Published in final edited form as:

Am J Med. 2022 July ; 135(7): e182–e193. doi:10.1016/j.amjmed.2022.02.036.

Clinician attitudes and beliefs associated with more aggressive diagnostic testing

Deborah Korenstein, MD[a], **Laura D. Scherer, PhD**[b,c,d], **Andrew Foy, MD**[e,f], **Lisa Pineles, MA**[g], **Alison D. Lydecker, MPH**[g], **Jill Owczarzak, PhDh**[h], **Larry Magder, PhD**[g], **Jessica P. Brown, PhD**[h], **Christopher D. Pfeiffer, MD, MHS**[i,j], **Christopher Terndrup, MD**[k], **Luci Leykum, MD, MBA**[l,m], **Deborah Stevens, LCSW-C, MPH**[g], **David A. Feldstein, MD**[n], **Scott A. Weisenberg, MD**[o], **Jonathan D. Baghdadi, MD, PhD**[g,p], **Daniel J Morgan, MD, MS**[g,p]

[a]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY

[b]Adult and Child Consortium of Health Outcomes Research and Delivery Science (ACCORDS), University of Colorado School of Medicine, Aurora, CO

[c]Division of Cardiology, University of Colorado School of Medicine, Aurora, CO

[d]Center of Innovation for Veteran-Centered and Value-Driven Care, VA Denver, Denver, CO

[e]Department of Medicine, Penn State College of Medicine, Hershey, PA

[f]Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA

[g]Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD

[h]Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

[i]Division of Infectious Diseases, Department of Medicine, Oregon Health & Science University, Portland, OR

[j]Division of Hospital and Specialty Medicine, VA Portland Health Care System, Portland, OR

Correspondence to Deborah Korenstein, korenstd@mskcc.org, Memorial Sloan Kettering Cancer Center, 485 Lexington Avenue, 2nd Floor, New York, NY 10017, Tel: 1-646-888-8147, Fax: 1-929-321-1518.

Author contributions: DK, DM, and LP planned the study; all authors contributed to survey development. LP, DS, and DM conducted the study. LM and AL designed and performed the analysis. DK drafted the manuscript and all authors contributed to revisions. DK and DM are the guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted, and that this manuscript represents a full and honest report of the findings.

Competing interests: Dr. Korenstein's work on this project was supported in part by a Cancer Center Support Grant from the National Cancer Institute to Memorial Sloan Kettering Cancer Center (P30 CA008748). CDF reports serving as the local site investigator for a Pfizer-sponsored C. difficile vaccine trial (protocol #B5091007) since July 2020 under a Cooperative Research and Development Agreement (CRADA) to VA Portland. DK discloses that her spouse serves on the scientific advisory board of Vedanta Biosciences and has equity interest, serves on the scientific advisory board of PFL-NYC and provides consulting for Fimbrion.
All other authors declare no financial relationships with any organizations that might have an interest in the submitted work in the previous three years and no other relationships or activities that could appear to have influenced the submitted work.

[k]Division of General Internal Medicine & Geriatrics, Department of Medicine, Oregon Health & Science University, Portland, OR

[l]Department of Medicine, Dell Medical School, the University of Texas at Austin, Austin, TX

[m]South Texas Veterans Health Care System, San Antonio, TX

[n]Department of Medicine, University of Wisconsin School of Medicine and Public Health, Madison, WI

[o]Department of Medicine, New York University Grossman School of Medicine, New York, NY

[p]VA Maryland Healthcare System, Baltimore, MD

Abstract

Background: Variation in clinicians' diagnostic test utilization is incompletely explained by demographics and likely relates to cognitive characteristics. We explored clinician factors associated with diagnostic test utilization

Methods: Self-administered survey of attitudes, cognitive characteristics and reported likelihood of test ordering in common scenarios; frequency of lipid and liver testing in patients on statin therapy. Participants were 552 primary care physicians, nurse practitioners (NP), and physician assistants [PA] from practices in 8 US states across 3 regions, from Measures included: 1 June 2018 to 26 November 2019. We measured Testing Likelihood Score: the mean of 4 responses to testing frequency and self-reported testing frequency in patients on statins.

Results: Respondents were 52.4% residents, 36.6% attendings, and 11.0% NP/PAs; most were White (53.6%) or Asian (25.5%). Median age was 32; 53.1% were female. Participants reported ordering tests for a median of 20% (stress tests) to 90% (mammograms) of patients; Testing Likelihood Scores varied widely (median 54%, IQR 43%−69%). Higher scores were associated with geography, training type, low numeracy, high malpractice fear, high medical maximizer score, high stress from uncertainty, high concern about bad outcomes, and low acknowledgment of medical uncertainty. More frequent testing of lipids and liver tests was associated with low numeracy, high medical maximizer score, high malpractice fear, and low acknowledgment of uncertainty.

Conclusions: Clinician variation in testing was common, with more aggressive testing consistently associated with low numeracy, being a medical maximizer, and low acknowledgment of uncertainty. Efforts to reduce undue variations in testing should consider clinician cognitive drivers.

Keywords

utilization; medical overuse; clinician characteristics

Introduction:

Appropriate utilization of diagnostic tests is critical to high-value care. Overly aggressive testing can lead to overdiagnosis and cascades of further tests and treatments, resulting in patient harm.[1,2] Conversely, failure to provide needed testing may miss diagnoses.

Recognizing the central role of screening and diagnostic tests, many high-value care efforts have focused on improving testing decisions.[3,4]

Higher levels of testing often do not improve outcomes, implying excess testing is low-value care.[5] Utilization rates vary widely across regions and among centers within regions, suggesting influence of both the broad community and the local environment on practice.[3,6–8] Within centers, there is also wide variation in practice patterns among individual clinicians.[9–11] Demographic factors explain little of this variation.[12] It is likely that clinician thought patterns, attitudes, and beliefs influence clinical decisions. For example, physician belief in evidence-based medicine, perceived patient preferences, and personality traits are all associated with medication prescribing.[13–15] More broadly, the ability to characterize clinicians as generally high or low utilizers across services supports the importance of these personal characteristics.[16–18] While efforts to optimize utilization of diagnostic tests must address drivers at all levels, [19] many efforts at improving healthcare value to date have targeted the health delivery system or the local environment[20], and even those targeting clinicians remain agnostic to specific clinician characteristics.[21,22] Despite widespread attention and multiple efforts, little progress has been made at improving healthcare value and reducing overuse.[23]

Better understanding of traits associated with clinical practice patterns can improve clinician self-awareness and inform interventions to optimize care delivery. To explore clinician factors associated with test utilization, we conducted a survey study of US primary care clinicians across 8 US states.

Methods:

Survey

We developed a survey to assess clinician test understanding, the process of making a diagnosis using probability, and perceptions of actions taken by practicing clinicians in similar scenarios. The survey included items regarding basic demographic characteristics, educational background, past malpractice suits, and practice setting, as well as other items related to perceptions around test and treatment efficacy that are presented elsewhere.[24]

A draft survey was developed by primary investigators based on previous surveys of risk understanding.[25–30] The survey was further revised by an expert panel during an in-person meeting and two conference calls and then piloted with ten clinicians for comprehension and interpretation questions.

Clinician self-reported test ordering

The survey assessed testing decisions commonly encountered by primary care clinicians in routine practice for which appropriateness was uncertain.[24] Individual questions assessed how often clinicians consider using mammograms for low-risk breast cancer screening, stress testing for low-risk cardiac ischemia, chest x-rays for high-risk pneumonia and urine cultures for very low-risk urinary tract infection (Appendix Figure 1). For each scenario, participants were asked how often they order the test for their own patients with similar presentations, estimated as a percentage between 0 and 100.

Testing Likelihood score

We calculated the mean of the responses from the 4 testing questions to create a *Testing Likelihood Score* for each participant. The combined score was the mean of these percentages across four questions and therefore could also range from 0 to 100. If a participant did not respond to each of the four questions, the mean was calculated over the available responses. In an additional sensitivity analysis, we created Standardized Testing Likelihood Scores to account for variable distributions among responses to separate questions. Each of the 4 testing questions was individually standardized by subtracting the mean of the testing question from the value, and then dividing the difference by the standard deviation of the testing question. This resulted in each standardized testing question having a mean of 0 and a standard deviation of 1. We then calculated the mean of these 4 standardized testing questions to create a Standardized Testing Likelihood Score.

Two additional questions about test-ordering were not linked to a clinical scenario; participants were asked how often they order lipid panels and liver function tests (LFT) in patients on statin therapy. Responses were open-ended in surveys (i.e., “every ___ months”) and were dichotomized to fewer than 12 months vs 12 or more months for analysis based on both common practice and on the distribution of responses.

Attitudes and cognitive characteristics

The survey included previously used questions and scales including assessment of level of burnout[31], degree of comfort with uncertainty,[32] fear of malpractice,[33] the Medical Maximizer-Minimizer Scale (Modified Version),[34] the Risk Taking Scale,[35] and a numeracy score.[36] The Medical Maximizer-Minimizer Scale was originally designed to assess an individual patient’s orientation toward healthcare. Maximizers tend to prefer medical action even when there is treatment burden and a low chance of benefit, and minimizers prefer to avoid medical intervention especially when there is treatment burden and a low chance of benefit. This scale was modified for this study its developer (LS) and uses 8 questions to assess clinician maximizing-minimizing orientation. Finally, the survey included a newly created item related to participant recognition of uncertainty in medicine. (Appendix Figure 2)

Enrollment procedure

After obtaining IRB approval at each of three coordinating sites, we enrolled clinicians as described elsewhere.[24] The survey was administered in paper format at sites in multiple states near Baltimore, MD (Mid-Atlantic), Portland, Oregon (Pacific Northwest), and San Antonio, TX (Texas). Respondents were provided with a US $50 gift card for completion, if permitted by their employer. Clinicians who did not complete the survey after three subsequent contacts were considered non-participants.

A sample size of 500 was planned to provide generalizable results across enrollment sites. The target sample was surpassed while we collected outstanding surveys.

Statistical analysis

Survey responses were entered into a REDCap database with double data entry. Several clinician characteristics -- including Risk Taking score, fear of malpractice, Medical Maximizer-Minimizer Scale, numeracy score, burnout score, stress from uncertainty and concern about bad outcomes -- were divided into Low, Medium and High groups based on tertiles.

We evaluated associations between clinician personal characteristics and 3 measures of testing aggressiveness: 1) the Testing Likelihood score, 2) the frequency of LFT testing in patients on statin therapy and 3) the frequency of lipid testing in patients on statin therapy. We characterized high utilizers by describing rates of responses in the top quintile of Testing Likelihood scores.

We compared those who completed all key survey questions and those who did not using Chi-square tests. Associations between clinician characteristics and Testing Likelihood score were measured using Analysis of Variance. Associations between clinical characteristics and frequency of routinely checking lipid panels and LFTs were measured using the Chi-square test or Fisher's exact test, as appropriate. All statistical tests were two-tailed and p values <0.05 were considered statistically significant. Because we sought to describe all factors associated with testing aggressiveness regardless of collinearity, we did not perform multivariable analysis.

All statistical analyses were performed with Stata 15.1 (Stata Corp., College Station, TX). Role of the funding source: This study was funded by a National Institutes of Health New Innovator award. The funder had no role in study design, completion or analysis. Patients and the public were not involved in this research. Institutional review board approval was obtained at each of the 3 coordinating sites (Baltimore, Maryland; San Antonio, Texas; and Portland, Oregon), for verbal informed consent with a waiver of documentation.

Results

Participant Demographics

The survey was offered to 723 primary care physicians, nurse practitioners (NP) and physician assistants (PA) practicing in Delaware, Maryland, Oregon, Pennsylvania, Texas, Virginia, Washington and the District of Columbia (Table 1). Overall response rate was 81% (585/723). Of the 585 clinicians who returned the survey, we excluded 33 who did not complete all questions necessary for analysis (final n=552)[24]. Respondents were 52.4% MD or DO residents (289/552), 36.6% MD or DO attendings (202/552), and 11.0% NPs or PAs (61/552). 53.6% (296/552) identified as White, 25.5% (141/552) Asian, 8.2% (45/552) Hispanic, 6.7% (37/552) Black, and 3.4% (19/552) more than one race; 2.2% (12/552) declined to self-identify. Median age was 32 (IQR 29–44) and 53.1% (292/550) were female (Table 1).

The survey required a self-reported median of 20 minutes to complete (IQR 15–25). Compared to those who completed all items, the 33 respondents who did not complete the survey were more likely to be female (81% non-completers vs. 53% final cohort, p=0.003),

have practiced >10 years (50% non-completers vs. 27% final cohort, p=0.019), or to be NPs or PAs (41% non-completers vs. 11% final cohort, p<0.001).

Testing practices:

Rates of test ordering varied by clinical scenario: clinicians reported ordering stress tests a median of 20% (IQR 10–50) of the time, chest x-rays a median of 90% (IQR 75– 100) of the time, mammograms a median of 50% (IQR 15–99) of the time, and urine cultures a median of 50% (IQR 10–90) of the time (Figure 1).

For patients on statin therapy, both LFTs and lipid panels were reported to be ordered a median of every 12 months (range 0.5 to never). Testing Likelihood scores varied, with a median of 54 (IQR 43–69). Fourteen of the 552 respondents (2.5%) were missing data from at least one clinical scenario.

Associations between personal characteristics and Testing Likelihood scores

Mean Testing Likelihood scores were significantly associated with study site region, with less testing aggressiveness in the Pacific Northwest compared to other regions. On average, NP/PAs had higher mean testing likelihood scores than MD/DOs and clinicians with post-graduate training in Family Medicine had higher scores than those with training in Internal Medicine. In addition, higher scores were associated with the following attitudes and cognitive traits: lower numeracy, higher fear of malpractice, higher medical maximizer score, higher stress from uncertainty, higher concern about bad outcomes, and low acknowledgment of uncertainty in medicine. (Table 2) Results were similar when outcomes were analyzed as continuous variables and using the standardized score analysis, with significant correlations with the same factors (Appendix Tables 1 and 2).

Associations between personal characteristics and ordering of lipids and LFTs

Self-reported frequency of checking lipids and LFTs in asymptomatic patients treated with statins had similar correlations with US region and degree program but not type of post-graduate training. (Table 3) More frequent ordering of both lipids and LFTs was associated with lower numeracy, higher medical maximizer score, and low acknowledgment of uncertainty in medicine (Table 3). In addition, malpractice fear was associated with ordering both LFTs and lipids. No other cognitive traits were significantly associated with ordering frequency.

We explored correlations between Testing Likelihood Scores and the frequency of lipid and LFT testing. We found higher Testing Likelihood Scores among participants who reported testing lipids (63 vs. 51) and LFTs (63 vs. 51) more often than in those who tested less often (P<0.01 for both comparisons), suggesting that some participants had a general predilection toward testing across different contexts.

Discussion:

In a survey of over 500 clinicians from eight US states, we found great variation in reported use of testing with more aggressive testing consistently associated with low numeracy, being

a medical maximizer, malpractice fear, and low acknowledgment of uncertainty in medicine. Risk intolerance and discomfort with uncertainty were also associated with aggressiveness. The observed correlation across different patient testing decisions suggests that individual clinician characteristics may drive general testing aggressiveness.

In our study, clinician attitudes and beliefs were associated with practice patterns. We, like others, found that malpractice fear was associated with more testing although notably, testing was not associated with actually having been sued for malpractice.[37–39] Among personal characteristics, poor numeracy and being a medical maximizer correlated consistently with testing aggressiveness. The relationship between these clinician characteristics and testing are not well described in the literature. Poor numeracy among medical students and doctors has been demonstrated in several studies and correlates with poor understanding of risk.[40–42] Misunderstanding risk may lead clinicians with poor numeracy to counsel patients sub-optimally[43], which may lead to more aggressive testing. Further, while clinicians generally overestimate the accuracy of diagnostic tests[30], those with poor numeracy may more grossly overestimate accuracy and consequently overvalue testing.

Our study represents the first evaluation of the modified Medical Maximizer-Minimizer Scale in clinicians. Patients with higher maximizer scores are more likely to pursue a range of health interventions, including high-value, preference-sensitive, and low-value care, and those with low scores may avoid care regardless of appropriateness. [44–46] It is likely that clinicians across the maximizer-minimizer spectrum manifest similar care preferences, for both themselves and their patients. While several important characteristics may underly being a medical maximizer, we suspect it reflects important biases that influence clinical decision-making. If upon further evaluation, the Medical Maximizer-Minimizer Scale for clinicians consistently predicts testing aggressiveness, it could potentially be leveraged in clinical contexts where leaders seek to minimize undue practice variations.

Our findings regarding clinician attitudes toward uncertainty merit exploration. Past studies of clinician uncertainty employing the Physicians Reactions to Uncertainty Scale have produced mixed findings regarding the association between uncertainty and utilization in terms of patient charges, specialty referrals, self-reported test ordering, and evaluations of incidental findings.[32,47–49] Our study findings were similarly mixed, with higher mean Testing Likelihood among participants with higher scores but no association of uncertainty scores with lipid and LFT ordering. However, the Physicians Reactions to Uncertainty Scale addresses only emotional reactions to uncertainty, while clinician responses to uncertainty can be emotional, cognitive and ethical[50]; the association between tolerance of uncertainty and clinician well-being emphasizes the importance of a range of reactions[51]. In contrast, our novel question assessing clinician acknowledgment of uncertainty in clinical practice was associated with less aggressive testing (and conversely, denying uncertainty in medicine was associated with more aggressive testing). Little is known about the cognitive underpinnings of this response[50]. Independent of emotional responses, it is possible that clinicians who recognize that uncertainty is part of medical practice are less driven to pursue elusive certainty through aggressive testing.[49] Cognitive responses to uncertainty may in fact drive testing and should be further investigated.

Some characteristics were inconsistently predictive of testing aggressiveness in our study. Personal risk intolerance has correlated with clinical practice in prior studies. Among emergency department physicians, the most risk averse had higher admission rates and more use of cardiac markers in patients with chest pain and more use of imaging in patients with abdominal pain [52,53]. In our study, however, risk-taking scores were not associated with testing outcomes, perhaps because testing in a primary care setting has different drivers than in other settings.

While our study was designed to evaluate individual characteristics, we found demographic factors that correlated with testing aggressiveness. Like others, we found differences in testing aggressiveness based on geographic area[54]. Age and years in practice were not associated with testing, except for more frequent LFT ordering for patients on statins by older clinicians. These findings, while mixed, suggest experience does not lead to less testing and may reflect shifts in training and the durability of early practice patterns[55].

We also found that NP and PA clinicians tested more aggressively than those with MD or DO degrees. While few studies have compared clinical practice patterns between physician and non-physician practitioners, they provide similar levels of low-value care.[56] The differences seen in our study may relate to our focus on services that are discretionary or of unclear value rather than on those that are broadly acknowledged as excessive. We studied these services to gain insight into diagnostic reasoning rather than to evaluate guideline adherence, which was the focus of older studies. Amidst growth in care delivery by non-physicians, future studies should further explore these differences.[57] Our findings regarding post-graduate training among physicians were inconsistent, with slightly higher Testing Likelihood scores among those trained in family medicine compared to internal medicine. These findings stand in contrast to other studies that found fewer referrals and less spending on low-value care and overall care by family medicine-trained physicians.[16,47,58] Again the difference in our results may relate to our focus on services of uncertain value, which may expose between-group differences in cognitive processes rather than guideline adherence.

Our study has limitations. While overall response rate was high, the small fraction of respondents who did not fully complete the survey were more likely to be female, NPs or PAs, or in practice more than 10 years. All measures of testing practice were self-reported, which may contain bias. Our vignettes were designed to explore nuanced clinical situations, so while more aggressive testing may reflect overuse, the appropriate level of testing is not known. Further, we created Testing Likelihood Scores for purposes of analysis by pooling responses from vignettes covering diverse topics for which responses varied widely, creating a number with inexact meaning that may be biased. However, analyzing the data in different ways led to similar correlations with determinants of testing aggressiveness, suggesting that this measure is consistent. Finally, while validity was extensively optimized via a multidisciplinary expert panel, reliability of our novel survey was not assessed; we used previously validated instruments to measure clinician characteristics, but also included a few novel items.

Conclusion:

In conclusion, testing aggressiveness among primary care clinicians correlates consistently with poor numeracy, being a medical maximizer, and lack of strong recognition of the uncertainty in medical practice. To improve healthcare value we must address the spectrum of drivers of overuse, including individual cognitive factors. Understanding clinician attitudes and beliefs associated with testing is an important first step toward change.

Acknowledgements

There were no contributors other than listed authors.

This project was funded by NIH NLM DP2LM012890 (New Innovator Award) PI Morgan. The funder had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

APPENDIX

Appendices: Clinician attitudes and beliefs associated with more aggressive diagnostic testing

Appendix Table 1:

Correlations of clinician characteristics with Mean Testing Likelihood Score. The likelihood test score is the average of 4 individual testing scores provided by the respondent.

Variable	Range	Correlation	P-value
Age (years)	24 to 75	0.08	0.05
Study Site	Pacific northwest Mid-Atlantic Texas	0.21	<0.01
Degree and training	MD or DO resident MD or DO attending NP or PA	0.11	<0.01
Years in Practice (since graduation)	0 to 46	0.07	0.09
Ever sued for malpractice	Yes No	−0.02	0.63
Numeracy Score	0 to 3	−0.16	<0.01
Risk Taking Score	6 to 32	−0.05	0.26
Fear of Malpractice	6 to 30	0.10	0.02
Medical Maximizer-Minimizer Scale	1 to 7	0.21	<0.01
Burnout score	1 to 5	−0.01	0.74

The Revised Physicians' Reactions to Uncertainty Scales

Variable	Range	Correlation	P-value
Stress from uncertainty subscale	3 to 18	0.07	0.12
Concern about bad outcomes subscale	3 to 18	0.13	<0.01
There is often uncertainty in the practice of medicine	1 to 6	−0.16	<0.01

Appendix Table 2:

Mean Standardized Testing Likelihood score and associations with clinician characteristics on bivariable analysis.

Four individual testing scores provided by the respondent were standardized separately and then the mean across the 4 scores was calculated.

Variable	Participants	Mean Standardized Testing Likelihood Score (SD)	P-value*
All	All (n=552)	0.00004 (0.58)	
Age (years)	<30 (n=171)	−0.02 (0.51)	0.50
	30–39 (n=208)	−0.03 (0.61	
	≥40 (n=167)	0.04 (0.59)	
Study Site	Pacific northwest (n=112)	−0.27 (0.53)	<0.01
	Mid-Atlantic (n=304)	0.06 (0.57)	
	Texas (n=136)	0.10 (0.56)	
Degree and training	MD or DO resident (n=289)	−0.01 (0.55)	0.02
	MD or DO attending (n=202)	−0.04 (0.61)	
	NP or PA (n=61)	0.20 (0.54)	
Type of post-graduate training (MDs and DOs)	Internal Medicine (n=335)	−0.05 (0.57)	0.19
	Family Medicine (n=142)	0.02 (0.59)	
Years in Practice (since graduation)	<3 years (n=239)	−0.03 (0.54)	0.37
	3–9 years (n=160)	−0.004 (0.58)	
	10+ years (n=145)	0.06 (0.62)	
Ever sued for malpractice	Yes (n=31)	−0.09 (0.51)	0.35
	No (n=520)	−0.006 (0.58)	
Numeracy Score	Low (n=64)	0.14 (0.54)	<0.01
	Medium (172)	0.07 (0.61)	
	High (n=306)	−0.08 (0.55)	
Risk Taking Score	Low (n=162)	0.02 (0.55)	0.65
	Medium (n=201)	0.02 (0.64)	
	High (n=189)	−0.03 (0.53)	

Variable	Participants	Mean Standardized Testing Likelihood Score (SD)	P-value*
Fear of Malpractice	Low (n=189)	−0.06 (0.61)	0.02
	Medium (n=176)	−0.04 (0.57)	
	High (n=183)	0.10 (0.54)	
Medical Maximizer-Minimizer Scale	Low (n=169)	−0.16 (0.60)	<0.01
	Medium (n=212)	0.01 (0.55)	
	High (n=164)	0.15 (0.54)	
Burnout Score	Very low (n=77)	0.11 (0.50)	0.21
	Low (n=277)	−0.02 (0.57)	
	Medium or High (n=190)	−0.001 (0.61)	
The Revised Physicians' Reactions to Uncertainty Scales			
Stress from uncertainty subscale	Low (n=202)	−0.02 (0.56)	<0.01
	Medium (n=218)	−0.06 (0.58)	
	High (n=128)	0.13 (0.59)	
Concern about bad outcomes subscale	Low (n=172)	−0.07 (0.57)	<0.01
	Medium (n=185)	−0.08 (0.57)	
	High (n=194)	0.13 (0.57	
There is often uncertainty in the practice of medicine	Disagree / agree slightly (n=131)	0.20 (0.59)	<0.01
	Agree moderately (n=188)	−0.06 (0.57)	
	Agree strongly (n=232)	−0.07 (0.55)	

*p-values are for comparisons across the category

1. **Ms. Smith, a previously healthy 35-year-old woman who smokes tobacco presents with five days of fatigue, productive cough, worsening shortness of breath, fevers to 102°F and decreased breath sounds in the lower right field. She has a heart rate of 105 but otherwise vital signs are normal. She has no particular preference for testing and wants your advice.**
IN YOUR PRACTICE, in what proportion of patients with symptoms like Ms. Smith do you order a chest X-ray? _____%
2. **You are seeing Ms. Johnson, 45-year-old woman, for an annual visit. She has no specific risk factors or symptoms for breast cancer. She has no particular preference for testing and wants your advice.**
IN YOUR PRACTICE, in what proportion of women aged 40-50 at average risk do you order mammograms? _____%
3. **You are seeing Mrs. Jones, a 43-year-old premenopausal woman with atypical chest pain and a normal ECG. She has no risk factors and normal vital signs/examination. She has no particular preference for testing and wants your advice.**
IN YOUR PRACTICE, in what proportion of women aged 40-50 with atypical chest pain do you order an exercise stress test or other type of stress test? _____%
4. **Mr. Williams, a 65-year-old man, comes to the office for follow up of his osteoarthritis. He has noted foul-smelling urine and no pain or difficulty with urination. A urine dipstick shows trace blood. He has no particular preference for testing and wants your advice.**
IN YOUR PRACTICE, in what proportion of patients at average risk with symptoms like Mr. Williams do you order a urine culture? _____%

Appendix Figure 1:
Clinical scenarios presented to clinicians to determine their likelihood to order tests.

	Disagree Strongly	Disagree Moderately	Disagree Slightly	Neutral	Agree Slightly	Agree Moderately	Agree Strongly
It is important to treat disease even when it does not make a difference in survival.	O	O	O	O	O	O	O
It is important to treat a disease even when it does not make a difference in quality of life.	O	O	O	O	O	O	O
Doing everything to fight illness is always the right choice.	O	O	O	O	O	O	O
When it comes to health care, the only responsible thing to do is to actively seek medical care.	O	O	O	O	O	O	O
When managing a health issue, my preference is to wait and see if the problem gets better on its own before doing anything about it.	O	O	O	O	O	O	O
When it comes to health care, watching and waiting is never an acceptable option.	O	O	O	O	O	O	O
When it comes to medical treatment, more is usually better.	O	O	O	O	O	O	O

	Disagree Strongly	Disagree Moderately	Disagree Slightly	Neutral	Agree Slightly	Agree Moderately	Agree Strongly
Diagnostic tests always provide helpful information even if their results do not directly impact care.	O	O	O	O	O	O	O

Appendix Figure 2:
Medical Maximizer Scale for Clinicians

References:

1. Ganguli I, Lupo C, Mainor AJ, et al. Prevalence and Cost of Care Cascades After Low-Value Preoperative Electrocardiogram for Cataract Surgery in Fee-for-Service Medicare Beneficiaries. JAMA Intern Med. 2019;179(9):1211–1219. [PubMed: 31158270]
2. Korenstein D, Chimonas S, Barrow B, Keyhani S, Troy A, Lipitz-Snyderman A. Development of a Conceptual Map of Negative Consequences for Patients of Overuse of Medical Tests and Treatments. JAMA Intern Med. 2018;178(10):1401–1407. [PubMed: 30105371]
3. Schwartz AL, Landon BE, Elshaug AG, Chernew ME, McWilliams JM. Measuring low-value care in Medicare. JAMA Intern Med. 2014;174(7):1067–1076. [PubMed: 24819824]
4. Segal JB, Nassery N, Chang HY, Chang E, Chan K, Bridges JF. An index for measuring overuse of health care resources with Medicare claims. Med Care. 2015;53(3):230–236. [PubMed: 25590675]
5. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. Ann Intern Med. 2003;138(4):288–298. [PubMed: 12585826]
6. Colla CH, Morden NE, Sequist TD, Schpero WL, Rosenthal MB. Choosing wisely: prevalence and correlates of low-value health care services in the United States. J Gen Intern Med. 2015;30(2):221–228. [PubMed: 25373832]
7. Radomski TR, Feldman R, Huang Y, et al. Evaluation of Low-Value Diagnostic Testing for 4 Common Conditions in the Veterans Health Administration. JAMA Netw Open. 2020;3(9):e2016445.

8. Sirovich BE, Lipner RS, Johnston M, Holmboe ES. The association between residency training and internists' ability to practice conservatively. JAMA Intern Med. 2014;174(10):1640–1648. [PubMed: 25179515]
9. Chen YA, Gray BG, Bandiera G, MacKinnon D, Deva DP. Variation in the utilization and positivity rates of CT pulmonary angiography among emergency physicians at a tertiary academic emergency department. Emerg Radiol. 2015;22(3):221–229. [PubMed: 25209190]
10. O'Neill L, Kuder J. Explaining variation in physician practice patterns and their propensities to recommend services. Med Care Res Rev. 2005;62(3):339–357. [PubMed: 15894708]
11. Shashar S, Ellen M, Codish S, Davidson E, Novack V. Medical Practice Variation Among Primary Care Physicians: 1 Decade, 14 Health Services, and 3,238,498 Patient-Years. Ann Fam Med. 2021;19(1):30–37. [PubMed: 33431388]
12. Schwartz AL, Jena AB, Zaslavsky AM, McWilliams JM. Analysis of Physician Variation in Provision of Low-Value Services. JAMA Intern Med. 2019;179(1):16–25. [PubMed: 30508010]
13. Eguale T, Buckeridge DL, Winslade NE, Benedetti A, Hanley JA, Tamblyn R. Drug, patient, and physician characteristics associated with off-label prescribing in primary care. Arch Intern Med. 2012;172(10):781–788. [PubMed: 22507695]
14. 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. Bmj. 2004;328(7437):444. [PubMed: 14966079]
15. Wu X, Jiang YN, Zhang YL, et al. Impact of Physicians' Personalities and Behavioral Traits on Treatment-Related Decision-making for Elderly Acute Myeloid Leukemia. J Gen Intern Med. 2021.
16. Barreto TW, Chung Y, Wingrove P, et al. Primary Care Physician Characteristics Associated with Low Value Care Spending. J Am Board Fam Med. 2019;32(2):218–225. [PubMed: 30850458]
17. Bouck Z, Ferguson J, Ivers NM, et al. Physician Characteristics Associated With Ordering 4 Low-Value Screening Tests in Primary Care. JAMA Netw Open. 2018;1(6):e183506.
18. Hong AS, Ross-Degnan D, Zhang F, Wharam JF. Clinician-Level Predictors for Ordering Low-Value Imaging. JAMA Intern Med. 2017;177(11):1577–1585. [PubMed: 28973068]
19. Morgan DJ, Leppin AL, Smith CD, Korenstein D. A Practical Framework for Understanding and Reducing Medical Overuse: Conceptualizing Overuse Through the Patient-Clinician Interaction. J Hosp Med. 2017;12(5):346–351. [PubMed: 28459906]
20. Abrams M, Nuzum R, Zezza M, Ryan J, Kiszla J, Guterman S. The Affordable Care Act's payment and delivery system reforms: a progress report at five years. Issue Brief (Commonw Fund). 2015;12:1–16. [PubMed: 26040019]
21. Born K, Kool T, Levinson W. Reducing overuse in healthcare: advancing Choosing Wisely. Bmj. 2019;367:l6317.
22. Colla CH, Mainor AJ, Hargreaves C, Sequist T, Morden N. Interventions Aimed at Reducing Use of Low-Value Health Services: A Systematic Review. Med Care Res Rev. 2017;74(5):507–550. [PubMed: 27402662]
23. Mafi JN, Reid RO, Baseman LH, et al. Trends in Low-Value Health Service Use and Spending in the US Medicare Fee-for-Service Program, 2014–2018. JAMA Netw Open. 2021;4(2):e2037328.
24. Morgan DJ, Pineles L, Owczarzak J, et al. Accuracy of Practitioner Estimates of Probability of Diagnosis Before and After Testing. JAMA Intern Med. 2021.
25. Casscells W, Schoenberger A, Graboys TB. Interpretation by physicians of clinical laboratory results. N Engl J Med. 1978;299(18):999–1001. [PubMed: 692627]
26. Fagerlin A, Sepucha KR, Couper MP, Levin CA, Singer E, Zikmund-Fisher BJ. Patients' knowledge about 9 common health conditions: the DECISIONS survey. Med Decis Making. 2010;30(5 Suppl):35s–52s. [PubMed: 20881153]
27. G G. Reckoning with Risk: Learning to Live With Uncertainty. Penguin Books; 2003.
28. Krouss M, Croft L, Morgan DJ. Physician Understanding and Ability to Communicate Harms and Benefits of Common Medical Treatments. JAMA Intern Med. 2016;176(10):1565–1567. [PubMed: 27571226]

29. Manrai AK, Bhatia G, Strymish J, Kohane IS, Jain SH. Medicine's uncomfortable relationship with math: calculating positive predictive value. JAMA Intern Med. 2014;174(6):991–993. [PubMed: 24756486]
30. Whiting PF, Davenport C, Jameson C, et al. How well do health professionals interpret diagnostic information? A systematic review. BMJ Open. 2015;5(7):e008155.
31. Dolan ED, Mohr D, Lempa M, et al. Using a single item to measure burnout in primary care staff: a psychometric evaluation. J Gen Intern Med. 2015;30(5):582–587. [PubMed: 25451989]
32. Allison JJ, Kiefe CI, Cook EF, Gerrity MS, Orav EJ, Centor R. The association of physician attitudes about uncertainty and risk taking with resource use in a Medicare HMO. Med Decis Making. 1998;18(3):320–329. [PubMed: 9679997]
33. Katz DA, Williams GC, Brown RL, et al. Emergency physicians' fear of malpractice in evaluating patients with possible acute cardiac ischemia. Ann Emerg Med. 2005;46(6):525–533. [PubMed: 16308068]
34. Scherer LD, Caverly TJ, Burke J, et al. Development of the Medical Maximizer-Minimizer Scale. Health Psychol. 2016;35(11):1276–1287. [PubMed: 27617512]
35. Pearson SD, Goldman L, Orav EJ, et al. Triage decisions for emergency department patients with chest pain: do physicians' risk attitudes make the difference? J Gen Intern Med. 1995;10(10):557–564. [PubMed: 8576772]
36. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med. 1997;127(11):966–972. [PubMed: 9412301]
37. Kanzaria HK, Hoffman JR, Probst MA, Caloyeras JP, Berry SH, Brook RH. Emergency physician perceptions of medically unnecessary advanced diagnostic imaging. Acad Emerg Med. 2015;22(4):390–398. [PubMed: 25807868]
38. Lyu H, Xu T, Brotman D, et al. Overtreatment in the United States. PLoS One. 2017;12(9):e0181970.
39. Studdert DM, Mello MM, Sage WM, et al. Defensive medicine among high-risk specialist physicians in a volatile malpractice environment. Jama. 2005;293(21):2609–2617. [PubMed: 15928282]
40. Garcia-Retamero R, Cokely ET, Wicki B, Joeris A. Improving risk literacy in surgeons. Patient Educ Couns. 2016;99(7):1156–1161. [PubMed: 26879804]
41. Johnson TV, Abbasi A, Schoenberg ED, et al. Numeracy among trainees: are we preparing physicians for evidence-based medicine? J Surg Educ. 2014;71(2):211–215. [PubMed: 24602712]
42. Sheridan SL, Pignone M. Numeracy and the medical student's ability to interpret data. Eff Clin Pract. 2002;5(1):35–40. [PubMed: 11874195]
43. Petrova D, Kostopoulou O, Delaney BC, Cokely ET, Garcia-Retamero R. Strengths and Gaps in Physicians' Risk Communication: A Scenario Study of the Influence of Numeracy on Cancer Screening Communication. Med Decis Making. 2018;38(3):355–365. [PubMed: 28884617]
44. Evron JM, Reyes-Gastelum D, Banerjee M, et al. Role of Patient Maximizing-Minimizing Preferences in Thyroid Cancer Surveillance. J Clin Oncol. 2019;37(32):3042–3049. [PubMed: 31573822]
45. Scherer LD, Shaffer VA, Caverly T, DeWitt J, Zikmund-Fisher BJ. Medical Maximizing-Minimizing Predicts Patient Preferences for High- and Low-Benefit Care. Med Decis Making. 2020;40(1):72–80. [PubMed: 31975657]
46. Scherer LD, Kullgren JT, Caverly T, et al. Medical Maximizing-Minimizing Preferences Predict Responses to Information about Prostate-Specific Antigen Screening. Med Decis Making. 2018;38(6):708–718. [PubMed: 29916784]
47. Franks P, Williams GC, Zwanziger J, Mooney C, Sorbero M. Why do physicians vary so widely in their referral rates? J Gen Intern Med. 2000;15(3):163–168. [PubMed: 10718896]
48. Ganguli I, Simpkin AL, Colla CH, et al. Why Do Physicians Pursue Cascades of Care After Incidental Findings? A National Survey. J Gen Intern Med. 2020;35(4):1352–1354. [PubMed: 31346910]
49. Schneider A, Löwe B, Barie S, Joos S, Engeser P, Szecsenyi J. How do primary care doctors deal with uncertainty in making diagnostic decisions? The development of the 'Dealing with Uncertainty Questionnaire' (DUQ). J Eval Clin Pract. 2010;16(3):431–437. [PubMed: 20604824]

50. Alam R, Cheraghi-Sohi S, Panagioti M, Esmail A, Campbell S, Panagopoulou E. Managing diagnostic uncertainty in primary care: a systematic critical review. BMC Fam Pract. 2017;18(1):79. [PubMed: 28784088]
51. Begin AS, Hidrue M, Lehrhoff S, Del Carmen MG, Armstrong K, Wasfy JH. Factors Associated with Physician Tolerance of Uncertainty: an Observational Study. J Gen Intern Med. 2021:1–7.
52. Pines JM, Hollander JE, Isserman JA, et al. The association between physician risk tolerance and imaging use in abdominal pain. Am J Emerg Med. 2009;27(5):552–557. [PubMed: 19497460]
53. Pines JM, Isserman JA, Szyld D, Dean AJ, McCusker CM, Hollander JE. The effect of physician risk tolerance and the presence of an observation unit on decision making for ED patients with chest pain. Am J Emerg Med. 2010;28(7):771–779. [PubMed: 20837253]
54. Adesoye T, Kimsey LG, Lipsitz SR, et al. Geographic variation in Medicare and the military healthcare system. Am J Manag Care. 2017;23(8):e259–e264. [PubMed: 29087149]
55. Chen C, Petterson S, Phillips R, Bazemore A, Mullan F. Spending patterns in region of residency training and subsequent expenditures for care provided by practicing physicians for Medicare beneficiaries. Jama. 2014;312(22):2385–2393. [PubMed: 25490329]
56. Mafi JN, Landon BE. Comparing Use of Low-Value Health Care Services Among U.S. Advanced Practice Clinicians and Physicians. Ann Intern Med. 2017;166(1):77.
57. Auerbach DI, Staiger DO, Buerhaus PI. Growing Ranks of Advanced Practice Clinicians - Implications for the Physician Workforce. N Engl J Med. 2018;378(25):2358–2360. [PubMed: 29924944]
58. Bertakis KD, Helms LJ, Azari R, Callahan EJ, Robbins JA, Miller J. Differences between family physicians' and general internists' medical charges. Med Care. 1999;37(1):78–82. [PubMed: 10413395]

Clinical significance:

Clinicians vary widely in testing practice; interventions to reduce unnecessary testing have had small impact. This large multi-state cross-sectional study of primary care clinicians found that cognitive characteristics including low numeracy, being a “medical maximizer”, and low acknowledgment of uncertainty were associated with testing aggressiveness. Identifying clinician characteristics associated with more aggressive testing could inform future interventions aimed at reducing overuse of tests.

Clinical significance:

- This large multi-state cross-sectional study found that primary care clinician testing aggressiveness was associated with cognitive characteristics.
- Low numeracy, being a “medical maximizer” and low acknowledgment of uncertainty in medicine were associated with testing aggressiveness across a range of tests.
- Risk-taking scores and discomfort with uncertainty were not consistently associated with testing aggressiveness.
- NP and PA clinicians tested more aggressively than MDs.

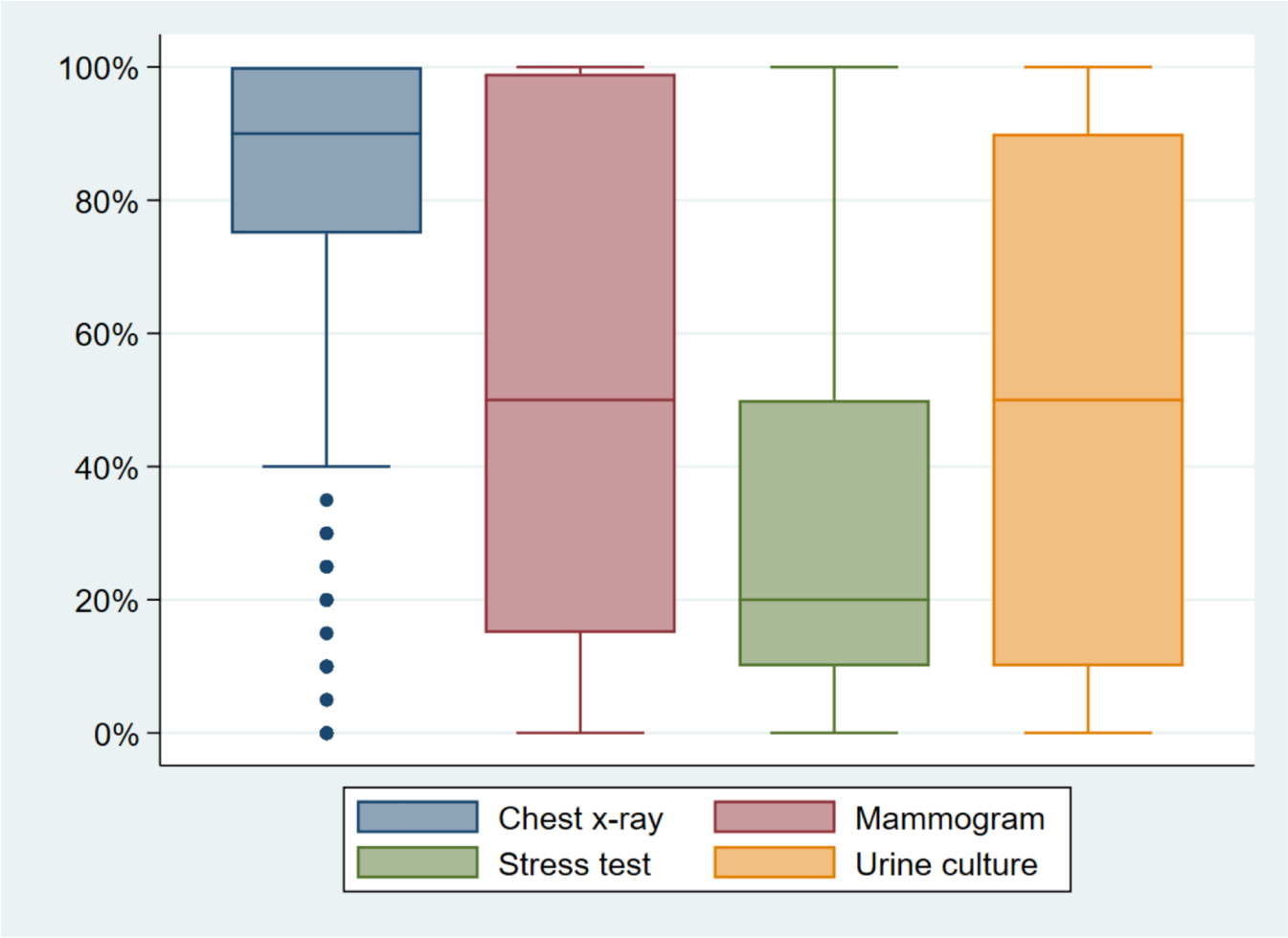

Figure 1:

Clinician variation in test ordering.

Clinician variation in self-reported frequency of test ordering for their own patients for mammograms for low-risk breast cancer screening, stress testing for low-risk cardiac ischemia, chest x-rays for high-risk pneumonia and urine cultures for very low-risk urinary tract infection. Shaded boxes denote the interquartile range and the line in the shaded boxes indicates the median value. Outer bars describe the minimum and maximum values, except in the case of outliers, which are represented as dots.

Table 1:

Participant demographics (n=552)

Variable	Characteristic	N (%)
Age (years)	<30	171 (31.3)
	30–39	208 (38.1)
	≥40	167 (30.6)
Gender	Male	258 (46.9)
	Female	292 (53.1)
Race	White	296 (53.6)
	Asian	141 (25.5)
	Other	103 (18.6)
Study Site	Pacific Northwest	112 (20.3)
	Mid-Atlantic	304 (55.1)
	Texas	136 (24.6)
Degree and training	MD or DO resident	289 (52.4)
	MD or DO attending	202 (36.6)
	NP or PA	61 (11.0)
Type of post-graduate training (MDs and DOs, n=491)	Internal Medicine	335 (68.2)
	Family Medicine	142 (28.9)
	Other	14 (2.9)
Years in Practice (since graduation)	<3 years	239 (43.9)
	3–9 years	160 (29.4)
	10+ years	145 (26.7)
Ever sued for malpractice	Yes	31 (5.6)
	No	520 (94.2)

Table 2:

Mean Testing Likelihood score and associations with clinician characteristics on bivariable analysis. The likelihood test score is the average of 4 individual testing scores provided by the respondent.

Variable	Participants	Mean Testing Likelihood Score, % (SD)	P-value*
All	All (n=552)	55 (19)	
Age (years)	<30 (n=171)	54 (17)	0.35
	30–39 (n=208)	54 (20)	
	≥40 (n=167)	56 (20)	
Study Site	Pacific northwest (n=112)	45 (17)	<0.01
	Mid-Atlantic (n=304)	57 (19)	
	Texas (n=136)	58 (19)	
Degree and training	MD or DO resident (n=289)	54 (19)	<0.01
	MD or DO attending (n=202)	54 (20)	
	NP or PA (n=61)	63 (18)	
Type of post-graduate training (MDs and DOs)	Internal Medicine (n=335)	52 (19)	0.04
	Family Medicine (n=142)	56 (19)	
Years in Practice (since graduation)	<3 years (n=239)	53 (18)	0.17
	3–9 years (n=160)	55 (19)	
	10+ years (n=145)	57 (20)	
Ever sued for malpractice	Yes (n=31)	53 (16)	0.63
	No (n=520)	55 (19)	
Numeracy Score	Low (n=64)	59 (18)	<0.01
	Medium (172)	58 (21)	
	High (n=306)	52 (18)	
Risk Taking Score	Low (n=162)	55 (18)	0.60
	Medium (n=201)	55 (21)	
	High (n=189)	54 (18)	
Fear of Malpractice	Low (n=189)	53 (20)	0.03
	Medium (n=176)	54 (19)	
	High (n=183)	58 (18)	
Medical Maximizer-Minimizer Scale	Low (n=169)	49 (20)	<0.01
	Medium (n=212)	55 (18)	
	High (n=164)	60 (18)	
Burnout Score	Very low (n=77)	58 (17)	0.26
	Low (n=277)	54 (19)	
	Medium or High (n=190)	55 (20)	

Variable	Participants	Mean Testing Likelihood Score, % (SD)	P-value*
The Revised Physicians' Reactions to Uncertainty Scales			
Stress from uncertainty subscale	Low (n=202)	54 (19)	<0.01
	Medium (n=218)	53 (19)	
	High (n=128)	59 (20)	
Concern about bad outcomes subscale	Low (n=172)	53 (19)	<0.01
	Medium (n=185)	52 (19)	
	High (n=194)	59 (19)	
There is often uncertainty in the practice of medicine	Disagree / agree slightly (n=131)	62 (20)	<0.01
	Agree moderately (n=188)	53 (19)	
	Agree strongly (n=232)	52 (18)	

* p-values are based on ANOVAs, for comparisons across the category

Table 3:

Frequency of clinicians reporting it is their practice to check routine lipid panels and LFTs in asymptomatic patients on statin treatment more than once per year

Variable	Participants	Check lipids <12 months N=167 of 532	P-value	Check LFTs <12 months N=175 of 532	P-value
Age (years)	<30	49 (29%)	0.09	42 (26%)	<0.01
	30–39	56 (28%)		63 (31%)	
	≥40	62 (38%)		66 (42%)	
Study Site	Pacific Northwest	5 (5%)	<0.01	7 (7%)	<0.01
	Mid-Atlantic	110 (37%)		103 (36%)	
	Texas	53 (41%)		65 (48%)	
Degree and training	MD or DO attending	52 (27%)	0.03	64 (33%)	0.06
	MD or DO resident	89 (32%)		84 (30%)	
	NP or PA	27 (45%)		27 (47%)	
Type of post-graduate training (MDs and DOs)	Internal Medicine	97 (30%)	0.55	101 (31%)	0.51
	Family Medicine	37 (27%)		40 (28%)	
Years in Practice (since graduation)	<3 years	81 (34%)	0.15	81 (35%)	0.02
	3–9 years	39 (25%)		37 (24%)	
	10+ years	45 (32%)		53 (38%)	
Ever sued for malpractice	Yes	13 (42%)	0.19	12 (39%)	0.48
	No	155 (31%)		163 (33%)	
Numeracy Score	Low	30 (48%)	<0.01	29 (47%)	0.01
	Medium	55 (33%)		57 (35%)	
	High	79 (26%)		83 (28%)	
Risk Taking Score	Low	49 (31%)	0.54	52 (33%)	0.44
	Medium	66 (34%)		69 (36%)	
	High	53 (29%)		54 (30%)	
Fear of Malpractice	Low	40 (22%)	<0.01	47 (26%)	0.03
	Medium	64 (37%)		58 (35%)	
	High	63 (35%)		69 (39%)	
Medical Maximizer-Minimizer Scale	Low	30 (18%)	<0.01	31 (19%)	<0.01
	Medium	59 (29%)		69 (34%)	
	High	76 (47%)		72 (46%)	
Burnout score	Low	34 (45%)	0.02	33 (45%)	0.05
	Medium	76 (28%)		78 (30%)	
	High	55 (30%)		62 (33%)	
The Revised Physicians' Reactions to Uncertainty Scales					
Stress from uncertainty subscale	Low	52 (27%)	0.16	61 (32%)	0.28
	Medium	68 (31%)		64 (30%)	
	High	47 (37%)		48 (38%)	

Variable	Participants	Check lipids <12 months N=167 of 532	P-value	Check LFTs <12 months N=175 of 532	P-value
Concern about bad outcomes subscale	Low	48 (29%)	0.28	57 (34%)	0.19
	Medium	52 (29%)		48 (28%)	
	High	68 (36%)		70 (36%)	
There is often uncertainty in the practice of medicine	Disagree / agree slightly	60 (48%)	<0.01	60 (49%)	<0.01
	Agree moderately	59 (32%)		65 (35%)	
	Agree strongly	49 (21%)		50 (22%)	