Reducing Unnecessary Use of High-Technology Imaging in Patients with Low Back Pain: A PI-CME Approach

Susan M. Szpunar, DrPH,1 Steven E. Minnick, MD,1 Louis D. Saravolatz, MD,2 Kenneth Bollin, MD,3 Jesse Heard, MD,1
Mary Chase, BA,1 Gayle Blackmer, MBA4
Departments of 1Medical Education and 1Internal Medicine, St. John Hospital and Medical Center, Grosse Pointe Woods, Michigan; 2St. John Family Medicine Center, St. Clair Shores, Michigan; 4St. John Providence Health System Physicians’ Network Information Technology, Warren, Michigan

ABSTRACT

Background: We conducted a Performance Improvement–Continuing Medical Education (PI-CME) project addressing the appropriate use of high-technology imaging in patients with low back pain.

Method: The project followed the 3 stages of PI-CME. Stage A consisted of a retrospective chart review of patients with low back pain seen at 2 internal medicine and 2 family medicine clinics from 1 July 2009 to 31 December 2009. Charts were reviewed for the presence of “red flags” (symptoms indicating a more severe back condition) and the imaging studies that were ordered. Stage B included an educational session, confidential physician progress reports, and an interactive patient tool. Stage C consisted of 2 periods of post-intervention chart review, presentation of post-intervention findings to physicians, and physician reflection.

Results: At baseline, only 3 of 157 patients presenting with 0 red flags had a high-technology image ordered (1.9%). More than 70% of patients with 3 or more red flags had no imaging ordered. The findings suggested that once the information from the medical record was taken into account, instead of over-utilization, physicians may have been under-utilizing imaging. At each follow-up period, there were 0 patients with no red flags who had a high-technology image ordered. The percent of patients with 3 or more red flags with no imaging ordered decreased over the project interval from 72% (n = 116/161) to 5.7% (n = 2/35, P < .0001).

Conclusions: PI-CME is an effective tool for addressing practice gaps. It engages physicians in an active, personal learning journey. Reliance on billing data alone can lead to erroneous findings with respect to utilization of services or guideline adherence.

INTRODUCTION

Among adults in the United States, back pain is a highly prevalent condition, resulting in significant lost time from work and other activities, disability, and health care expenditures. The 1-year prevalence of back pain is approximately 15% to 20%. Approximately 80% of Americans experience at least 1 episode of back pain in their lifetime [1,2]. Back pain is a leading reason for visits to physicians, diagnostic tests, and prescriptions for analgesics and other medications.

Using data from the 1998 Medical Expenditure Panel Survey (MEPS), Luo et al reported that a total of 25.9 million adults reported back pain sometime in 1998 [3]. Total health care expenditures for individuals suffering from back pain reached $90.7 billion.

Because of the burden of illness and expenditure resulting from back pain, the US Department of Health and Human Services developed a clinical practice guideline concerning back pain in 1994 [4]. Since then, states and other agencies have constructed similar guidelines concerning the diagnosis and care of patients presenting with low back pain (LBP). These endeavors have indicated that uncomplicated acute
LBP tends to be a benign self-limited condition that does not warrant any imaging studies [5]. Most of these patients are able to resume normal activities within 30 days. Clinicians need to evaluate patients carefully to determine the presence of indications for more complicated illness, often termed “red flags” (Table 1).

### Table 1. “Red Flags” That May Indicate the Presence of Complicated Low Back Pain*

<table>
<thead>
<tr>
<th><strong>Red Flag</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Recent significant trauma, or milder trauma patient age 50 or older</td>
</tr>
<tr>
<td>Unexplained weight loss</td>
</tr>
<tr>
<td>Unrelenting night pain or pain at rest</td>
</tr>
<tr>
<td>Pain with distal numbness or leg weakness</td>
</tr>
<tr>
<td>Loss of bowel or bladder control (retention or incontinence)</td>
</tr>
<tr>
<td>Unexplained fever [38°C or 100.4°F for &gt; 48 hours]</td>
</tr>
<tr>
<td>Immunosuppression or on immunosuppressive medication</td>
</tr>
<tr>
<td>History of cancer</td>
</tr>
<tr>
<td>IV drug use</td>
</tr>
<tr>
<td>Prolonged use of corticosteroids, osteoporosis</td>
</tr>
<tr>
<td>Focal neurological or sensory deficit progressive or disabling symptoms</td>
</tr>
<tr>
<td>Duration &gt; 6 weeks</td>
</tr>
<tr>
<td>Previous back surgery</td>
</tr>
<tr>
<td>Saddle anesthesia</td>
</tr>
<tr>
<td>HIV</td>
</tr>
<tr>
<td>History of any fracture after the age of 50</td>
</tr>
<tr>
<td>Kyphosis of the spine</td>
</tr>
</tbody>
</table>


The National Committee for Quality Assurance (NCQA) Healthcare Effectiveness Data and Information Set (HEDIS®) 2009 measures include a metric on the use of imaging studies for LBP [6]. This measure is used to assess the percentage of patients with a primary diagnosis of uncomplicated LBP that did not have an imaging study (plain radiograph, magnetic resonance imaging [MRI], or computed tomography [CT] scan). The NCQA notes, in their rationale for this measure, that

1. Uncomplicated (ie, no red flags) back pain is usually benign and self-limiting and does not call for imaging studies.
2. Imaging studies are frequently overused in the evaluation of patients with LBP.
3. Abnormalities found when imaging people without back pain are just as prevalent as those found in patients with back pain.

The NCQA does note that a small portion of the large patient population with back pain—patients with the “red flags”—will need to be evaluated further to investigate for more serious health problems.

The St. John HealthPartners, a provider organization (PO) including faculty staff physicians within the St. John Health System, is a participant in the Physician Group Incentive Program (PGIP) Radiology Management Initiative, sponsored by Blue Cross/Blue Shield (BC/BS) of Michigan [7]. The PGIP Radiology Initiative is a program designed to moderate costs associated with imaging by reducing the use of inappropriate procedures. As part of this initiative, each practice participating in the initiative receives an annual unadjusted and risk-adjusted dashboard report that summarizes the practice’s performance and compares it with other participating practices within Michigan. Based on data provided by BC/BS, the St. John group was well above benchmark performance on the use of high-technology imaging, particularly for spinal imaging [8]. Based upon this perceived performance gap, we designed a Performance Improvement—Continuing Medical Education (PI-CME) project designed to foster appropriate use of high-technology imaging in patients with LBP.

### METHODS

As defined by the American Medical Association, PI-CME is a certified CME activity in which an accredited CME provider structures a long-term 3-stage process by which a physician or group of physicians learn about specific performance measures, assess their practice using the selected performance measures, implement interventions to improve performance related to these measures over a useful interval of time, and then reassess their practice using the same performance measures [9].

For our project, we focused on 4 clinics: 2 internal medicine clinics located at St. John Hospital and Medical Center (SJH&MC) (IMD, the faculty practice, and IMSC, the resident specialty clinic) and 2 Family Medicine clinics also associated with SJH&MC (FMC, Family Medicine Clinic, and MMC, Masonic Medical Center). We obtained permission to conduct the project at these sites from the respective department chairs. The project was reviewed by the SJH&MC Institutional Review Board, and, as a quality improvement project, it was deemed exempt from further IRB review.

Our project consisted of the 3 phases outlined in the PI-CME approach. We followed the steps outlined by the American Medical Association [9] for designing PI-CME programs. The program was also modeled on a PI-CME program conducted at SJH&MC that addressed the standing orders for thromboprophylaxis in patients at risk for venous thromboembolism (VTE) (unpublished data available upon request). The 3 stages are as follows.

### Stage A: Performance Assessment

The first step in our project was baseline assessment and corresponds to Stage A of the PI-CME activity. Stage A was accomplished by querying the data warehouse to obtain a list of all adult patients (ages 18 years and older) seen during the period 1 July 2009 through 31 December 2009 with an ICD-9 (International Classification of Disease, Ninth Revision) diagnosis code consistent with low back pain (724.0-724.9) at any of the 4 clinics.

To determine whether high-technology imaging was appropriate or not, we relied on the St. John Ambulatory Care Guideline...
for patients with LBP. The guideline lists the set of 17 “red flags” that indicate a need for more in-depth follow-up of the patient presenting with LBP (Table 1). The “red flags” in the St. John Guideline follow from the evidence-based guidelines as presented by the Agency for Healthcare Research and Quality (AHRQ) guideline on LBP [10] and the NCQA HEDIS measures [6].

A research assistant completed chart review for each of the patients to collect data on demographic factors, major complaint, the presence of any of the red flags, any imaging ordered by the physician (plain x-ray, MRI, CT scan, myelography, or CT-myelography), and any other treatment prescribed for the patient.

The data were entered into a Microsoft Access™ (Microsoft, Redmond, WA, USA) database, and random checks of the data were done for consistency. From these baseline data, we determined the following for each physician in each clinic: the number of LBP patients seen in the 6-month baseline period, the number and percentage of those patients who had 0 red flags, the number and percentage of patients with 0 red flags who nevertheless had high-technology imaging ordered, the number and percentage of patients with 3 or more red flags, and finally, any patients with 3 or more red flags who had no imaging ordered.

In designing the project, we chose to use the criterion of “3 or more” red flags because some of the red flags may indicate a simple low back pain condition that has a long recovery period (ie, a chronic pain problem). For example, duration of pain for more than 6 weeks does not, by itself, indicate that the patient has a more serious back condition—recovery times are variable. Thus, the use of 3 or more red flags was chosen to identify the patient population who were most likely to have a more serious LBP condition present.

These data were used to generate a confidential report for each physician that allowed them to see if they had ordered any high-technology imaging procedures for patients with 0 red flags. If imaging for a patient with no red flags had been ordered, the report included the patient’s name, date of service, and ICD-9 code so the physician could review the medical record to determine why the imaging was ordered. Similarly, the confidential report also listed the name, service date and ICD-9 codes of any patients who presented with 3 or more red flags who had no imaging ordered (including plain x-rays).

**Stage B: Educational Intervention and Coaching**

At each clinic, an educational session was held during a regular staff meeting to introduce the project and to discuss the care of patients with simple LBP. This 1-hour session included information about the natural history of LBP, the social and economic burden of LBP, the indications for high-technology imaging (ie, the red flags), the most common treatments for LBP, and a time for questions.

A new, interactive patient tool was also introduced and provided to each of the clinics. This tool is a handout about LBP that can be given to the patient. The handout describes the common causes of LBP, simple self-care techniques, the signs that indicate that the patient may have a more serious problem and should contact the doctor immediately, and a place for the patient to write down notes, questions for the doctor, and reflections on steps they could take to prevent LBP in the future (eg, weight loss, proper body mechanics, etc.). The handout was used as a communication device where physicians could write personalized recommendations about self-care and where the patient could take notes or jot down questions for their physician. It also served as a reminder to patients that most LBP is self-limited and benign, while also providing information about the symptoms to watch for that would indicate a worsening condition.

At the end of the educational session, the LBP PI-CME “coach” for each clinic was given the confidential reports for each of the physicians in their office. Each confidential report was in a sealed envelope and included the data described above as well as a list of the red flags as noted in the St. John Ambulatory Care Guideline. These reports, for the baseline period, were then distributed to each of the participating physicians. Each physician was directed to address any questions or concerns to the designated coach for the clinic.

**Stage C: Post-Intervention Data Collection and Monitoring**

Following the educational sessions and the distribution of the baseline performance data to each physician, we conducted 2 follow-up chart reviews. For the periods 1 August 2010 through 30 September 2010 and 1 October 2010 through 30 November 2010, we again queried the data warehouse to obtain a list of all adult patients seen at each clinic with an ICD-9 code indicative of LBP. A research assistant again performed chart reviews to collect the data points as described above. After chart review was completed for each 2-month interval, physicians again received a confidential letter that provided information as in Stage A. These confidential letters were placed in sealed envelopes along with a list of the “red flags” and distributed to each participating physician by the respective clinic coach.

Following the distribution of the last set of confidential letters (for the follow-up period of 1 October 2010 through 30 November 2010), each participating physician was asked to complete an anonymous online survey about the PI-CME project that asked each physician to reflect upon their performance, what they had learned, and what they had liked and disliked about the project.

**Statistical Analysis**

Data were analyzed using chi-squared techniques and analysis of variance (ANOVA). Although in the report cards, physicians were provided with their own personalized...
data, in the statistical analysis all data were aggregated across physician, and the patient experience was assessed. The statistical analysis did not focus on individual physician performance. Thus, for the comparison of patient characteristics at the 3 time points, we compared the aggregate (not per physician) patient groups. Similarly, in the comparison of the use of imaging, aggregate comparisons were made. A P value less than .05 was considered to indicate statistical significance. Data were analyzed using SPSS v. 19.0 (IBM, Armonk, NY, USA).

RESULTS
In total, 18 physicians completed all 3 stages of the PI-CME process. Physicians who joined any of the practices during the project and had not been present for the baseline data collection or the educational intervention received reports about the use of high-technology imaging in their LBP patients, but they were not awarded CME credit for the project. Only those physicians who completed all 3 stages were eligible for credit.

Table 2 displays the general characteristics of the baseline, first follow-up, and second follow-up populations. Because the baseline period was in 2009, if a patient from baseline was also in the first follow-up dataset, this was counted as a new episode of LBP. Because the first follow-up and second follow-up datasets were back-to-back chronologically, if a patient in the second follow-up dataset was also in the first follow-up dataset, that patient was removed from the second follow-up dataset as the visit was likely a continuation of the original episode of care. There was no difference in mean age or distribution by clinic among the 3 periods; however, there were significantly more individuals classified as “other race” in period 3 (P < .0001).

Table 3 demonstrates that, even at baseline, there was no evidence of overuse of high-technology imaging. Out of 157 patients who presented with no red flags, only 3 high-technology images (all MRIs) were ordered. Interestingly, 72% of patients who were found to have 3 or more red flags had no imaging ordered (not even a plain x-ray).

During both follow-up periods, the number of patients with 0 red flags was higher than during the baseline period (P < .0001). During both follow-up periods there were 0 patients with no red flags who received an order for a high-technology image (the chi-squared test is invalid for this comparison because of 0 cells.) The percentage of patients with 3 or more red flags was lower in the first follow-up period when compared to the baseline period (P < .0001), but the percentage in the 2 follow-up periods was not significantly different. We found, however, that the percentage of patients with 3 or more red flags who had no imaging ordered (including plain x-rays) improved steadily from a baseline of 72% to 5.7% by the second follow-up period (P < .0001).

DISCUSSION
The Disappearing Practice Gap
The baseline finding that only 1.9% of patients with 0 red flags had a high-technology image ordered clearly indicated that our physicians were not over-utilizing high-technology imaging in this patient group; on the contrary, the finding that 72% of patients with 3 or more red flags who had no imaging ordered suggests that physicians may have been reluctant to order imaging even when red flags were present. These findings were surprising given that the PGIP data suggested over-utilization of high-technology imaging in low back pain. This difference in findings is likely a result of our careful review of the patient’s medical record as opposed to the use of billing codes only.

From administrative or billing data alone, the most useful piece of information that can be gleaned is the ICD-9 diagnosis code. In
the absence of the medical history data from the medical record, however, the ICD-9 code alone does not allow the determination of whether an imaging study was appropriate or inappropriate. For example, a 54-year-old female patient has an office visit with the ICD-9 code of 724.2 (lumbago). Based on that information alone, one cannot determine if a high-technology image, if ordered, is appropriate. By looking at the medical record, however, we find that the patient had 7 of the red flags listed in Table 1. Based upon the diagnosis of lumbago alone, it is likely the test would be considered inappropriate; however, when the data from the medical record are combined with the ICD-9 diagnosis code it is clear that a high-technology image was an appropriate diagnostic choice.

Thus, data derived solely from claims or billing data do not show the complete picture of the patient's condition—these data can only be gleaned from the medical record. When the patient's full medical history is taken into account, we found that our physicians were not over-utilizing high-technology imaging.

### The Power of PI-CME

As reflected in Table 3, our PI-CME project had a positive impact on physician practice. The low number of patients with 0 red flags at baseline who had a high-technology image ordered was reduced to 0 in both follow-up periods. Additionally, the number of patients with 3 or more red flags who had no imaging ordered was reduced to 5.7% by the second follow-up period. It is clear that the provision of the education session and the patient report cards helped physicians to provide more appropriate care to their patients.

In the literature, most articles about physician report cards concern report cards that are provided to patients about specific physicians or report cards that are provided to physicians about financial benchmarks. In our program, each physician received personalized report cards and no mention was made of any financial incentives. Roberts et al found that individual quality improvement “report cards” in the context of a larger quality improvement initiative led to significant improvements in an academic pulmonary department [11]. Ruzicka and Leenen also found that individual physician monitoring and feedback helped to improve hypertension control rates [12]. From a previous PI-CME project concerning thromboprophylaxis in patients at risk of VTE, we found that regular reports to physicians were useful tools in gaining improved compliance with the VTE standing orders (unpublished data, available upon request). PI-CME is an inherently data-driven approach to continuing medical education, so the methods of providing physician-specific performance data are likely to evolve from simple report cards as used in this project to more technologically advanced systems such as electronic dashboards.

### Did the Use of Imaging Increase?

If patients with 3 or more red flags were more likely to have images ordered during the follow-up period than at baseline, there is a possibility that the overall use of imaging may have increased. Overall, at baseline, 6.8% of all LBP patients had at least 1 high-technology image ordered. Some individuals had more than 1 image ordered, so the overall percentage of high-technology images ordered was 8.5%. At our first follow-up, 12% of all patients had at least 1 high-technology image ordered. As some individuals had more than 1 type of image ordered, the overall percentage of high-technology images ordered remained steady. Thus, our data indicate that even though individuals with 3 or more red flags were more likely to have images ordered after the intervention, the overall percentage of high-technology images ordered remained steady. These data indicate more effective targeting of imaging to individuals with red flags.

### Table 3. Use of High-Technology Imaging in the Pre-Intervention and Follow-Up Periods

<table>
<thead>
<tr>
<th></th>
<th>Baseline (7/1/09–12/31/09)</th>
<th>First Follow-Up</th>
<th>Second Follow-Up</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of charts reviewed</td>
<td>796</td>
<td>376</td>
<td>299</td>
<td></td>
</tr>
<tr>
<td>Percent [n] of patients with no red flags</td>
<td>19.7% [157]</td>
<td>44.9% [169]</td>
<td>43.1% [129]</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>Percent [n] of patients with 1 or 2 red flags</td>
<td>60.1% [478]</td>
<td>38.6% [145]</td>
<td>45.2% [135]</td>
<td></td>
</tr>
<tr>
<td>Percent [n] of patients with 3 or more red flags</td>
<td>20.2% [161]</td>
<td>16.5% [62]</td>
<td>11.7% [35]</td>
<td></td>
</tr>
<tr>
<td>Percent [n] of patients with no red flags who had a high-technology image ordered</td>
<td>1.9% [3/157]</td>
<td>0.0% [0/169]</td>
<td>0.0% [0/129]</td>
<td>No test computed</td>
</tr>
<tr>
<td>Percent [n] of patients with 3 or more red flags who had no imaging ordered (including plain x-rays)</td>
<td>72.0% [116/161]</td>
<td>50.0% [31.62]</td>
<td>5.7% [2/35]</td>
<td>&lt; .0001*</td>
</tr>
</tbody>
</table>

*Computed from chi-squared analyses comparing all 3 time periods.*
Physician Feedback

During the PI-CME project, we received communication from many physicians about the process. Frequently, after a set of report cards were distributed, physicians reported that they went to the electronic medical record to review those patients for whom a high-technology image was or was not ordered according to the presence of red flags. This verbal feedback from the physicians, coupled with the changes noted in Table 3, clearly indicates that the physicians were reviewing their individual data and were changing their practice based upon their own individual practice data. From the SurveyMonkey® (Palo Alto, CA, USA) anonymous evaluation, physicians were routinely positive about the program, made suggestions for future programs, and indicated that they had engaged in the reflective process as prescribed in Stage C of the PI-CME process.

Limitations

One notable finding is that the percentage of individuals with no red flags was higher in the 2 follow-up periods than at baseline. This finding may be a result of differences in how far back the research assistant looked in the chart to ascertain red flags. This difference in red flag ascertainment would lead to an overestimation of the number of patients with no red flags in the 2 follow-up periods. Despite this potential overestimation, the rate of use of high-technology imaging in patients who may have been incorrectly classified as “no red flags” did not go up—0 high-technology images were ordered on this group in the 2 follow-up periods. Although the number of patients with 0 red flags increased, there was not a large difference in the number of patients with 3 or more red flags—the metric we used to identify patients who may have more serious LBP. Thus, although there may have been a difference in red flag assessment, our results still suggest that high-technology imaging in patients with less than 3 red flags was not over-utilized, and in patients with 3 or more red flags, physicians were more likely to order an image (often a plain x-ray).

As this project was designed solely as a PI-CME activity and not as an investigation of concept of PI-CME itself, no control group was included. The use of a control group would have changed the nature of this activity to more of a research study about PI-CME than a PI-CME activity itself.

CONCLUSIONS

PI-CME is a powerful tool in the armamentarium of continuing medical education. The participating physicians’ involvement in the 3 stages of our PI-CME project engaged them in a learning journey that ultimately led to more appropriate patient care and utilization of imaging in LBP patients. Our data also indicated the hazards of solely using billing data to determine compliance with guidelines and benchmarks. Without the fullness of information found in the medical record, erroneous conclusions can be made about utilization. In this project, we relied on chart review, an intensive and time-consuming process, but with the advent and use of population management programs, physicians will be able to quickly review the status of their patient groups with respect to guideline adherence and for purposes of PI-CME.

PI-CME is an exciting, innovative approach to CME that requires active physician participation and that can lead to individualized learning and improvement in patient care. Through PI-CME, patients, physicians, payers, and health systems may all benefit from this unique learning experience.

ACKNOWLEDGEMENTS

This activity was funded through an educational grant from Pfizer to the Association of Hospital Medical Education (AHME) to support PI-CME activities for AHME member institutions. AHME retains full control over the distribution of individual grants under this collaborative grant program.

None of the authors have any conflicts of interest.

Dr. Szpunar has full access to the data and can verify the integrity of the data.

REFERENCES


