


Clinician-Level Variation in Three Measures Representing Overuse Based on the American Geriatrics Society Choosing Wisely Statement



Theresa A. Rowe, DO, MS¹ , Tiffany Brown, MPH¹, Ji Young Lee, MS¹, Jeffrey A. Linder, MD, MPH¹, Mark W. Friedberg, MD, MPP^{2,3,4}, Jason N. Doctor, PhD⁵, Daniella Meeker, PhD⁵, Jody D. Ciolino, PhD¹, and Stephen D. Persell, MD, MPH^{1,6}

¹Division of General Internal Medicine and Geriatrics, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; ²RAND Corporation, Boston, MA, USA; ³Blue Cross Blue Shield of Massachusetts, Boston, MA, USA; ⁴Division of General Internal Medicine, Brigham and Women's Hospital, Boston, MA, USA; ⁵University of Southern California Keck School of Medicine, Los Angeles, CA, USA; ⁶Center for Primary Care Innovation, Institute for Public Health and Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA.

IMPORTANCE: The extent of clinician-level variation in the overuse of testing or treatment in older adults is not well understood.

OBJECTIVE: To examine clinician-level variation for three new measures of potentially inappropriate use of medical services in older adults.

DESIGN: Retrospective analysis of overall means and clinician-level variation in performance on three new measures.

SUBJECTS: Adults aged 65 years and older who had office visits with outpatient primary or immediate care clinicians within a single academic medical center health system between July 1, 2016, and June 30, 2017.

MEASURES: Two electronic clinical quality measures representing potentially inappropriate use of medical services in older adults: prostate-specific antigen testing against guidelines (PSA) in men aged 76 and older; urinalysis or urine culture for non-specific reasons in women aged 65 and older; and one intermediate outcome measure: hemoglobin A1c less than 7.0 in adults aged 75 and older with diabetes mellitus treated with insulin or oral hypoglycemic medication.

RESULTS: Sixty-nine clinicians and 2009 patients contributed observations to the PSA measure, 144 clinicians and 5933 patients contributed to the urinalysis/urine culture measure, and 42 clinicians and 665 patients contributed to the diabetes measure. Meaningful clinician-level performance variation was greatest for the PSA measure (intraclass correlation coefficient [ICC] = 0.27), followed by the urinalysis/urine culture measure (ICC = 0.18), and the diabetes measure (ICC = 0.024). The range of possible overuse across clinician quartiles was 8–54% for the PSA measure, 3–35% for the urinalysis/urine culture measure, and 13–49% for the diabetes measure. The odds ratios of overuse in the highest quartile compared with the lowest for the PSA, urinalysis/urine culture, and

diabetes measures were 99.3 (95% CI 43 to 228), 15.7 (10 to 24), and 6.0 (3.3 to 11), respectively.

CONCLUSIONS: Within the same health system, rates of potential overuse in elderly patients varied greatly across clinicians, particularly for the process measures examined.

J Gen Intern Med 35(6):1797–802

DOI: 10.1007/s11606-020-05748-8

© Society of General Internal Medicine 2020

BACKGROUND

Overuse is care in which potential harms outweigh potential benefits. This applies both to over-testing that may lead to downstream actions that cause harm or overtreatment where therapeutics are used under circumstances where the potential harms likely outweigh benefits.^{1, 2} The Choosing Wisely Campaign was developed to address over-testing and overtreatment for commonly encountered conditions and is gaining momentum internationally.³ A working group from the American Geriatrics Society applied these principles to older adults.⁴ Several of the topics identified were well-established examples for which clinicians in ambulatory practice often do not follow best practices thereby leading to patient harm.^{2, 5–10} Three specific examples relevant to general ambulatory care in which clinicians deviate from Choosing Wisely recommendations include the following:⁴ (1) prostate-specific antigen (PSA) testing to screen for prostate cancer in older men,¹¹ (2) overuse of urine studies (e.g., urinalysis and/or urine cultures) to detect pyuria and/or bacteriuria in older women without specific genitourinary signs and/or symptoms or conditions,¹² and (3) overtreatment of diabetes mellitus in older adults to achieve a hemoglobin A1C level (HbA1c) of <7.0% with drugs that may cause hypoglycemia.^{13, 14}

There are several factors that might influence clinician decision making regarding testing and treatment related to overuse in older adults, including patient, clinician, and

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11606-020-05748-8>) contains supplementary material, which is available to authorized users.

Received August 23, 2019

Accepted February 14, 2020

Published online March 3, 2020

environmental characteristics.^{15, 16} A few studies have shown that clinician-level variation may influence whether a test or treatment is performed.¹⁷ To our knowledge, the extent of clinician-level variation in metrics of overuse of testing or treatment in older adults has not been well established. Understanding clinician-level variation may inform the development of targeted interventions aimed to improve adherence to evidence-based practices.

The objective of this paper is to examine the clinician-level variation for measures of three areas of potential overuse in older adults: (1) PSA screening in older men (PSA), (2) overuse of urine studies in older women without specific genitourinary signs and/or symptoms (UA/UC), and (3) diabetes mellitus overtreatment in older adults (DM).

METHODS

Design

The study was a retrospective analysis of process and intermediate outcome measures in adults aged 65 years and older evaluated in outpatient primary care or immediate care within two regions (north and central) of a health system affiliated with an academic medical center between July 1, 2016, and June 30, 2017. Northwestern University's Institutional Review Board approved the study.

Subjects

Clinicians in this sample included physicians, physician assistants, advanced practice nurses in primary care specialties (e.g., internal medicine, family practice, geriatrics) providing outpatient care for patients who were eligible for any of the three overuse measures below. Clinicians practicing in immediate care were included only in the UA/UC measure. Only clinicians with ten or more eligible patients attributed to them were included in the clinician-level analyses for any given measure. We describe the three specific older adult patient populations we examined during the study period below.

Measurements

The primary measures of interest were adherence to three Choosing Wisely recommendations, operationalized as electronic clinical quality measures (eCQM) calculated from electronic health record (EHR) data that represent potentially inappropriate use of medical services in older adults.

PSA testing against guidelines was defined as the presence of a PSA laboratory result in the EHR during the study period (numerator) among men aged 76 years and older without a preceding diagnosis or procedure suggesting a history of prostate cancer (denominator) (eTable 1). Patients on androgenic agents or anabolic steroids were excluded (eTable 1). Laboratory orders that constituted PSA testing against guidelines are provided in eTable 2. Patients were attributed to the primary

care clinician with whom they had the greatest number of visits during the measurement period.

Urine Testing for Non-specific Reasons was defined as the proportion of urinalyses (UA) and/or urine cultures (UC) that were not associated with a visit diagnostic code indicating a specific genitourinary sign, symptom¹⁸ or other potentially relevant indication among instances where a UA/UC was obtained in the interval 24 h before to 48 h after a face-to-face ambulatory care visit by a woman aged 65 years or older with a qualifying clinician. Patients with a UC order were excluded if they had a visit diagnostic code indicating a genitourinary sign or symptom (e.g., hematuria, urgency, frequency, hematuria, dysuria, suprapubic tenderness, and/or costovertebral angle pain or tenderness)¹⁸, genitourinary condition (e.g. nephrolithiasis), or other indications of infection (e.g., fever) as described in eTable 3. Patients with a UA order were excluded if they had a genitourinary tract specific sign or symptom, genitourinary condition, other indication of infection, or a systemic illness or comorbidity for which a UA may be appropriate (e.g., hypertension, rheumatologic condition) as described in eTable 3. Testing episodes were attributed to the clinician who ordered the test. Only the first testing episode was included in patients with multiple qualifying visits.

Diabetes Overtreatment in the Elderly was defined as the having the most recent hemoglobin A1C during the study period or the year prior to the study period of less than 7.0 among adults aged 75 years and older with a diagnosis of diabetes mellitus who had insulin or an oral hypoglycemic (e.g., sulfonylurea or meglitinide, eTable 4) on their active medication list at the end of the study period. Patients only on medications without much potential for hypoglycemia (e.g., metformin) were excluded. Patients were attributed to the primary care clinician with whom they had the greatest number of visits during the measurement period.

Additional details of the measure criteria and measure development process are available in the supplemental criteria (eMethods). For each measure, a physician examined approximately 100 charts to confirm the accuracy and completeness of the measure criteria.

We obtained discrete EHR data for all participants. Characteristics included age, race and ethnicity, insurance, and marital status.

Statistical Analysis

For each measure, we examined the proportion of patients meeting the measure in the population overall and by an individual clinician. We calculated the intraclass correlation coefficient (ICC) from a mixed logistic regression for each measure to assess the degree of variation at the clinician level. We constructed caterpillar plots with 95% confidence intervals to show the overall variation of individual clinicians' performance and to visualize the variability across clinicians and help determine which clinicians tended to differ in performance from the rest. We also sought to quantify the odds of

Table 1 Rates and Within-Clinician Correlation for Three Geriatric Overuse Measures

Measure	Clinicians, N*	Patients, N	Proportion with overuse (lower is preferred)	Clinician-level ICC	Number of cases per clinician (IQR)
PSA testing against guidelines	69	2009	0.23	0.27	21 (15, 35)
Urine testing without specific signs/symptoms or conditions	144	5933	0.23	0.18	23 (14, 51)
Diabetes overtreatment in elderly	42	665	0.32	0.024	14 (11, 19)

ICC: intraclass correlation coefficient; IQR: interquartile range; PSA: prostate-specific antigen

*Clinicians with 10 or more eligible patients attributed were included

an individual patient receiving inappropriate testing or treatment based on the treating clinician. To accomplish this, for each measure, we divided the clinicians into quartiles based on their performance rates. Then, we applied logistic regression models for the patient-level indicator of a potentially inappropriate test, with the primary predictor of interest corresponding to their clinician's quartile membership. We calculated odds ratios for the 2nd–4th quartiles (e.g., higher inappropriate use) compared with the 1st quartile (e.g., lowest inappropriate use). For the DM model, we adjusted for age group. For PSA and UA/UC we adjusted for both age group and race. We examined whether clinician-level performance on one measure was correlated with performance on another by calculating Pearson correlation coefficients. SAS 9.4 (SAS Institute Inc. Care, NC) was used for all analyses and R version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria) was used for the plots.

RESULTS

Table 1 shows overall rates for each overuse eCQM. The mean age of patients in each group was 81.6 years for PSA testing, 73.4 years for urine testing without specific signs/symptoms or conditions, and 81.0 years for overtreatment of DM. Patient characteristics are listed in eTable 5. The overall proportion of patients with potentially inappropriate use for each measure were 0.23 for PSA, 0.23 for UA/UC, and 0.32 for DM. Table 2 describes the estimated odds of obtaining an inappropriate test

for patients attributed to the clinician in the higher overusing quartiles compared with the lowest overusing quartile. The estimated odds of receiving a PSA test against guidelines for patients attributed to a clinician in the highest overusing quartile were 99.3 times greater than for patients attributed to a clinician in the lowest quartile. The estimated odds that a UA/UC was done for non-specific reasons for patients evaluated by a clinician in the highest overusing quartile were 15.7 times greater than for patients seeing a clinician in the lowest quartile. The estimated odds of DM overtreatment for patients attributed to a clinician in the highest overusing quartile were 6.0 times greater than for patients attributed to a clinician in the lowest quartile. Performance by individual clinicians along with corresponding 95% confidence limits are shown in Figures 1, 2, and 3. Clinician performance on the PSA measure was positively correlated with performance on the diabetes measure (eTable 6). Correlation at the clinician level was highest for the PSA measure (ICC = 0.27), followed by UC/UA (ICC = 0.18), and lowest for DM (ICC = 0.024).

DISCUSSION

We examined three eCQMs to assess potential overuse for topics described in the geriatrics Choosing Wisely Campaign⁴ and applied them to a population of patients and clinicians within a single large health system. We observed several notable findings: (1) overuse, as assessed by these measures, was not uncommon, with measured rates ranging from 23 to

Table 2 Rates of Geriatric Overuse Measures by Quartile of Clinician Performance

Measure	Clinicians, N	Patients, N	Proportion with overuse (lower is preferred)	Odds ratio of overuse compared with first quartile (95% CI)
PSA testing against guidelines				
1st quartile	17	432	0.014	Reference
2nd quartile	17	535	0.08	6.7 (2.8–16.0)
3rd quartile	17	457	0.22	20.1 (8.7–46.6)
4th quartile	18	585	0.54	99.3 (43.3–228)
Urine testing without specific signs/symptoms or conditions				
1st quartile	34	747	0.03	Reference
2nd quartile	38	1282	0.13	4.5 (2.9–7.0)
3rd quartile	31	1560	0.23	8.8 (5.8–13.4)
4th quartile	41	2344	0.35	15.7 (10.4–23.6)
Diabetes overtreatment in elderly				
1st quartile	10	134	0.134	Reference
2nd quartile	10	197	0.269	2.3 (1.3–4.2)
3rd quartile	12	167	0.347	3.3 (1.8–6.0)
4th quartile	10	167	0.491	6.0 (3.3–10.9)

CI: confidence interval; PSA: prostate-specific antigen

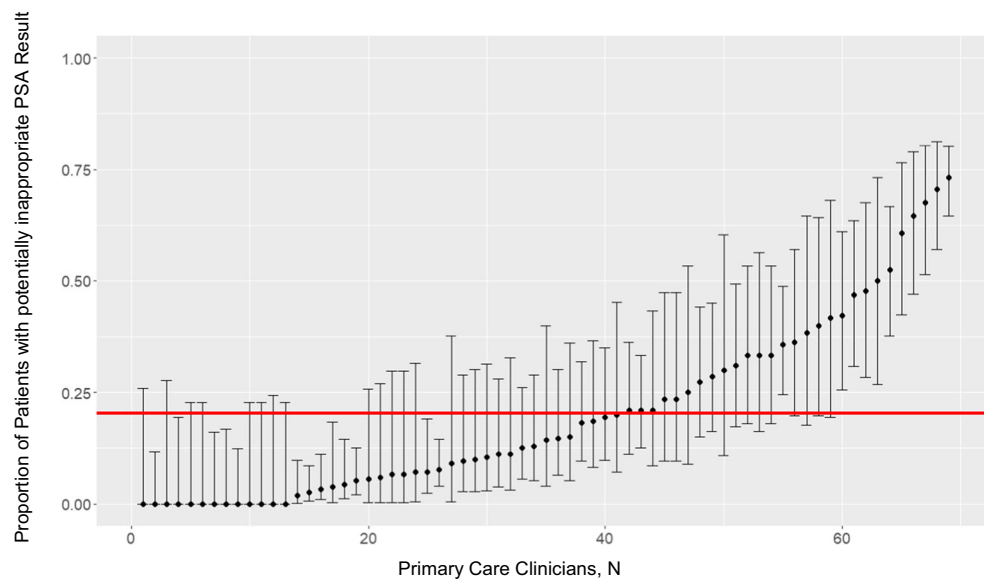


Figure 1 Proportion of men 76 and older receiving screening PSA against guidelines by primary care clinician. Circles represent the observed proportion of patients with PSA result for each clinician. Vertical lines represent the 95% CIs. Bold (red) horizontal line represents the proportion of patients with a screening PSA result for the entire group of patients attributed to these primary care clinicians ($n = 69$). Results are limited to only clinicians with 10 or more eligible patients.

32%; (2) the performance on these measures varied considerably across clinicians, (3) the amount of variation occurring at the clinician level was much greater for the two process measures (urine testing and PSA screening) than for the intermediate outcome measure (diabetes overtreatment), and (4) PSA and DM were positively correlated, but UA/UC was not correlated with either of the other measures. For the PSA and UA/UC—both related to diagnostic overuse—there were clinicians whose rates were zero or close to zero and others with potential overuse more than a majority of the time. These results suggest that for these two quality measures there are some clinicians whose behavior is highly consistent across

patients in their practice, but may differ greatly from other clinicians in their medical group.

We are not aware of prior work that has examined multiple eCQMs related to overuse in older patients in the ambulatory care setting, but several investigators have used administrative data. A prior study that used Medicare claims data from the state of Texas in 2010 to evaluate PSA screening in men over age 75 found a screening rate from primary care clinicians of 28.8% with similarly wide variation in practice across clinicians and an intraclass correlation at the clinician level of 0.27 (identical to the ICC we observed).¹⁹ In another study of Texas Medicare patients, variation in the early imaging for low back

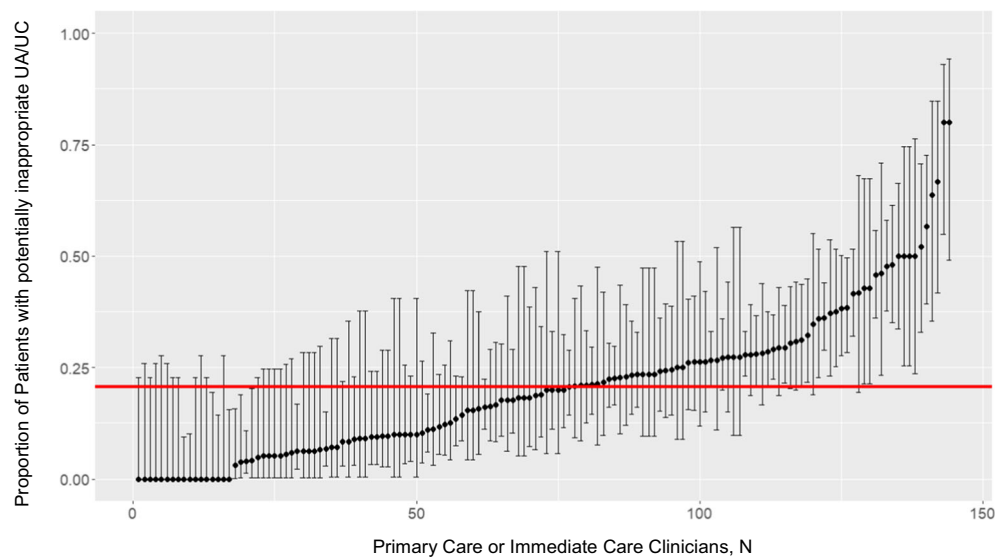


Figure 2 Proportion of UA and/or UC done for women ≥ 65 years of age without an appropriate diagnosis by clinician. Note: Circles represent the observed proportion of UA and/or UC done for women ≥ 65 years of age without an appropriate diagnosis for each clinician. Vertical lines represent the 95% CIs. Bold (red) horizontal line represents the mean proportion for all eligible cases attributed to the group of primary care and immediate care clinicians ($n = 144$). Results are limited to only clinicians with 10 or more eligible patients.

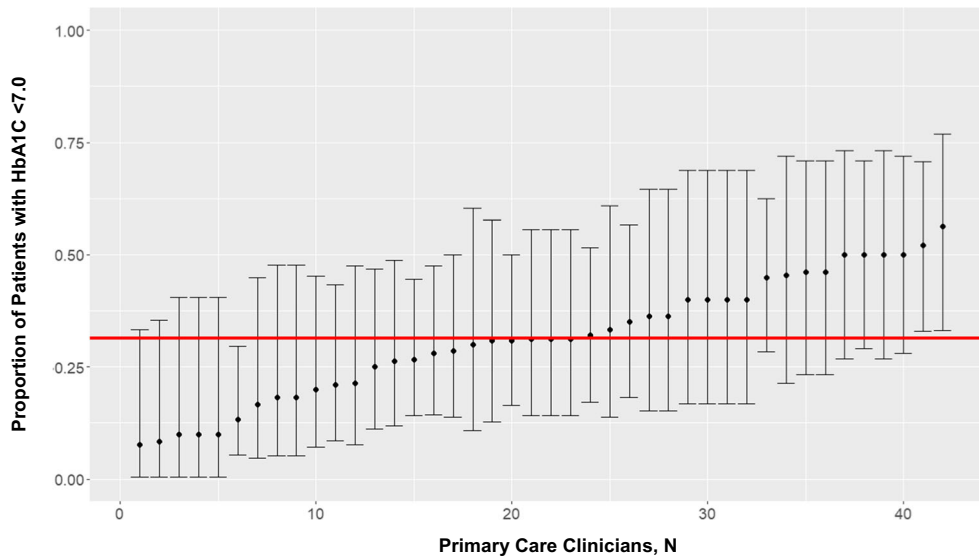


Figure. 3 Proportion of diabetes patients 75 or older on insulin or oral hypoglycemic medication with most recent HbA1c < 7.0 by clinician
Note: Circles represent the observed proportion of patients on insulin and/or an oral hypoglycemic with a HbA1c < 7.0. Vertical lines represent the 95% CIs. Bold (red) horizontal line represents the mean proportion for all eligible patients attributed to the group of primary care clinicians ($n = 42$). Results are limited to only clinicians with 10 or more eligible patients.

pain among primary care clinicians was substantial (ICC 0.25) and displayed a distribution similar to what we observed for the two process measures we examined.¹⁷ Pendrith et al. examined three low-value services (specific low-value uses of advanced imaging for low back pain, repeated dual-energy X-ray absorptiometry, and cervical cancer screening) among primary care physicians in Ontario Canada and found considerable variation at both the practice and regional level.²⁰

The positive correlation between two of the measures examined here suggests that these two measures may be assessing an underlying common clinical quality domain at the clinician level in this population. Prior work on clinician-level quality measurement that examined diabetes quality measures did show a good degree of association across some quality measures at the clinician level which provided support for aggregating measures to represent a single construct. However, these were not measures of overuse.²¹ Bouck and colleagues used administrative data to examine four low-value tests and found that among primary care clinicians from Ontario, there were physicians who could be identified as overusing more than one test. In particular, the same physicians tended to be over-users of low-value ECGs and chest X-rays.²² Another study analyzing data on low-value care among Medicare beneficiaries also found significant variation in use of low-value services among primary care physicians. However, in contrast to our study, Schwartz et al. found substantial overuse even among the least wasteful physicians.²³ They also did not find that physician characteristics contributed significantly to the variation seen in the use of low-value care, though other studies have found physician characteristics to have important associations with overuse.²² Further work is needed to determine whether a meaningful composite measure of overuse can be constructed from individual quality measures at the clinician level. Additionally, future work should seek to identify the key drivers that lead clinicians towards over-testing and overtreatment. Insights into

not only the prevalence of overuse but also its causes will be needed to inform the development of future interventions that aim to change this behavior. This study adds to our understanding of how some clinicians (even in the absence of an improvement intervention) are not overusing the tests we examined, Exploring the knowledge and attitudes of these low-using physicians may also help identify viable ways to reduce overuse.

LIMITATIONS

First, this study was conducted within one health system and may not be generalizable to all settings. We expect there might be more variation at the clinician level if a wider range of clinicians were examined. Second, the measures of overuse relied on discretely coded data and may not have captured all the information that justifies clinical decision making. Thus, it is possible that in some cases there was inconsistency between what was documented using discrete diagnostic codes and the clinical rationale provided in the free text of the chart. We acknowledge that some patients always pose exceptions to quality measures and the optimal rate for these measures is probably not zero. This is of concern particularly for the urine testing measure. Third, some variation could be due to variation in coding practices rather than true differences in clinical care. Fourth, for the urine testing measure, we used the proportion of urine tests that were inappropriate, not the proportion of patients, because of the high frequency for which urine tests are ordered in this population. Lastly, the diabetes measure we used may underestimate the overuse of hypoglycemic treatment in older adults. For some individuals with advanced age, serious comorbidities, or frequent or serious hypoglycemia, greater de-intensification of insulin or oral hypoglycemics may be warranted.²⁴

CONCLUSIONS

We were able to design and implement several electronic clinical quality measures of overuse pertaining to ambulatory care of older adults. We found considerable room for improvement in all three measures and significant variability in process measures at the individual clinician level. Further studies should explore these measures in other populations and examine the clinical outcomes of interventions aimed both at lowering the use of these practices overall and at reducing unjustified clinician-level variability.

Acknowledgments:

Dr. Persell, Dr. Rowe, and Ms. Lee had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Dr. Persell receives unrelated research support from Omron Healthcare Co, LTD, and has previously received unrelated research support from Pfizer, Inc.

Dr. Friedberg, since 2016, has received financial support for research from the RAND Corporation, Agency for Healthcare Research and Quality, American Board of Medical Specialties Research and Education Foundation, American Medical Association, Center for Medicare & Medicaid Innovation, Centers for Medicare & Medicaid Services, Cedars-Sinai Medical Center, Commonwealth Fund, Milbank Memorial Fund, National Institute on Aging, National Institute on Drug Abuse, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute on Minority Health and Health Disparities, Patient-Centered Outcomes Research Institute, and Washington State Institute for Public Policy. Since 2016, Dr. Friedberg has received payments from Consumer Reports for consulting services, from Wolters Kluwer for co-authorship of an UpToDate article about hospital quality measurement, and from Harvard Medical School for tutoring medical students in health policy. Since 2016, Dr. Friedberg has received support to attend meetings from the American Medical Association, Gordon and Betty Moore Foundation, and the United States Department of Veterans Affairs. Dr. Friedberg also has a clinical practice in primary care at Brigham and Women's Hospital and thus receives payment for clinical services, via the Brigham and Women's Physician Organization, from dozens of commercial health plans and government payers, including but not limited to Medicare, Medicaid, Blue Cross and Blue Shield of Massachusetts, Tufts Health Plan, and Harvard Pilgrim Health Plan, which are the most prevalent payers in Massachusetts.

Funding/Support: The study was funded by grant R21AG057383 from the National Institute on Aging, National Institutes of Health.

Role of the Funding Source: The content is solely the responsibility of the authors and does not necessarily represent views of the National Institutes of Health. The funder had no role in the design, collection, or interpretation of data or in the decision to submit for publication.

Corresponding Author: Theresa A. Rowe, DO, MS; Division of General Internal Medicine and Geriatrics, Department of Medicine Northwestern University Feinberg School of Medicine, Chicago, IL, USA (e-mail: theresa.rowe@northwestern.edu).

Compliance with Ethical Standards:

Conflict of Interest: The authors declare that they do not have a conflict of interest.

REFERENCES

- Baker DW, Gaseem A, Reynolds PP, Gardner LA, Schneider EC, American College of Physicians Performance Measurement C. Design and use of performance measures to decrease low-value services and achieve cost-conscious care. *Ann Intern Med.* 2013;158(1):55–59.
- Morgan DJ, Dhruva SS, Wright SM, Korenstein D. 2016 Update on Medical Overuse: A Systematic Review. *JAMA Intern Med.* 2016;176(11):1687–1692.
- Levinson W, Kallewaard M, Bhatia RS, et al. 'Choosing Wisely': a growing international campaign. *BMJ Qual Saf.* 2015;24(2):167–174.
- Choosing Wisely. Ten Things Clinicians and Patients Should Question. American Geriatrics Society <http://www.choosingwisely.org/wp-content/uploads/2015/02/AGS-Choosing-Wisely-List.pdf>. Accessed October 31, 2018.
- Gordon LB, Waxman MJ, Ragsdale L, Mermel LA. Overtreatment of presumed urinary tract infection in older women presenting to the emergency department. *J Am Geriatr Soc.* 2013;61(5):788–792.
- Lipska KJ, Ross JS, Miao Y, Shah ND, Lee SJ, Steinman MA. Potential overtreatment of diabetes mellitus in older adults with tight glycemic control. *JAMA Intern Med.* 2015;175(3):356–362.
- Nicolle LE, Bradley S, Colgan R, et al. Infectious Diseases Society of America guidelines for the diagnosis and treatment of asymptomatic bacteriuria in adults. *Clin Infect Dis.* 2005;40(5):643–654.
- Etzioni R, Penson DF, Legler JM, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. *J Natl Cancer Inst.* 2002;94(13):981–990.
- Finucane TE. "Tight control" in geriatrics: the emperor wears a thong. *J Am Geriatr Soc.* 2012;60(8):1571–1575.
- Morgan DJ, Dhruva SS, Coon ER, Wright SM, Korenstein D. 2017 Update on Medical Overuse: A Systematic Review. *JAMA Intern Med.* 2018;178(1):110–115.
- Moyer VA, Force USPST. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med.* 2012;157(2):120–134.
- Trautner BW, Petersen NJ, Hysong SJ, Horwitz D, Kelly PA, Naik AD. Overtreatment of asymptomatic bacteriuria: identifying provider barriers to evidence-based care. *Am J Infect Control.* 2014;42(6):653–658.
- Arnold SV, Lipska KJ, Wang J, Seman L, Mehta SN, Kosiborod M. Use of Intensive Glycemic Management in Older Adults with Diabetes Mellitus. *J Am Geriatr Soc.* 2018;66(6):1190–1194.
- LeRoith D, Biessels GJ, Braithwaite SS, et al. Treatment of Diabetes in Older Adults: An Endocrine Society* Clinical Practice Guideline. *J Clin Endocrinol Metab.* 2019;104(5):1520–1574.
- McKinlay JB, Link CL, Freund KM, Marceau LD, O'Donnell AB, Lutfey KL. Sources of variation in physician adherence with clinical guidelines: results from a factorial experiment. *J Gen Intern Med.* 2007;22(3):289–296.
- Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL. Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. *Ann Intern Med.* 2002;136(2):111–121.
- Tan A, Zhou J, Kuo YF, Goodwin JS. Variation among Primary Care Physicians in the Use of Imaging for Older Patients with Acute Low Back Pain. *J Gen Intern Med.* 2016;31(2):156–163.
- Mody L, Juthani-Mehta M. Urinary tract infections in older women: a clinical review. *JAMA.* 2014;311(8):844–854.
- Jaramillo E, Tan A, Yang L, Kuo YF, Goodwin JS. Variation among primary care physicians in prostate-specific antigen screening of older men. *JAMA.* 2013;310(15):1622–1624.
- Pendrith C, Bhatia M, Ivers NM, et al. Frequency of and variation in low-value care in primary care: a retrospective cohort study. *CMAJ Open.* 2017;5(1):E45–E51.
- Kaplan SH, Griffith JL, Price LL, Pawlson LG, Greenfield S. Improving the reliability of physician performance assessment: identifying the "physician effect" on quality and creating composite measures. *Med Care.* 2009;47(4):378–387.
- 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.
- 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.
- American Diabetes A. 6. Glycemic Targets: Standards of Medical Care in Diabetes-2019. *Diabetes Care.* 2019;42(Suppl 1):S61–S70.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.