactions, nonlinear effects, and discarded predictors. We estimated that we had approximately 539 degrees of freedom for analysis, with fewer for cluster-level variables depending on model structure (described below).

For measures of perioperative risk, we tested Revised Cardiac Risk Index (RCRI), Myocardial Infarction or Cardiac Arrest (MICA), and MICA's categorization of surgeries using a previously-published crosswalk.[2,10,11] (Although different procedures likely have different intrinsic cardiac risk, we used the MICA categorization of surgeries to avoid overfitting.[11,12]) Upon finding that few surgical categories were associated with different stress testing rates, we replaced that multinomial variable with indicator variables for each category associated with different testing rates in our data (aortic, peripheral vascular, and urologic surgeries). As a separate sensitivity analysis, we excluded patients seen in advance of cardiac or vascular surgery, who are generally evaluated elsewhere in our institution. We tested both documented and calculated RCRI, which may differ for a variety of reasons including lab results between the clinic visit and documentation, erroneous diagnoses/chart lore, outside records unavailable in the electronic medical record, and misconceptions about how RCRI is calculated. We treated both estimates of RCRI as continuous to force a monotonic relationship (no theory would support lower testing rates at higher predicted cardiac risk). We tested MICA-predicted probability both as a continuous variable and dichotomized at 1%. Although we used the American Society of Anesthesiologists (ASA) physical status to calculate MICA, we did not test that separately because it is assigned at the time of surgery.

For measures of functional status, we tested estimated metabolic equivalents (METs), which in this clinic is based on a semi-quantitative questionnaire, and the physician's subjective global assessment of function, which is comparable to the Eastern Cooperative Oncology Group (ECOG) score.[13] For measures of social and financial support, we tested area deprivation index (a measure of socioeconomic disadvantage based on education, employment, housing quality, and poverty measures by census block group), race, ethnicity, marital status, and age (here dichotomized at age 65 to reflect changes in access to care with universal Medicare eligibility).[14] For measures of medical comorbidities and illness, we considered age, vital signs at the clinic visit, diagnoses of coronary artery disease, cerebrovascular disease, or congestive heart failure, diabetes, use of insulin, creatinine, tobacco use, and predicted probability of obstructive coronary artery disease.[15]

To accurately capture the predicted probability of obstructive coronary artery disease among patients without an existing diagnosis, we applied natural language processing to extract three pain characteristics from the full-text clinic note, when patients were documented to have chest pain: (1) substernal, (2) provoked by exertion, and (3) relieved by either rest or nitroglycerin. Notes with all three chest pain criteria were considered to document "typical" chest pain, notes with two criteria were considered to document "atypical" chest pain, and other documentation of chest pain was considered to represent non-specific chest pain.[15] For visits without documented chest pain, we estimated pretest probability as though patients had non-specific chest pain. We used an interaction with our variable for a previous diagnosis of coronary artery disease, such that this estimated probability was considered a predictor only when patients did not have an existing diagnosis.

For measures of physician tendencies and experience, we tested (on the date of each visit) years of post-residency practice (a proxy for overall experience) and the number of previous encounters the physician had completed in our dataset (a proxy for experience in preoperative risk assessment more specifically). For measures of time, we theorized that patients who had previous cardiac stress tests would be less likely to be referred for preoperative stress testing, and that physicians would give greater weight to more recent tests (i.e., the relationship would be time-dependent). We used the date of the visit as a continuous variable to test for changing stress test rates over time. To assess for changes related to publication of the current ACC/AHA guideline, we created a dichotomous variable for whether the visit occurred before or after said guideline's December 9, 2014 publication, and tested for interactions between that term and other predictors.[1]

Because unstable angina would be a potential indication for cardiac testing regardless of upcoming surgery, we investigated the frequency of angina in a subset of notes between June 2013 and July 2016; an EHR template used during this period included a structured questionnaire of symptoms. One such symptom was "Angina within 30 days".

Model structure

Conceptually, visits could be thought of as clustered by patient (with physician-level variables at the level of the visit) or by physician (with patient-level variables at the level of the visit). We tested different structures using empty models and calculated intraclass correlation coefficients to estimate the proportion of variance explained by unmeasured patient-level or physician-level factors.

With visits clustered by physician, approximately 0.4% of variance in stress test ordering was at the level of the physician. When we clustered visits by patient, approximately 4.9% of variance in stress test ordering was at the level of the patient. We therefore developed our model using physician- and visit-level variables clustered by patient, including physician ID as a visit-level indicator ("dummy") variable. This approach drops some low volume providers for whom outcomes are overfitted but should capture more unmeasured variance at both the patient and provider levels.

Multivariable modeling

Finally, we added the remaining predictor variables from each submodel into a multivariable logistic regression model. We again pruned predictors based on BIC and examined for nonlinear or categorical relationships. We revisited our model structure using the final predictor variables, comparing models clustered by patient or physician based on BIC.

Because results of a multilevel logistic regression with interaction terms have limited intuitive meaning, we calculated marginal effects for reporting. Holding all other variables at their medians, we estimated the effect of changing one predictor variable at a time.

Data extraction and missingness

Our methods for extracting data from the electronic medical record have been described previously.[16] We considered patients as having each considered diagnosis if it had been documented at any time before or at the analyzed visit. For creatinine and other lab testing, we used the most recent measurement up to and including the day of the clinic visit. We used multiple imputation by chained equations to address missing data and previously described standards to ensure multiple imputation did not introduce Monte Carlo error.[17] We imputed predictor variables other than those with negligible missing data: age, sex, vital signs, and binary variables indicating previous diagnoses.

Patient and Public Involvement

Patients were not involved in the design of this study.

Results

Overall, 5.2% of visits to the preoperative clinic led to a cancelled or completed preoperative stress test (8,303/159,795), with 5.1% (8,085; 97.4% of those referred) completing the test. Patient demographics, selected risk factors, and proportions of missing data are shown in Tables 1 and 2. Unadjusted physician referral rates are shown in Figure 1.

Marginal testing rates across each predictor variable, with other variables at their respective medians, are shown in Table 3. In general, patients were more likely to be referred for preoperative stress testing as estimated perioperative risk increased and for specific categories of surgeries (aortic, peripheral vascular, and urologic). Of those, patients undergoing aortic surgery were most likely to be referred for stress testing, though our dataset included relatively few aortic surgeries and confidence intervals for that predictor were wide. Even after adjusting for all other factors, different providers were more likely to refer for preoperative stress testing than others: a visit to the 95th percentile physician in this clinic would result in stress testing 3.8% of the time, while a visit to the 5th percentile physician in this clinic would result in testing around 1.2% of the time.

Other important patient variables included the physician's subjective assessment of global patient function, METs, socioeconomic advantage or disadvantage compared to the median, BMI, diastolic blood pressure, existing diagnoses of ischemic heart disease or congestive heart failure, estimated probability of obstructive coronary artery disease, and tobacco use. Visits later in our dataset were less likely to result in a preoperative stress test compared with earlier visits. Each of these variables, while significant, appeared to exert less influence than surgical categories, estimated surgical risk, and provider. Fully adjusted provider marginal rates are shown in Figure 2.

Results were very similar for models clustered by patient or physician. Information criteria would slightly favor a model clustered by physician (BIC: 51040) compared to a model clustered by patient with a physician indicator variable (BIC: 52009). Meanwhile, the model clustered by physician had a slightly lower R^2 (0.1896) compared with a model clustered by patient with physician as an indicator variable (0.1907).

Between June 2013 and July 2016, 23,034 visits used our EHR template that included structured entry of “Angina within 30 days”. Of those, 48 (0.2%) were marked as “Yes”. Of 107 other visits flagged by natural language processing as potentially including unstable angina, manual chart review of a random sample of 50 visits showed 3 negations missed by natural language processing, 32 descriptions of historical symptoms that had prompted testing or intervention previously, 2 quotations of test reports (coronary catheterization reports that included unstable angina as the indication for testing), and 13 cases of unstable angina.

Discussion

In this cohort of patients seen for preoperative risk assessment at a single health system, we have identified key drivers of preoperative stress testing, which include type of surgery, estimated surgical risk, and patient functional status. Our results demonstrate use of preoperative stress testing in a real-world cohort.

Current guidelines recommend preoperative stress testing for patients whose predicted perioperative adverse cardiac event risk exceeds 1% and whose functional capacity is poor or unknown, when such testing would change clinical management. Although we cannot determine from these data whether physicians thought testing would change management, and predicted surgical risk scores have poor concordance across the 1% threshold, patients able to perform four or more metabolic equivalents of activity made up nearly one-third of all stress test referrals.[2,10,18] Our data suggest that a substantial fraction of preoperative stress tests were inconsistent with current guidance.

Predicted surgical risk was nonetheless a key driver of preoperative stress testing. Testing rates increased with increasing RCRI, without a clear dichotomization at any particular value of RCRI. And interestingly, although MICA was essentially never documented, a MICA-predicted surgical risk of greater than 1% appears to be a better single predictor variable than RCRI. Physicians could be trying to incorporate the guideline-recommended threshold into their decision-making while relying on cohorts with different calibration, or could be deliberately avoiding a stark dichotomization of risk at 1%.[10,18]

As with predicted surgical risk, physicians appeared to consider functional status as something between the dichotomy of current guidance and continuum of risk encountered in clinical practice. Testing rates were higher among less functional patients and lower among patients able to achieve higher METs, but neither were especially important predictors in our model. It seems probable that both clinical skill and clinical decision-making are more nuanced than we can discern from our data source. For example, physicians likely vary in both their ability to elicit anginal equivalents and their interpretation of potentially ambiguous symptoms. And even if they were presented with equivalent information, various physicians might reasonably make different decisions about testing based on factors we are not able to investigate. For example, physicians could reasonably be more inclined to test before a pancreaticoduodenectomy than a laparoscopic cholecystectomy in view of those procedures' very different metabolic demands, but we lack sufficient power to test individual surgical procedures without overfitting.[12] In any case, functional status and patient variables other than predicted perioperative risk explained little variance in testing rates.

Surgical category also offers insights into testing rationale. Patients were tested more frequently before aortic or peripheral vascular interventions, perhaps reflecting persistent beliefs that patients with coronary artery disease should be identified and revascularized before vascular surgery.[19–21] However, patients undergoing vascular surgery are generally evaluated elsewhere at our institution, leaving our sample small and confidence intervals wide. We also note that patients are more likely to be referred for stress testing before urologic surgery, after controlling for patient risk factors and despite the fact that urologic procedures are not associated with higher intrinsic cardiac risk than other common surgical categories. Anecdotally, physicians practicing in this clinic have reported that a number of urologists at our institution are reluctant to operate on high-risk patients unless those patients first undergo preoperative testing. While investigating such a hypothesis would require a different approach than ours, clearly every physician in a preoperative clinic functions within a larger system of care and must build consensus among a team of treating physicians.

Other significant predictor variables include tobacco use, BMI, diastolic blood pressure, ischemic heart disease, and a patient's census block group. Because each of these are correlated with risk of obstructive coronary artery disease, one possibility is that data unavailable to us (such as outside records) led to some portion of the preoperative stress testing we observed. Although that remains possible, multiple observations argue against a simple explanation that these variables are proxies for coronary disease risk. First, higher probability of obstructive coronary artery disease, calculated based on available data, was associated with lower likelihood of preoperative testing. Second, patients residing in the wealthiest and poorest census tracts were approximately as likely to be referred for stress testing, with patients in the middle of the socioeconomic range less likely. Finally, diabetes was not associated with testing. It would seem that either physicians in our dataset did not incorporate patients' pretest probability of obstructive coronary disease in their decision to refer for testing, or that their assessments were poorly calibrated.

Angina or its equivalents do not appear to be a frequent rationale for testing in this cohort. Around 0.2% of notes that used a templated review of pertinent symptoms noted angina within 30 days, parsing of free text notes did not identify unstable angina with appreciable frequency, and many cases identified through natural language processing appeared not to be unstable angina on manual chart review of sampled visits. Although this could represent a failure to document findings that were present during the visit, it would seem more likely that a preoperative visit before elective noncardiac surgery is an inefficient tool to screen for angina.

The frequency of stress testing declined over time in our dataset, in contrast with increasing testing rates suggested in other contexts.[22,23] A recent cross-sectional analysis of claims data from patients who had total hip or knee arthroplasty also identified decreasing testing frequency over a similar period.[24] Our cohort study begins with a visit to a preoperative risk assessment clinic, incorporates detailed clinical data, and is not limited to patients who have completed orthopedic surgery. Our analysis thus extends previous understanding by showing that the reduction is not limited to orthopedic procedures, not a result of selecting patients not referred for stress testing for elective surgery, and not consequent to lower predicted cardiac risk. Taken together, these two analyses with different limitations suggest a shift in practice away from preoperative cardiac stress testing. Neither analysis suggests a clear change in testing frequency after the release of current guidelines. Although there can certainly be time lags between publication and consequent practice change, our findings argue against a causal relationship between publication of the current guideline and near-term changes in testing rates.[25]

Our model demonstrates physician practice variation: with all other predictors held at their medians, the 95th percentile physician was around three times more likely to order preoperative stress testing than the 5th percentile physician. But we caution against using our results, or others, for profiling individual providers, which is generally a low-reliability exercise and prone to gaming.[26] Our dataset is among the largest clinical datasets of preoperative risk assessment, but true outliers are rare and most physicians are not detectably different from the mean after adjustment (see Figure 2). As with other observations of physician practice variation, ours suggests a deeper failure: that we do not yet understand how best to use preoperative stress testing to mitigate perioperative cardiac risk.[27]

Surgery carries inherent cardiac risk, and stress testing may reflect physician discomfort with the malpractice or cognitive liabilities that cardiac risk entails.[28–32] Stress testing can inform diagnosis and prognosis, but outcomes will only improve if testing results in interventions that reduce perioperative risk. Such interventions have proven elusive: preoperative revascularization did not reduce cardiac risk in the largest randomized trial to date, beta blockers are more likely harmful than helpful, and other interventions (e.g., statins) that may be allocated differently based on stress testing likely have modest effects, if any.[33–36] Intraoperative care or postoperative testing patterns could differ based on whether a preoperative stress test was performed, but what practices in those settings might reduce the risk of major adverse cardiac events remain equally unclear. While estimated perioperative cardiac risk appears to drive stress testing, it remains to be seen how stress testing might reduce perioperative cardiac risk.

Although we have made every effort to ensure the internal validity of our data, analysis, and results, our data may not adequately represent drivers of or variation in preoperative stress testing before some common types of surgery, including ophthalmologic surgery, which is evaluated elsewhere in our institution.[37] Our IMPACT clinic has made substantial efforts to provide uniform care, which could have reduced physician variation in our dataset, and we cannot analyze variation across health system or region, which can also be substantial.[38] As with any single-center study, results should be extrapolated to other settings with caution. For example, we observed higher testing frequency before urologic surgery than would be expected for cardiac risk; other centers may have different surgical categories with testing out of proportion to surgical risk. The need to build a consensus plan of care among a treatment team is true across institutions, but the particulars of our institution's consensus may not be.

Additionally, as with other observational studies, our analytical choices are difficult to separate from our theoretical framework, and may influence our results in various ways.[39] For example, we rejected physician experience as a predictor of testing in favor of a random effect for each physician and the date of each visit due to our prespecified analytic criteria (rejecting predictor variables that worsened AIC and BIC). Still, experience differs by physician and necessarily accrues over time. A reasonable investigator with a different theoretical model could assume broadly stable testing rates over time and conclude that testing decreases as physicians gain experience (see Supplemental Appendix). Time in particular is rife with potential confounders of this sort. Recent work using other datasets also identified reduced stress testing over time (described above), offering reassurance that our analytic criteria led to the best conclusion. Still, the effect of experience on appropriate testing could be an avenue for further investigation.

But while the limitations of our study reflect the limitations of any single center observational study, the detailed clinical data available to us offers distinct advantages over earlier work. We have demonstrated real-world use of preoperative stress testing before a wide range of possible surgical interventions, using visit-level data to comprehensively assess variation in and predictors of preoperative cardiac stress testing.

In summary, the frequency of preoperative stress testing varied with estimated surgical risk, patient functional status, socioeconomic status, ischemic heart disease, congestive heart failure, body mass index, diastolic blood pressure, surgical category, and provider. The fraction of patients referred for stress testing appears to be declining over time, but testing remains common and highly dependent on the provider. The value of preoperative stress testing remains to be established.

Funding Sources:

This work was funded by NHLBI 1K08HL141598.

Conflict of Interest Disclosures:

None

Author Contributions:

Dr. Pappas: This author conceived the study, collected the data, performed the analysis, and wrote the initial draft of the manuscript.
Dr. Sessler: This author helped improve the analysis and revised the manuscript.
Dr. Auerbach: This author helped improve the analysis and revised the manuscript.
Dr. Kattan: This author helped improve the analysis and revised the manuscript.
Mr. Milinovich: This author helped obtain and validate data.
Dr. Blackstone: This author helped improve the analysis and revised the manuscript.
Dr. Rothberg: This author helped improve the analysis and revised the manuscript.

Ethical Approval:

This study was approved by the Cleveland Clinic Institutional Review Board (IRB #18-1076).

Data Availability Statement

Data used are from our electronic health record and are not available for outside access.

References

1. Fleisher LA, Fleischmann KE, Auerbach AD, et al. 2014 ACC/AHA guideline on perioperative cardiovascular evaluation and management of patients undergoing noncardiac surgery: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;130(24):2215-2245. doi:10.1161/CIR.0000000000000105

2. Pappas MA, Sessler DI, Rothberg MB. Anticipated Rates and Costs of Guideline-Concordant Preoperative Stress Testing. *Anesth Analg*. 2019;128(2):241-246. doi:10.1213/ANE.0000000000003754

3. Glance LG, Faden E, Dutton RP, et al. Impact of the choice of risk model for identifying low-risk patients using the 2014 American college of cardiology/American Heart Association perioperative guidelines. *Anesthesiology*. 2018;129(5):889-900. doi:10.1097/ALN.0000000000002341

4. Hlatky MA, Boineau RE, Higginbotham MB, et al. A brief self-administered questionnaire to determine functional capacity (The Duke Activity Status Index). *Am J Cardiol*. 1989;64(10):651-654. doi:10.1016/0002-9149(89)90496-7

5. Reilly DF, McNeely MJ, Doerner D, et al. Self-reported exercise tolerance and the risk of serious perioperative complications. *Arch Intern Med*. 1999;159(18):2185-2192. doi:10.1001/archinte.159.18.2185

6. Melon CC, Eshtiaghi P, Luksun WJ, Wijeysundera DN. Validated questionnaire vs physicians' judgment to estimate preoperative exercise capacity. *JAMA Intern Med*.

2014;174(9):1507-1508. doi:10.1001/jamainternmed.2014.2914

7. Wijeysundera DN, Pearse RM, Shulman MA, et al. Assessment of functional capacity before major non-cardiac surgery: an international, prospective cohort study. *Lancet*. 2018;391(10140):2631-2640. doi:10.1016/S0140-6736(18)31131-0

8. Eddy DM. Variations in physician practice: The role of uncertainty. *Health Aff*. 1984;3(2):74-89. doi:10.1377/hlthaff.3.2.74

9. Mark DB, Federspiel JJ, Cowper PA, et al. Economic Outcomes With Anatomical Versus Functional Diagnostic Testing for Coronary Artery Disease. *Ann Intern Med*. 2016;165(2):94-102. doi:10.7326/M15-2639

10. Lee TH, Marcantonio ER, Mangione CM, et al. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. *Circulation*. 1999;100(10):1043-1049. doi:10.1161/01.CIR.100.10.1043

11. Gupta PK, Gupta H, Sundaram A, et al. Development and validation of a risk calculator for prediction of cardiac risk after surgery. *Circulation*. 2011;124(4):381-387. doi:10.1161/CIRCULATIONAHA.110.015701

12. Liu JB, Liu Y, Cohen ME, Ko CY, Sweitzer BJ. Defining the Intrinsic Cardiac Risks of Operations to Improve Preoperative Cardiac Risk Assessments. *Anesthesiology*. 2018;128(2):283-292. doi:10.1097/ALN.0000000000002024

13. Oken MM, Creech RH, Davis TE. Toxicology and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol*. 1982;5(6):649-655. doi:10.1097/00000421-198212000-00014

14. Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible - The neighborhood atlas. *N Engl J Med*. 2018;378(26):2456-2458. doi:10.1056/NEJMp1802313

15. Genders TSS, Steyerberg EW, Hunink MGM, et al. Prediction model to estimate presence of coronary artery disease: Retrospective pooled analysis of existing cohorts. *BMJ*. 2012;344(7862). doi:10.1136/bmj.e3485

16. Milinovich A, Kattan MW. Extracting and utilizing electronic health data from Epic for research. *Ann Transl Med*. 2018;6(3):42-42. doi:10.21037/atm.2018.01.13

17. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30(4):377-399. doi:10.1002/sim.4067

18. Duceppe E, Parlow J, MacDonald P, et al. Canadian Cardiovascular Society Guidelines on Perioperative Cardiac Risk Assessment and Management for Patients Who Undergo Noncardiac Surgery. *Can J Cardiol*. 2017;33(1):17-32. doi:10.1016/j.cjca.2016.09.008

19. Eagle KA, Singer DE, Brewster DC, Darling RC, Mulley AG, Boucher CA. Dipyridamole-thallium scanning in patients undergoing vascular surgery. Optimizing preoperative evaluation of cardiac risk. *JAMA*. 1987;257(16):2185-2189. http://www.ncbi.nlm.nih.gov/pubmed/3560400. Accessed November 11, 2019.

20. Hertzer NR, Beven EG, Young JR, et al. Coronary artery disease in peripheral vascular patients. A classification of 1000 coronary angiograms and results of surgical management. *Ann Surg*. 1984;199(2):223-233. doi:10.1097/00000658-

198402000-00016

21. Boucher CA, Brewster DC, Darling RC, Okada RD, Strauss HW, Pohost GM. Determination of Cardiac Risk by Dipyridamole-Thallium Imaging before Peripheral Vascular Surgery. *N Engl J Med*. 1985;312(7):389-394. doi:10.1056/NEJM198502143120701

22. Sheffield KM, McAdams PS, Benarroch-Gampel J, et al. Overuse of preoperative cardiac stress testing in medicare patients undergoing elective noncardiac surgery. *Ann Surg*. 2013;257(1):73-80. doi:10.1097/SLA.0b013e31826bc2f4

23. Sigmund AE, Stevens ER, Blitz JD, Ladapo JA. Use of preoperative testing and physicians' response to professional society guidance. *JAMA Intern Med*. 2015;175(8):1352-1359. doi:10.1001/jamainternmed.2015.2081

24. Rubin DS, Hughey R, Gerlach RM, Ham SA, Ward RP, Nagele P. Frequency and Outcomes of Preoperative Stress Testing in Total Hip and Knee Arthroplasty from 2004 to 2017. *JAMA Cardiol*. 2021;6(1):13-20. doi:10.1001/jamacardio.2020.4311

25. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: The National Academies Press; 2001. doi:10.17226/10027

26. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *J Am Med Assoc*. 1999;281(22):2098-2105. doi:10.1001/jama.281.22.2098

27. Wennberg JE, Freeman JL, Culp WJ. Are Hospital Services Rationed in New Haven

or Over-Utilized in Boston? *Lancet*. 1987;329(8543):1185-1189. doi:10.1016/S0140-6736(87)92152-0

28. Saint S, Vaughn VM, Chopra V, Fowler KE, Kachalia A. Perception of Resources Spent on Defensive Medicine and History of Being Sued Among Hospitalists: Results from a National Survey. *J Hosp Med*. 2018;13(1):26-29. doi:10.12788/jhm.2800

29. Feinstein AR. The 'Chagrin Factor' and Qualitative Decision Analysis. *Arch Intern Med*. 1985;145(7):1257-1259. doi:10.1001/archinte.1985.00360070137023

30. Rolfe A, Burton C. Reassurance after Diagnostic Testing with a Low Pretest Probability of Serious Disease: Systematic Review and Meta-analysis. *JAMA Intern Med*. 2013;173(6):407-416. doi:10.1001/jamainternmed.2013.2762

31. Van der Weijden T, Van Velsen M, Dinant GJ, Van Hasselt CM, Grol R. Unexplained complaints in general practice: Prevalence, patients' expectations, and professionals' test-ordering behavior. *Med Decis Mak*. 2003;23(3):226-231. doi:10.1177/0272989X03023003004

32. Little P, Dorward M, Warner G, Stephens K, Senior J, Moore M. Importance of patient pressure and perceived pressure and perceived medical need for investigations, referral, and prescribing in primary care: Nested observational study. *Br Med J*. 2004;328(7437):444-446. doi:10.1136/bmj.38013.644086.7c

33. McFalls EO, Ward HB, Moritz TE, et al. Coronary-Artery Revascularization before Elective Major Vascular Surgery. *N Engl J Med*. 2004;351(27):2795-2804. doi:10.1056/NEJMoa041905

34. Blessberger H, Kammler J, Domanovits H, et al. Perioperative beta-blockers for preventing surgery-related mortality and morbidity. *Cochrane Database Syst Rev*. March 2018. doi:10.1002/14651858.CD004476.pub3

35. Berwanger O, Manach Y Le, Suzumura EA, et al. Association between pre-operative statin use and major cardiovascular complications among patients undergoing non-cardiac surgery: The VISION study. *Eur Heart J*. 2016;37(2):177-185. doi:10.1093/eurheartj/ehv456

36. Berwanger O, de Barros e Silva PGM, Barbosa RR, et al. Atorvastatin for high-risk statin-naïve patients undergoing noncardiac surgery: The Lowering the Risk of Operative Complications Using Atorvastatin Loading Dose (LOAD) randomized trial. *Am Heart J*. 2017;184:88-96. doi:10.1016/j.ahj.2016.11.001

37. Kerr EA, Chen J, Sussman JB, Klamerus ML, Nallamothu BK. Stress Testing Before Low-Risk Surgery: So Many Recommendations, So Little Overuse. *JAMA Intern Med*. 2015;175(4):645. doi:10.1001/jamainternmed.2014.7877

38. Home - Dartmouth Atlas of Health Care. https://www.dartmouthatlas.org/. Accessed November 11, 2019.

39. Palpacuer C, Hammas K, Duprez R, Laviolle B, Ioannidis JPA, Naudet F. Vibration of effects from diverse inclusion/exclusion criteria and analytical choices: 9216 different ways to perform an indirect comparison meta-analysis. *BMC Med*. 2019;17(1):174. doi:10.1186/s12916-019-1409-3

29

Table 1. Patient and Surgical Characteristics.

		Total	Percent of category	Percent of all visits	Completed preoperative stress test	Percent of all preoperative stress tests
Age		159,795	100.0%	100.0%	8,303	100.0%
Sex	Female	88,738	55.5%	55.5%	4,079	49.1%
	Male	71,055	44.5%	44.5%	4,224	50.9%
Previous diagnosis of ischemic heart disease	No	128,505	80.4%	80.4%	4,831	58.2%
	Yes	31,290	19.6%	19.6%	3,472	41.8%
Previous diagnosis of congestive heart failure	No	146,556	91.7%	91.7%	6,922	83.4%
	Yes	13,239	8.3%	8.3%	1,381	16.6%
Previous diagnosis of cerebrovascular disease	No	141,519	88.6%	88.6%	6,567	79.1%
	Yes	18,276	11.4%	11.4%	1,736	20.9%
Systolic Blood Pressure		159,488	100.0%	99.8%	8,285	99.8%
Diastolic Blood Pressure		159,481	100.0%	99.8%	8,284	99.8%
Body Mass Index		157,473	100.0%	98.5%	8,155	98.2%
Creatinine (RCRI categorization)	≤ 2.0 mg/dL	151,885	97.1%	95.1%	7,695	92.7%
	> 2.0 mg/dL	4,487	2.9%	2.8%	542	6.5%
Creatinine (MICA categorization)	≤ 1.5 mg/dL	144,369	90.3%	90.3%	7,128	85.8%
	> 1.5 mg/dL	12,003	7.5%	7.5%	1,109	13.4%
	Unknown	3,423	2.1%	2.1%	66	0.8%
Prescribed insulin	No	147,610	92.4%	92.4%	7,136	85.9%
	Yes	12,185	7.6%	7.6%	1,167	14.1%
RCRI surgical category	High risk	27,709	23.8%	17.3%	872	10.5%
	Other	88,929	76.2%	55.7%	360	4.3%
Area Deprivation Index		126,076	100.0%	78.9%	7,091	85.4%
RCRI (documented)	0	50,785	75.1%	31.8%	1,548	18.6%
	1	12,642	18.7%	7.9%	988	11.9%
	2	3,321	4.9%	2.1%	402	4.8%
	3	742	1.1%	0.5%	108	1.3%
	4	151	0.2%	0.1%	21	0.3%
	5	15	0.0%	0.0%	2	0.0%
RCRI (calculated)	0	56,879	52.0%	35.6%	355	4.3%
	1	36,020	32.9%	22.5%	393	4.7%
	2	11,143	10.2%	7.0%	245	3.0%
	3	4,006	3.7%	2.5%	98	1.2%
	4	1,172	1.1%	0.7%	33	0.4%
	5	204	0.2%	0.1%	7	0.1%
	6	12	0.0%	0.0%	0	0.0%
MICA risk estimate	≤ 1%	71,448	44.7%	44.7%	608	7.3%
	> 1%	88,347	55.3%	55.3%	7,695	92.7%
Physician subjective assessment of patient global function	1	17,991	19.1%	11.3%	489	5.9%
	1-2	13,386	14.2%	8.4%	634	7.6%
	2	40,829	43.4%	25.6%	1,739	20.9%

	2-3	8,588	9.1%	5.4%	646	7.8%
	3	9,310	9.9%	5.8%	732	8.8%
	3-4	1,922	2.0%	1.2%	253	3.0%
	4	1,999	2.1%	1.3%	175	2.1%
Insurance	Medicare	20,744	52.2%	13.0%	1,192	14.4%
	Medicaid	2,384	6.0%	1.5%	83	1.0%
	Private	14,764	37.1%	9.2%	307	3.7%
	Other listed insurer	1,881	4.7%	1.2%	53	0.6%
ASA Class	1	3,325	3.0%	2.1%	9	0.1%
	2	34,026	30.9%	21.3%	181	2.2%
	3	65,298	59.3%	40.9%	854	10.3%
	4	7,454	6.8%	4.7%	157	1.9%
MICA surgical category	Anorectal	1,213	1.0%	0.8%	3	0.0%
	Aortic	115	0.1%	0.1%	34	0.4%
	Bariatric	703	0.6%	0.4%	6	0.1%
	Brain	4,780	4.1%	3.0%	15	0.2%
	Breast	8,541	7.4%	5.3%	17	0.2%
	Cardiac	237	0.2%	0.1%	10	0.1%
	Ear, nose, throat	3,640	3.1%	2.3%	19	0.2%
	Foregut/hepatopancreatobiliary	3,252	2.8%	2.0%	37	0.4%
	Gallbladder, appendix, adrenals, or spleen	1,974	1.7%	1.2%	17	0.2%
	Gynecologic	7,458	6.4%	4.7%	51	0.6%
	Hernia	2,778	2.4%	1.7%	16	0.2%
	Intestines	16,081	13.9%	10.1%	87	1.0%
	Neck	3,076	2.7%	1.9%	18	0.2%
	Nonesophageal thoracic	290	0.2%	0.2%	4	0.0%
	Orthopedic	26,005	22.4%	16.3%	208	2.5%
	Other abdomen	1,970	1.7%	1.2%	19	0.2%
	Peripheral vascular	1,355	1.2%	0.8%	110	1.3%
	Skin	7,126	6.1%	4.5%	78	0.9%
	Spinal	11,994	10.3%	7.5%	59	0.7%
	Urologic	13,423	11.6%	8.4%	414	5.0%
	Vein	48	0.0%	0.0%	3	0.0%
Before or after release of current guideline	Before	98,465	61.6%	61.6%	5,911	71.2%
	After	61,330	38.4%	38.4%	2,392	28.8%
Able to perform activities of at least 4 METs	No	90,260	56.5%	56.5%	5,713	68.8%
	Yes	69,535	43.5%	43.5%	2,590	31.2%
Tobacco use	Current smoker	18,806	12.4%	11.8%	1,028	12.4%
	Former smoker	62,067	40.9%	38.8%	3,834	46.2%
	Never smoker	70,966	46.7%	44.4%	2,976	35.8%
Probability of obstructive CAD		159,793	100.0%	100.0%	8,303	100.0%

Table 2: Summary of continuous variables.

	Mean ± SD
Age (years)	58.8 ± 15.2
Systolic Blood Pressure	129 ± 19.1
Diastolic Blood Pressure	73 ± 11.3
Body Mass Index	30.1 ± 7.6
Creatinine	1.02 ± 0.85
Area Deprivation Index	54.2 ± 24.6
Estimated METs of activity	5.22 ± 1.35
MICA risk estimate	0.021 ± 0.025
Previous patients seen by physician in clinic	1,587 ± 1,536
Predicted probability of obstructive CAD	0.129 ± 0.136

Table 3. Marginal results for each variable in our final model, with all other variables held at their medians.

For example, with all other variables at their respective medians, a visit on June 30, 2008 would have resulted in preoperative stress testing approximately 3.5% of the time, while a visit on June 30, 2018 would have resulted in preoperative stress testing approximately 1.3% of the time. Provider effects are summarized for space considerations; full marginal results by physician are included in the Supplemental Appendix.

Predictor	Value	Marginal Rate	95% CI	
MICA estimate > 1%	0	0.6%	0.6%	0.7%
	1	7.1%	6.7%	7.5%
RCRI (documented)	0	2.2%	2.0%	2.3%
	1	2.7%	2.5%	2.9%
	2	3.4%	3.0%	3.8%
	3	4.2%	3.5%	5.0%
	4	5.1%	4.0%	6.5%
	5	6.4%	4.7%	8.6%
Subjective assessment of patient function	1	2.1%	1.9%	2.2%
	2	2.4%	2.2%	2.5%
	3	2.8%	2.6%	3.0%
	4	3.2%	2.9%	3.5%
Estimated METs	2	3.3%	2.9%	3.7%
	4	2.7%	2.5%	2.9%
	8	1.8%	1.6%	1.9%
Body mass index	20	2.1%	1.9%	2.3%
	30	2.4%	2.2%	2.5%
	40	2.7%	2.5%	2.9%
Diastolic blood pressure	70	2.3%	2.2%	2.5%
	90	2.5%	2.3%	2.7%
	110	2.7%	2.4%	3.0%
Ischemic heart disease	No	2.1%	2.0%	2.3%
	Yes	3.6%	3.3%	3.9%
Congestive heart	No	2.4%	2.2%	2.5%

failure	Yes	2.1%	1.9%	2.3%
Area deprivation index	10	2.8%	2.5%	3.1%
	50	2.2%	2.0%	2.3%
	90	2.6%	2.4%	2.8%
Predicted probability of obstructive CAD	5%	2.6%	2.4%	2.7%
	10%	2.4%	2.3%	2.6%
	20%	2.2%	2.1%	2.4%
Tobacco use	Current smoker	2.6%	2.3%	2.8%
	Former smoker	2.5%	2.3%	2.7%
	Neither	2.2%	2.1%	2.4%
Date	2008.06.30	3.5%	3.2%	3.8%
	2013.06.30	2.6%	2.4%	2.8%
	2018.06.30	1.3%	1.2%	1.4%
Surgical category	Aortic	23.4%	6.0%	91.1%
	Peripheral vascular	8.7%	6.7%	11.3%
	Urologic	9.2%	8.3%	10.2%
	Other	1.9%	1.7%	2.0%
Physician (summary)	Lowest	1.0%	0.1%	4.4%
	5th percentile	1.2%	0.6%	2.6%
	Median	2.3%	2.1%	2.6%
	95th percentile	3.8%	3.2%	4.5%
	Highest	6.1%	2.7%	13.5%

Figure 1: Unadjusted rates of preoperative stress testing, by physician.

[Attached separately as Figure_01.tif]

Figure 2: Mean marginal rates of preoperative stress testing, by physician.

[Attached separately as Figure_02.tif]

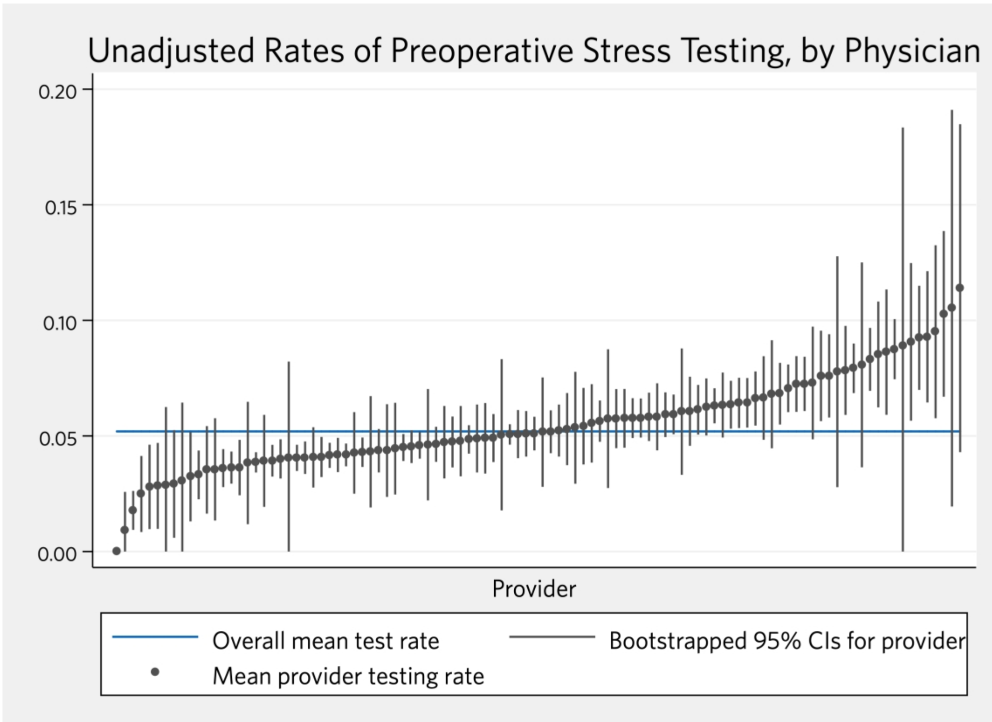


Figure 1: Unadjusted rates of preoperative stress testing, by physician

569x413mm (144 x 144 DPI)

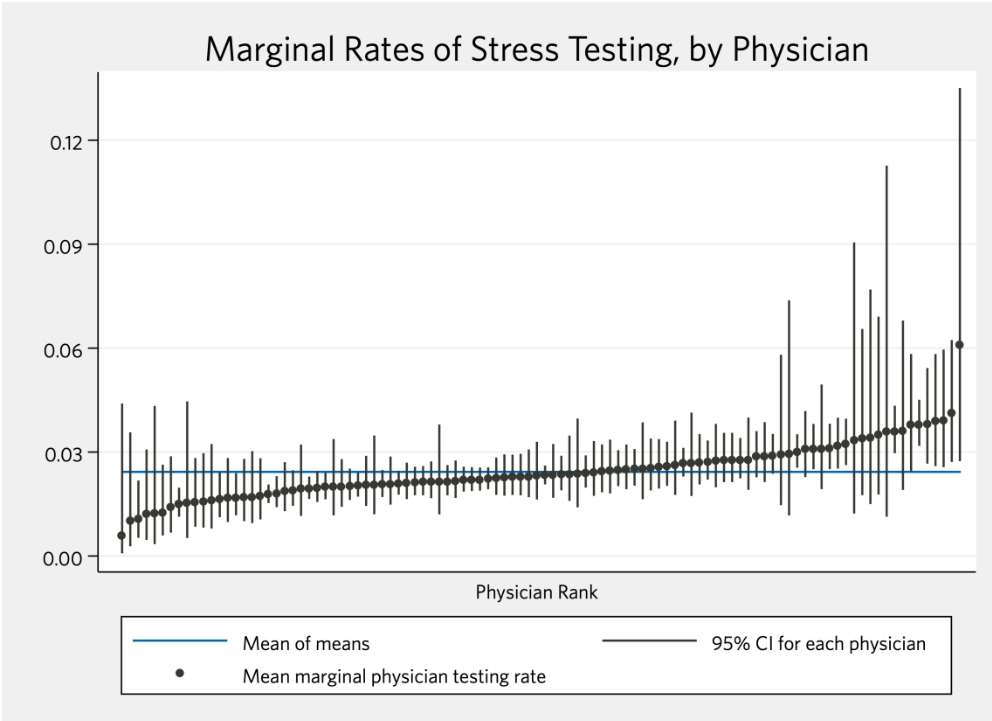


Figure 2: Mean marginal rates of preoperative stress testing, by physician

569x413mm (144 x 144 DPI)

Supplemental Material for: Variation in preoperative stress testing by patient, physician, and surgical type

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

1. Regression results, clustered by physician

Results have been edited/trimmed for clarity and brevity. Interaction terms have been replaced in the table below with exponents (^2) when applicable. All dichotomous variables (those prepended with “1.” in the table) use a value of 0 as the referent.

```
Random-effects logistic regression            Number of obs     =    154,171
Group variable: Physician_ID                  Number of groups  =        104
Observations per group:                                    min =          7
                                                           avg =    1,482.4
                                                           max =      8,245
```

Completed or cancelled test	Beta Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
1.aortic_surgery	2.522167	.6078339	4.15	0.002	1.166378	3.877956
1.peripheral_vascular_surgery	1.531945	.1152814	13.29	0.000	1.285334	1.778556
1.urologic_surgery	1.5871	.0657605	24.13	0.000	1.446111	1.728088
documented_RCRI	.2135885	.03032	7.04	0.000	.1506825	.2764944
1.Gupta_greater_than_1pct	2.422218	.0472362	51.28	0.000	2.329492	2.514944
Estimated METs	-.1030793	.0128133	-8.04	0.000	-.1284818	-.0776767
functional_class	.1444786	.0214833	6.73	0.000	.1020898	.1868674
ADI_national	-.0141123	.0025089	-5.62	0.000	-.0190783	-.0091463
ADI_national^2	.0001325	.0000221	5.99	0.000	.0000888	.0001763
BMI	.0120129	.0016333	7.36	0.000	.0088076	.0152182
DBP	.0037949	.0011446	3.32	0.001	.0015493	.0060406
1.ischemic_heart_disease	.5187915	.0390133	13.30	0.000	.4412007	.5963823
predicted_prob_of_CAD	-.9799052	.1105385	-8.86	0.000	-1.198045	-.761765
1.congestive_heart_failure	-.1196815	.0476752	-2.51	0.014	-.2142709	-.0250922
date	.002128	.0004881	4.36	0.000	.0011708	.0030853
date^2	-6.15e-08	1.25e-08	-4.92	0.000	-8.61e-08	-3.70e-08
1.current_smoker	.1436825	.0424903	3.38	0.001	.0603125	.2270526
1.former_smoker	.1288456	.0290712	4.43	0.000	.0718283	.1858628
constant	-23.48573	4.748773	-4.95	0.000	-32.79838	-14.17308
/lnsig2u	-3.592461	.2658187			-4.113895	-3.071026
sigma_u	.1659232	.0220527			.1278436	.2153451
rho	.0082988	.0021877			.0049434	.0138999

2. Physician marginal rates of stress testing

Due to space constraints, only selected physician marginal results are displayed in Table 3. Here we present full marginal results by physician. These are also displayed visually in Figure 2.

Physician Rank	Marginal Rate	95% CI
1	0.58%	0.08% – 4.40%
2	1.00%	0.28% – 3.57%
3	1.07%	0.52% – 2.18%
4	1.20%	0.47% – 3.08%
5	1.22%	0.34% – 4.33%
6	1.25%	0.59% – 2.64%
7	1.39%	0.68% – 2.88%
8	1.50%	1.13% – 1.97%
9	1.52%	0.52% – 4.46%
10	1.55%	0.85% – 2.83%
11	1.56%	0.82% – 2.97%
12	1.60%	0.79% – 3.23%
13	1.64%	1.12% – 2.42%
14	1.67%	0.98% – 2.83%
15	1.68%	1.17% – 2.39%
16	1.68%	1.01% – 2.80%
17	1.70%	0.95% – 3.03%
18	1.73%	1.05% – 2.82%
19	1.78%	1.53% – 2.06%
20	1.80%	1.41% – 2.30%
21	1.87%	1.29% – 2.71%
22	1.89%	1.45% – 2.47%
23	1.93%	1.16% – 3.22%
24	1.95%	1.65% – 2.29%

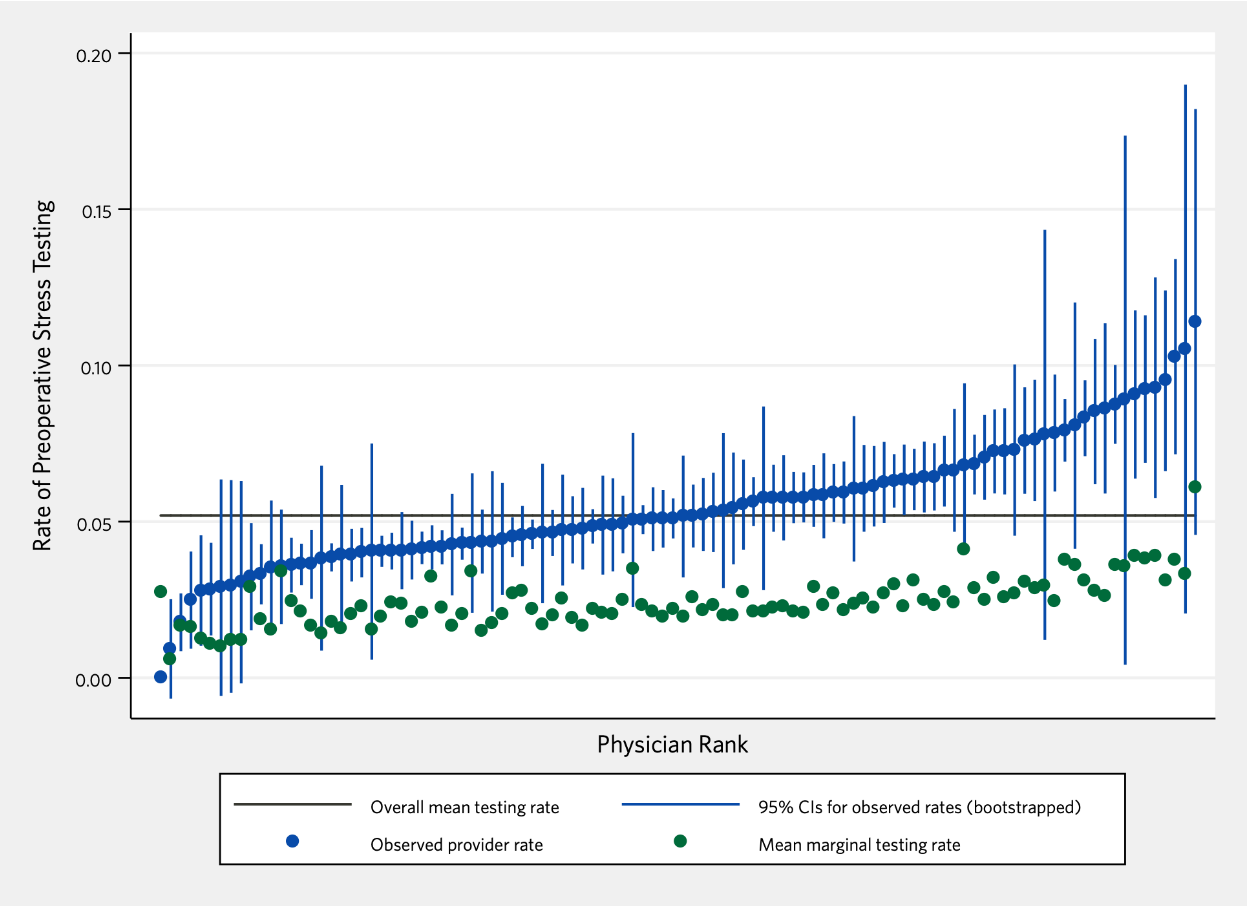
Physician Rank	Marginal Rate	95% CI
25	1.96%	1.56% – 2.45%
26	1.99%	1.63% – 2.42%
27	1.99%	1.17% – 3.38%
28	1.99%	1.42% – 2.80%
29	2.02%	1.62% – 2.52%
30	2.03%	1.71% – 2.42%
31	2.04%	1.45% – 2.89%
32	2.05%	1.20% – 3.48%
33	2.06%	1.70% – 2.48%
34	2.06%	1.48% – 2.87%
35	2.09%	1.77% – 2.46%
36	2.10%	1.64% – 2.69%
37	2.12%	1.75% – 2.57%
38	2.13%	1.75% – 2.60%
39	2.13%	1.67% – 2.73%
40	2.13%	1.20% – 3.80%
41	2.14%	1.75% – 2.62%
42	2.15%	1.68% – 2.76%
43	2.19%	1.86% – 2.58%
44	2.19%	1.87% – 2.56%
45	2.20%	1.89% – 2.55%
46	2.22%	1.93% – 2.56%
47	2.24%	1.76% – 2.84%
48	2.26%	1.74% – 2.93%
49	2.27%	1.76% – 2.93%
50	2.27%	1.75% – 2.95%
51	2.29%	1.70% – 3.08%

Physician Rank	Marginal Rate	95% CI
52	2.32%	1.63% – 3.30%
53	2.32%	2.07% – 2.61%
54	2.33%	1.68% – 3.23%
55	2.34%	1.89% – 2.89%
56	2.35%	1.59% – 3.48%
57	2.36%	1.40% – 3.97%
58	2.39%	1.97% – 2.90%
59	2.40%	1.74% – 3.32%
60	2.43%	1.84% – 3.22%
61	2.46%	1.81% – 3.36%
62	2.48%	2.02% – 3.05%
63	2.49%	1.93% – 3.22%
64	2.50%	2.03% – 3.09%
65	2.51%	1.64% – 3.86%
66	2.53%	1.89% – 3.39%
67	2.56%	1.95% – 3.37%
68	2.58%	2.03% – 3.30%
69	2.62%	1.76% – 3.91%
70	2.68%	2.29% – 3.12%
71	2.68%	1.73% – 4.14%
72	2.70%	2.07% – 3.52%
73	2.71%	2.20% – 3.33%
74	2.75%	1.98% – 3.81%
75	2.75%	2.13% – 3.55%
76	2.75%	2.13% – 3.55%
77	2.76%	2.24% – 3.41%
78	2.76%	1.91% – 4.00%

Physician Rank	Marginal Rate	95% CI
79	2.86%	2.27% – 3.60%
80	2.87%	2.14% – 3.86%
81	2.89%	2.38% – 3.52%
82	2.92%	1.47% – 5.81%
83	2.94%	1.17% – 7.38%
84	2.99%	2.55% – 3.51%
85	3.09%	2.28% – 4.19%
86	3.09%	2.50% – 3.81%
87	3.09%	1.93% – 4.95%
88	3.09%	2.51% – 3.82%
89	3.17%	2.52% – 3.99%
90	3.22%	2.62% – 3.96%
91	3.33%	1.23% – 9.05%
92	3.39%	1.76% – 6.55%
93	3.40%	1.50% – 7.69%
94	3.50%	1.77% – 6.91%
95	3.57%	1.13% – 11.27%
96	3.59%	2.97% – 4.34%
97	3.60%	1.91% – 6.79%
98	3.77%	2.44% – 5.83%
99	3.79%	3.18% – 4.51%
100	3.80%	2.67% – 5.43%
101	3.89%	2.60% – 5.83%
102	3.91%	2.56% – 5.96%
103	4.11%	2.71% – 6.23%
104	6.08%	2.74% – 13.51%

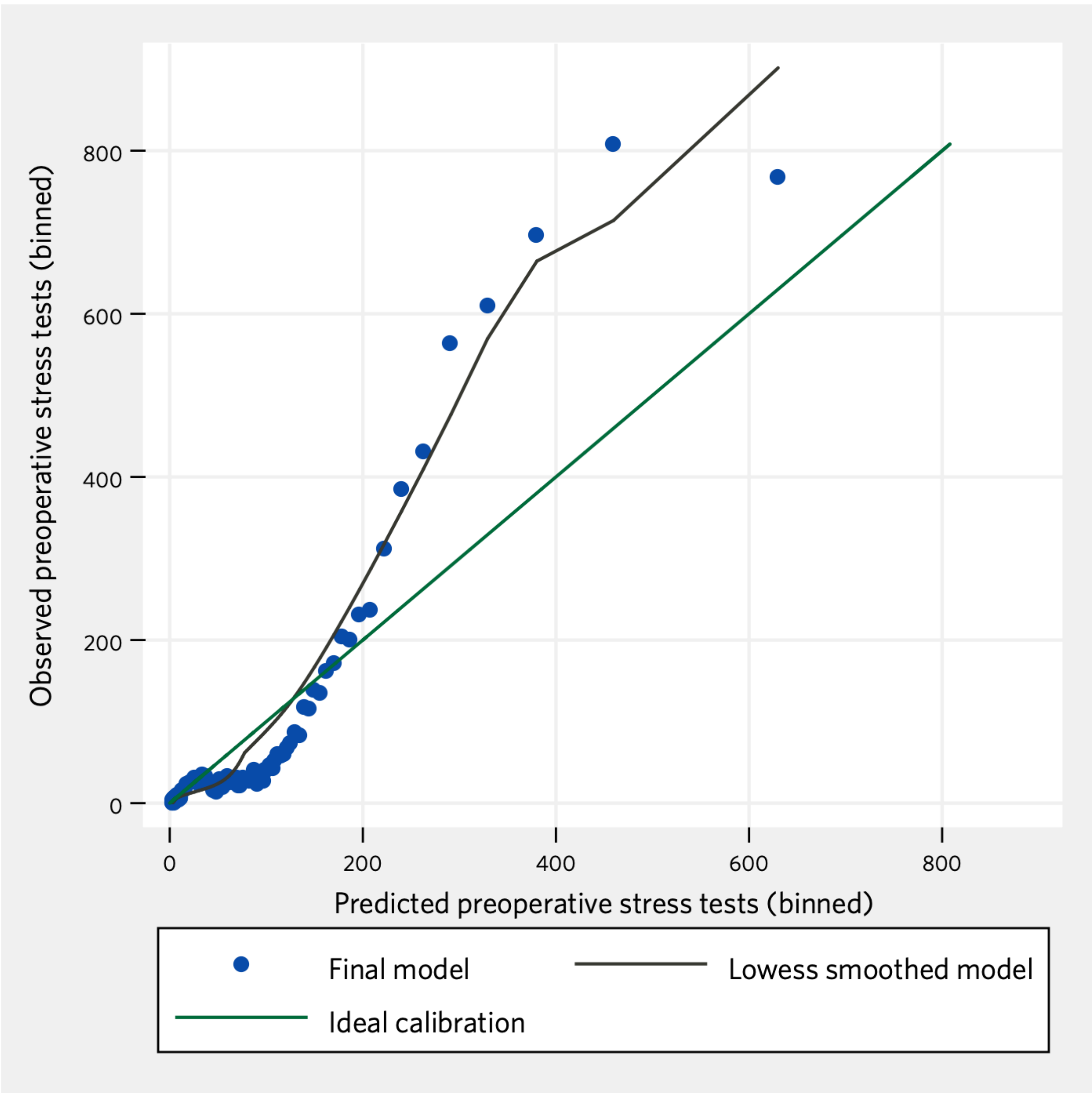
3. Physician rates of stress testing, unadjusted and marginal

Figure 1 demonstrates unadjusted rates of stress testing, and Figure 2 demonstrates marginal predictions for each physician, controlling for all other factors. Here we overlay the marginal results on the unadjusted results to demonstrate the effect of adjustment.



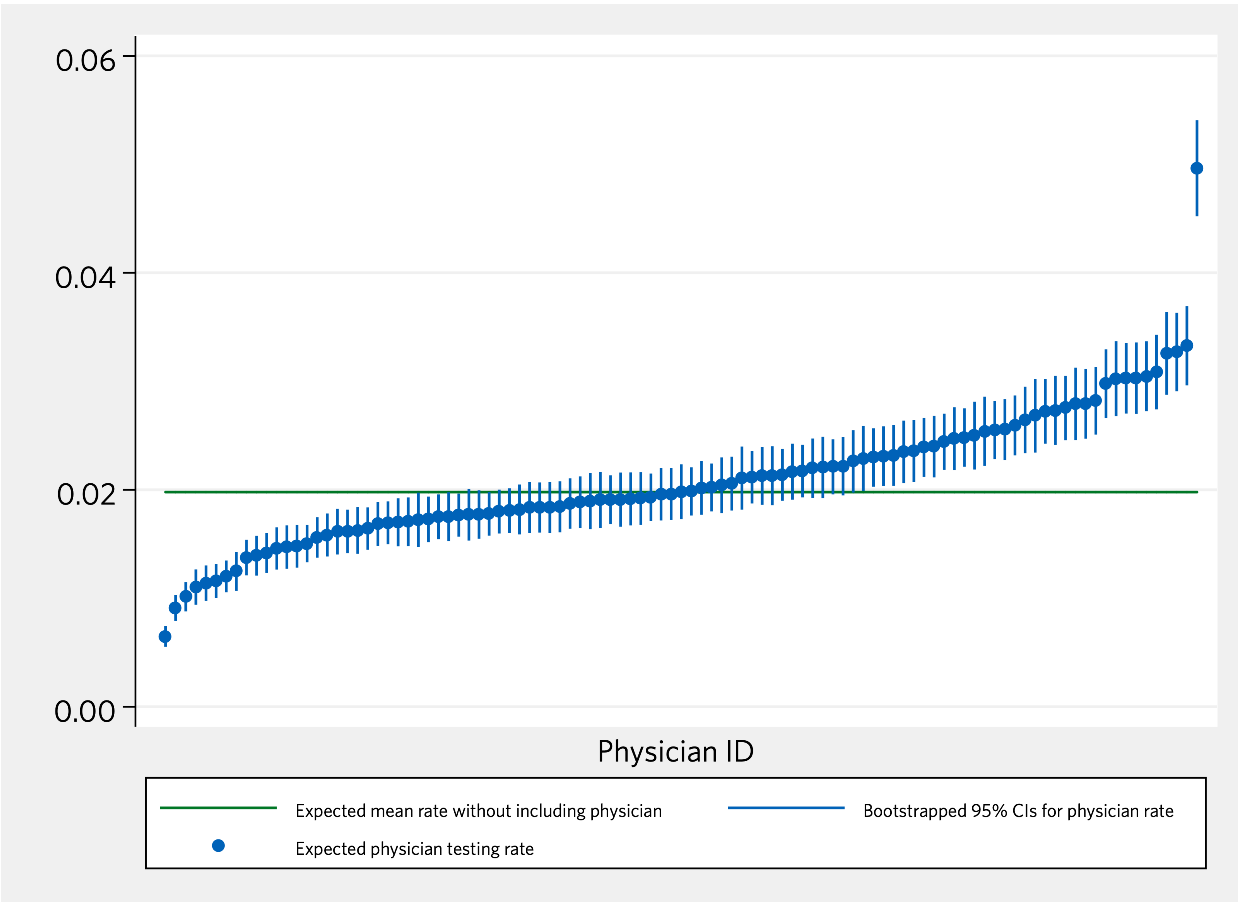
4. Calibration of final model

Based on the plot above, we suspected less-than-ideal calibration of our final model. We do not know of a universally accepted method to assess the calibration of a multilevel model on multiply-imputed data, but in most of our assessments this model fails Hosmer-Lemeshow Goodness-of-Fit testing, and the calibration plot shown here (binned into centiles) indeed suggests poor calibration. We emphasize again that our goal here is to explain variance in testing, not to guide future physicians in who should be referred for stress testing or to enable individual physician profiling.



5. Expected testing rates with an identical population

One way to contextualize provider effects is to imagine that each provider sees an identical panel of patients and estimate the consequent differences in outcomes. Here, we sampled 1,000 patients from our original population and estimated rates of stress testing if that same cohort were seen by each physician in our dataset. The overall mean is the expected rate for this small cohort without controlling for physician ID.



6. Sensitivity analysis excluding patients planned for aortic or vascular surgery

We repeated our analysis while excluding patients who were considered for aortic or vascular surgery. As in all models with dichotomous outcomes, the fixed variance leads to different effect sizes when using a different list of predictors. All effects are in the same direction as in our base-case analysis, as shown below. All changes in effect size are smaller than the smallest effect size in our base-case model (congestive heart failure). We have highlighted results where marginal rates differ from the base case by 0.02% or greater. This is an arbitrary threshold based on the intuition that a difference of less than 1 test per 500 visits is small. Due to rounding, some cells with less than a 0.2% absolute difference are displayed as differences of 0.2% in the cells below.

This analysis includes a total of 151,213 visits.

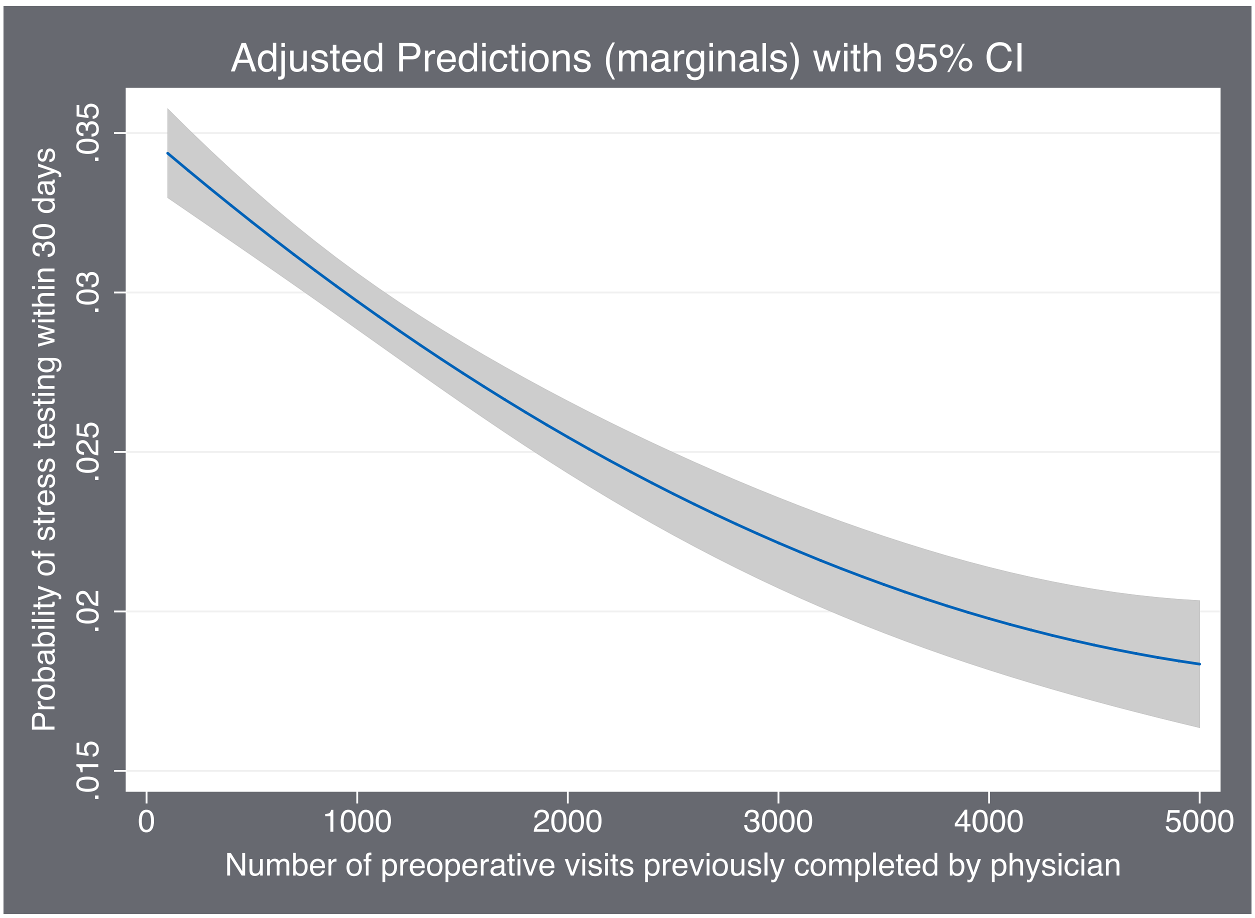
Predictor	Value	Mean marginal rate, base case	95% CI	Mean marginal rate with aortic and vascular surgery patients excluded	95% CI
MICA > 1%	0	0.6%	(0.6% – 0.7%)	0.6%	(0.5% – 0.7%)
	1	7.1%	(6.7% – 7.5%)	6.8%	(6.4% – 7.1%)
Documented RCRI	0	2.2%	(2.0% – 2.3%)	2.1%	(1.9% – 2.2%)
	1	2.7%	(2.5% – 2.9%)	2.6%	(2.4% – 2.8%)
	2	3.4%	(3.0% – 3.8%)	3.2%	(2.8% – 3.6%)
	3	4.2%	(3.5% – 5.0%)	3.9%	(3.3% – 4.7%)
	4	5.1%	(4.0% – 6.5%)	4.9%	(3.8% – 6.2%)
	5	6.4%	(4.7% – 8.6%)	6.0%	(4.4% – 8.2%)
Functional class	1	2.1%	(1.9% – 2.2%)	2.0%	(1.8% – 2.1%)
	2	2.4%	(2.2% – 2.5%)	2.2%	(2.1% – 2.4%)
	3	2.8%	(2.6% – 3.0%)	2.6%	(2.4% – 2.8%)
	4	3.2%	(2.9% – 3.5%)	2.9%	(2.6% – 3.3%)
Estimated metabolic equivalents	2	3.3%	(2.9% – 3.7%)	3.0%	(2.7% – 3.4%)
	4	2.7%	(2.5% – 2.9%)	2.5%	(2.3% – 2.7%)
	8	1.8%	(1.6% – 1.9%)	1.7%	(1.5% – 1.9%)
Body mass index	20	2.1%	(1.9% – 2.3%)	2.0%	(1.8% – 2.1%)
	30	2.4%	(2.2% – 2.5%)	2.2%	(2.1% – 2.4%)
	40	2.7%	(2.5% – 2.9%)	2.5%	(2.3% – 2.7%)

Predictor	Value	Mean marginal rate, base case	95% CI	Mean marginal rate with aortic and vascular surgery patients excluded	95% CI
Diastolic blood pressure	70	2.3%	(2.2% – 2.5%)	2.2%	(2.1% – 2.4%)
	90	2.5%	(2.3% – 2.7%)	2.4%	(2.2% – 2.6%)
	110	2.7%	(2.4% – 3.0%)	2.6%	(2.3% – 2.9%)
Ischemic heart disease	0	2.1%	(2.0% – 2.3%)	2.0%	(1.9% – 2.2%)
	1	3.6%	(3.3% – 3.9%)	3.3%	(3.1% – 3.7%)
Congestive heart failure	0	2.4%	(2.2% – 2.5%)	2.3%	(2.1% – 2.4%)
	1	2.1%	(1.9% – 2.3%)	2.0%	(1.8% – 2.2%)
Area deprivation index	10	2.8%	(2.5% – 3.1%)	2.6%	(2.4% – 2.9%)
	50	2.2%	(2.0% – 2.3%)	2.1%	(1.9% – 2.2%)
	90	2.6%	(2.4% – 2.8%)	2.4%	(2.2% – 2.7%)
Predicted probability of obstructive coronary artery disease	5%	2.6%	(2.4% – 2.7%)	2.4%	(2.3% – 2.6%)
	10%	2.4%	(2.3% – 2.6%)	2.3%	(2.2% – 2.5%)
	20%	2.2%	(2.1% – 2.4%)	2.1%	(1.9% – 2.2%)
Tobacco use	Current smoker	2.6%	(2.3% – 2.8%)	2.4%	(2.2% – 2.7%)
	Former smoker	2.5%	(2.3% – 2.7%)	2.4%	(2.2% – 2.6%)
	Neither	2.2%	(2.1% – 2.4%)	2.1%	(1.9% – 2.2%)
Date	2008.06.30	3.5%	(3.2% – 3.8%)	3.3%	(3.0% – 3.6%)
	2013.06.30	2.6%	(2.4% – 2.8%)	2.5%	(2.3% – 2.7%)
	2018.06.30	1.3%	(1.2% – 1.4%)	1.2%	(1.1% – 1.3%)
Surgical category	Aortic	23.4%	(6.0% –91.1%)	–	
	Peripheral vascular	8.7%	(6.7% –11.3%)	–	
	Urologic	9.2%	(8.3% –10.2%)	8.8%	(7.9% – 9.9%)
	Other	1.9%	(1.7% – 2.0%)	1.8%	(1.7% – 2.0%)
Physician (summary)	Lowest	1.0%	(0.1% – 4.4%)	0.5%	(0.1% – 3.6%)
	5th percentile	1.2%	(0.6% – 2.6%)	0.8%	(0.3% – 1.8%)
	Median	2.3%	(2.1% – 2.6%)	1.6%	(1.2% – 2.0%)

Predictor	Value	Mean marginal rate, base case	95% CI	Mean marginal rate with aortic and vascular surgery patients excluded	95% CI
	95th percentile	3.8%	(3.2% – 4.5%)	2.7%	(1.8% – 4.1%)
	Highest	6.1%	(2.7% –13.5%)	4.5%	(1.8% –11.3%)

7. Marginal testing rate as a function of physician experience

As with all datasets, our conclusions are a product of many decisions. For example, although we rejected physician experience as a predictor of testing rate in favor of date and a physician-specific random effect, reasonable investigators could disagree. To generate the graph below, we replaced date in our model with the number of visits each physician had completed between the beginning of our dataset and the visit in question (a proxy for preoperative clinic experience). We then computed and graphed marginal probabilities as described in our primary results.



STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2-3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4-5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	4-5
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5-8
Bias	9	Describe any efforts to address potential sources of bias	9-10
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5-10
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-10
		(*b*) Describe any methods used to examine subgroups and interactions	5-10
		(*c*) Explain how missing data were addressed	9-10
		(*d*) If applicable, explain how loss to follow-up was addressed	5
		(*e*) Describe any sensitivity analyses	5
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	5,10
		(b) Give reasons for non-participation at each stage	17
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	26-27
		(b) Indicate number of participants with missing data for each variable of interest	26
		(c) Summarise follow-up time (eg, average and total amount)	5
Outcome data	15*	Report numbers of outcome events or summary measures over time	10

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10-11, 26, 28
		(*b*) Report category boundaries when continuous variables were categorized	26
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	28
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11-12
Discussion			
Key results	18	Summarise key results with reference to study objectives	12-15
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	16-18
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	18
Generalisability	21	Discuss the generalisability (external validity) of the study results	17
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	1,19

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.