most vulnerable times in their lives—as in any other specialty. Not only are minority physicians more likely to work in underserved areas when their training is complete, but the quality of health care may be better when clinicians and patients are of the same race or ethnicity.5,6

The findings also raise questions, however. For example, it is not clear what makes obstetrics and gynecology so attractive to black and Hispanic trainees. I am unaware of any organized program to attract students or retain trainees who are underrepresented minorities. Data from the Association of American Medical Colleges show that between 1980 and 2012, the number of black women graduates of US medical schools increased 4-fold while the number of black men graduating declined.7 Most of the gains for underrepresented minorities in obstetrics and gynecology are likely among black women, partly owing to the high percentage of women entering the specialty.

We lack data on the number of black and Hispanic trainees who complete the 4-year obstetrics and gynecology residency or further subspecialty training in urogynecology, maternal fetal medicine, gynecologic oncology, reproductive endocrinology, and infertility or family planning. These subspecialists account for the majority of researchers and medical school faculty and care for many underserved women with complex medical needs. Not only are black and Hispanic physicians underrepresented among medical school faculty, but they are less likely to be promoted, less likely to hold senior faculty and administrative positions, and less likely to be funded by the National Institutes of Health.8,9 Similarly, the number of women in leadership positions within academic institutions is low when compared with the increasing number of women in obstetrics and gynecology and other specialties.

It is important to learn what factors contributed to the increase in numbers of women, blacks, and Hispanics in obstetrics and gynecology and whether those factors could improve representation in other specialties. It is also important, however, for obstetrics and gynecology and other specialties in which diversity is improving to monitor their workforce, the quality of postgraduate training, and the advancement of women, blacks, and Hispanics to senior and leadership roles. Progress will stall if women or underrepresented minorities do not advance. Ensuring a diverse physician workforce will require the continuing attention of medical school leadership and health care systems, and interventions to provide opportunities for diverse physicians to join the leadership ranks. Increasing physician diversity is yet another opportunity to improve the quality of care for all of our patients, particularly the most disadvantaged and those with a disproportionate burden of disease.

Laura E. Riley, MD

Author Affiliation: Department of Obstetrics and Gynecology, Massachusetts General Hospital, Boston.

Corresponding Author: Laura E. Riley, MD, Department of Obstetrics and Gynecology, Massachusetts General Hospital, 32 Fruit St, Boston, MA 02114 (rilley@mgh.harvard.edu).


Conflict of Interest Disclosures: None reported.


Undermeasuring Overuse—An Examination of National Clinical Performance Measures

Clinical performance measures, designed both to evaluate and motivate clinician and institutional performance, have assumed a central role in efforts to improve the quality of US health care. Concerns have been raised, however, about the collective power of such measures to influence practice on a large scale.1,2 In particular, some worry that if measures predominantly target underuse of care—and incentives tend to reward clinicians for doing more—this could inadvertently contribute to the problem of overuse.3 We sought to determine whether and to what extent outpatient process measures preferentially target underuse compared with overuse.

| Box. Examples of Clinical Performance Measures, by Target Issue |
| Measures targeting underuse: “Has too little care been provided?” |
| • The percentage of patients for whom a lipid panel is ordered within 3 months after being prescribed lipid-lowering medication (goal: high percentage) |
| • The percentage of patients with deep vein thrombosis prescribed anticoagulation for at least 3 months after the diagnosis (goal: high percentage) |
| Measures targeting overuse: “Has too much care been provided?” |
| • The percentage of patients undergoing back imaging within 28 days of a visit for new low back pain (goal: low percentage) |
| • The percentage of patients dispensed an antibiotic within 3 days of diagnosis with bronchitis (goal: low percentage) |
| Measures targeting misuse: “Has care been provided incorrectly?” |
| • The median time from emergency department arrival to time of transfer to another facility for acute coronary intervention (goal: shorter time) |
| • The percentage of patients 18 years or older with pneumonia who receive their first dose of antibiotics within 6 hours after arrival at the hospital (goal: high percentage) |
Methods  |  We identified all outpatient and emergency department (ED) process measures contained in major national measure programs and clearinghouses (eg, National Quality Forum) as of mid-2012. We excluded measure collections addressing exclusively inpatients and collections of other measure collections, and excluded measures if they pertained to outcomes, settings other than outpatient office or ED, or non-clinical aspects of care. Across 16 measure collections, we identified 1037 unique measures, of which 521 (50.2%) met inclusion criteria. Institutional review board approval was not sought because no human subjects were involved.

Three coders independently categorized each measure according to target issue (underuse, overuse, or misuse; $\kappa = 0.73$) and clinical service addressed (eg, laboratory testing). Measures targeting underuse were defined as those asking “Has too little care been provided?”; overuse, “Has too much care been provided?”; and misuse, “Has care been provided incorrectly?” (Box). Coding rules were specified a priori. Coding discrepancies were reconciled by periodic consensus conference or, rarely, arbitrated by a fourth coder. We determined measure frequencies by target issue—overall, within each measure collection, and within each clinical service category. Analyses were performed using Stata statistical software (version 11.2; StataCorp).

Results  |  Of 521 unique measures that met inclusion criteria, 477 (91.6%) targeted underuse while 34 (6.5%) targeted overuse; 14 (2.7%) addressed misuse (4 measures addressed 2 target is-
Of 16 measure collections, just 3 contained an appreciable (≥10%) representation of overuse measures; nearly half (7 of 16) contained no overuse measures (Figure).

Most overuse measures (82.4%) addressed either diagnostic imaging or medication prescription (Figure). By comparison, underuse was well represented (over half of measures) as a target of measures across all categories of clinical service.

**Discussion** Providing high-quality health care requires both providing beneficial care and reducing nonbeneficial care. Increasingly, primary care clinicians and others worry that performance measurement may, through an emphasis on identifying and penalizing underuse, foster a culture of “more is better” and inadvertently encourage overuse. To our knowledge, our study is the first to systematically examine and quantify the existence of such an emphasis: current outpatient clinical process measures, both overall and within nearly all major measure collections, overwhelmingly target underuse of clinical services.

Performance measurement is well positioned to address both underuse and overuse—if, in the aggregate, clinicians are encouraged to ask themselves, “Am I doing enough for this patient without doing too much?” We believe our findings highlight the need to anticipate and monitor the aggregate effects—both intended and unintended—of measure program implementation. We would, moreover, advocate the development and implementation of a prospective underuse/overuse taxonomy as one means by which to promote greater balance across measure collections—or within individual measures—that simultaneously address underuse and overuse. Such a Goldilocks approach to performance measurement, as has been previously proposed, could encourage clinicians and institutions to target a balance of care that is just right.

Notwithstanding certain limitations of the present study, which focuses on measures themselves rather than on the putative connection between measure balance and physician behavior, we have shown that the current state of outpatient clinical performance measurement fails to address overuse—and may inadvertently reward it.

**LESS IS MORE**

**Urinalysis Orders Among Patients Admitted to the General Medicine Service**

Urinalysis (UA) is a frequently ordered rapid screening test to exclude the presence of a urinary tract infection (UTI) among patients admitted to the general medicine (GM) service from the emergency department. Despite its excellent negative predictive value, a positive UA result is nonspecific because it occurs in as many as 90% of asymptomatic elderly patients. We hypothesized that overuse of UA in the emergency department contributes to overdiagnosis and excessive use of antibiotics for UTI among patients admitted to the GM service.

**Methods** We conducted a prospective cohort study of consecutive adult patients to assess the appropriateness of UA orders on admission to the GM service of a large tertiary care center for 4 consecutive weeks in September to October 2014 and 3 consecutive weeks in January 2015. Each patient was assessed within 24 hours for indications for UA, including symptoms of UTI based on guidelines for patients with and without urinary catheters or acute kidney injury, defined as a 2-fold rise in serum creatinine levels. We recorded the frequency of empirical therapy for UTI, orders for urine culture (UC), and antimicrobial prescriptions based on UC results. We compared the proportion of patients who underwent UC or received antibiotic treatment (empirical or by UC result) using χ² tests based on a positive or negative UA result. We assessed predictors of UA orders without indication (UTI or acute kidney injury) using a multivariable logistic regression model with 75 years or older, sex, residence in long-term care, diabetes mellitus, dementia, or presence of 3 or more comorbidities as variables. We obtained approval from the research ethics board of Stony Brook Health Sciences Center. All data were deidentified and informed consent was waived.

**Author Affiliations:** Department of Emergency Medicine, Stony Brook University Medical Center, Stony Brook, New York (Newton); Division of Trauma Surgery, Department of Surgery, Stony Brook University Medical Center, Stony Brook, New York (Zazzera); Stamford Health System, Stamford, Connecticut (Van Moorsel); Outcomes Group, VA Medical Center, White River Junction, Vermont (Sirovich); The Dartmouth Institute for Health Policy and Clinical Practice, Hanover, New Hampshire (Sirovich).

**Accepted for Publication:** June 17, 2015.

**Corresponding Author:** Erika H. Newton, MD, MPH, Department of Emergency Medicine, Stony Brook University Medical Center, HSC Level 4, Room 080, Stony Brook, NY 11794-8350 (erika.newton@stonybrookmedicine.edu).

**Published Online:** August 10, 2015. doi:10.1001/jamainternmed.2015.4025.

**Author Contributions:** Dr Newton had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study concept and design:** Newton, Sirovich.

**Acquisition, analysis, or interpretation of data:** All authors.

**Drafting of the manuscript:** Newton.

**Critical revision of the manuscript for important intellectual content:** All authors.

**Statistical analysis:** Newton, Sirovich.

**Administrative, technical, or material support:** Newton.

**Study supervision:** Sirovich.

**Conflict of Interest Disclosures:** None reported.

**Disclaimer:** The views expressed herein do not necessarily represent the views of the Department of Veterans Affairs or the US Government.


