Impact of Providing Fee Data on Laboratory Test Ordering

A Controlled Clinical Trial

Leonard S. Feldman, MD; Hasan M. Shihab, MBChB, MPH; David Thiemann, MD; Hsin-Chieh Yeh, PhD; Margaret Ardolino, RN, MS; Steven Mandell, MS; Daniel J. Brotman, MD

Importance: Inpatient care providers often order laboratory tests without any appreciation for the costs of the tests.

Objective: To determine whether we could decrease the number of laboratory tests ordered by presenting providers with test fees at the time of order entry in a tertiary care hospital, without adding extra steps to the ordering process.

Design: Controlled clinical trial.

Setting: Tertiary care hospital.

Participants: All providers, including physicians and nonphysicians, who ordered laboratory tests through the computerized provider order entry system at The Johns Hopkins Hospital.

Intervention: We randomly assigned 61 diagnostic laboratory tests to an “active” arm (fee displayed) or to a control arm (fee not displayed). During a 6-month baseline period (November 10, 2008, through May 9, 2009), we did not display any fee data. During a 6-month intervention period 1 year later (November 10, 2009, through May 9, 2010), we displayed fees, based on the Medicare allowable fee, for active tests only.

Main Outcome Measures: We examined changes in the total number of orders placed, the frequency of ordered tests (per patient-day), and total charges associated with the orders according to the time period (baseline vs intervention period) and by study group (active test vs control).

Results: For the active arm tests, rates of test ordering were reduced from 3.72 tests per patient-day in the baseline period to 3.40 tests per patient-day in the intervention period (8.59% decrease; 95% CI, −8.99% to −8.19%). For control arm tests, ordering increased from 1.15 to 1.22 tests per patient-day from the baseline period to the intervention period (5.64% increase; 95% CI, 4.90% to 6.39%) (P < .001 for difference over time between active and control tests).

Conclusions and Relevance: Presenting fee data to providers at the time of order entry resulted in a modest decrease in test ordering. Adoption of this intervention may reduce the number of inappropriately ordered diagnostic tests.


F

ROM 1990 TO 2007, HOSPITAL expenditures increased an average of 7.2% per year.¹ Medicare fee-for-service inpatient spending totaled $114 billion in 2009. Whereas major procedures and evaluation and management services increased by approximately 30% from 2000 to 2009, imaging and diagnostic tests increased by 85% during that same period.² To encourage the profession to improve its stewardship of health care resources, health care leaders have proposed adding a seventh general competency, “cost-consciousness and stewardship of resources,” to the current 6 competencies defined by the Accreditation Council for Graduate Medical Education and the American Board of Medical Specialties.³ Many agree that one of the first steps toward a more cost-effective health care system is to “cut waste.”⁴

See Invited Commentary at end of article

Empirical evidence supports that not all tests ordered are needed to provide high-quality care. PricewaterhouseCoopers’ Health Research Institute identified $1.2 trillion in health care spending waste. According to their survey, two-thirds of respondents said “that overused diagnostic
testing was driving up healthcare costs.” A 2006 study from Australia found that 67.9% (2.01 tests per patient-day) of inpatient laboratory tests ordered during a 6-month intervention period did not contribute to patient care. Over-treatment and excessive use of diagnostic tests likely resulted in upward of $226 billion in waste to the US Health Care System in 2011. Nine medical specialty societies, the American Board of Internal Medicine Foundation, and Consumer Reports recently highlighted this theme as they introduced the Choosing Wisely campaign. The first phase of the campaign centers on identifying testing “prone to overuse.”

Although some of the overuse of diagnostic tests is due to physicians’ practice of defensive medicine, it is also clear that most physicians do not know how much tests cost. Other factors, such as patient expectations, insufficient understanding of the limitations (operating characteristics) of tests, inability to retrieve the results of a test already performed, learned behaviors, and economic incentives, may also influence ordering behavior.

Cost containment is not a new concern. In 1983, Grossman reviewed the 5 intervention strategies that had been used to contain costs up to that point: educational strategies, feedback strategies to compare actual ordering behavior with ordering protocols, cost-awareness strategies, rationing strategies, and market-oriented financial incentives and risk-sharing plans. Solomon et al reviewed 49 studies in 1998 and found that more successful interventions were usually “based on multiple behavioral factors.” Their review included the 1990 study by Tierney et al that displayed charges at the time of test ordering in an outpatient academic clinic. Tierney et al found that the number of laboratories and costs decreased but the differences did not persist after the intervention ended. Solomon et al also reviewed the 1997 study by Bates et al who also displayed the charges at the time of order entry, using a computer-based system, but found no changes in testing volume.

The array of studies has continued to grow, showing variable success. In a 1999 study, Bates et al attempted to reduce redundant laboratory orders using a computerized reminder at the time of order entry, which had a limited overall effect. Hampers et al found that when providers saw charges listed and calculated the total charges for each diagnostic workup when laboratory tests were ordered in a pediatric emergency department, overall charges dropped by 27% compared with a control period. Wang et al were able to reduce test ordering in a coronary care unit by developing evidence-based guidelines, incorporation of those guidelines into the computer admission orders, and education efforts. Other interventions have included feedback to primary care physicians, small peer group quality improvement meetings, and administrative policies that limited repeated orders.

We tested an exportable and simple method to reduce laboratory test ordering that would not add extra steps to the ordering process. In lieu of traditional educational interventions, feedback sessions, and report cards, we hypothesized that we could influence ordering behavior by simply displaying the Medicare allowable fees of some laboratory tests in the computerized provider order entry system.

METHODS

SETTING AND INTERVENTION

The study was conducted in Baltimore, Maryland, at The Johns Hopkins Hospital, a 1051-bed tertiary care hospital. The clinical laboratories perform an average of 3.6 million inpatient tests annually. Using data from fiscal year 2007 at our hospital, we compiled lists of the 35 laboratory tests that were most frequently ordered throughout the hospital and the 35 that were most expensive. To be included, the most expensive tests needed to be ordered at least 30 times during fiscal year 2007. For all tests, we defined the displayed fee as the 2008 Medicare allowable charge. Each test was then randomly assigned to be an “active” test or a control test. The randomization was performed at the test level for both practical purposes (individual diagnostic tests must have the same display throughout the institution) and to prevent contamination between groups (a provider who saw the fee of a complete blood cell count presented for a given patient is likely to remember that information when ordering the same test on a different patient and might communicate this information to a colleague). The institutional review board approved the study and waived the requirement for informed consent.

During a 6-month baseline period (November 10, 2008, through May 9, 2009), we did not display any fee information to providers. During a 6-month intervention period that took place exactly 1 year later (November 10, 2009, through May 9, 2010), we displayed the fees of only the active tests to ordering providers via the computerized provider order entry (cPOE) system, Sunrise Clinical Manager (Allscripts Corp), as shown in Figure 1. All cPOE users who ordered laboratory tests saw the displayed fees of only the active tests. Providers were unaware that the fees were displayed as part of an institutional review board–approved study unless the research team was specifically questioned; in such instances, those who inquired were informed only that the fees were being displayed as part of a research project but were not informed of the purpose of the study or of the existence of a concurrent set of control tests.

OUTCOMES AND MEASUREMENTS

Outcomes included the total number of orders placed, the frequency of ordered tests (per patient-day), and total charges as-
associated with the orders. For all outcomes, we examined changes according to time period (baseline vs intervention period) and by study group (active test vs control). We calculated patient-days using hospital administrative data for all hospital nursing units using the cPOE system during the study periods. Additional data captured from the cPOE system included the ordering provider, patient demographic information, nursing unit, and whether the order was entered as part of an order set (vs as a stand-alone order). We categorized ordering providers as physicians (vs nonphysicians) and used the nursing units to determine the service the patient was on (medical non-intensive care unit [ICU], surgical non-ICU, ICU, or other service). We also performed a subgroup analysis of the expensive tests.

STATISTICAL ANALYSIS

Statistical analyses were conducted at the order level and at the test level; data are presented as frequency of orders and aggregate charges. All hypothesis tests were 2 tailed. First, treating each order as an independent event, χ² tests were used to compare the total number of orders placed during the baseline and intervention periods, according to the intervention group (control vs active). The percentage of change in total orders entered during the 6-month intervention period (relative to the baseline period) was calculated for both groups. Then, to account for the slight difference in number of patient-days between the intervention and baseline periods, the percentage of change in order frequency was expressed on a per-patient-day basis and the change in order frequency per patient-day was calculated for both active and control arms. Rate ratios were also calculated to estimate the relative change in order frequency per patient-day between the baseline and intervention periods according to arm, using Poisson regressions with over-dispersion. Adjusted rate ratios were then calculated, including the following adjustor variables: ordering provider (physician vs other), clinical service (medical non-ICU, surgical non-ICU, ICU, or other service), order as part of order set (yes vs no), patient age, and patient sex. To calculate the difference in order frequency attributable to the intervention arm, we calculated the P value for the interaction term between the study group (control vs active) and the time period (baseline vs intervention) in predicting the order frequency. The aforementioned order- and patient-specific adjustor variables were also included in this model. Next, we analyzed the data at the test level rather than the order level. For these analyses, we examined the absolute change in the total number of orders for each test (such as basic metabolic panel) during the baseline vs the intervention period, as well as the corresponding percentage of change. Similarly, we examined the change in total charges at the test level. Since the adjustor variables we used applied to the order (eg, ICU patient vs medical non-ICU patient) rather than the test (basic vs comprehensive metabolic panel), we did not perform multivariable adjustments for the test-level analyses.

We performed all analyses using Stata/SE software, version 12 (StataCorp).

RESULTS

TESTS INCLUDED

A total of 70 diagnostic tests met inclusion criteria and were randomized to active and control arms. Of these, 31 were ultimately included in the control arm and 30 in the active arm as shown in the eTable (http://www.jamainternalmed.com). Nine tests, 3 active and 4 controls, were excluded from analysis owing to technical challenges, including an inability to present data in a clearly visible fashion within the cPOE system, display formatting changes occurring during the study periods, or protocol changes for ordering specific tests during the study periods. For instance, the hospital began monitoring anti-Xa heparin instead of activated partial thromboplastin time in between the baseline and intervention study periods.

SAMPLE CHARACTERISTICS

Table 1 and Table 2 summarize the order, patient, and hospital service characteristics of the tests ordered in the active and control arms during the baseline period. The laboratory tests in both arms during the baseline period were similar with respect to which hospital service ordered the tests and the type of provider (physician, nurse, other) who ordered them. The patients in the 2 arms were also similar in sex and age. More than 50% of the orders originated in the medicine service, which includes hematology-oncology. Orders were placed as order sets for 47% of orders in the active arm and 38% in the control arm.
Table 3. Change in Order Frequency and Fees by Study Arm

<table>
<thead>
<tr>
<th></th>
<th>Active Arm (30 Tests)</th>
<th>Control Arm (31 Tests)</th>
<th>P Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of orders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline period</td>
<td>458 297</td>
<td>142 196</td>
<td></td>
</tr>
<tr>
<td>Intervention period</td>
<td>416 805</td>
<td>149 455</td>
<td></td>
</tr>
<tr>
<td>Change</td>
<td>−9.1</td>
<td>5.1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Total fees (in US dollars), $b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline period</td>
<td>4 865 019</td>
<td>2 200 910</td>
<td></td>
</tr>
<tr>
<td>Intervention period</td>
<td>4 375 636</td>
<td>2 254 178</td>
<td></td>
</tr>
<tr>
<td>Change</td>
<td>−10.1</td>
<td>2.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Orders per patient-day, No. b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline period</td>
<td>3.72</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>Intervention period</td>
<td>3.40</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td>Change, % (95% CI)</td>
<td>−8.59 (−8.99 to −8.19)</td>
<td>5.64 (4.90 to 6.39)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Rate ratio (95% CI)</td>
<td>0.91 (0.86 to 0.92)</td>
<td>1.04 (1.02 to 1.06)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Period-arm interaction b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted rate ratio c</td>
<td>0.89 (0.87 to 0.91)</td>
<td>1.03 (1.01 to 1.05)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Period-arm interaction b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fees (in US dollars), per patient-day, $b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline period</td>
<td>39.49</td>
<td>17.87</td>
<td></td>
</tr>
<tr>
<td>Intervention period</td>
<td>35.70</td>
<td>18.39</td>
<td></td>
</tr>
<tr>
<td>Change, % (95% CI)</td>
<td>−9.60 (−9.72 to −9.48)</td>
<td>2.94 (2.75 to 3.13)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*P values were calculated at the order level, treating each order as an independent event, as discussed in the Methods section.

b The total number of patient-days was 123 192 for the baseline period and 122 566 for the intervention period.

c Adjusted for provider (physician vs others), service (medicine, intensive care units, surgery, other service), order set (yes vs no), patient sex, and patient age.

(\(P < .001\) for difference between active and control tests in the change in order frequency). This pattern was not altered by adjustment for order, provider, and patient characteristics.

The total charge difference in the active arm was a decrease of $3.79 per patient-day (−9.60% [95% CI, −9.72% to −9.48%]), with a charge increase in control tests of $0.52 per patient-day (2.94% [95% CI, 2.75% to 3.13%]), translating to a net hospital-wide charge decrease of $436 115 during the 6-month intervention period for the 2 groups combined (Figure 2). Notably, there was a marked decrease in the ordering frequency of comprehensive metabolic panels (randomized to the active arm) and an increase in ordering frequency for basic metabolic panels (randomized to the control arm), accounting for some of the cost shifting apparent in our data. In all, 7 of the diagnostic tests in the active arm and none in the control arm exhibited a total charge decrease of more than $25 000 between the intervention and baseline periods (\(P = .005\); Fisher exact test). The frequencies of the orders during the baseline and intervention periods, by test, are shown in the eTable.

Table 4 shows more conservative estimates of the effect of the intervention, analyzing the data at the level of the test rather than at the level of the order, resulting in a total sample size of 61 (30 intervention tests and 31 control tests). Although all trends point toward a reduction in order frequency and charges in the intervention arm relative to the control arm, statistical significance is attenuated, particularly when data are presented as a percentage change, and particularly when the data include infrequently ordered tests (ordered <135 times during the baseline period or resulting in <$5000 in total expenditures).

A subgroup analysis of the expensive tests revealed that the overall statistically significant differences be-

ORDER FREQUENCY

In all, 1 166 753 orders were placed during the 12 months of the study (baseline and intervention periods combined). The numbers of patient-days were similar in the baseline and intervention periods (123 192 and 122 566 patient-days, respectively). During the baseline period, 458 297 orders were entered for tests in the active arm and 142 196 for those in the control arm (Table 3). During the intervention period, 416 805 orders were entered for active tests, a 9.1% reduction from baseline. The number of tests per patient-day decreased as well, from 3.72 to 3.40, an 8.59% drop (95% CI; −8.99 to −8.19%). In contrast, 149 455 orders were entered for control tests, a 5.1% increase from baseline in total tests ordered. This was reflected in an increase in tests per patient-day from 1.15 vs 1.22 (5.6% increase; 95% CI, 4.90% to 6.39%)
tween the active and control arms in ordering and charges were driven by the changes in the frequent tests. During the intervention period, 7124 expensive tests were ordered. Both the expensive active and expensive control arms had small but similar decreases from baseline ($P = .83$).

**COMMENT**

Our findings suggest that simply displaying the Medicare allowable fee of diagnostic laboratory tests at the time of order entry can affect physician ordering behavior, even without any additional educational interventions. We observed a 9.1% reduction in tests ordered in the active arm, which was partially offset by an increase in ordering of control tests. This created a net charge reduction of more than $400,000 at the hospital level during the 6-month intervention period, assuming no overall change in diagnostic tests that were not included in this study. Although the overall financial impact is modest, our study offers evidence that presenting providers with associated test fees as they order is a simple and unobtrusive way to alter behavior. Unlike the process in previous studies, no extra steps were added to the ordering process and no large-scale educational efforts accompanied this exportable intervention.

The behavioral and possibly financial impact of broadening the intervention to target a larger percentage of diagnostic tests remains unknown. If the intervention is broadened, targeting the most frequent tests would probably yield the most success because the expensive tests are ordered too infrequently to affect outcomes significantly. It is notable that we offered no direct or indirect incentives for changing ordering behavior, suggesting that physicians can act in a cost-conscious manner even without direct incentives.

A major strength of our study was the use of concurrent controls and randomization at the test level (rather than at the order level). Had we randomized the fee display to the individual order, providers would have been able to learn how much each test costs by seeing this information displayed when ordering tests for other patients or at different times for the same patient. Although providers may have made assumptions about the reasons for the fee display, there was no systematic educational intervention undertaken, and, to our knowledge, no providers other than the authors were aware of the specific goals of the study or the particular control tests chosen. Another strength of the study was that we included all tests ordered throughout the hospital, by providers in different departments and levels of care; there were no major changes in the findings resulting from adjustment for the type of provider, the department, the acuity of care, or whether the test was ordered as part of an order set.

The study does have multiple limitations. First, the extent of asymmetric randomization of tests to the active and control arms was unexpected. Because some tests are ordered far more frequently than others, some degree of asymmetry was anticipated. However, we did not anticipate that more than 3 times more orders would occur in the active arm than in the control arm. Although we stratified by frequent and expensive tests before randomizing, many of the most commonly ordered tests were randomized by chance to the active arm. Conversely, almost twice as many expensive tests were ordered in the control arm as in the active arm. Although the overall asymmetry would not be expected to affect findings with regard to the percentage change in test ordering from the baseline to the intervention period, it may have affected our estimate of the net change in charges.

Second, the study lasted 6 months, and the durability of the results over years is not known. Third, it remains unknown whether displaying fees of all tests would lead to a more dramatic reduction in test ordering or desensitize providers to the displayed fee information. Fourth, we do not know whether the practice setting (an academic teaching institution where most orders are placed by residents) affected the impact of this intervention. Providers in other settings may have differing degrees of cost consciousness at baseline or varying susceptibility to changing behavior in response to a simple educational intervention.

Fifth, we assume, but cannot prove, that the decrease in tests ordered did not affect clinical care unfavorably.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Active Arm</th>
<th>Control Arm</th>
<th>$P$ Value$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tests, No.</td>
<td>30</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Mean change in total orders per test (intervention — baseline), No.</td>
<td>$-1383$ ($-2367$ to $-399$)$^b$</td>
<td>$234$ ($-404$ to $872$)</td>
<td>.006</td>
</tr>
<tr>
<td>Mean change in order frequency, %</td>
<td>$-8.60$ ($-12.45$ to $-4.74$)</td>
<td>$1.60$ ($-11.28$ to $14.47$)</td>
<td>.13</td>
</tr>
<tr>
<td>Change in order frequency for tests ordered $\geq$135 times (25th percentile) during baseline period, %</td>
<td>$-10.34$ ($-13.78$ to $-6.91$)</td>
<td>$4.92$ ($-10.98$ to $20.83$)</td>
<td>.058</td>
</tr>
<tr>
<td>Fees (in US dollars), $</td>
<td>$4,865,019$</td>
<td>$2,209,910$</td>
<td></td>
</tr>
<tr>
<td>Baseline period</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention period</td>
<td>$4,375,636$</td>
<td>$2,254,178$</td>
<td></td>
</tr>
<tr>
<td>Total absolute change</td>
<td>$-489,383$</td>
<td>$53,259$</td>
<td></td>
</tr>
<tr>
<td>Mean change, $</td>
<td>$-16,313$ ($-29,305$ to $-3321$)</td>
<td>$1,718$ ($-6372$ to $9809$)</td>
<td>.02</td>
</tr>
<tr>
<td>Mean change, %</td>
<td>$-8.60$ ($-12.45$ to $-4.74$)</td>
<td>$1.60$ ($-11.28$ to $14.47$)</td>
<td>.13</td>
</tr>
<tr>
<td>Change for tests with total fees $\geq$5000 at baseline, %</td>
<td>$-10.86$ ($-14.22$ to $-7.49$)</td>
<td>$4.15$ ($-11.13$ to $19.42$)</td>
<td>.069</td>
</tr>
</tbody>
</table>

$^a$ $P$ value comparing active and control arm difference of changes from baseline to intervention periods.

$^b$ Parenthetical ranges with means are 95% CIs.
and that we effectively cut waste, rather than rationing care. If adverse events increased as a result of fee-conscious testing, any potential savings could be lost. Finally, the overall financial impact of the behavior changes we observed are likely to be affected by factors such as fixed costs of testing equipment and laboratory technicians, which may tend to attenuate the true cost savings for the institution and the health care system unless behavior changes are widespread and prolonged.

We conclude that displaying the Medicare allowable fees of diagnostic tests at the time of ordering can modestly affect provider ordering behavior. Whether broadening this intervention and coupling it with educational interventions related to cost consciousness and stewardship of resources will increase its effect on clinical practice deserves further study, provided that providers are not inappropriately incentivized to limit needed care.

Accepted for Publication: January 10, 2013.

Published Online: April 15, 2013. doi:10.1001/jamainternalmed.2013.232

Correspondence: Leonard S. Feldman, MD, Division of General Internal Medicine, The Johns Hopkins University School of Medicine, 600 N Wolfe St, Nelson 215, Baltimore, MD 21287 (LF@jhmi.edu).

Author Contributions: Drs Feldman, Yeh, and Brotman had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Feldman and Brotman. Acquisition of data: Feldman, Thiemann, Ardolino, Mandell, and Brotman. Analysis and interpretation of data: Feldman, Shihab, Thiemann, Yeh, and Brotman. Drafting of the manuscript: Feldman, Shihab, and Brotman. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: Shihab, Thiemann, Yeh, and Brotman. Obtained funding: Feldman. Administrative, technical, and material support: Ardolino, Mandell, and Brotman. Study supervision: Brotman.

Conflict of Interest Disclosures: None reported.

Funding/Support: This work was supported in part by The Johns Hopkins Hospitalist Scholars Program.

Role of the Sponsors: The Scholars Program, per se, did not have any role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript.

Additional Contributions: We appreciate the support of the General Internal Medicine Methods Core and our information technology partners who made the intervention possible.

Online-Only Material: An eTable is available at http://www.jamainternalmed.com.

REFERENCES


