Using Unannounced Standardized Patients to Explore Variation in Care for Patients With Depression

Sondra Zabar, MD
Kathleen Hanley, MD
Amanda Watsula-Morley, MA
Jenna Goldstein, BA
Lisa Altshuler, PhD
Heather Dumorne, BA
Andrew Wallach, MD
Barbara Porter, MD
Adina Kalet, MD, MPH
Colleen Gillespie, PhD

ABSTRACT

Background Physicians across specialties need to be skilled at diagnosing and treating depression, yet studies show underrecognition and inadequate treatment. Understanding the reasons requires specifying the influence of patient presentation, screening, and physician competence.

Objective We deployed an unannounced standardized patient (SP) case to assess clinic screening and internal medicine (IM) residents' practices in identifying, documenting, and treating depression.

Methods The SP represented a new patient presenting to the outpatient clinic, complaining of fatigue, with positive Patient Health Questionnaire (PHQ) items 2 and 9 and a family history of depression. The SPs assessed clinic screening and IM resident practices; appropriate treatment was assessed through chart review and defined as the resident doing at least 1 of the following: prescribing a selective serotonin reuptake inhibitor (SSRI), making a referral, or scheduling a 2-week follow-up.

Results Of 129 IM residents, 85 (66%) provided appropriate treatment, 79 (61%) appropriately referred, 59 (46%) prescribed an SSRI, and 49 (38%) scheduled a 2-week follow-up, while 40 (31%) did not add depression to the problem list. The IM residents who used PHQ-2 and PHQ-9 were more likely to appropriately (89%) versus inappropriately (50%) treat ($P < .001$). Compared with those who did not, residents who treated appropriately assessed depression symptoms more ($P < .001$) and had better communication (73% versus 50%, $P = .02$), patient centeredness (74% versus 42%, $P = .03$), and patient activation skills (35% versus 11%, $P < .001$).

Conclusions The use of unannounced SPs helps identify targets for training residents to provide evidence-based treatment of depression.

Introduction

Diagnosing depression is often complicated by the initial clinical presentation. Approximately two-thirds of depressed individuals present to primary care with somatic complaints,[1–4] which reduce the likelihood of recognition, and therefore is the leading contributor to missed depression diagnoses.[4] Missed diagnoses may result in a search for other explanations of the symptoms, causing unnecessary medical testing.[5] Even with a positive screening Patient Health Questionnaire (PHQ)-2, only 5% of patients receive a follow-up PHQ-9,[5] and only 34% have these results addressed.[6] Even with a correct diagnosis, only half of the patients with diagnosed depression receive treatment.[5]

DOI: http://dx.doi.org/10.4300/JGME-D-17-00736.1

Editor's Note: The online version of this article contains all details of the unannounced standardized patient case and the fatigue standardized patient checklist.

Internal medicine (IM) residents' accuracy at assessing depression in patients has been shown to be poor[7,8] but can be improved by a screening or inventory tool.[8] While depression cannot be managed effectively without being identified, screening alone is not sufficient. Systems-based approaches to ensure comprehensive screening, diagnosis, and collaborative management should contribute to patients receiving appropriate care,[9,10] yet disaggregating the contribution of the individual physician from that of the care system to inadequate depression care is difficult.

Unannounced standardized patients (SPs) have been used to assess clinical skills, residents' professionalism, physician compliance with clinical guidelines, response to requests for direct-to-consumer advertised antidepressants, and efficacy of educational interventions.[11,12] Studies have explored the acceptable feasibility and validity of unannounced SPs to assess the quality of care and have shown SPs to be an effective means for obtaining detailed information on practice in context.[13] Since 2009, our institution has used unannounced SPs to

unobtrusively assess learners and the clinical microsystem.[14] Currently, residents in our IM program may see between 4 and 6 unannounced SP cases (including the depression case described in this article) designed to assess the care provided to a range of common clinical primary care issues.

The study took advantage of the controlled nature of unannounced SPs to deliver a consistent portrayal of a depressed patient and to explore whether residents' ability to diagnose and manage depression is associated with clinic screening processes and trainees' patient interaction skills.

What was known and gap
Studies have shown underrecognition and inadequate treatment of depression in primary care settings.

What is new
Unannounced standardized patients presented with symptoms of depression to an internal medicine ambulatory clinic to assess the percentage of cases where residents appropriately diagnosed and initiated treatment.

Limitations
Single site, single specialty study limits generalizability; impact of supervising faculty on treatment decisions was not assessed.

Bottom line
Use of standardized patients helped identify areas for improvement in diagnosing and treating depression in ambulatory internal medicine.

Methods

Unannounced SPs were seen by IM residents (including categorical and primary care tracks) during their continuity clinic at 2 urban primary care clinics. Residents were informed that they would see an unannounced SP as part of their own assessment. Data are reported only for residents who consented to include their data in a research registry (92%, 129 of 140 residents). Residents who consented did not differ from those who did not consent in terms of postgraduate year [PGY], sex, or age. The program's curriculum on depression included a 2-hour session during ambulatory block, practice of skills and feedback in annual objective structured clinical examinations with depression cases, and ongoing just-in-time teaching during precepting.

The SPs were men in their early 20s, and they were trained to present to the clinic as the same new patient with fatigue and insomnia. If asked or screened, they described symptoms positive for depression based on the PHQ-2 and PHQ-9, widely used depression screening tools with evidence of reliability and validity.[15] Full case details are provided as online supplemental material. The SPs (actors) received 6 hours of training in assessment and character portrayal, including how to meet criteria for a positive response on PHQ-2 and PHQ-9 (scores of 14–16).

In both clinics, medical assistants are expected to screen all new patients for depression using the PHQ-2 and, if positive, conduct the full PHQ-9. Results are expected to be communicated to the physician via the electronic health record (EHR) prior to the patient encounter.

Case fidelity was maintained through ongoing review of audio recordings of the SP-physician encounter. The documented PHQ-9 scores of the SPs averaged 14.6 over 129 visits (SD = 2.3), suggesting consistent, accurate portrayal of symptoms. Based on resident feedback, the estimated detection rate for this case was 10%, with almost all detections occurring after the visit. Overall scores did not differ by whether the visit was detected.

We used 2 assessments to describe residents' management of this case: (1) a comprehensive SP checklist to capture specific practices associated with the visit and competence in clinical patient interaction skills using behaviorally anchored items (provided as online supplemental material),[16] and (2) a systematic review of the EHR notes written by residents. Data were collected and managed using REDCap electronic data capture tools hosted at the New York University Langone Medical Center.[17]

The SP checklist assessed whether the medical assistant screened for depression using the PHQ-2 and followed up with PHQ-9, and it evaluated residents' exploration of depressive symptom criteria (sleep, energy, concentration, appetite, psychomotor retardation, guilt, suicidal thoughts). The unannounced SPs indicated the degree to which the resident explored the patient's social/family support, family medical history, unhealthy habits (alcohol, smoking, drug use), and past medical history, and reported on the extent of education they received from the resident about the connection between the presenting symptoms and depression.

Residents' clinical skills, as assessed through the checklist, included 12 items in the domains of communication skills (4 information gathering, 5 relationship development, 3 patient education and counseling); 4 items on patient centeredness (answered all questions, took a personal interest in the patient); and 4 patient activation items (encounter helped the patient understand the condition and feel confident he or she could manage it). Items used a *not done*, *partly done*, and *well done* scale with behavioral anchors. Domain scores were calculated as percentage of items rated *well done*. Internal

consistency for each domain, measured using Cronbach's alpha, exceeded 0.70.

We reviewed residents' EHR notes for treatment recommendations, prescriptions, referrals, recommended time frames for follow-up visits, and whether depression was included on problem lists. Chart reviews were completed by 2 research assistants, who were blind to the study hypotheses. Initial checks on interrater reliability using a random sample of 20 charts showed strong agreement between raters (>85%) for all abstracted items. Appropriate treatment was defined as 1 or more of the following: prescribing a selective serotonin reuptake inhibitor (SSRI), making a referral, and scheduling a follow-up visit in less than 2 weeks.

Residents were grouped into those who treated the depression case appropriately and those who did not. The 2 groups were compared in terms of resident visit practices and overall clinical skills. We also explored whether treatment patterns differed by PGY, residency track (categorical or primary care), or site. We used independent sample *t* tests to compare group means on scores (percentage well done) and chi-square statistics to compare distributions.

This study was conducted through a resident medical education research registry[18] approved by the New York University School of Medicine Institutional Review Board. Residents consented to including their education data in a research database.

Results

A total of 129 IM residents saw the unannounced SP depression case. From walking in the front door to leaving the clinic, mean visit length was 43 minutes (SD = 18 minutes, range 14–125 minutes).

Of the SP cases, 92% (119 of 129) were screened for depression by the medical assistant. Screening rates did not differ among clinic locations or over time. In the 8% (10 of 129) of cases where the SP was not screened by the medical assistant, no resident listed depression as the primary diagnosis. Overall, 29% (37 of 129) of residents listed depression/dysthymia as the primary diagnosis in the EHR problem list, and 35% (45 of 129) ranked insomnia as the primary diagnosis. Only 48% of residents (62 of 129) included depression/dysthymia on the problem list at all (64% [83 of 129] documented it as a routine examination, and 54% [70 of 129] documented it as insomnia). In addition, 70% (90 of 129) fully explored the patients' depression symptoms, and 53% (68 of 129) connected the presenting symptom (fatigue) with depression.

The FIGURE describes how residents managed this depression case, classifying practice into appropriate (66%, 85 of 129 residents) or inappropriate treatment (34%, 44 of 129 residents). Residents made a psychiatric or collaborative care referral in 61% (52 of 85) of visits, prescribed an SSRI in 46% (39 of 85) of visits, and scheduled a follow-up appointment in 38% (32 of 85) of visits. The FIGURE shows the distinct subgroups of treatment combinations.

Characteristics for the appropriate and inappropriate treatment groups are shown in TABLES 1 and 2. Both groups had similar rates of assessing patients' medical history and substance use, but significantly more of those who provided appropriate treatment fully explored the depressive symptoms (83%, 71 of 85) than those who did not (41%, 18 of 44; X^2 (1, N = 128) = 25.93; $P < .001$). Significantly more residents who treated appropriately had the patients' PHQ-9 (in addition to PHQ-2) data (89% [76 of 85] versus 50% [22 of 44]; chi-square [1, N = 129] = 24.67; $P < .001$) and included depression on the problem list (69% [59 of 85] versus 14% [6 of 44]; chi-square [1, N = 129] = 36.08; $P < .001$). The 2 groups did not significantly differ in terms of visit length, average return visit, PGY, attending physician, or clinic location.

Residents who provided appropriate treatment had significantly higher scores in communication domains and subdomains (information gathering, relationship development, patient education, and counseling). This group also had higher scores in patient centeredness (74% [63 of 85] versus 42% [18 of 44], $P = .003$) and patient activation (35% [30 of 85] versus 11% [5 of 44], $P < .001$). Residents who provided appropriate treatment in the depression case did not consistently have better performance in other unannounced SP cases compared with those who did not.

Discussion

We found substantial variation of care for a common presentation of depression in a resident continuity clinic. Residents were more likely to provide appropriate treatment when a PHQ-9 was completed by medical assistants, supporting the proposition that physicians' clinical reasoning may be influenced by how the resident and clinical microsystem work together.

We found that the majority of visits with inappropriate treatment were associated with a positive PHQ-2, suggesting that these residents did not know, or were reluctant to follow, depression management guidelines. Equally striking was the fact that when making appropriate treatment decisions, only two-thirds of residents committed to a depression diagnosis in the medical record. While it is understandable that a resident may not want to commit to this

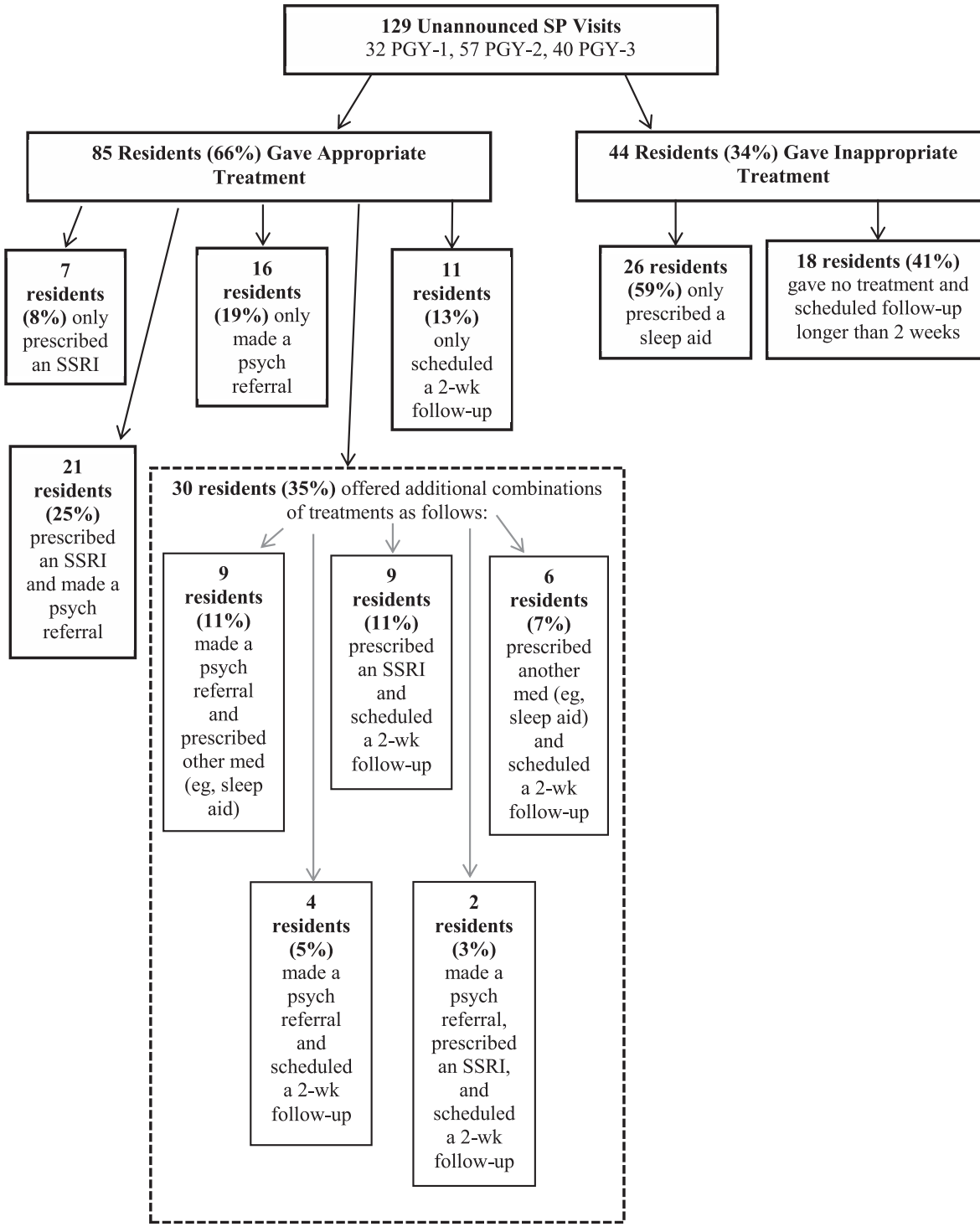

FIGURE
Description of Appropriate and Inappropriate Treatment for Standardized Depression Case

diagnosis on the first visit, it is troubling that return visits were not consistently scheduled to enable proper monitoring of the patient.

Residents with more effective communication, patient centeredness, and patient activation skills were more likely to diagnose depression and initiate appropriate treatment. This is consistent with previous literature, which found positive associations between patient satisfaction, physician empathy, and treatment adherence for depressed patients,[19] and

TABLE 1
Differences in Context, Clinic, and Resident Characteristics by Appropriate Versus Inappropriate Treatment

Characteristic	Provided Appropriate Treatment (n = 85 visits)	Did Not Provide Appropriate Treatment (n = 44 visits)	Test Statistic (*P* value)
Length of visit, min			
Mean (SD)	44.54 (19.08)	39.60 (15.30)	*t* test = –1.59 (.06)
Range	14–125	15–80	
Clinic location, No. (%)			
Bellevue	49 (57)	29 (66)	chi-square = 0.83 (.36)
Gouverneur	36 (43)	15 (34)	
Internal medicine residency track, No. (%)			
Primary care	52 (61)	18 (41)	chi-square = 3.02 (.21)
Categorical	33 (39)	26 (59)	
Postgraduate year (PGY), No. (%)			
PGY-1	24 (28)	8 (18)	chi-square = 4.84 (.09)
PGY-2	40 (47)	17 (39)	
PGY-3	21 (25)	19 (43)	
Screening, No. (%)			
No PHQ-2 or PHQ-9	4 (5)	6 (14)	chi-square = 24.67 (< .001)
PHQ-2	5 (6)	16 (36)	
PHQ-2 and PHQ-9	76 (89)	22 (50)	
Assessment, No. (%)			
Fully assessed depression	71 (83)	18 (41)	chi-square = 25.93 (< .001)
Problem list included, No. (%)			
Depression/dysthymia	59 (69)	46 (14)	chi-square = 36.08 (< .001)
Insomnia	35 (41)	35 (80)	chi-square = 18.54 (< .001)
Routine examination	52 (61)	33 (75)	chi-square = 3.14 (.21)
Other	21 (25)	5 (11)	chi-square = 4.52 (.11)
Return visit, wk			
Mean (SD)	8.40 (13.61)	11.41 (13.12)	*t* test = 1.21 (.12)
Range	1–53	3–54	

Abbreviation: PHQ, Patient Health Questionnaire.

showed that shared decision-making increases the probability that depressed patients see an improvement in symptoms.[20] Since patient activation is critical in improving health outcomes for depressed patients,[21] the low scores in this domain are an area for improvement.[22]

Our study has limitations. The unannounced SP methodology is limited to new patient visits and does not provide insight into what would transpire during follow-up visits. Our determination of appropriate versus inappropriate treatment was based on chart reviews, which may not fully capture residents'

TABLE 2
Differences Between Residents Who Provided Appropriate Treatment and Those Who Did Not: Core Clinical Skills

Core Clinical Skills	Provided Appropriate Treatment (n = 85), Mean % Well Done (SD)	Did Not Provide Appropriate Treatment (n = 44), Mean % Well Done (SD)	*t* Test (*P* Value)
Overall communication	73 (25)	50 (32)	–4.43 (.002)
Information gathering	73 (29)	52 (37)	–3.45 (.006)
Relationship development	77 (26)	57 (40)	–3.24 (.001)
Patient education and counseling	64 (37)	35 (36)	–4.29 (.001)
Patient centeredness	74 (32)	42 (38)	–5.21 (.003)
Patient activation	35 (40)	11 (30)	–3.47 (.001)

clinical decisions. The sample was limited to residents from 1 program and, thus, may not generalize; it is also unclear how attending supervisors influenced treatment decisions.

The unannounced SP case helped uncover deficits in residents' diagnosis and treatment of depression, while reinforcing the importance of depression screening protocols in the clinical microsystem. We plan to do more resident education and faculty development on depression management and documentation. In future studies, we will explore both why residents were not comfortable in making a diagnosis even when following appropriate treatment guidelines and their reluctance to activate depression management protocols in the face of a positive screening.

Conclusion

The use of an unannounced SP presenting with a somatic complaint and underlying depression reinforced the importance of effective clinic screening and helped target areas for improvement in residents' depression curriculum and training to ensure that residents effectively treat depression in the primary continuity care setting.

References

1. Stadje R, Dornieden K, Baum E, et al. The differential diagnosis of tiredness: a systematic review. *BMC Fam Pract*. 2016;17(1):147.
2. Petterson S, Miller BF, Payne-Murphy JC, et al. Mental health treatment in the primary care setting: patterns and pathways. *Fam Syst Health*. 2014;32(2):157–166.
3. Tylee A, Gandhi P. The importance of somatic symptoms in depression in primary care. *Prim Care Companion J Clin Psychiatry*. 2005;7(4):167–176.
4. Timonen M, Liukkonen T. Management of depression in adults. *BMJ*. 2008;336(7641):435–439.
5. Goldman LS, Nielsen NH, Champion HC. Awareness, diagnosis, and treatment of depression. *J Gen Intern Med*. 1999;14(9):569–580.
6. Tiemstra JD, Fang K. Depression screening in an academic family practice. *Fam Med*. 2017;49(1):42–45.
7. Medow MA, Borowsky SJ, Dysken S, et al. Internal medical residents' ability to diagnose and characterize major depression. *West J Med*. 1999;170(1):35–40.
8. Sliman RJ, Donohue TA, Jarjoura D, et al. Recognition of depression by internal medicine residents. *J Community Health*. 1992;17(3):143–152.
9. Lin EH, Simon GE, Katon WJ, et al. Can enhanced acute-phase treatment of depression improve long-term outcomes? A report of randomized trials in primary care. *Am J Psychiatry*. 1999;156(4):643–645.
10. Ludman E, Von Korff M, Katon W, et al. The design, implementation, and acceptance of a primary care-based intervention to prevent depression relapse. *Int J Psychiatry Med*. 2000;30(3):229–245.
11. Kravitz RL, Epstein RM, Feldman MD, et al. Influence of patients' requests for direct-to-consumer advertised antidepressants: a randomized controlled trial. *JAMA*. 2005;293(16):1995–2002.
12. Krane NK, Anderson D, Lazarus CJ, et al. Physician practice behavior and practice guidelines: using unannounced standardized patients to gather data. *J Gen Intern Med*. 2009;24(1):53–56.
13. Peabody JW, Luck J, Glassman P, et al. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA*. 2000;283(13):1715–1722.
14. Zabar S, Ark T, Gillespie C. Can unannounced standardized patients assess professionalism and communication skills in the emergency department? *Acad Emerg Med*. 2009;16(9):915–918.
15. DeJesus RS, Vickers KS, Melin GJ, et al. A system-based approach to depression management in primary care using the Patient Health Questionnaire-9. *Mayo Clin Proc*. 2007;82(11):1395–1402.
16. Zabar S, Kachur E, Hanley K, et al. *Objective Structured Clinical Examinations: 10 Steps to Planning and Implementing OSCEs and Other Standardized Patient Exercises*. New York, NY: Springer; 2012.
17. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377–381.
18. Gillespie C, Zabar S, Altshuler L, et al. The Research on Medical Education Outcomes (ROMEO) Registry: addressing ethical and practical challenges of using "bigger," longitudinal educational data. *Acad Med*. 2016;91(5):690–695.
19. Frankel RM. Emotion and the physician-patient relationship. *Motiv Emot*. 1995;19(3):163–173.
20. Clever SL, Ford DE, Rubenstein LV, et al. Primary care patients' involvement in decision-making is associated with improvement in depression. *Med Care*. 2006;44(5):398–405.
21. Alegría M, Polo A, Gao S, et al. Evaluation of a patient activation and empowerment intervention in mental health care. *Med Care*. 2008;46(3):247.
22. Hibbard JH, Mahoney ER, Stock R, et al. Do increases in patient activation result in improved self-management behaviors? *Health Serv Res*. 2007;42(4):1443–1463.

Sondra Zabar, MD, is Director, Division of General Internal Medicine and Clinical Innovation, Director, Standardized Patient

Program, and Professor, Department of Medicine, New York University School of Medicine; **Kathleen Hanley, MD,** is Associate Director, Primary Care Residency Program, Site Director, Agency for Healthcare Research and Quality R18 Patient Safety, and Associate Professor, Department of Medicine, New York University School of Medicine; **Amanda Watsula-Morley, MA,** is Research Coordinator, Program for Medical Education Innovations and Research, New York University School of Medicine; **Jenna Goldstein, BA,** is Research Assistant, Program Medical Education Innovation and Research, New York University School of Medicine; **Lisa Altshuler, PhD,** is Assistant Professor, Department of Medicine, New York University School of Medicine; **Heather Dumorne, BA,** is Research Data Associate, Program for Medical Education Innovations and Research, New York University School of Medicine; **Andrew Wallach, MD,** is Section Chief, Division of General Internal Medicine, Bellevue Hospital, and Associate Professor, Department of Medicine, New York University School of Medicine; **Barbara Porter, MD,** is Associate Director of Outpatient Experience, Internal Medicine Residency Program, and Assistant Professor, Department of Medicine, New York University School of Medicine; **Adina Kalet, MD, MPH,** is Director, Research on Medical Education Outcomes Unit, Division of General Internal Medicine and Clinical Innovation, and Arnold P. Gold Professor of Humanism and Professionalism, Department of Medicine, New York University School of Medicine; and **Colleen Gillespie, PhD,** is Director, Division of Education Quality, and Associate Professor, Department of Medicine, New York University School of Medicine.

Funding: This project was supported by Agency for Healthcare Research and Quality Grant 11-01865, Health Resources and Services Adminstration Primary Care Grant 11-02247, and REDCap.

Conflict of interest: The authors declare they have no competing interests.

Preliminary results from this study were presented at the Society of General Internal Medicine Annual Scientific Meeting, Toronto, Ontario, Canada, April 22–25, 2015.

Corresponding author: Sondra Zabar, MD, New York University School of Medicine, 550 First Avenue, New York, NY 10016, 212.263.1138, sondra.zabar@nyumc.org

Received October 4, 2017; revisions received January 31, 2018, and March 9, 2018; accepted March 14, 2018.