Background: Worldwide efforts are under way to improve the quality of clinical practice. Most quality measurements, however, are poorly validated, expensive, and difficult to compare among sites.

Objective: To validate whether vignettes accurately measure the quality of clinical practice by using a comparison with standardized patients (the gold standard method), and to determine whether vignettes are a more or less accurate method than medical record abstraction.

Design: Prospective, multisite study.

Setting: Outpatient primary care clinics in 2 Veterans Affairs medical centers and 2 large, private medical centers.

Participants: 144 of 163 eligible physicians agreed to participate, and, of these, 116 were randomly selected to see standardized patients, to complete vignettes, or both.

Measurements: Scores, expressed as the percentage of explicit quality criteria correctly completed, were obtained by using 3 methods.

Results: Among all physicians, the quality of clinical practice as measured by the standardized patients was 73% correct (95% CI, 72.1% to 73.4%). By using exactly the same criteria, physicians scored 68% (CI, 67.9% to 68.9%) when measured by the vignettes but only 63% (CI, 62.7% to 64.0%) when assessed by medical record abstraction. These findings were consistent across all diseases and were independent of case complexity or physician training level. Vignettes also accurately measured unnecessary care. Finally, vignettes seem to capture the range in the quality of clinical practice among physicians within a site.

Limitations: Despite finding variation in the quality of clinical practice, we did not determine whether poorer quality translated into worse health status for patients. In addition, the quality scores are based on measurements from 1 patient–provider interaction. As with all other scoring criteria, vignette criteria must be regularly updated.

Conclusions: Vignettes are a valid tool for measuring the quality of clinical practice. They can be used for diverse clinical settings, diseases, physician types, and situations in which case-mix variation is a concern. They are inexpensive and easy to use. Vignettes are particularly useful for comparing quality among and within sites and may be useful for longitudinal evaluations of interventions intended to change clinical practice.

This instrument design obviates the need to adjust quality scores for the variation in disease severity and comorbid conditions found in actual patient populations. Our vignettes are also distinct from other quality measurements of clinical practice because they do not focus on a single task, or even a limited set of tasks, but instead comprehensively evaluate the range of skills needed to care for a patient.

Vignettes are particularly well-suited for quality assessments of clinical practice that are used for large-scale (8, 9), cross-system comparisons (10, 11) or for cases in which ethical issues preclude involving patients or their records (7, 12, 13). They are also ideal for evaluations that require holding patient variation constant (14, 15) or manipulating patient-level variables (15–17). The appeal of vignettes has resulted in their extensive use in medical school education (18, 19), as well as various studies that explicitly evaluate the quality of clinical practice in real-life settings and comparative analysis among national health care systems (10, 20–23).

Before vignette-measured quality can be used confidently in these settings, however, 2 important questions...
must be answered: How valid are vignettes as a measure of actual clinical practice? Can vignettes discriminate among variations in the quality of clinical practice? This has led to a search to define a gold standard for validation (24–26). We and others have used standardized patients as this standard. Standardized patients are trained actors who present unannounced to outpatient clinics as patients with a given clinical condition. Immediately after meeting with a physician, the standardized patient records on a checklist what the physician did during the visit (26–28). Rigorous methods, which we have described in detail elsewhere (29), ensure that standardized patients can be considered a gold standard. In addition, we have demonstrated the validity of standardized patients as a gold standard by concealing audio recorders on standardized patients during visits. The overall rate of agreement between the standardized patients’ checklists and the independent assessment of the audio transcripts was 91% (26).

We previously used paper-and-pen vignettes in a study limited to only 1 health care system, the Veterans Administration, and found that they seemed to be a valid measure of the quality of clinical practice according to their rate of agreement with standardized patient checklists (26). For this study, we wanted to confirm the validity of vignettes by using a more complex study design that introduced many more randomly assigned physicians, a broader range of clinical cases, and several sites representing different health care systems. We also wanted to test a refined, computerized version of vignettes, which we believe are more realistic and streamline data collection and scoring. We were particularly interested in determining whether the vignettes accurately capture variation in the quality of clinical practice, which has become increasingly prominent in the national debate on quality of care (30, 31). We hoped that vignettes could contribute to this debate by providing a low-cost measure of variation across different health care systems.

**Methods**

**Sites**

The study was conducted in 4 general internal medicine clinics: 2 Veterans Affairs (VA) medical centers and 2 large, private medical centers. One private site is a closed group model, and the other, primarily staffed by employed physicians, contracts with managed care plans. All sites are located in California, and each has an internal medicine residency training program. One VA medical center and 1 private site are located in 1 of 2 cities. The 2 VA medical centers are large, academically affiliated hospitals with large primary care general internal medicine practices. We chose the 2 private sites that were generally similar to the VA medical centers and to each other; each had large primary care practices and capitated reimbursement systems that provide primary care general internists with a broad scope of clinical decision-making authority.

**Study Design**

At each site, all attending physicians and second- and third-year residents who were actively engaged in the care of general internal medicine outpatients were eligible to participate in the study. We excluded only interns. Of 163 eligible physicians, 144 agreed to participate. We informed consenting physicians that 6 to 10 standardized patients might be introduced unannounced into their clinics over the course of a year and that they might be asked to complete an equal number of vignettes.

Sixty physicians were randomly selected to see standardized patients: 5 physicians from each of the 3 training levels at each of the 4 sites (Figure 1). We assigned standardized patients to each selected physician for 8 clinical cases—simple and complex cases of chronic obstructive pulmonary disease, diabetes, vascular disease, and depression. We abstracted the medical records from the 480 standardized patient visits. Each selected physician also completed a computerized clinical vignette for each of the 8 cases. For standardized patient visits that a selected physician did not complete, a replacement physician, who was randomly selected from the same training level at the same site, completed the visit. Eleven physicians required replacements. The 11 replacement physicians completed 24 standardized patient visits. Each replacement physician completed vignettes for all 8 cases. Finally, we randomly selected 45 additional physicians to serve as controls and complete vignettes (only) for all 8 cases. A total of 116 physicians participated in the study by seeing standardized patients, completing vignettes, or both. Standardized patients presented to the clinics between March and July 2000, and physicians completed vignettes between May and August 2000.

**Vignette Data Collection**

We developed the vignettes by using a standardized protocol. We first selected relatively common medical conditions frequently seen by internists. All selected conditions had explicit, evidence-based quality criteria and accepted standards of practice that could be used to score the vignettes, as well as be measured by standardized patients and chart abstraction. We developed written scenarios that described a typical patient with 1 of the same 4 diseases (chronic obstructive pulmonary disease, diabetes, vascular disease, or depression). For each disease, we developed a simple (uncomplicated) case and a more complex case with a comorbid condition of either hypertension or hypercholesterolemia. This yielded a total of 8 clinical cases. (A sample vignette and scoring sheet are available at www.annals.org.)

The physician completing the vignette “sees the patient” on a computer. Each vignette is organized into 5 sections, or domains, which, when completed in sequential order, recreate the normal sequence of events in an actual patient visit: taking the patient’s history, performing the physical examination, ordering radiologic or laboratory
tests, making a diagnosis, and administering a treatment plan.

For example, the computerized vignette first states the presenting problem to the physician and prompts the physician to “take the patient’s history” (that is, ask questions that would determine the history of the present illness; past medical history, including prevention; and social history). Physicians can record components of the history in any order without penalty. The entire format is open-ended: The physician enters the history questions directly into the computer and, in the most recent computerized versions, receives real-time responses. When the history is completed, the computer confirms that the physician has finished and then provides key responses typical of a patient with the specific case. The same process is repeated for the four remaining domains.

In addition to the open-ended format, we have taken three steps to avoid potential inflation of vignette scores. First, physicians are not allowed to return to a previous domain and change their queries after they have seen the computerized response. Second, the number of queries is limited in the history and physical examination domains. For example, in the physical examination domain, physicians are asked to list only the 6 to 10 essential elements of the examination that they would perform. Third, they are given limited time to complete the vignette (just as time is limited during an actual patient visit).

Each physician completed 8 vignettes: 4 per physician given in a random order during two 60- to 70-minute sittings conducted during daytime work hours at the hospital or clinic. Physicians who saw standardized patients were given the vignette for a particular case only after completing the standardized patient visit. A trained abstractor reviewed the physician’s responses for each completed vignette and assessed whether each explicit scoring criterion was met.

**Standardized Patient Data Collection**

The standardized patients were experienced actors from standardized patient medical education programs. Because 2 of the sites were VA medical centers, we hired only adult male standardized patients. We followed established protocols for standardized patient training, including practice in presenting 1 of the 8 clinical cases and in

---

1 = simple case; 2 = complex case; COPD = chronic obstructive pulmonary disease; DM = diabetes mellitus; MCO = managed care organization; MD = medical doctor (physician); SP = standardized patient; VAMC = Veterans Affairs medical center.
scoring physician performance on a checklist. Training for each case followed a script that exactly corresponded to the clinical scenario presented in the vignette for that case. After training the standardized patients, we enrolled them unannounced into clinics. This enrollment process involved creating fictitious names; personal data; insurance status; and brief medical histories, including laboratory results when indicated. We entered these data into a paper-based and computerized record systems in the specific formats required for each site, including the necessary entries into the electronic medical records, paper clinical records, and administrative systems.

The standardized patient completed a closed-ended checklist immediately after each visit, indicating whether the physician took specific actions corresponding to each scoring criterion. This checklist generated a standardized patient score that could be compared directly with the score of the corresponding vignette completed by the same physician for that clinical case.

**Medical Record Data Collection**

Each standardized patient visit also generated a medical record that we retrieved from the clinic. A trained abstractor examined the information entered into the medical record to determine what specific actions the physician took or did not take during the visit. These actions corresponded exactly with the quality criteria that we used to score the vignettes. The resulting chart abstraction score for each visit was therefore directly comparable to the standardized patient score for that visit and the physician’s vignette score for that clinical case.

**Scoring**

We conceptualize “good-quality” clinical practice as the comprehensive provision of services for a particular clinical case in a manner that leads to better outcomes for patients. We therefore determined what set of actions a physician should take—or should not take—during a patient visit to treat a clinical case in a manner that has been shown to lead to better health outcomes. This yielded a comprehensive set of explicit quality criteria, rather than simply a few selected measures of care that might be gained, such as determining whether a medication was appropriately prescribed or whether the patient was screened for a comorbid condition. Our criteria captured whether the physician 1) determined the entire relevant history; 2) performed the relevant physical examination items; 3) ordered the necessary laboratory or imaging tests; 4) made the correct diagnosis, including etiology; and 5) prescribed a complete treatment (management) plan.

We identified quality criteria from 3 sources: an evidence-based literature search on the clinical practices that lead to better health outcomes, U.S. and international clinical guidelines, and local expert panels of academic and community physicians consisting of both generalists and specialists. We used the recommendations by the expert panels to modify and finalize the master criteria list derived from the literature and guidelines. We used identical criteria to score the vignettes, standardized patients, and charts for a particular clinical case.

We assigned a weight of 1.0 to individual criteria that expert reviewers believed to be most critical for quality care. We grouped criteria that were deemed less important, such as several physical examination items subsidiary to 1 clinical construct, into categories and implicitly assigned them weights less than 1.0.

Most of the explicit criteria measured whether the physician took specific actions necessary to provide good-quality care for a particular case. In addition, we also measured whether the physician took any action that she or he should not have for that case, such as prescribing antibiotics for an infection highly likely to be of viral cause. We measured these items of unnecessary care, including inappropriate tests, procedures, medications, or referrals, only for the domains of testing and treatment because the marginal time cost and risk of unnecessary history or physical examination items is negligible.

Each physician who saw standardized patients had a total of 3 scores (1 for each method) for each case. We compiled a physician’s scores into a set so that the scores could be compared across cases by method. The score for all 3 methods—vignette, standardized patient checklist, and chart abstraction—was equal to the number of categories correct divided by the number of total possible categories (an average of 30 per case). We scored unnecessary care separately by counting the number of unnecessary items for each vignette or standardized patient visit.

To ensure the accuracy of scoring, physician investigators audited randomly selected standardized patient checklists, chart abstraction forms, and vignette scoring forms from each site.

**Statistical Analysis**

We compared mean scores for each of the 3 methods to determine how well vignettes and chart abstraction measured actual quality compared with the standardized patient benchmark. We disaggregated these comparisons by disease, site, case complexity, and physician training level. We evaluated the statistical significance of the differences in mean scores among the 3 methods by using the F test from an analysis of variance (ANOVA) model that considered the matching of vignette, standardized patient, and chart abstraction scores for each physician for each case. Specifically, the 3-way crossed, 1-way nested model included factors for site, physician training level, quality measurement method, and physician (nested within site), plus a site-by-method interaction. Where differences among means for the 3 methods were statistically significant, we used the Tukey–Kramer multiple comparison procedure to evaluate the significance of comparisons between pairs of methods by using a global 5% significance level. We also considered other interaction terms (method by disease, method by case complexity, and method by...
physician training level) in the ANOVA model to assess the consistency of the results across these factors. We estimated the 95% CIs by using adjusted errors to account for the nested study design.

To explore the ability of vignettes to detect variation in performance among sites, we created a box plot of the total scores for each physician for each method for each site. We constructed this plot to show the site median bar, the interquartile range, and the 5th to 95th percentile range. We evaluated variation among sites by comparing site median scores for each method. We measured variation within a site by the distance from the 5th to the 95th percentile. We compared within-site variation across sites by normalizing each site’s 5th to 95th percentile range by its SD for each method.

In addition, we wanted to comprehensively assess how well vignettes measure unnecessary care. Because standardized patients cannot be trained to accurately report unnecessary care, we therefore developed a combined “standardized patient plus chart” measure that added some items from the medical record (including tests, referrals, and medications) to the standardized patient checklist items, yielding a complete set of criteria for a visit that exactly matched the vignette scoring criteria for that case.

We also used the “standardized patient plus chart” to assess the ability of vignettes to detect the poorest performers. To do this, we aggregated the scores from all cases completed by a physician. We reported results as the percentage of times the vignettes and standardized patient scores identified the same physicians performing in the lowest quartile. Because results were the same whether we used lowest quintile, quartile, or tertile, we report the results for only the lowest quartile.

To assess whether seeing a standardized patient before completing the vignette for that case “cued” the physician in a way that affected vignette performance, we compared mean scores for vignettes matched with a standardized patient visit to the mean scores for vignettes without an associated standardized patient visit (mostly those completed by physicians who had not seen standardized patients). We compared vignette scores in the 2 groups by using a linear regression model that accounted for the clustering of physicians within site and training level.

We conducted all statistical analyses by using Stata software, versions 6.0, 7.0, and 8.2 (StataCorp, College Station, Texas).

Role of the Funding Source

The funding source had no role in the design, conduct, or reporting of the study or in the decision to submit the manuscript for publication. We had open and full access to all the data files for the study and had full control over the data.

RESULTS

Figure 2 shows that vignette scores more closely reflected the measured quality of care by the gold standard standardized patients than did the abstracted medical record: 73% (95% CI, 72.1% to 73.4%) of the criteria were correctly done as measured by standardized patients, compared with 68% (CI, 67.9% to 68.9%) for vignettes and 63% (CI, 62.7% to 64.0%) for chart abstraction ($P < 0.001$ per ANOVA model; all pairwise comparisons between methods significant at a global 5% level). Figure 2 also shows that vignette performance with respect to standardized patients was consistent across all 4 diseases evalu-
ated. This consistent ranking of vignettes with respect to standardized patients was also observed across the 4 sites (data not shown). **Figure 3** shows that vignettes performed similarly well for uncomplicated versus complex cases and at all 3 levels of physician training, with scores deviating from the standardized patient encounter by 1% to 6%. Chart abstraction scores less closely reflected the standardized patient scores across all diseases, case complexity, or training levels.

**Figure 4** indicates that vignettes also measured differences in quality among the 4 sites. Specifically, vignettes ranked the sites in the same order, as did the standardized patients. **Figure 4** also indicates that the variation within sites (that is, among providers at the same site) greatly exceeded the variation among sites. Within-site variation (that is, the 5th to the 95th percentile range) ranged from 46.2% for standardized patients to 36.0% for vignettes. Vignettes compared favorably to standardized patients in detecting this variation. When the 5th to 95th percentile ranges were normalized by SD, the variation among sites was 1.8% and 1.0% for standardized patients and vignettes, respectively, while the corresponding within-site variation was 3.3% and 3.2%, respectively.

We investigated 4 other measurement characteristics of vignettes. First, we measured the ability of vignettes to detect physicians who performed in the lowest quartile, according to the comprehensive standardized patient plus chart results. Vignettes identified 8 of 15 (53% [CI, 27% to 79%]) physicians. Second, seeing a standardized patient before completing a vignette seemed to have a small “cueing” effect. Physicians who were cued by previously seeing a standardized patient for a particular case had an absolute average vignette score that was 2.6% higher ($P = 0.031$) than those who only completed the vignette but saw no corresponding standardized patient. Third, vignettes and standardized patients were almost identical in their ability...
to detect unnecessary care. Figure 5 shows that the number of unnecessary items ranged from 0 to 7 per physician per case and that vignettes demonstrated the same distribution of unnecessary tests, treatments, and referrals as standardized patients. Fourth, we compared the time it took personnel to obtain and score the vignettes compared with locating, abstracting, and scoring a medical record. We found that, with trained personnel, it typically takes 40 minutes in total for each medical record versus 25 minutes for a vignette.

**DISCUSSION**

Vignettes can be a valuable tool for measuring the quality of clinical practice (13, 32). Unaddressed concerns about their validity and ability to discern variations in quality, however, have limited their use. The hope has been that vignettes could be constructed in such a way that they would measure the process of care in various clinical settings and even allow for cross-system or cross-national comparisons (8, 10, 14, 23, 33–35). These expectations arise from the observation that vignettes are less expensive to administer than other measurements and eliminate the need for difficult case-mix adjustments. If validated, vignettes also offer the prospect of a more thorough measurement of clinical practice than typical instruments that measure few quality criteria per case. In particular, we have been interested in the possibility of using vignettes to measure the effectiveness of policy interventions that are designed to improve care for an array of clinical conditions and physician characteristics in diverse patient populations (9, 17, 36–45).

We report results from a large validation study of computerized vignettes. We compared the quality of clinical practice, as measured by vignettes, with standardized patient checklists for almost 500 outpatient visits, as well as the medical records of these visits. We made these comparisons across different health care systems, a range of clinical conditions, and different levels of training among randomly sampled general internists in an outpatient setting.

We found that vignettes provided consistently better measurements of the quality of clinical care than did medical record abstraction when we compared both with the standardized patient checklists (the gold standard). This measurement capability of vignettes was robust across 4 different clinical conditions, simple and complex cases, different levels of physician training, and 4 sites. Vignettes mirrored the differences in quality of clinical care both among sites in different health care systems and within the same site found by the standardized patients. When we compared the differences in the scores among and within sites, the rank order of overall quality measured by vignettes was the same as that measured by standardized patients. Other studies have observed this ability to detect differences between sites (32, 46, 47).

We were particularly struck by the 3-fold greater magnitude of the variation within sites compared with the variation among sites. This variation has been described elsewhere but has always been confounded by questions of whether differences in measured quality or differences in the health of the underlying population cause this variation (23). This problem did not complicate our study since vignettes present the identical “case mix” at each site (48, 49). This implies that institutions can decrease quality variation at the provider level by targeting both low-end performers and specific clinical areas or cases in which providers perform poorly by using education and reward incentives.

In this study, vignettes accurately captured unnecessary care. This suggests that vignettes not only have a high
correlation with actual practice but also have a capacity to measure the efficiency of clinical care. Thus, among physicians who score well, vignettes might be useful to distinguish between physicians who are “careful and smart” (ordering a parsimonious set of necessary tests and referrals) and those who are “just cautious” (ordering an array of unnecessary and unhelpful tests, referrals, or treatments in addition to the correct ones).

Computerization reduces the time and money required to score either handwritten responses to vignettes or to abstract charts. Administering the computerized vignettes also seems more realistic than our earlier paper-and-pen vignettes. The computerization allows for real-time responses that more closely simulate the patient-physician interaction. In the future, we expect to enhance the cost-effectiveness of vignettes to include electronic scoring, thereby reducing costs and facilitating access to vignette technology by health plans and medical groups (50, 51).

Although previous research has documented that vignettes are, at the very least, a measure of knowledge (52, 53), our findings indicate that, when properly constructed, vignettes are a valid measure of what physicians do during actual clinical encounters with patients. We believe that to maximize their validity, vignettes must incorporate several features that we included. Specifically, they should be open-ended (54); impose realistic temporal constraints; provide online, real-time responses where necessary; reflect clinical complexity (9, 12, 50); use evidence-based scoring criteria that are linked to improved health outcomes; and measure both unnecessary and necessary care.

Of interest, the abstracted medical record seems to be less representative of the actual visit than the vignette. We and others have noted that the chart is an imperfect measure of what occurs during a patient visit (2–5, 55). Physicians, for example, often perceive the chart more as a billing device, legal document, or communication tool than as a recording device (56).

As promising as these results seem, we must point out that vignettes identified only half of the worst performances across sites. Vignettes, by construct, also do not capture the important elements of caring and collegial rapport that are critical to overall patient well-being. These shortcomings underscore our belief that vignettes, like other measures of clinical practice, should not be used as a solitary measure of individual clinical competence (57). Clearly, different measurement methods capture different elements of practice. We believe, however, that some of the measured differences also reflect the variation in an individual’s day-to-day performance. The cueing effect found among physicians who had previously seen a standardized patient supports this notion. The enhanced score of these physicians is small but suggests that variation does exist within a group’s performance and that vignette scores can be improved if effort level (or motivation) is increased.

Our study has several limitations. While vignettes may capture variation in clinical care, we have not demonstrated that this variation translates into equivalent better (or worse) health outcomes. We confined our study to general internists in teaching programs and conditions in men in which the physical findings can be simulated. While gynecologic and pediatric conditions have been studied by using vignettes (12, 14, 21, 41, 54), these conditions cannot be simulated by standardized patients, making it difficult to compare methods. Vignettes in this study are also limited because they captured the quality of clinical care during only 1 visit, and we observed a cueing effect in physicians who saw standardized patients before the vignettes. In addition, the 4 vignettes require between 45 minutes and 1 hour to complete, which may be difficult when the physician’s time is already limited. To reduce this burden, we recommend giving the vignettes only every 6 to 12 months when they are used in longitudinal studies. Finally, vignette criteria, like all scoring criteria, must be regularly updated and linked to improvements in patients’ health status if they can truly be useful (58).

The discouraging levels of and the wide variations in the quality of clinical care in the United States and other countries (30, 59) can no longer be overlooked. Experience indicates, however, that physician practice can be improved, but only if it is measured (31, 60). This study describes an innovative method—clinical vignettes—that seems to measure the level of and variation in clinical practice for a defined set of conditions among different sites. Our reported validation of vignettes is a first step toward their wider application. However, more work is needed to link vignette scores to policy interventions or improvements in patient outcomes. While vignettes capture actual practice, control for case mix, can measure necessary and unnecessary care, and are easy to administer at a low cost, they have limitations that must be considered and are best used in conjunction with other measures.

From San Francisco Veterans Affairs Medical Center, Institute for Global Health, University of California, San Francisco, and California Pacific Medical Center, San Francisco, California; University of California, Los Angeles, Veterans Affairs Greater Los Angeles Healthcare System, and Permanente Medical Group, Los Angeles, California; RAND, Santa Monica, California; and Veterans Affairs Center for the Study of Healthcare Provider Behavior, Sepulveda, California.

Acknowledgments: The authors thank Elizabeth O’Gara (University of California, Los Angeles), Julianne Arnall (Stanford University), and Molly Bates Efrous and Olig Yeretsian (University of California, San Francisco), who trained and coordinated the standardized patient visits; the staff of the hospitals that worked so diligently to simulate the visits, provide their laboratory data, and introduce the simulated medical records; Brett Lewis, Ed La Calle, and Dan Bertenthal for their programming assistance and keen insights; Anne Sunderland and Miriam Polon for their assistance with the manuscript; and the actors for their fine performances.

Grant Support: By grant 11R 98118-1 from the Veterans Affairs Health Service Research and Development Service, Washington DC. Dr. Peabody is
Potential Financial Conflicts of Interest: None disclosed.

Requests for Single Reprints: John W. Peabody, MD, PhD, Institute for Global Health, 74 New Montgomery, Suite 508, San Francisco, CA 94105; e-mail, peabody@psg.ucsf.edu.

Current author addresses are available at www.annals.org.

References


38. Rothera I, Jones R, Gordon C. An examination of the attitudes and practice of general practitioners in the diagnosis and treatment of depression in older
60. McGlynn EA. There is no perfect health system. Health Aff (Millwood). 2004;23:100-2. [PMID: 15160807]
Current Author Addresses: Dr. Peabody: Institute for Global Health, 74 New Montgomery Street, Suite 508, San Francisco, CA 94105. Dr. Luck: Department of Health Services, University of California, Los Angeles, School of Public Health, PO Box 951772, Los Angeles, CA 90095-1772. Mr. Glassman: Division of General Internal Medicine (111G), Veterans Affairs Greater Los Angeles, 11301 Wilshire Boulevard, Los Angeles, CA 90073.

Dr. Jain: Medical Service (111), Veterans Affairs Medical Center, 4150 Clement Street, San Francisco, CA 94121. Dr. Hansen: 3801 Sacramento Street, Suite 309, San Francisco, CA 94118. Dr. Spell: 4950 Sunset Boulevard, Los Angeles, CA 90027. Dr. Lee: Center for the Study of Healthcare Provider Behavior, Veterans Affairs Medical Center (152), 16111 Plummer Street, Building 25, Sepulveda, CA 91343-2036.