BMJ Open

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

Matthew A Pappas [iD],[1,2,3] Daniel I Sessler,[3,4] Andrew D Auerbach,[5] Michael W Kattan,[6] Alex Milinovich,[6] Eugene H Blackstone,[7] Michael B Rothberg[2,8]

For numbered affiliations see end of article.

Correspondence to
Dr Matthew A Pappas;
pappasm@ccf.org

ABSTRACT

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

Setting A dedicated preoperative risk assessment and optimisation clinic at a large integrated medical centre 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 optimisation 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 8303 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 ischaemic heart disease, tobacco use and body mass index. Patients living in either the most-deprived 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.

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 judgement 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 non-cardiac 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 non-cardiac surgery. This study was approved by the Cleveland Clinic Institutional Review Board (IRB number 18–1076).

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

Predictor variables

We theorised 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 non-linear or categorical relationships using visual examination of binned scatter plots. To avoid overfitting, we limited candidate predictors to fewer than one predictor variable per 15 preoperative stress tests, including tested interactions, non-linear 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 categorisation of surgeries using a previously published crosswalk.[2 10 11] Although different procedures likely have different intrinsic cardiac risk, we used the MICA categorisation of surgeries to avoid overfitting.[11 12] On 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 dichotomised at 1%. Although we used the American Society of Anesthesiologists, 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 semiquantitative 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 dichotomised 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 postresidency practice (a proxy for overall experience) and the number of previous encounters the physician had completed in our data set (a proxy for experience in preoperative risk assessment more specifically). For measures of time, we theorised 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 (ie, 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. 9 December 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-level and visit-level variables clustered by patient, including physician identifier (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 modelling

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 non-linear 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 analysed 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 that 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 data set included relatively few aortic surgeries and CIs 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 with the median, body mass index (BMI), diastolic blood pressure, existing diagnoses of ischaemic heart disease or congestive heart failure, estimated probability of obstructive coronary artery disease and tobacco use. Visits later in our data set 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 favour a model clustered by physician (BIC: 51040) compared with 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 three negations missed by natural language processing, 32 descriptions of historical symptoms that had prompted testing or intervention previously, 2 quotations of test reports (coronary catheterisation reports that included unstable angina as the indication for testing) and 13 cases of unstable angina.

Table 1 Patient and surgical characteristics

	Total	Per cent of category	Per cent of all visits	Completed preoperative stress test	Per cent of all preoperative stress tests
Age	159 795	100.0	100.0	8303	100.0
Sex					
Female	88 738	55.5	55.5	4079	49.1
Male	71 055	44.5	44.5	4224	50.9
Previous diagnosis of ischaemic heart disease					
No	128 505	80.4	80.4	4831	58.2
Yes	31 290	19.6	19.6	3472	41.8
Previous diagnosis of congestive heart failure					
No	146 556	91.7	91.7	6922	83.4
Yes	13 239	8.3	8.3	1381	16.6
Previous diagnosis of cerebrovascular disease					
No	141 519	88.6	88.6	6567	79.1
Yes	18 276	11.4	11.4	1736	20.9
Systolic blood pressure	159 488	100.0	99.8	8285	99.8
Diastolic blood pressure	159 481	100.0	99.8	8284	99.8
Body mass index	157 473	100.0	98.5	8155	98.2
Creatinine (RCRI categorisation)					
≤ 2.0 mg/dL	151 885	97.1	95.1	7695	92.7
> 2.0 mg/dL	4 487	2.9	2.8	542	6.5
Creatinine (MICA categorisation)					
≤ 1.5 mg/dL	144 369	90.3	90.3	7128	85.8
> 1.5 mg/dL	12 003	7.5	7.5	1109	13.4
Unknown	3423	2.1	2.1	66	0.8
Prescribed insulin					
No	147 610	92.4	92.4	7136	85.9
Yes	12 185	7.6	7.6	1167	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	7091	85.4
RCRI (documented)					
0	50 785	75.1	31.8	1548	18.6
1	12 642	18.7	7.9	988	11.9
2	3321	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	4006	3.7	2.5	98	1.2
4	1172	1.1	0.7	33	0.4

Table 1 Continued

	Total	Per cent of category	Per cent of all visits	Completed preoperative stress test	Per cent of all preoperative stress tests
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	7695	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	1739	20.9
2–3	8588	9.1	5.4	646	7.8
3	9310	9.9	5.8	732	8.8
3–4	1922	2.0	1.2	253	3.0
4	1999	2.1	1.3	175	2.1
Insurance					
Medicare	20 744	52.2	13.0	1192	14.4
Medicaid	2384	6.0	1.5	83	1.0
Private	14 764	37.1	9.2	307	3.7
Other listed insurer	1881	4.7	1.2	53	0.6
ASA class					
1	3325	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	7454	6.8	4.7	157	1.9
MICA surgical category					
Anorectal	1213	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	4780	4.1	3.0	15	0.2
Breast	8541	7.4	5.3	17	0.2
Cardiac	237	0.2	0.1	10	0.1
Ear, nose, throat	3640	3.1	2.3	19	0.2
Foregut/hepatopancreatobiliary	3252	2.8	2.0	37	0.4
Gallbladder, appendix, adrenals, or spleen	1974	1.7	1.2	17	0.2
Gynecologic	7458	6.4	4.7	51	0.6
Hernia	2778	2.4	1.7	16	0.2
Intestines	16 081	13.9	10.1	87	1.0
Neck	3076	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	1970	1.7	1.2	19	0.2
Peripheral vascular	1355	1.2	0.8	110	1.3
Skin	7126	6.1	4.5	78	0.9

Table 1 Continued

	Total	Per cent of category	Per cent of all visits	Completed preoperative stress test	Per cent of all preoperative stress tests
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	5911	71.2
After	61 330	38.4	38.4	2392	28.8
Able to perform activities of at least four METs					
No	90 260	56.5	56.5	5713	68.8
Yes	69 535	43.5	43.5	2590	31.2
Tobacco use					
Current smoker	18 806	12.4	11.8	1028	12.4
Former smoker	62 067	40.9	38.8	3834	46.2
Never smoker	70 966	46.7	44.4	2976	35.8
Probability of obstructive CAD	159 793	100.0	100.0	8303	100.0

ASA, American Society of Anesthesiologists; CAD, coronary artery disease; MET, estimated metabolic equivalent; MICA, myocardial infarction or cardiac arrest calculator; RCRI, revised cardiac risk index.

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 METs 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 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 dichotomisation 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,

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±1536
Predicted probability of obstructive CAD	0.129±0.136

CAD, coronary artery disease; MET, estimated metabolic equivalent; MICA, myocardial infarction or cardiac arrest calculator.
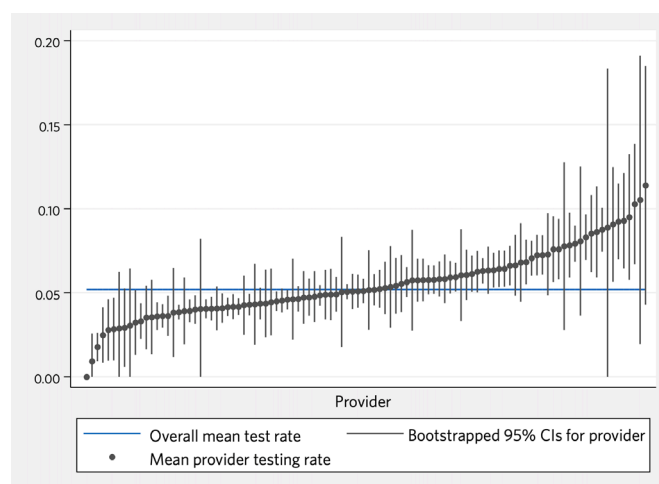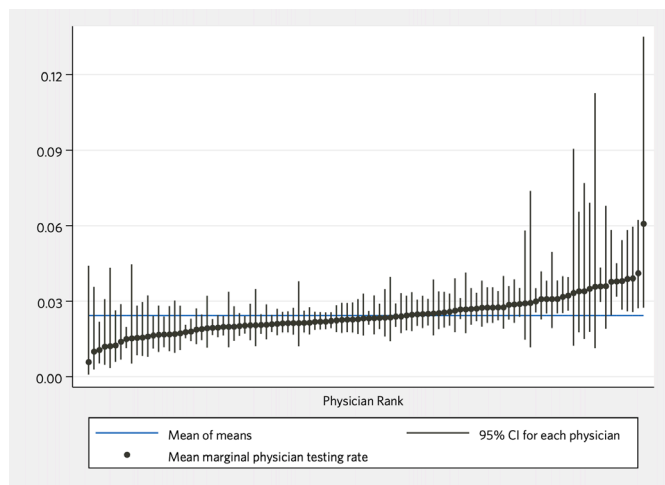

Figure 1 Unadjusted rates of preoperative stress testing, by physician.

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

Predictor	Value	Marginal rate	95% CI
MICA estimate>1%	0	0.6%	0.6 to 0.7
	1	7.1%	6.7 to 7.5
RCRI (documented)	0	2.2%	2.0 to 2.3
	1	2.7%	2.5 to 2.9
	2	3.4%	3.0 to 3.8
	3	4.2%	3.5 to 5.0
	4	5.1%	4.0 to 6.5
	5	6.4%	4.7 to 8.6
Subjective assessment of patient function	1	2.1%	1.9 to 2.2
	2	2.4%	2.2 to 2.5
	3	2.8%	2.6 to 3.0
	4	3.2%	2.9 to 3.5
Estimated METs	2	3.3%	2.9 to 3.7
	4	2.7%	2.5 to 2.9
	8	1.8%	1.6 to 1.9
Body mass index	20	2.1%	1.9 to 2.3
	30	2.4%	2.2 to 2.5
	40	2.7%	2.5 to 2.9
Diastolic blood pressure	70	2.3%	2.2 to 2.5
	90	2.5%	2.3 to 2.7
	110	2.7%	2.4 to 3.0
Ischaemic heart disease	No	2.1%	2.0 to 2.3
	Yes	3.6%	3.3 to 3.9
Congestive heart failure	No	2.4%	2.2 to 2.5
	Yes	2.1%	1.9 to 2.3
Area deprivation index	10	2.8%	2.5 to 3.1
	50	2.2%	2.0 to 2.3
	90	2.6%	2.4 to 2.8
Predicted probability of obstructive CAD	5%	2.6%	2.4 to 2.7
	10%	2.4%	2.3 to 2.6
	20%	2.2%	2.1 to 2.4
Tobacco use	Current smoker	2.6%	2.3 to 2.8
	Former smoker	2.5%	2.3 to 2.7
	Neither	2.2%	2.1 to 2.4
Date	30 June 2008	3.5%	3.2 to 3.8
	30 June 2013	2.6%	2.4 to 2.8
	30 June 2018	1.3%	1.2 to 1.4
Surgical category	Aortic	23.4%	6.0 to 91.1
	Peripheral vascular	8.7%	6.7 to 11.3
	Urologic	9.2%	8.3 to 10.2
	Other	1.9%	1.7 to 2.0
Physician (summary)	Lowest	1.0%	0.1 to 4.4
	5th percentile	1.2%	0.6 to 2.6
	Median	2.3%	2.1 to 2.6
	95th percentile	3.8%	3.2 to 4.5
	Highest	6.1%	2.7 to 13.5

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

CAD, coronary artery disease; MET, estimated metabolic equivalent; MICA, myocardial infarction or cardiac arrest calculator; RCRI, revised cardiac risk index.

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

while relying on cohorts with different calibration, or could be deliberately avoiding a stark dichotomisation 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 predictor was especially important 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 revascularised before vascular surgery.[19–21] However, patients undergoing vascular surgery are generally evaluated elsewhere at our institution, leaving our sample small and CIs 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 prasticing 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, ischaemic heart disease and a patient's census block group. Because each of these is 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 data set 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 template 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 non-cardiac surgery is an inefficient tool to screen for angina.

The frequency of stress testing declined over time in our data set, 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 orthopaedic surgery. Our analysis, thus, extends previous understanding by showing that the reduction is not limited to orthopaedic 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 data set is among the largest clinical data sets 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 revascularisation did not reduce cardiac risk in the largest randomised trial to date, beta blockers are more likely harmful than helpful and other interventions (eg, 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 data set, and we cannot analyse variation across health system or region, which can also be substantial.[38] As with any single-centre 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 centres 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 favour of a random effect for each physician and the date of each visit due to our prespecified analytic criteria (rejecting predictor variables that worsened 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 online supplemental appendix). Time, in particular, is rife with potential confounders of this sort. Recent work using other data sets 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-centre 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, ischaemic heart disease, congestive heart failure, BMI, 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.

Author affiliations
[1]Department of Hospital Medicine, Cleveland Clinic, Cleveland, Ohio, USA
[2]Center for Value-based Care Research, Cleveland Clinic, Cleveland, Ohio, USA
[3]Outcomes Research Consortium, Cleveland, Ohio, USA
[4]Department of Outcomes Research, Anesthesiology Institute, Cleveland Clinic, Cleveland, Ohio, USA
[5]Department of Hospital Medicine, University of California San Francisco, San Francisco, California, USA
[6]Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA
[7]Miller Family Heart and Vascular Institute, Cleveland Clinic, Cleveland, Ohio, USA
[8]Department of Internal Medicine, Cleveland Clinic, Cleveland, Ohio, USA

ORCID iD
Matthew A Pappas http://orcid.org/0000-0002-0353-1785

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;130:2215–45.

2 Pappas MA, Sessler DI, Rothberg MB. Anticipated rates and costs of Guideline-Concordant preoperative stress testing. *Anesth Analg* 2019;128:241–6.

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;129:889–900.

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;64:651–4.

5 Reilly DF, McNeely MJ, Doerner D, *et al*. Self-Reported exercise tolerance and the risk of serious perioperative complications. *Arch Intern Med* 1999;159:2185–92.

6 Melon CC, Eshtiaghi P, Luksun WJ, *et al*. Validated questionnaire vs physicians' judgment to estimate preoperative exercise capacity. *JAMA Intern Med* 2014;174:1507–8.

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;391:2631–40.

8 Eddy DM. Variations in physician practice: the role of uncertainty. *Health Aff* 1984;3:74–89.

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;165:94–102.

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;100:1043–9.

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;124:381–7.

12 Liu JB, Liu Y, Cohen ME, *et al*. Defining the intrinsic cardiac risks of operations to improve preoperative cardiac risk assessments. *Anesthesiology* 2018;128:283–92.

13 Oken MM, Creech RH, Tormey DC, Davis TE, *et al*. Toxicity and response criteria of the eastern cooperative Oncology Group. *Am J Clin Oncol* 1982;5:649–56.

14 Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible - The neighborhood atlas. *N Engl J Med* 2018;378:2456–8.

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:e3485.

16 Milinovich A, Kattan MW. Extracting and utilizing electronic health data from EPIC for research. *Ann Transl Med* 2018;6:42.

17 White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;30:377–99.

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;33:17–32.

19 Eagle KA, Singer DE, Brewster DC, *et al*. Dipyridamole-Thallium scanning in patients undergoing vascular surgery. optimizing preoperative evaluation of cardiac risk. *JAMA* 1987;257:2185–9.

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;199:223–33.

21 Boucher CA, Brewster DC, Darling RC, *et al*. Determination of cardiac risk by dipyridamole-thallium imaging before peripheral vascular surgery. *N Engl J Med* 1985;312:389–94.

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;257:73–80.

23 Sigmund AE, Stevens ER, Blitz JD, *et al*. Use of preoperative testing and physicians' response to professional Society guidance. *JAMA Intern Med* 2015;175:1352–9.

24 Rubin DS, Hughey R, Gerlach RM, *et al*. Frequency and outcomes of preoperative stress testing in total hip and knee arthroplasty from 2004 to 2017. *JAMA Cardiol* 2021;6:13–20.

25 Institute of Medicine. *Crossing the quality chasm: a new health system for the 21st century*. Washington, DC: The National Academies Press, 2001.

26 Hofer TP, Hayward RA, Greenfield S. The Unreliability of Individual Physician "Report Cards" for Assessing the Costs and Quality of Care of a Chronic Disease. *JAMA* 1999;281:2098–105.

27 Wennberg JE, Freeman JL, Culp WJ. Are hospital services rationed in New Haven or over-utilised in Boston? *Lancet* 1987;1:1185–9.

28 Saint S, Vaughn VM, Chopra V, *et al*. Perception of resources spent on defensive medicine and history of being Sued among hospitalists: results from a national survey. *J Hosp Med* 2018;13:26–9.

29 Feinstein AR. The 'chagrin factor' and qualitative decision analysis. *Arch Intern Med* 1985;145:1257–9.

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;173:407–16.

31 van der Weijden T, van Velsen M, Dinant G-J, *et al*. Unexplained Complaints in General Practice: Prevalence, Patients' Expectations, and Professionals' Test-Ordering Behavior. *Med Decis Making* 2003;23:226–31.

32 Little P, Dorward M, Warner G, *et al*. Importance of patient pressure and perceived pressure and perceived medical need for investigations, referral, and prescribing in primary care: nested observational study. *BMJ* 2004;328:444–6.

33 McFalls EO, Ward HB, Moritz TE, *et al*. Coronary-Artery revascularization before elective major vascular surgery. *N Engl J Med* 2004;351:2795–804.

34 Blessberger H, Kammler J, Domanovits H, *et al*. Perioperative beta-blockers for preventing surgery-related mortality and morbidity. *Cochrane Database Syst Rev* 2018;3:CD004476.

35 Berwanger O, Le Manach Y, 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;37:177–85.

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.

37 Kerr EA, Chen J, Sussman JB, *et al*. Stress testing before low-risk surgery: so many recommendations, so little overuse. *JAMA Intern Med* 2015;175:645.

38 Home - Dartmouth Atlas of Health Care. Available: https://www.dartmouthatlas.org/ [Accessed 11 Nov 2019].

39 Palpacuer C, Hammas K, Duprez R, *et al*. Vibration of effects from diverse inclusion/exclusion criteria and analytical choices: 9216 different ways to perform an indirect comparison meta-analysis. *BMC Med* 2019;17:174.