Watchful Waiting Strategy May Reduce Low-Value Diagnostic Testing

Larissa May, MD, MSPH, MSHS, Peter Franks, MD, MPH, Anthony Jerant, MD, and Joshua Fenton, MD, MPH

Background: PCPs need effective communication strategies to address patient requests for low-value testing while sustaining patient-provider partnerships. Watchful waiting – allowing a negotiated period of time to pass before making a firm testing decision – shows promise as a tool for addressing patient requests for low-value testing.

Methods: Observational analysis of data from a randomized controlled trial of a communication intervention designed to boost patient-centeredness and reduce low-value test ordering among 61 resident primary care physicians. Intervention effectiveness was assessed during follow-up encounters of unannounced standardized patients (SPs) who requested low-value tests. We examined associations between five physician counseling behaviors and overall patient-centeredness (Measure of Patient-Centered Communication) and requested test ordering.

Results: During 155 SP encounters, residents most commonly used reassurance (96% of encounters), evidence-based recommendations (97%), and watchful waiting (68%). Resident advice to pursue watchful waiting was associated with 39% lower likelihood of test ordering (adjusted marginal effect of \(-38.6\% [95\% CI -43.6 to -33.6]\)). When all communication behaviors were examined together, only watchful waiting was significantly associated with test ordering (marginal effect of \(-38\% [95\% CI -44.3 to -31.7]\%)). Overall patient-centeredness was not associated with low-value testing.

Conclusion: Resident physician counseling to pursue watchful waiting was associated with less ordering of requested low-value diagnostic tests, while overall patient-centeredness was not. (J Am Board Fam Med 2016;29:710–717.)

Keywords: Communication; Counseling; Diagnostic Tests, Routine; Follow-up Studies; Physical Examination; Physicians, Primary Care; Probability; Risk; Watchful Waiting
testing as a means of providing that reassurance, even though such testing does little to quell patient anxiety or reassure. Further, patients whose test requests are denied are less satisfied with their visits, a matter of increasing concern to PCPs, who often receive incentives based on patient experience scores. There is a need to uncover strategies PCPs can use to address patient requests for low-value testing while maintaining patient and provider satisfaction.

One such approach is to enhance the overall patient-centeredness of PCP communication. In patient-centered interactions, physicians explore patients’ concerns and requests within a psychosocial context, with the goal of finding common ground from which evidence-based recommendations can be discussed and negotiated. Patient-centered counseling has been associated with reduced diagnostic testing without compromising patient satisfaction, suggesting that a more patient-centered approach may enable PCPs to meet the emotional and informational needs of patients who request low-value tests without acceding to testing.

However, the full patient-centered communication paradigm is multifaceted, complex, and associated with longer visits, and therefore is not always feasible to apply in toto. Among discrete patient-centered communication behaviors, recommending watchful waiting—allowing a negotiated period of time to pass before making a firm testing decision—shows promise as a tool for addressing patient requests for low-value testing. Most studies of watchful waiting have concerned treatment decisions, but 1 randomized controlled trial (RCT) examined its impact on low-value blood testing for unexplained medical symptoms. In the trial, conducted in the Netherlands, PCPs trained to recommend watchful waiting ordered fewer blood tests than controls, with no adverse impact on patient satisfaction or anxiety. However, evidence that advice to pursue watchful waiting is associated with less testing in primary care settings in the Netherlands, we conducted an observational analysis examining the association of watchful waiting, other specific physician counseling behaviors (normalization, risks outweighing benefits, reassurance, and evidence-based recommendations for no testing), and patient-centered communication with low-value test ordering.

**Methods**

*Design and Setting*

This observational analysis used data from an RCT of an educational intervention delivered by standardized patient instructors (SPIs) during 2 simulated office visits with in primary care residents at the University of California, Davis Medical Center, Sacramento, CA. During these visits, SPIs spent about 20 minutes acting as a patient requesting low-value tests. SPIs then broke out of the patient role and provided personalized feedback to intervention physicians. Residents in the control group had visits with SPIs without personalized feedback. The clinical scenarios for the 2 SPI visits were a 48-year-old man with subacute back pain requesting a spinal MRI and a perimenopausal woman at low risk for osteoporotic fracture requesting DXA screening.

The overall goal of the intervention was to enhance the patient-centeredness of residents’ responses to patient requests for low-value tests. In addition to enhancing overall patient-centeredness, the intervention emphasized 5 specific communication behaviors: (1) normalization; (2) reassuring patients by pointing to specific features of their history and physical examination; (3) explaining that the risks of testing outweigh the benefits in their situation; (4) suggesting watchful waiting such that testing will be ordered only if symptoms persist; and (5) advising evidence-based strategies to address patient concerns instead of requesting test-
ing. Following the intervention, test ordering by physicians was measured during 3 unannounced standardized patient visits over the subsequent 3 to 12 months. Because we found no intervention effect on study outcomes or on specific communication behaviors, we conducted this post hoc observational analysis to assess associations between the 5 specific communication behaviors, overall patient-centeredness, and test ordering during the unannounced standardized patient visits. The study was approved by our institutional review board.

Participants
Participants were family medicine or internal medicine resident physicians who had completed ≥1 year of residency and provided regular primary care at 1 of 2 primary care clinics at the university medical center. Residents were invited to participate in a study of “patient-doctor communication.” Participants provided informed verbal consent and were randomly assigned to intervention and control groups.

Standardized Patient Measurement Visits
After SPI visits, residents saw up to 3 unannounced standardized patients (SPs) scheduled during regular clinic hours over a 3- to 12-month follow-up period, including visits with (1) a male patient with subacute back pain requesting spinal MRI; (2) a postmenopausal woman with fatigue requesting DXA screening; and (3) a 30-year-old woman with recent-onset headache requesting neuroimaging (to assess the generalization of intervention effects to other low-value tests). Using detailed case histories, 9 SPs were trained to convincingly portray patients, to request tests early during visits, and to accept omission of testing if residents persisted in declining their request. Using standardized forms, an SP supervisor prospectively assessed fidelity by listening to audio-recordings of selected visits, assessing role fidelity using a checklist, which was used to provide corrective feedback to SPs.

Residents in the 2 clinics routinely precept patients with attending physicians. We repeatedly informed staff attending physicians about the study design and requested that they be nondirective when advising residents if they suspected a resident was seeing an SP. Residents at each clinic are authorized to order diagnostic tests without an attending physician’s cosignature.

We monitored SP detection using an E-mail survey sent 2 to 4 weeks after SP visits; this survey asked residents whether they suspected seeing an SP recently. If residents suspected SPs, we asked them to describe the SP and whether their clinical decisions differed from what they would have done for a real patient.

Measures
The main study outcome was whether residents ordered requested low-value tests during unannounced SP visits (eg, spinal MRI in the patient with low-back pain). We assessed test ordering by standardized electronic medical record review.

Patient-centered communication was assessed using the Measure of Patient-Centered Communication (MPCC), a validated measure ranging from 0 to 100 (the least to most patient-centered) based on transcribed audio-recordings of the encounters. Each component has a theoretical range of 0 to 100, with the total score being an average of the 3 component scores. The MPCC scores physicians on their exploration of the patient’s experience of illness (component 1), the psychosocial context (component 2), and physician efforts to find common ground on diagnosis and treatment (component 3). Two trained research assistants coded audiotaped recordings of the visits, resolving disagreements by consensus. Coders were trained by a doctorate-level qualitative researcher over a 3-month period by discussing and building consensus on blinded coding results for the initial 10 to 15 SP encounters.

Coders also rated the extent to which doctors engaged in the targeted specific physician counseling behaviors: (1) normalization, (2) informing patients about reassuring features of the history and physical examination, (3) explaining that the risks of testing outweighed the benefits, (4) advising watchful waiting, and (5) recommending evidence-based strategies instead of immediate testing. Coding for advising watchful waiting was based on 2 elements: (1) whether the resident recommended that the patient not undergo testing at this time, and (2) whether the resident stated that he or she would reconsider testing at a later point if symptoms or concerns persisted. We specified acceptable evidence-based strategies for each case (eg, dietary calcium intake for women requesting DXA). Each behavior was assessed as present or absent during the encounter.
We collected data on resident characteristics, including age, sex, postgraduate year, and specialty (family or internal medicine). Residents completed a baseline questionnaire including measures of stress from clinical uncertainty (theoretical score range, 13 to 78) and reluctance to disclose such uncertainty (theoretical score range, 9 to 39); higher scores reflected greater stress and reluctance, respectively. These measures capture the extent to which physicians feel anxiety, uneasiness, discomfort, or emotional turmoil when faced with uncertainty in patient care.

Analyses
Data were analyzed using Stata 14.1 (StataCorp, College Station, TX). Descriptive analyses used t tests or χ² tests to examine the relationship between test ordering and other study variables. Relationships between communication behaviors and test ordering were assessed using a series of logistic regression analyses within a generalized estimating equation framework, which adjusts standard errors for the nesting of SPs within physicians. A base model (excluding the communication variables) included only the intervention group and case scenario. Subsequent analyses added individual specific communication behaviors or components of the MPCC. Models also included clusters of the specific communication behaviors or the total MPCC score. To facilitate interpretation of the logistic regression parameter estimates, these are reported as adjusted marginal effects (the adjusted prevalence of test ordering associated with each level of the predictor). For each analysis, we report the amount of variance explained in test ordering; explained variance was assessed using McFadden’s pseudo-R². Our aim was to identify the contribution made by each of the communication measures in explaining the variance in test ordering.

Results
Of 64 potentially eligible residents, 61 agreed to participate and were randomized, and 59 had at least 1 follow-up visit with an SP. In the 155 encounters with unannounced SPs who requested low-value tests, PCPs ordered spinal MRI for 27% of the SPs with subacute back pain, DXA for 45% of the low-risk women requesting DXA screening, and neuroimaging for 9% of the SPs with uncomplicated headache. Overall, low-value tests were ordered during 26.5% of unannounced SP visits and during a similar percentage of visits with intervention versus control physicians (27.3 vs 25.6%; P = .82).

Of the communication behaviors, physicians most commonly provided reassurance and recommended evidence-based recommendations, both of which occurred during the most encounters regardless of whether tests were ordered (Table 1). Physicians explained that the risks of tests outweighed the potential benefits in 65% of visits when no tests were ordered versus 37% of visits when tests were ordered (P = .002). Physicians advised watchful waiting in 68% of visits overall but advised watchful waiting more commonly during those encounters for which a test was not ordered (88% vs 12%; P < .001). In bivariate analyses, overall patient-centeredness was not associated with test ordering, nor were individual components of the MPCC. Physician characteristics were not significantly associated with test ordering, including postgraduate year, stress from uncertainty, and reluctance to disclose uncertainty.

In adjusted analyses assessing the contribution of individual and combined communication behaviors to test ordering, physician advice to pursue watchful waiting, an explanation that the risks of testing outweigh the potential benefits, and a discussion of reassuring features of the history and physical examination were each significantly associated with low-value test ordering (Table 1). While the base model (including the study arm and case scenario) explained 9% of test ordering behavior, the suggestion of watchful waiting by itself produced substantial additional explained variance (pseudo-R² = 53%) (Table 1, series 1). Patients advised to pursue watchful waiting were 39% less likely to have low-value tests ordered (adjusted marginal effect, −38.6%; 95% confidence interval [CI], −43.6% to −33.6%). When all communication behaviors except watchful waiting were included, only discussion that the risks of testing outweighed the potential benefits was associated with test ordering (marginal effect, −22%; 95% CI, −32.5% to −11.2%) (Table 2, series 2), but the total amount of variance explained was much less (pseudo-R² = 16%). When watchful waiting was included along with other communication behaviors (Table 2, series 3), only watchful waiting was significantly associated with test ordering (marginal effect, −38%; 95% CI, −44.3% to −31.7%), and the inclusion of
all communications behaviors explained little additional variance compared with the model that included watchful waiting alone (pseudo-$R^2 = 56\%$ vs 53\%, respectively).

We similarly assessed the influence of patient-centeredness on low-value test ordering (Table 3). Compared with the variance explained by the base model alone, neither individual components of the MPCC (series 4) nor the overall patient-centeredness of the visit (series 5) explained significant additional variance in test ordering.

Based on surveys conducted within 4 weeks of SP visits, residents suspected seeing a recent SP in 59\% of visits (60 of 101 responses to surveys after the visits). In 53 of the 60 visits (88\%) in which residents suspected seeing SPs, they responded that they managed the patient exactly as they would a similar real patient; rates of test ordering were similar for the 7 visits in which residents reported “minor differences” in management compared with visits when they reported managing SPs “exactly alike” real patients (14.3\% vs 15.4\%).

### Discussion
In the context of patient requests for low-value diagnostic tests, we found that a physician’s suggestion to pursue a watchful waiting strategy was associated with a substantially lower likelihood of test ordering, and that this single communication behavior explained most of the variance in test ordering. After accounting for advice to pursue watchful
waiting, other communication behaviors had no additional impact on test ordering, nor did the overall patient-centeredness of the interaction. Our results suggest that watchful waiting shows promise as a simple counseling strategy that physicians can use to avert low-value test ordering.

While “watchful waiting” has been advocated as a strategy to reduce antibiotic prescribing for acute otitis media and in the management of prostate cancer, we assessed its potential influence in the context of diagnostic testing. In a cluster-randomized trial from the Netherlands, patients with unexplained symptoms were randomized to immediate blood testing versus 4 weeks of watchful waiting. In general, patients and physicians found watchful waiting acceptable, without adverse effects on patient anxiety. While specific aspects of physician-patient communication may reduce patient anxiety, test ordering does not seem to affect patient satisfaction, and diagnostic testing for patients at low risk of serious disease does little to reassure patients or resolve symptoms. Our findings suggest that a watchful waiting approach might also be an effective means of reducing low-value diagnostic testing in US primary care settings.

The potential effectiveness of watchful waiting as a counseling strategy is supported by theories of patient autonomy and control. Locus of control theory suggests that patient perception of control may be associated with positive health outcomes. When patients request tests, a negative response from physicians may threaten patients’ sense of autonomy or control, whereas a watchful waiting approach may validate patients’ concerns and bolster their sense of control over their health and symptoms. In contrast to watchful waiting, reassurance by itself may be perceived by patients as invalidating their concerns. By boosting patient autonomy, a watchful waiting strategy may also allow physicians to maintain patient satisfaction, which is major quality improvement target. A watchful waiting approach may also improve the

Table 2. Incremental Variance in Low-Value Test Ordering Explained by Specific Physician Communication Behaviors

<table>
<thead>
<tr>
<th>Variables, by Model</th>
<th>Adjusted Probability of Test Ordering*</th>
<th>P Value</th>
<th>Pseudo-R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base model†</td>
<td>—</td>
<td>—</td>
<td>9.3%</td>
</tr>
<tr>
<td>Series 1‡</td>
<td>Normalization 4.7% (−23.6 to 32.9%)</td>
<td>.75</td>
<td>9.3%</td>
</tr>
<tr>
<td></td>
<td>Risks outweigh benefits −22.6% (−33.1 to 12.0)</td>
<td>&lt;.001</td>
<td>15.8%</td>
</tr>
<tr>
<td></td>
<td>Reassurance −19.2% (−37.7 to −0.8)</td>
<td>.04</td>
<td>20.3%</td>
</tr>
<tr>
<td></td>
<td>Watchful waiting −38.6% (−43.6 to −33.6)</td>
<td>&lt;.001</td>
<td>53.1%</td>
</tr>
<tr>
<td></td>
<td>Evidence-based recommendations −15.7% (−40.3 to 8.9)</td>
<td>.21</td>
<td>9.6%</td>
</tr>
<tr>
<td>Series 2§</td>
<td>Normalization 6.4% (−23.4 to 36.2%)</td>
<td>.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Risks outweigh benefits −21.9% (−32.5 to −11.2%)</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reassurance −14.7% (−31.9 to 2.6%)</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evidence-based recommendations 1.8% (−19.9% to 23.5%)</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td>Series 3∥</td>
<td>Normalization 9.8% (−11.1 to 30.6%)</td>
<td>.67</td>
<td>55.7%</td>
</tr>
<tr>
<td></td>
<td>Risks outweigh benefits −6.7% (−15.2 to 1.7%)</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reassurance 5.4 (−12.9 to 23.6)</td>
<td>.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Watchful waiting −38.0% (−44.3 to −31.7%)</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evidence-based recommendations 10.3% (−2.2 to 22.7%)</td>
<td>0.11</td>
<td></td>
</tr>
</tbody>
</table>

*Adjusted for the randomized controlled trial intervention.
†The base model included study arm and standardized patient (patient with back pain requesting magnetic resonance imaging, woman requesting dual-energy x-ray absorptiometry, patient with headache requesting magnetic resonance imaging). Headache was associated with significantly less test ordering than back pain.
‡In series 1, each communication behavior was included individually in separate models with base model variables.
§In series 2, all communication behaviors, except watchful waiting, were included simultaneously together with base model variables.
∥In series 3, all communication behaviors were included simultaneously together with the base model.

CI, confidence interval.
physician experience by reducing the greater perceived difficulty of visits with patients who request diagnostic tests.\textsuperscript{8}

Although previous observational analyses suggested that patient-centered communication may be associated with lower overall diagnostic testing,\textsuperscript{1,14} more patient-centered communication was not associated with test ordering in our study. While prior studies have examined practice patterns among physicians with more versus less patient-centered communication styles, our study examined the association between patient-centeredness and test ordering within a specific visit when a low-value test was requested. It is conceivable that more patient-centered physicians may order fewer diagnostic tests overall, yet within the subset of visits in which low-value tests are requested, physician patient-centeredness may not influence test ordering.

Discretionary care, including test ordering, varies substantially among PCPs.\textsuperscript{27} The national Choosing Wisely\textsuperscript{TM} initiative seeks to reduce low-value testing by disseminating informational materials to providers and the public.\textsuperscript{3} However, the initial impact of the campaign on practice patterns seems to be limited,\textsuperscript{28} and Choosing Wisely is purposefully limited in scope, focusing on short lists of care elements, only some of which involve diagnostic testing. Individually targeted provider education, audit and feedback, and financial incentives have stronger evidence of effectiveness in reducing low-value testing, as do systems-based approaches (eg, computerized clinical decision support).\textsuperscript{29} However, these approaches are again most feasibly applied to short lists of tests and are relatively resource intensive. Teaching physicians how to use watchful waiting may be a generally applicable strategy to reduce the use of a broad range of low-value tests.

This study was limited by several constraints. First, the study had an observational design, and unmeasured confounding is possible. Second, the study included resident physicians at 2 academic practices and may have limited generalizability to other primary care practices or to urgent care or emergency settings. Third, residents often suspected that they were seeing SPs, which may have influenced communication behaviors. However, the similar rate of test ordering among residents who did and did not report differences in SP management imply that SP detection did not substantially confound the observed associations between watchful waiting and lower test ordering. Fourth, attending physicians may have influenced resident communication or testing behavior, although we asked attending physicians to be nondirective when they suspected an SP encounter.

**Conclusion**

In this observational study, counseling of patients by resident physicians who used a watchful waiting strategy was associated with less ordering of low-value diagnostic tests requested by SPs. This single communication strategy explained a large proportion of the total variance in testing. These findings

### Table 3. Incremental Variance in Low-Value Test Ordering Explained by Patient Centered Communication

<table>
<thead>
<tr>
<th>Variables, by Model</th>
<th>Adjusted Effect on the Probability of Test Ordering\textsuperscript{*} (95% CI)</th>
<th>P Value</th>
<th>Pseudo-$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base model\textsuperscript{1}</td>
<td>---</td>
<td>---</td>
<td>9.3%</td>
</tr>
<tr>
<td>Series 4\textsuperscript{4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Component 1: patients’ experience of illness</td>
<td>0.1% (−1.0%, 0.8%)</td>
<td>.84</td>
<td>9.3%</td>
</tr>
<tr>
<td>Component 2: psychosocial context</td>
<td>0.1% (−0.3%, 0.4%)</td>
<td>.71</td>
<td>9.3%</td>
</tr>
<tr>
<td>Component 3: attempt to find common ground</td>
<td>0.0% (−0.5%, 0.4%)</td>
<td>.86</td>
<td>9.3%</td>
</tr>
<tr>
<td>Series 5\textsuperscript{5}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total MPCC</td>
<td>0.0% (−0.8%, 0.8%)</td>
<td>.94</td>
<td>9.3%</td>
</tr>
</tbody>
</table>

\textsuperscript{*}Adjusted for the randomized controlled trial intervention.

\textsuperscript{1}The base model included study arm and standardized patient (patient with back pain requesting magnetic resonance imaging, woman requesting dual-energy x-ray absorptiometry, patient with headache requesting magnetic resonance imaging). Headache was associated with significantly less test ordering than back pain.

\textsuperscript{2}In series 4, each Measure of Patient-Centered Communication (MPCC) component was included individually in separate models with base model variables.

\textsuperscript{3}Series 5 included the total MPCC together with base model variables.

CI, confidence interval.

---

716 JABFM November–December 2016 Vol. 29 No. 6 http://www.jabfm.org
suggest that watchful waiting may be a simple, effective communication strategy for reducing low-value testing in primary care.

References