

Tracking Low Back Problems in a Major Self-Insured Workforce

Toward Improvement in the Patient's Journey

Harris Allen, PhD, Marcia Wright, PharmD, Terri Craig, PharmD, Jack Mardekian, PhD,
Raymond Cheung, PhD, Robert Sanchez, PhD, William B. Bunn III, MD, JD, MPH, and William Rogers, PhD

Objective: To assess the cost outcomes of treatment approaches to care for back problems in a major self-insured workforce, using published guidelines to focus on low back pain. **Methods:** Longitudinally tracked episodes of three types of International Classification of Diseases, Ninth Revision diagnosis code-identified back problems ($n = 14,787$) during 2001 to 2009. Identified five patterns of care on the basis of the first 6 weeks of claims and compared their total costs per episode with tests that included splits by episode type and duration, use of guidelines, and propensity-derived adjustments. **Results:** Care congruent with 10 of 11 guidelines was linked to lower total costs. Of the five patterns, complex medical management and chiropractic reported the highest and lowest rates, respectively, of guideline-incongruent use of imaging, surgeries, and medications, and the highest and lowest total costs. **Conclusions:** Approaches marked by higher resource utilization and lower guideline congruence are linked to greater low back pain total costs. Total cost is a needed input for guideline development.

Up to 84% of the general population has been found to report low back pain (LBP) symptoms at some point during their lifetime.¹ Back pain is the second most common symptom-related reason for clinician visits in the United States.² Yet, a high level of variability in practice patterns suggests much clinician uncertainty about how to treat LBP. Indeed, despite sharp increases in the use of various treatments, ranging from imaging, opioid pain medications, injections, and complementary/alternative medicine to surgery, there is no clear evidence reported to date of improved functional status and declining work disability related to LBP.³

These developments have raised LBP as a priority for employers focusing on workforce health and productivity (H&P). These employers are seeking to maximize the ratio of outcomes achieved relative to costs incurred (ie, value) for the investments that they are making in their employees (EEs). This search includes ways that preserve the viability of these investments for the future (ie, sustainability).⁴ It is leading these employers to identify targets for

a broadening array of measurement and management initiatives. Low back pain is becoming one such target because of the evidence emerging from recent studies. Across work that has ranged from estimating the unique contributions of LBP⁵ to head-to-head comparisons with other conditions,⁶ the magnitude of the total health burden, that is, the health care transactions driving direct costs and the lost productivity, short-term disability, long-term disability, and workers' compensation (WC) driving indirect costs, with which LBP is being linked has become a source of concern.

For these employers, the previously described picture is serving to reveal their "skin in the game" because clinicians and professional societies disseminate guidelines for improving the effectiveness and reducing the variability and costs of LBP treatment. A key contribution that these employers can make—not only to the continuous quality improvement of these guidelines but also to the refinement of their own H&P approaches—is the cultivation of evidence as to what is actually taking place on the ground.⁷ What treatment choices do EEs and their clinicians make? Do these decisions lead to differences in outcomes that are traceable to these choices? To what extent does congruence with guidelines alter the effect of LBP management choices on EE/patient outcomes?

An employer with well-developed and proven evidence-gathering and analytical capabilities is uniquely positioned to tackle these questions. Herein, we report a study conducted to obtain initial answers using an integrated database for a major self-insured employer with an accumulated track record of articles addressing a diversity of H&P issues in this journal.⁸⁻¹⁷

TREATING LOW BACK PAIN

Most patients (greater than 85%) seen in primary care for LBP have "nonspecific low back pain," which cannot be attributed to pathophysiological causes, while the balance are linked to a specific disease or spinal pathology (eg, radiculopathy).¹⁸ Although the long-term prognosis for LBP is generally favorable—with the norm being rapid improvement in pain, disability, and return to work in the first month and further improvement that often planes out at the 3-month mark—the duration of the condition and its effect can be substantial. In a large prospective study, although 90% of general practice patients with LBP stopped seeing their doctors within 3 months, only 25% were reported to be completely free of pain and disability after 1 year.¹⁹

Consistent with the generally favorable long-term prognosis for the LBP, roughly only a third of patients seek medical care, with the balance apparently improving on their own.³ Of the former, one third are substantially improved after 1 week and two thirds after 7 weeks.²⁰ Even so, recurrences may affect 40% to 50% of patients within 6 months²¹ and 70% within 12 months.²⁰ For many, the natural history is akin to chronic conditions such as asthma, marked by chronic mild symptoms and intermittent exacerbations. In one prospective study of patients with acute LBP, chronic LBP was diagnosed for 20% of patients within 2 years of their first visit.²²

These features have led to the differentiation of three categories for classifying the duration of treatments for back pain in general and LBP specifically. Acute LBP refers to LBP that lasts 4 weeks or less, subacute LBP refers to LBP that lasts between 4 and

From the Harris Allen Group, LLC (Dr Allen), Brookline, Mass; US Medical Affairs (Dr Wright), Pfizer Integrated Health, Overland Park, Kans; Pfizer Primary Care Medical Affairs (Dr Craig), Lincoln, Nebr; Pfizer Inc (Dr Mardekian), New York; Pfizer Integrated Health (Drs Cheung and Sanchez), New York; Health, Safety, Security & Productivity, Navistar, Inc (Dr Bunn), Lisle, Ill; and Tufts Medical Center (Dr Rogers), Boston, Mass.

This study was sponsored by Pfizer Inc. Harris Allen and William Rogers are principal and a paid associate, respectively, of the Harris Allen Group, LLC, which served as a paid consultant organization to Pfizer in connection with this study and the development of this manuscript. William Bunn served as an unpaid adviser to the Harris Allen Group. Marcia Wright, Terri Craig, Jack Mardekian, and Raymond Cheung are employees of Pfizer Inc. Robert Sanchez was a full-time employee of Pfizer Inc at the time this manuscript was drafted.

Supplemental digital content is available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.joem.org).

The authors declare no conflicts of interest.

Address correspondence to: Harris Allen, PhD, Harris Allen Group, LLC, 150 Walnut St Rd #2, Brookline, MA 02445 (harris@harrisallengroup.com).

Copyright © 2014 by American College of Occupational and Environmental Medicine

DOI: 10.1097/JOM.0000000000000210

12 weeks, and chronic LBP refers to LBP that persists longer than 12 weeks. Wide variations in the use of diagnostic tests, pain medications, injections, and surgery suggest much clinician variability about optimal therapy for each category. Indeed, many patients with chronic LBP may not be receiving evidence-based care. A recent study, for example, linked 732 North Carolina adults with chronic LBP to overuse of unproven interventions (eg, traction), high use of second-line medications (eg, opioids), and underuse of exercise therapy and, for depressed patients, antidepressants.²³

Such outcomes have led to growing worries that the condition has become “overmedicalized.”²⁴ Pertinent evidence is emerging in several areas, for example, as follows:

1. Early use of magnetic resonance imaging (MRI) has been linked to prolonged disability, higher medical costs, and greater use of surgery^{21,25} and to reviews finding no benefit on health, function of disability outcomes for LBP.²⁶
2. Patients receiving chiropractic care have been found to record lower associations of probability of disability recurrence than patients of physicians and physical therapists.²⁷ On the contrary, chiropractic care with shorter duration has also been found to be associated with lower rates of disability recurrence and shorter disability duration.²⁸

The frequency and the strength of opioid dosage in LBP treatment have been positively linked to claim duration.²⁹ Likewise, the odds ratios for a catastrophic claim (total cost of \$100,000 or more) when spinal surgical procedures were performed increased 10-fold when treatment included opioid use.³⁰ Yet, close provider supervision can apparently have mitigating effects, as suggested by a third study, which found that each additional week between the filling of opioid prescriptions predicted 14% longer disability and time off work.³¹

Guidelines

To promote effective and efficient care for LBP patients, clinical guidelines have been published in the United States³² and Europe.³³ Country-specific guidelines have also surfaced.^{34–36} These encompass recommendations for diagnosis, screening, care pathways, and treatment algorithms and, in general, specify that the criteria for evaluation (eg, ordering diagnostic studies³⁷) and treatment (eg, performing surgeries³⁸) for patients with possible work-related LBP should not be different from the criteria for patients with nonoccupational LBP.

A synthesized reading of these guidelines suggests that they have elements that can be examined using the criteria pertinent to employers and their capacity to manage the value/sustainability challenge that they face as a stakeholder group. A key issue is the direct and indirect cost outcomes of the care that transpires under EE benefit coverage whose cost they and their EEs share. Although the delivery of care and its comportment with guidelines are beyond the province of employers who look to providers and payers to assume such responsibilities, outcomes are of keen interest. In particular, what is the effect on cost outcomes when care is incongruent with guidelines?

As with other clinical areas, guideline development and implementation for LBP continue to be marked by much flux and lack of consensus. Yet, when evaluating care and total cost outcomes for employer purposes, it matters less that the guidelines selected be the best available in terms of performance on scientific criteria. What matters more is that these guidelines are (1) sufficiently grounded in empirical work that they can serve as a credible foundation for this partner dialogue and collaboration and (2) readily testable in the context of the data and analytical resources to which employers have access. The following four areas taken from the recently published literature on LBP guidelines focus on timing aspects that meet these two criteria.

Imaging

Barring progressive neurological findings or suspicions of systematic etiology, routine x-rays have been recommended only after 4 weeks postonset if clinical improvement has not occurred.³⁹ Similarly, barring evidence at clinical examination of emergent conditions (eg, cauda equina syndrome), MRIs or computed tomographic (CT) scans may be appropriate after a minimum of 4 weeks of radicular symptoms.³⁹ Otherwise, excluding prior indications of progressive neurological deficits or a high suspicion of cancer or infection, CTs or MRIs have been recommended only after 12 weeks of persistent back pain.²

Manipulation

The use of short-term therapeutic courses of manipulation treatment may likewise be indicated, but guidelines have proposed that physical therapy (PT) referrals should not be made within—the least—the first 2 weeks of onset.³

Surgery

For patients with specific LBP etiology (eg, lumbar radiculopathy), a waiting period of at least 6 weeks has been recommended before they are to be considered candidates for surgery. For non-specific LBP, surgery has been recommended only for patients who have experienced persistent symptoms and associated disability for at least 1 year despite nonsurgical interventions.⁴⁰

Medications

Prescription (Rx) medications are of particular relevance to care for patients with acute LBP, where the focus for therapy is on short-term treatment for temporary symptomatic relief to maximize patient comfort, although they are also applicable to subacute and chronic LBP cases.³ Opioids, nonsteroidal anti-inflammatory drugs (NSAIDs), other pain medications, muscle relaxants, anxiolytics, and sedatives prescribed as monotherapy on a fixed schedule *or* in appropriate combinations have all been the subject of evidence assessment and recommendations in this regard.⁴¹ In general, timeframes used in recent LBP clinical trials have linked “short term” for these pharmacologic therapies to use not exceeding 12 weeks, although for benzodiazepines and muscle relaxants, short term has been linked to 4 weeks.⁴¹

Study Objectives

Previous work on the integrated database tapped for this study showed its capacity to identify specified patient groups and monitor their direct and indirect cost outcomes during 2001 to 2009. This study built on this precedent by examining the actuarial consequences of clinical treatment choices for back problems, using the previously described guidelines to focus on LBP. The objectives were the following:

1. Identify all active EEs reporting a back problem diagnosis during the study period.
2. Define and classify their initial patterns of medical care and Rx medication use.
3. Track the effect of these patterns on direct and indirect cost outcomes.
4. Further stratify these treatment patterns by measures of congruence with the previously described guideline aspects for LBP care and determine the effect on cost outcomes.

METHODS

Study Design, Company, and Data Sources

This study was a retrospective time-series analysis of an extract taken from an integrated database focusing on the continental

US active EEs of an international heavy manufacturer of transportation equipment. In 2009, this company was ranked 175th on the Fortune 500 list.

This database tracks measures of employee personnel characteristics as well as medical, behavioral health, pharmaceutical, WC, disability, absenteeism, and lost productivity during the 2001 to 2009 period. The sources tapped for these measures include health plan, pharmaceutical benefit manager and behavioral health manager claims obtained through the company's data manager vendor; WC, short-term disability and long-term disability data obtained from the company's WC/disability vendor via the company; payroll and human resource data obtained from the company's human resources department; and three rounds of "special topic" company-wide employee surveys administered by an external vendor using an electronic/manual approach in the spring and fall of 2001 and the fall of 2009. In 2011, this database was approved by the Western Institutional Review Board for compliance with the Health Insurance Portability and Accountability Act.

Study Population

As reported elsewhere,¹⁵ this workforce is largely older, male, and hourly, and averages relatively long company tenures and relatively mid-level annual income. But, it became significantly less so on each of these characteristics from 2001–2002 to 2008–2009: average age dropped from 48.5 to 46.2 years; percentage male, from 83.2% to 77.6%; percentage hourly, from 64.9% to 53.6%; average tenure, from 20.2 to 14.5 years; and average income, from \$66,791 to \$64,822 (2009 dollars).

Disease Sample

All EEs in the extract with at least 6 months of recorded continuous coverage on a company health plan during the study period were considered eligible for selection into the disease sample unless they developed a cancer diagnosis, retired, or died, after which they were dropped from the sample. A total of 21,080 individuals met these initial requirements. Employees with back pain were then identified on the basis of Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)⁴² diagnoses reported in medical claims. Back pain status was differentiated by the presence of one or more eligible diagnoses on one or more medical claims lines. Those identified were grouped as follows:

- #1: LBP with neurologic findings (LBP/neuro): ICD-9-CM codes 721.42, 721.91, 722.73, 722.80, 724.3, 724.4 (ie, diagnoses indicative of specific etiology; $n = 1837$)
- #2: LBP with no neurologic findings (LBP/nonneuro): ICD-9-CM codes 724.2, 724.5, 846.0-9, 847.2, 847.3, 847.9 (ie, diagnoses indicative of nonspecific etiology; $n = 8569$)
- #3: Other back: all ICD-9-CM codes in categories 13.3.-1-3 of the Clinical Classifications Software for the Healthcare Cost and Utilization Project⁴³ (CCS-HCUP) other than those in #1 and #2, including cervical and thoracic back pain diagnoses and all codes in the 720 to 724 range not listed previously ($n = 4381$).

Of the 21,080 EEs, 8300 (39.4%) had at least one back pain episode and many had more than one. Table 1 describes these three groups on measures ranging from demographic and job characteristics to health and utilization characteristics 6 months prior to disease onset, several of which were assessed by survey in spring 2001 and in fall 2009. All three groups were on average older, more likely to be hourly, and slightly more likely to be female than the aggregate sample. They were also likely to have somewhat longer job tenure but less annual income.

Units of Analysis

The units of analysis were episodes. An episode was defined by a period of medical claims with back pain diagnoses (all types,

inclusive), separated from previous episodes by at least 6 months without a back pain claim, and followed by at least 6 months without a back pain claim. After the episode was defined in this way, it was classified into one of the three groups described previously. If any claim in the episode had an LBP/neuro or LBP/nonneuro diagnosis, the episode was classified as such based on the first diagnosis reported. If there was never any such diagnosis in the stream of diagnoses listed, the episode was classified in the other back group.

A total of 14,787 back pain episodes were identified, of which 1837 were LBP/neuro, 8569 were LBP/nonneuro, and 4381 were other back. These counts included episodes that were left censored, whose first observation date was unclear because there was less than 6 months of enrolled history. With left-censored episodes excluded, there were 13,224 back pain episodes, of which 1623 were LBP/neuro, 7650 were LBP/nonneuro, and 3951 were other back. These counts included right-censored episodes, whose observation period ended too early (less than 6 months before the end of enrollment or the first cancer date) to know if the episode had truly ended.

Episode lengths were short. With left- and right-censored episodes included, the median LBP/neuro episode lasted 42 days, whereas the medians for LBP/nonneuro and other back episodes were 6 and 0 days, respectively (the group means were 148, 91.3, and 51.4 days, respectively). The breakdown by episode duration was as follows:

1. *LBP/neuro*: 44% acute, 18% subacute, and 38% chronic
2. *LBP/nonneuro*: 61% acute, 12% subacute, and 27% chronic
3. *Other back*: 70% acute, 13% subacute, and 17% chronic

The vast majority of episodes were defined on the first day of the episode. Of the LBP/neuro episodes, 84.6% were defined on the first day compared with 90.4% for LBP/nonneuro. By construction, all other back episodes were defined on the first day.

Of the LBP/neuro diagnoses, 52.3% were lumbosacral neuritis not otherwise specified, and 44.0% were sciatica. None of the other diagnoses in this category recorded greater than 1.9% prevalence. Of the LBP/nonneuro diagnoses, 52.2% were lumbago, 26.6% were backache not otherwise specified, 12.1% were sprains in the lumbar region, and 5.1% were lumbosacral sprains. None of the other diagnoses in this second category had greater than 2.2% prevalence.

Exploratory tests suggested notable "bleed" across these diagnostic groups over time. For those with an initial LBP/neuro diagnosis, 40% received an LBP/nonneuro diagnosis and 58% received another back diagnosis by 90 days after the first visit. For those with an initial LBP/nonneuro diagnosis, 6.8% received an LBP/neuro diagnosis and 35% received an other back diagnosis. By definition, all episodes initially classified as other back reported no additional LBP/neuro or LBP/nonneuro diagnoses in the first 90 days.

In this context, only the 40% and 6.8% had substantive meaning because the other back designation included many episodes that started with a noncommittal diagnosis defined only after some investigation or episodes in which the diagnosis associated with some procedures was ambiguous with respect to etiology. Nonetheless, this extent of change was not unexpected. With respect to LBP specifically, a key function of the initial clinical evaluation is to enable patients to be triaged appropriately by the type of diagnosis.² Yet, the use at initial visits for LBP of commonly recommended "red flag" questions developed from decision rules designed to detect malignancy, infection, cauda equina syndrome, fracture, and inflammatory disorders has been linked to high false-positive rates. One recent study of 1172 patients at the time of first consultation for new episodes of LBP in primary care settings, for example, linked 80% to one or more 25 "red flag" symptoms when the detected incidence of spinal fracture and malignancy turned out to be 0.7% and 0.0%, respectively.⁴⁴ Most significant pathologies will tend to become more clinically obvious over time because of either the clinical course or exacerbations stemming from usual care for nonspecific LBP.

TABLE 1. Sample Characteristics by Type of Back Pain Episode

Characteristic	Low Back/Neuro (n = 1,837)	Low Back/ Nonneuro (n = 8,569)	Other Back (n = 4,381)	F	P
Demographic					
Age, yrs	51.7	49.1	50.7	53.0	0.00
Male, %	75.6	76.3	75.2	1.1	0.34
High school or less, %*	11.0	12.6	13.3	0.4	0.65
Job					
Hourly (vs salaried), %	62.3	66.6	61.5	19.1	0.00
Company tenure, yrs	20.3	18.2	19.6	22.9	0.00
Annualized earnings, \$	45,392	43,540	44,330	2.4	0.10
Factors of interest					
Previous 6 mos					
Anxiety dx, %	1.7	1.0	1.2	0.5	0.60
Depression dx, %	5.3	6.2	4.7	2.5	0.09
Comorbid CCS categories, avg	7.4	6.8	6.9	9.6	0.00
Total direct costs, \$	8,295	7,344	9,485	15.7	0.00
Total workers' compensation costs, \$	544	258	457	2.1	0.13
Total short-term disability costs, \$	253	212	193	2.5	0.08
Total controllable absenteeism costs, \$	109	119	102	10.7	0.00
Spring 2001					
BMI >30, %*	35.7	34.2	34.5	0.2	0.83
Smoked, %*	19.2	23.5	19.5	4.5	0.01
Physical health ≤ 25th percentile*	37.7	31.8	32.5	3.4	0.03
Mental health ≤ 25th percentile*	29.6	26.1	29.4	2.7	0.07
Fall 2009					
BMI >30, %*	32.8	33.7	31.9	0.2	0.79
Currently smokes, %*	4.9	8.7	10.6	3.5	0.03
Physical health ≤ 25th percentile*	29.3	27.3	29.1	0.4	0.68
Mental health ≤ 25th percentile*	28.5	25.9	22.7	1.8	0.17
Dissatisfied with personal life, %*	12.6	10.8	13.4	1.1	0.34

*Obtained from survey; otherwise obtained from claims.

BMI, body mass index; CCS, Clinical Classifications Software; dx, diagnosis.

Procedures and Medications

Procedures for this study were identified on the basis of the *Current Procedural Terminology-4* fields holding either Healthcare Common Procedure Coding System codes or ICD-9-CM procedure codes found in medical claims. These procedures were counted as back-pain related if one of the diagnoses in the corresponding claims line item aligned with one of the diagnoses that made up the previously described definition of back pain. A total of 215,295 procedures were reported as such for the three back pain groups during the 9-year period. For these, the following categories were identified:

1. Physician or staff visit: evaluation and management; staff/technician (eg, laboratory team members like phlebotomists); mental/behavioral health; emergency department visit
2. Surgery and injections: major (invasive) primary low-back-related surgery; nerve blocks; and other injections and palliative procedures
3. Imaging: ultrasound; CT scan; MRI; other x-ray or imaging, including dye contrast
4. Chiropractor or PT visits
5. Other: hospital; ambulance

Medications were identified on the basis of the national drug category and ingredient entries found in pharmaceutical claims. In contrast to procedures, these entries were not attached to back pain

diagnoses specifically, and additional tests were, therefore, required to make these linkages. Specifically, these entries were grouped into drug classes that (1) the American Hospital Formulary Service Pharmacologic-Therapeutic Classification System⁴⁵ has categorized as approved for the treatment of central nervous system conditions (cf American Hospital Formulary Service class number 28:00 Central Nervous System Agents) and (2) were observed as having Rx fill patterns that covaried with the date of initial back pain diagnosis.

In the absence of knowing what was actually written for back pain, tests of this covariation were conducted to examine the timing of Rx fills relative to the episode start date. Consider opioids, for which it was expected that Rx filling would be highly related to episode starts, and topical steroids, for which little such association was expected. During the period from 9 days prior and 30 days after episode start date, 52% of all Rx's for opioids were filled either on or up to 3 days after the start date. The corresponding number for topical steroids was less than 20%. Thus, opioids showed a strong relationship to episode start date and were included as a medication class in the analyses. Topical steroids showed a much weaker association and were discarded as a class for these analyses. (Full results for these assessments are available on request.)

The total number of prescribed medications filled by three back pain groups sample across the 9-year study period was 44,100.

Of these, the following classes of medications showed sufficient covariation to be included in the analyses:

1. Opioids: short acting; long acting
2. Nonsteroidal anti-inflammatory drugs, including COX-2 selective inhibitors
3. Other pain medications (eg, antiepileptic)
4. Muscle relaxants
5. Selective serotonin reuptake inhibitors (SSRIs; eg, fluoxetine), serotonin norepinephrine reuptake inhibitors (SNRIs; eg, duloxetine), and tricyclics (eg, amitriptyline)
6. Oral steroids (eg, prednisone and methylprednisolone)
7. Anxiolytics, sedatives, and hypnotics (eg, benzodiazepines and barbiturates)

Table 2 gives the percentage of episodes, reporting the use of these procedures (with a pain diagnosis) and medications (with or without a diagnosis) by episode type within timeframes pertinent to LBP guidelines—by the 14th day (ie, from episode start date through day 14); by the 28th day; by the 90th day; by the 365th day; and after the 365th day. As shown, the more costly, complex procedures like major surgery, injections, and MRIs tended to be used at greater rates earlier on in LBP/neuro episodes. Rates of emergency department use were highest among LBP/nonneuro episodes. Rates of medication use for both LBP types tended to be on par and greater than those for the other back group. Otherwise, rates across the three episode types tended to be comparable.

Total Costs

The ultimate outcome for this study was annualized total costs per EE, which summed the following direct and indirect cost components (all expressed in 2009 dollars):

1. *Direct*: Inpatient admits, outpatient visits, emergency department visits, behavioral visits, laboratory tests, equipment orders, doctor's office injections (from medical claims), and prescriptions filled (from Rx claims). All components included the shares borne by the company and EEs, with the latter including all EE payouts except plan premiums. Dental care was excluded.
2. *Indirect*: WC, short- and long-term disability and "controllable absenteeism" (ie, time away from work for individual health reasons). The dollar amounts lost to absenteeism were derived from the company's pay check and job history records.

The total costs reported in a previous study on this database, it can be noted, included lost productivity because of presenteeism. Presenteeism was not included in this study because of the complexities that it would have introduced (eg, the claims-based episode start dates spanning 2001 through 2009 vs survey-based estimates pertaining to 2001 and 2009 only). With presenteeism excluded, this previous report recorded annual total costs of \$7859 and \$5869 per employee (unadjusted in 2009 dollars) for the 2001–2002 and 2008–2009 timeframes, respectively.¹⁵

Analytic Strategy

Analyses were first undertaken to classify the three types of back pain episodes by approach to medical treatment on the basis of claims activity. This classification was predicated on which providers were seen and what procedures were received during the first 6 weeks after episode start date. To investigate the extent to which the resulting taxonomy was merely a tautological reflection of case mix, the five initial treatment approaches thus identified were examined in relation to the previous treatment patterns with controls for episode type and duration. The relationships between these approaches and initial medication use were also explored.

Then, guidelines were brought into the mix with measures defined in terms of incongruence vis-à-vis aspects drawn from the

forementioned literature synthesis. These guideline measures focused on LBP episodes and were likewise examined in relation to initial treatment patterns. Last, the effect of initial treatment pattern (for all three back problem groups) and guideline incongruence (for the two LBP groups only) on cost outcomes was assessed. The longitudinal data set enabled these probes to be conducted across years 1, 2, and 3 of episodes.

To make adjustments, logistic regression was used to calculate probability scores on the basis of the regression of binary treatment variables on predictors that might make the selection of treatments nonrandom. Specifically, propensity-scoring techniques that others^{46–48} have developed to adapt Rubin's approach to multiple treatment choice scenarios were used to adjust for the following: age, sex, episode type and start date year, comorbidities, the total number of claims lines, had prior episode, and left-censored episode status (details available on request).

Although discrepancies between the unadjusted and adjusted results were of interest, the adjusted version was chosen as the basis for hypothesis testing because of its presumed greater accuracy in this context. These tests focused on two general hypotheses. First, greater rates of guideline incongruence will be associated with greater total cost outcomes over time. Second, the positive relationship between initial treatment patterns that are more medically complex in orientation and total costs will be exacerbated by greater rates of guideline incongruence. All analyses used the STATA 9 software program (STATA Corp, College Station, TX).

RESULTS

Initial Treatment Patterns

The first 6 weeks of medical claims were searched for patterns that reflected the presence or absence of clear treatment choices. Counts of the various procedures found were used to establish the following five overall patterns (overall sample percentage in parentheses):

1. *Information and Advice (TalkInfo)*: Episodes whose procedures during the first 6 weeks reflected information gathering or advice seeking but otherwise no overriding pattern: simple office visits, laboratory tests, emergency department or hospital visits, talk therapy, or visits involving imaging (x-ray, ultrasound, CT, or MRI) but no other procedures (59%).
2. *Complex Medical Management (Complex MM)*: Episodes in which the number of physician visits for nerve blocks, surgeries, or comparable procedures was greater than 1 and comprised at least the plurality of treatments. Any ties with other categories went to this category (2%).
3. *Chiropractic (Chiro)*: Episodes in which the number of visits to a chiropractor was greater than 1 and comprised the plurality or greater of procedures. This included cases involving manipulation billed as PT if the manipulation occurred on the same day (11%).
4. *Physical therapy (PT)*: Episodes in which the number of visits to a PT was greater than 1 and comprised the plurality or greater of procedures. Physical therapy by itself (no chiropractor) sometimes included devices or other palliative treatments (11%).
5. *Dabble*: Episodes with at most one visit for physician, chiropractic, or PT care, or at most one visit to two or more of these categories (17%).

Across the five patterns, when selected, the Complex MM, PT, and Chiro approaches were most likely to be initiated in LBP/neuro cases, whereas the TalkInfo approach was least likely to be initiated in LBP/neuro cases (Table 3). In contrast, the TalkInfo approach was most likely to be initiated in LBP/nonneuro cases, whereas the Dabble approach was slightly more likely to be initiated in other back cases. Breakouts by episode duration showed little change in

TABLE 2. Procedures and Medications by Episode Type and Incurred Date Within Episode

	Low Back/Neuro				Low Back/Nonneuro				Other Back						
	By 14th Day	By 28th Day	By 90th Day	By 365th Day	By 14th Day	By 28th Day	By 90th Day	By 365th Day	By 14th Day	By 28th Day	By 90th Day	By 365th Day	>365th Day		
	Day	Day	Day	Day	Day	Day	Day	Day	Day	Day	Day	Day	Day		
Procedures															
Physician/staff visit	74.5*	77.0	80.6	85.3	50.0	74.1	75.6	78.4	82.3	49.1	57.3	58.9	61.4	66.2	45.6
Evaluation and management															
Staff/technician	8.3	10.2	14.2	19.4	26.0	5.5	6.3	8.3	11.6	23.6	5.9	6.7	8.0	10.3	22.9
Mental/behavioral health	0.0	0.0	0.1	0.1	2.3	0.0	0.0	0.0	0.0	2.2	0.0	0.0	0.0	0.1	2.1
Emergency department visit	4.8	4.9	5.6	7.6	9.6	8.5	8.7	9.3	10.9	10.1	2.8	2.9	3.0	3.6	8.4
Surgery, injections															
Major low back surgery	0.7	1.2	3.3	6.6	4.0	0.2	0.4	1.1	2.3	2.1	0.3	0.5	0.9	1.5	1.6
Nerve blocks and injections	6.9	10.0	15.9	21.2	12.0	1.6	2.4	4.5	7.9	7.7	1.4	1.6	2.3	3.5	5.1
Imaging															
Ultrasound	0.3	0.4	0.5	0.8	4.7	0.6	0.7	0.8	1.1	4.3	0.6	0.6	0.6	0.7	5.0
CT	2.3	2.9	4.0	5.9	6.4	2.4	2.7	3.2	4.5	5.9	3.3	3.4	3.8	4.2	5.1
MRI	18.4	21.8	27.2	33.4	15.5	7.4	9.3	12.2	17.4	12.6	13.4	14.8	16.8	19.8	11.6
Other x-ray	38.1	41.2	46.9	55.1	35.8	31.7	34.0	37.8	44.9	31.1	37.7	38.9	41.0	45.4	29.8
Chiropractor and PT visits															
Chiropractor	23.6	24.3	25.3	27.4	12.3	11.7	12.1	13.0	15.2	9.3	19.7	20.0	20.4	21.4	11.3
PT	25.9	28.5	34.3	40.1	21.6	14.2	16.5	19.8	25.3	16.0	20.4	21.6	24.0	27.3	16.4
Other															
Hospital	5.8	6.4	8.8	12.5	13.7	5.6	6.1	7.2	9.8	12.6	3.1	3.5	4.0	5.2	11.4
Ambulance	0.2	0.2	0.3	0.4	2.3	1.2	1.3	1.4	1.6	2.0	0.5	0.5	0.6	0.8	1.9
Medications															
Opioids	28.0	33.7	42.4	58.0	51.7	28.6	32.2	40.1	55.4	50.9	17.3	20.9	28.9	44.4	46.4
Short acting	27.4	33.1	42.0	57.6	51.7	28.1	31.8	39.7	55.2	50.7	16.9	20.4	28.5	44.0	46.3
Long acting	1.2	2.0	3.5	5.6	6.0	1.0	1.3	2.0	3.6	5.1	0.9	1.4	1.7	2.5	3.8
NSAIDs	18.8	24.8	35.5	48.1	39.7	22.5	26.7	34.3	45.8	38.4	15.5	19.8	27.2	38.4	35.2
Other pain medications	2.9	4.3	7.1	11.9	11.7	1.0	1.8	3.3	5.6	8.0	1.4	2.3	3.7	5.6	7.0
Muscle relaxants	16.4	19.1	23.1	31.5	25.7	25.0	26.9	30.6	38.0	24.2	13.0	14.8	18.1	24.1	21.8
SSRI, SNRI, tricyclics	5.6	8.9	14.4	19.5	20.5	5.0	8.6	14.0	20.0	21.2	5.0	8.3	13.0	19.2	20.3
Oral steroids	3.8	6.6	11.7	15.3	17.6	3.4	5.8	9.6	13.8	14.6	3.4	5.9	10.1	14.2	15.2
Anxiolytic/sedative/hypnotics	1.5	2.8	5.4	9.6	14.2	2.0	3.1	5.4	9.8	13.2	1.8	2.8	4.9	8.9	13.2

*Cell entry is the percentage of episodes having any of the procedures (with a pain diagnosis) or drugs (with or without a back pain diagnosis) within the number of days after the episode start date that is specified in the column heading.
 CT, computed tomography; MRI, magnetic resonance imaging; NSAIDs, nonsteroidal anti-inflammatory drugs; PT, physical therapy; SSRI, selective serotonin reuptake inhibitor; SNRI, serotonin-norepinephrine reuptake inhibitor.

TABLE 3. Episode Type and Duration by Initial Treatment Pattern

Type of Back Pain	Duration	Initial Treatment Strategy*					Total
		TalkInfo	Dabble	Complex MM	Chiro	PT	
Low back/neuro	Acute	457†	166	18	125	48	814
	Subacute	88	53	26	75	86	328
	Chronic	209	132	48	144	162	695
	All	754 (41.1%)	351 (19.1%)	92 (5.0%)	344 (18.7%)	296 (16.1%)	1,837 (100.0%)
Low back/nonneuro	Acute	3,995	794	22	243	205	5,259
	Subacute	402	158	40	143	311	1,034
	Chronic	1,020	426	53	359	418	2,276
	All	5,417 (63.2%)	1,378 (16.1%)	95 (1.1%)	745 (8.7%)	934 (10.9%)	8,569 (100.0%)
Other back	Acute	2,107	601	18	220	113	3,059
	Subacute	155	97	14	145	157	568
	Chronic	246	153	18	218	119	754
	All	2,508 (57.3%)	851 (19.4%)	50 (1.1%)	583 (13.3%)	389 (8.9%)	4,381 (100.0%)
	Total	8,679 (58.7%)	2,580 (17.5%)	237 (1.6%)	1,672 (11.3%)	1,619 (11.0%)	14,787 (100.0%)

*Each episode is assigned to only one category. Categories are mutually exclusive.

†Unless otherwise indicated, cell entries are counts of episodes.

PT, physical therapy.

TABLE 4. Initial Treatment Pattern by Medication Class Percentage Filling a Prescription in the First 4 Weeks

Initial Treatment Pattern	Medication Class						
	Opioids	NSAIDs	Other Pain	Muscle Relaxants	SSRI/SNRI/Tricyclics	Oral Steroids	Anxiolytics/Sedatives/Hypnotics
TalkInfo	28.4	27	2.1	24.2	8.7	5.5	3.1
Dabble	33.5	21.3	2.2	21.9	8.5	6.5	3.4
Complex MM	59.1	28.7	8	29.1	10.5	5.9	5.5
Chiro	12.3	9.1	0.8	8.9	6.2	5.2	1.5
PT	38.6	30.8	3.7	44	9.7	7.8	3.2
F ^a	110.3**	75.2**	17.3**	53.5**	4.1*	4.2*	4.9**

* $P < 0.01$; ** $P < 0.001$.

^aF statistics for analysis of variance of column entries.

^bUnless otherwise indicated, entries are percentage "yes."

NSAIDs, nonsteroidal anti-inflammatory drugs; PT, physical therapy; SNRI, serotonin-norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor.

these patterns across acute, subacute, and chronic cases within each episode type.

Of the five patterns, the Complex MM approach was linked to the highest rates of Rx fills for four of the seven drug classes—opioids, other pain medications, SSRI/SNRI/tricyclics, and anxiolytics/sedatives/hypnotics (Table 4). The PT group was highest in NSAIDs, muscle relaxants, and oral steroids, whereas the Chiro group had the lowest Rx rates in all seven classes.

Treatment Choices and Costs

These differences in rates of medication use and the longitudinal differences in rates of use of complex procedures noted previously—both of which linked the more intensive treatment approaches to LBP/neuro cases—raised the question, to what extent was the fivefold taxonomy of treatment approaches a result of the type and duration of episodes? If the approaches were merely a reflection of the presenting diagnosis and length of the subsequent treatment incurred, then the taxonomy could be seen as a construction of circular reasoning, entirely reducible to the ICD-9 and the *Current Procedural Terminology-4* codes used to construct it and of little explanatory value in its own right.

Yet, this turned out not to be the case. When those EEs with more than one episode were examined, their previous initial treatment choice was highly predictive of their current initial treatment choice. In fact, they tended to start out by choosing the same approach that they had previously—those who had selected the chiropractic approach previously were most likely to again select the chiropractic approach for the current episode, those who had previously selected the PT and Complex MM approaches again selected the PT and Complex MM approaches, etc (Table 5). The selection effects remained intact even when episode type and duration were controlled. The initial treatment choices made by EEs and their providers exerted systematic influences above and beyond the presenting clinical features of the episodes that ensued.

These choices, in turn, had cost consequences. The Complex MM approach recorded the highest total unadjusted costs for the three back pain groups (Table 6), especially in year 1 but remained as such in years 2 and 3. The Chiro approach, in contrast, was consistently linked to the lowest costs. Its per-employee total stayed at the lowest level across all three groups for each year. Physical therapy was associated with relatively high costs in year 1 but tapered off in years 2 and 3.

TABLE 5. Initial Treatment Approach as a Function of Previous Treatment Approach and Episode Type and Duration

		Previous Treatment Approach Only			Previous Approach + Episode Type and Duration		
		RRR	SE	P	RRR	SE	P
Dabble	Previous Dabble	2.76 ^a	0.23	***	2.74	0.23	***
	Previous Complex MM	1.91	0.46	**	1.80	0.43	*
	Previous Chiro	3.97	0.46	***	3.97	0.46	***
	Previous PT	1.51	0.17	***	1.49	0.17	***
	Episode type		—		0.88	0.05	*
	Episode duration		—		1.00	0.00	***
Complex MM	Previous Dabble	2.41	0.49	***	2.38	0.59	***
	Previous Complex MM	6.34	2.54	***	5.20	2.12	***
	Previous Chiro	1.28	0.61		1.33	0.63	
	Previous PT	3.00	0.75	***	2.95	0.74	***
	Episode Type		—		0.43	0.06	***
	Episode Duration		—		1.00	0.00	***
Chiro	Previous Dabble	3.82	0.44	***	3.76	0.44	***
	Previous Complex MM	1.45	0.58		1.29	0.53	
	Previous Chiro	20.66	2.44	***	20.48	2.46	***
	Previous PT	1.83	0.29	***	1.80	0.29	***
	Episode Type		—		0.81	0.06	**
	Episode Duration		—		1.00	0.00	***
PT	Previous Dabble	1.47	0.17	**	1.46	0.17	**
	Previous Complex MM	2.35	0.60	**	2.08	0.54	**
	Previous Chiro	1.41	0.26		1.41	0.26	
	Previous PT	2.59	0.28	***	2.56	0.29	***
	Episode Type		—		00.73	0.05	***
	Episode Duration		—		1.00	0.00	***

^a*P* < 0.05; ***P* < 0.01; ****P* < 0.001.

³All estimates are calculated in relation to TalkInfo, which was the holdout group for this analysis.

PT, physical therapy; RRR, relative-risk ratio; SE, standard error.

Similarly, medication choices had cost consequences. With rare exception, medication use within 4 weeks of episode start date was linked to higher unadjusted cost totals than no use. Subtracting the “yes” and “no” rows for each drug class measure in Table 6 gives the differentials. These differentials were the largest in the first year but generally stayed intact through year 3 for all three groups. The greatest differentials tended to occur in the other pain class, although the differentials for SSRI/SNRI/tricyclics and opioids were nearly equivalent.

Guideline Incongruence

The analyses of guideline aspects were conducted on the LBP/neuro and LBP/nonneuro groups only (*n* = 10,406) because the guidelines dealt with care for LBP specifically. The 11 measures selected for assessment are given in Appendix 1 (see Supplemental Digital Content, available at <http://links.lww.com/JOM/A162>). Six focused on procedures—three on imaging, one on providers, and two on surgeries—whereas five focused on medications, with the guidelines for two procedures (MRI use and surgeries) distinguishing between LBP types. Each of these measures addressed a timing issue for the procedure or medication referenced, with incongruence scored high and congruence low. In each case, the intent was to set a relatively conservative criterion in terms, enabling it to be applied broadly with little fine-tuning.

Prevalence

The rates of guideline incongruence across both LBP groups proved substantial. Across each of the six procedure aspects, the rate of incongruent “hits” was significantly greater for low back/neuro cases than for low back/nonneuro cases (Table 7), especially with respect to imaging. At more than 24%, for example, the LBP/neuro group’s rate of receiving MRIs at or before the 28-day mark doubled LBP/nonneuro’s rate. Breakouts by episode duration showed differentials in the same direction and of similar magnitude in this pattern vis-à-vis acute, subacute, and chronic cases across both LBP episode types (Table 7).

Comparable differences between the two types—all in the 40% to 60% range and statistically significant—were found on the “PT/Chiro visits within 2 weeks” and surgery measures. Splits by duration found that 13.3% of chronic LBP/neuro episodes reported surgery on or before the 365-day mark compared with 6.6% for chronic LBP/nonneuro cases. Of note, these latter cases were incongruent with the surgical guideline.

Rates of incongruent medication use, in contrast, were more similar across the two LBP types. LBP/neuro’s rate of incongruent Rx filling was modestly significantly greater for opioids, NSAIDs, and other pain medications, whereas the LBP/nonneuro’s rate on muscle relaxants was slightly greater. All group differences between the two types did not exceed 3%, with one exception—a 5% difference in the rate of incongruent opioid use among subacute cases. Roughly,

TABLE 6. Total Costs* by Episode Type and Year and Initial Treatment Pattern and Rx Use

Description	Low Back/Neuro			Low Back/Nonneuro			Other Back		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Initial treatment									
TalkInfo	4,661.05†	3,612.51	2,789.85	3,582.36	2,576.95	2,413.77	3,922.35	3,050.98	1,909.62
Dabble	6,327.86	3,604.40	2,296.41	4,696.64	3,357.36	2,808.93	3,159.15	1,956.33	2,202.07
Complex MM	18,793.74	4,887.91	4,549.85	21,221.16	3,915.13	4,207.96	19,804.66	3,971.10	3,804.63
Chiro	2,870.46	2,420.05	1,693.31	2,910.39	1,594.73	2,263.31	2,314.32	1,225.93	1,387.41
PT	10,156.89	3,527.13	3,509.90	7,360.73	2,984.93	3,103.16	5,521.81	1,838.13	2,001.92
Initial Rx use‡									
Opioids									
Yes	9,360.56	5,448.75	3,631.53	6,040.53	3,544.50	3,182.84	5,181.77	2,378.68	2,959.47
No	4,580.21	2,364.12	2,224.52	3,492.71	2,259.47	2,264.77	3,554.39	2,543.46	1,670.05
NSAIDs									
Yes	7,671.04	4,544.36	3,391.88	5,120.87	3,143.77	3,056.04	3,912.45	2,516.82	2,280.35
No	5,765.56	3,050.49	2,460.77	4,010.50	2,481.04	2,345.36	3,886.22	2,507.69	1,821.99
Other pain medications									
Yes	13,919.22	6,923.55	5,071.67	9,166.98	4,094.73	5,067.42	8,261.26	2,924.91	2,625.71
No	5,933.48	3,308.97	2,612.83	4,226.26	2,644.40	2,508.43	3,793.54	2,500.19	1,913.20
Muscle relaxants									
Yes	8,226.57	3,825.14	2,160.82	4,417.46	2,422.39	2,627.76	4,778.02	2,744.14	3,062.73
No	5,790.59	3,368.14	2,860.44	4,276.02	2,763.65	2,527.35	3,732.33	2,468.11	1,729.37
SSRI/SNRI/tricyclics									
Yes	10,703.09	5,142.65	3,595.49	6,903.82	5,294.45	3,562.84	5,581.60	2,834.20	3,048.42
No	5,820.38	3,287.96	2,633.91	4,070.10	2,418.74	2,460.58	3,736.52	2,479.53	1,825.10
Oral steroids									
Yes	7,420.03	2,921.44	1,774.45	6,469.61	4,136.35	3,754.80	3,622.40	2,430.85	3,852.88
No	6,178.07	3,497.70	2,804.68	4,175.65	2,572.94	2,472.40	3,908.97	2,514.57	1,808.88
Anxiolytics/sedatives/hypnotics									
Yes	6,040.33	4,186.41	10,454.28	8,076.12	4,353.25	4,534.41	5,290.96	2,099.24	3,849.90
No	6,267.05	3,435.64	2,501.74	4,199.74	2,621.99	2,498.10	3,852.74	2,521.70	1,880.63

*Total costs = direct costs (medical/Rx) + indirect costs (workers' compensation/disability/absenteeism).

†Entries are unadjusted and expressed in 2009 dollars.

‡Prescriptions filled within 4 weeks of episode start date.

NSAIDs, nonsteroidal anti-inflammatory drugs; PT, physical therapy; SNRI, serotonin-norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor.

a fifth of opioid and NSAID use continued beyond the 84-day mark across both types.

Initial Treatment Pattern Effects

Across the five patterns, the greatest rates of incongruence were posted by the Complex MM approach (Table 8). Notwithstanding the small number of cases on which these rates for the Complex MM approach were calculated, the consistency of their greater magnitude relative to that for the other approaches was striking.

The 70% of Complex MM cases reporting their first x-ray on or before the 28-day mark and the nearly 52% reporting their first MRI or CT on or before the 28-day mark were both a highly significant 19+ percentage points greater than the next highest approach, PT. Checks across type and duration of LBP episode replicated the direction and size of these differentials. Similarly, the 58% of Complex MM cases reporting their first MRI or CT on or before the 84-day mark were a highly significant 22 percentage points greater than PT. This same difference held when the test included only LBP/nonneuro cases, for whom the 84-day period was specifically set.

Especially striking were the surgical results, where the Complex MM's rates of guideline incongruence far exceeded the other approaches (Table 8). Of note, 43% of LBP/nonneuro cases linked to

Complex MM reported major back surgeries on or before the 365th day, the waiting period set for this type of back pain. Similarly, almost 24% of eligible LBP/neuro cases reported major back surgeries on or before the 42nd day, the waiting period set for this other LBP type. By contrast, the surgical rates for the other approaches were in the low single digits.

The Complex MM approach also recorded the greatest incongruent rates in the medication area. More than 58% of Complex MM cases reported a timeline-excessive course of therapy on at least one of the five measures, with opioids and NSAIDs being the two classes with the greatest rates. This compared to more than 48% of cases reporting at least one timeline-excessive course of therapy for the PT approach, which otherwise reported a similar medication profile.

The one guideline aspect where the Complex MM approach posted the lowest score was the rate of Chiro/PT visits—the measure on which Chiro and PT posted the highest rates. More than 95% of all Chiro cases and nearly half of all PT cases reported their first such visit within 14 days of the start date. Equally striking were the results at the other end of the spectrum. TalkInfo and Complex MM approaches, and to a lesser extent the Dabble approach, all tended to record lower rates.

TABLE 7. Guideline Aspect Incongruence by LBP Episode Type and Duration^a

LBP Guideline Aspect	Low Back/Neuro				Low Back/Nonneuro				t ^b
	Acute	Subacute	Chronic	All	Acute	Subacute	Chronic	All	
Procedure									
Had first x-ray within 28 days of episode start	32.8 ^c	56.7	43.7	41.2	30.2	47.2	36.9	34.0	5.9***
Had first MRI/CT within 28 days of episode start	14.6	36.6	29.8	24.3	7.5	21.0	17.3	11.7	14.2***
Had first MRI/CT within 84 days of episode start	14.7	45.9	38.4	29.4	7.4	30.5	23.9	14.7	15.1***
Had first Chiro/PT session within 14 days of episode start	27.1	37.8	33.5	31.5	11.9	28.3	27.2	17.9	13.2***
Had major back surgery within 42 days of episode start	1.4	1.8	2.3	1.8	0.2	0.8	0.9	0.5	6.3***
Had major back surgery within 365 days of episode start	1.7	3.7	13.3	6.6	0.3	2.3	6.6	2.3	9.0***
Medication									
Filled opioids Rx before and after 84-day mark	14.3	18.3	38.6	24.2	15.6	23.3	36.0	22.0	2.1*
Filled NSAIDs Rx before and after 84-day mark	15.2	17.7	25.9	19.7	13.4	17.0	25.7	16.9	2.9**
Filled other pain Rx before and after 84-day mark	1.7	3.4	6.5	3.8	1.2	2.4	4.0	2.1	4.4***
Filled muscle relaxant Rx before and after 28-day mark	3.8	6.7	13.8	8.1	6.1	9.3	17.0	9.4	1.7
Filled benzodiazepine Rx before and after 28-day mark	3.2	4.6	6.6	4.7	3.4	4.6	6.9	4.5	0.50

*P < 0.05; **P < 0.01; ***P < 0.001.

^aLow back/neuro and low back/nonneuro groups only.

^bt values for comparison of Low Back/Neuro "All" vs. Low Back/Nonneuro "All."

^cUnless otherwise indicated, entries are percentage "yes."

CT, computed tomography; LBP, lower back pain; MRI, magnetic resonance imaging; NSAIDs, nonsteroidal anti-inflammatory drugs; PT, physical therapy.

TABLE 8. Guideline Aspect Incongruence by Initial Treatment Pattern^a

LBP Guideline Aspect	Initial Treatment Pattern					F ^b
	TalkInfo	Dabble	Complex MM	Chiro	PT	
Procedure						
Had first x-ray within 28 days of episode start	28.9 ^c	38.2	70.0	43.1	50.8	96.2***
Had first MRI/CT within 28 days of episode start	10.4	17.2	52.3	5.3	28.9	162.1***
<i>LBP/neuro only</i>	22.6	27.3	54.3	5.3	37.8	39.2***
Had first MRI/CT within 84 days of episode start	13.1	20.4	57.8	8.3	35.4	169.7***
<i>LBP/nonneuro only</i>	11.1	17.7	54.7	7.9	32.5	117.7***
Had first Chiro/PT session within 14 days of episode start	0	27.1	3.2	95.2	49.1	3,623.4***
Had major back surgery within 42 days of episode start	0	0.5	26.7	0.1	1.1	573.6***
<i>LBP/neuro only</i>	0	0.9	23.9	0.3	2.4	79.9***
Had major back surgery within 365 days of episode start	1.6	3.1	39.1	2.1	5.9	220.0***
<i>LBP/nonneuro only</i>	1.3	2.2	43.0	2.0	4.5	182.2***
Had 1+ guideline-incongruent imaging or surgical procedure	35.6	46.4	86.1	46.2	65.2	142.1***
Medication						
Filled opioids Rx before and after 84-day mark	20.8	25.7	44.4	12.4	30.8	47.0***
Filled NSAIDs Rx before and after 84-day mark	17.5	16.7	26.7	8.6	24.1	27.5***
Filled other pain Rx before and after 84-day mark	2.1	2.2	8.6	1.0	4.4	15.9***
Filled muscle relaxant Rx before and after 28-day mark	8.9	10.4	16.0	4.8	11.5	11.8***
Filled benzodiazepine Rx before and after 28-day mark	4.2	4.5	7.0	3.7	6.3	3.7**
Had 1+ guideline-incongruent Rx regimen	36.9	38.5	58.8	21.9	48.9	55.7***

*P < 0.05; **P < 0.01; ***P < 0.001.

^aLow back/neuro and low back/nonneuro groups only.

^bF statistics for analysis of variance of row entries.

^cUnless otherwise indicated, entries are percentage "yes."

CT, computed tomography; LBP, lower back pain; MRI, magnetic resonance imaging; NSAIDs, nonsteroidal anti-inflammatory drugs; PT, physical therapy.

Cost Effects

Similarly, with one exception, higher rates of guideline incongruence were linked to greater costs (Table 9). For both LBP groups, considerable increments differentiated the average total costs of episodes reporting incongruence with one or more of the 11 guideline aspects relative to those episodes reporting no incongruence at

year 1. In years 2 and 3, these increments tapered off somewhat but remained substantial.

Incongruence on the major surgery measures led the way in generating these cost differentials for both LBP groups, although both imaging measures produced notable differentials in the same direction (Table 9). Incongruence on the medication measures

TABLE 9. Total Cost Differentials by Guideline Aspect Incongruence, Episode Type, and Year^a

LBP Guideline Aspect	Low Back/Neuro			Low Back/Nonneuro		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Had first x-ray within 28 days of episode start	2,221.44 ^b	− 718.22	328.95	1,771.80	− 351.79	− 62.09
Had first MRI/CT within 28 days of episode start	5,199.54***	− 157.57	964.52	5,588.01***	847.12	1,192.05**
Had first MRI/CT within 84 days of episode start	5,535.49***	− 398.31	710.74	6,061.87***	940.97*	1,185.89**
Had first Chiro/PT session within 14 days of episode start	− 2,346.03	− 1,825.18	− 1,133.63	− 1,116.55	− 1,135.09	− 173.20
Had major back surgery within 42 days of episode start	22,396.26***	506.88	− 143.50	30,622.22***	719.09	− 820.19
Had major back surgery within 365 days of episode start	28,152.95	3,760.89	227.14	31,905.48	3,630.67	913.41
Filled opioids Rx before and after 84-day mark	8,656.80***	4,450.23***	3,023.08***	5,178.27***	2,112.94***	1,773.70***
Filled NSAIDs Rx before and after 84-day mark	3,861.33***	3,236.18***	1,140.37	3,115.84***	2,219.71***	1,571.15***
Filled other pain Rx before and after 84-day mark	5,624.52**	5,381.85**	3,778.32**	6,715.80***	2,460.78*	3,039.88**
Filled muscle relaxant Rx before and after 28-day mark	9,396.56***	2,409.22	605.29	3,650.72***	1,281.89**	1,539.49***
Filled benzodiazepine Rx before and after 28-day mark	4,646.43**	4,930.63**	1,166.54	4,265.70***	3,986.92***	1,420.89*
Had 1+ guideline-incongruent imaging or surgical procedure OR Rx regimen	5,930.44***	1,864.12*	1,503.10*	3,972.91***	1,731.19***	1,157.82***

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

^aLow back/neuro and low back/nonneuro groups only.

^bCell entries are unadjusted average cost differences for episodes with aspect minus episode without aspect, expressed in 2009 dollars.

CT, computed tomography; LBP, lower back pain; MRI, magnetic resonance imaging; NSAIDs, nonsteroidal anti-inflammatory drugs; PT, physical therapy.

reported similar cost differentials, consistently at levels that proved even more sustained over time. The lone exception occurred on the “PT/Chiro within 2 weeks” measure, where episodes that reported the first visit within the 2-week mark posted, on average, notably less costs than those episodes that refrained from having the first visit within the 2-week mark. On this measure as well, comparable differentials were recorded for both types of episodes across all 3 years.

Hypothesis Tests

The propensity-derived adjustments for hypothesis testing were framed to yield two types of comparisons (Table 10). For initial treatment pattern predictions, each pattern other than TalkInfo was assessed relative to the TalkInfo (the holdout group). For guideline aspect predictions, the comparisons assessed incongruent relative to congruent episodes.

Considered altogether, the adjusted results were highly similar to the unadjusted results for both sets of predictors (Table 6). Comparisons with the previously described corresponding results indicated that the adjustments for age, sex, episode type and start date year, claims lines, comorbidities, prior episode, and left-censored episode status via the propensity techniques made little difference in the direction and magnitude of the estimates relative to the unadjusted totals.

Of the initial treatment patterns, the total costs linked to the Complex MM approach far and away outstripped the total costs for the other approaches, particularly in year 1 but also in years 2 and 3 (Table 10). In year 1, the total costs of the average Complex MM approach episode exceeded the average TalkInfo episode by

just more than \$14,000 for LBP/neuro cases and more than \$17,000 for LBP/nonneuro cases. The corresponding total costs for the next-ranked approach, PT, were approximately \$5000 and \$3600, respectively. In year 3, the average Complex MM episode still incurred the highest total cost, exceeding the average TalkInfo episode by nearly \$2000 for LBP/neuro and nearly \$1800 for LBP/nonneuro. Physical therapy again ranked second for both types.

At the other end of the cost continuum, the Chiro approach alone averaged lower total costs than the TalkInfo approach for both types of LBP across all 3 years. Deficits for Chiro relative to the TalkInfo approach ranged from just more than \$1700 in year 1 to just more than \$1300 in year 3 per episode for LBP/neuro and from \$600 to nearly \$200 per episode of LBP/nonneuro cases. The Dabble approach posted totals in the mid-ranges across all 3 years for both types.

The adjusted cost totals for the guideline comparisons confirmed one set of clues to these differences (Table 10)—the considerable costs linked to incongruence on one or more of the 10 imaging, surgical, and medication guideline aspects selected for this analysis. The year 1 average total for LBP/neuro episodes was nearly \$5500 more per episode for when they reported incongruence with one or more of the 10 imaging, surgical, or medication aspects than those LBP/neuro episodes reporting no such incongruence. This differential remained intact with episodes incongruent on 1+ aspects averaging approximately \$1600 more in years 2 and 3. The corresponding totals for LBP/nonneuro episodes were just more than \$3200, \$1200, and \$920 per incongruent episode in years 1, 2, and 3, respectively. The incongruent use of surgeries, imaging, and medications

TABLE 10. Total Cost Differentials by Initial Treatment Pattern, Guideline Aspect Incongruence, Episode Type, and Year With Propensity Adjustments^a

	Low Back/Neuro			Low Back/Non-Neuro		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Initial treatment pattern						
Dabble	1,499.87 ^{*b}	11.62	− 521.30	1,019.74 ^{***}	597.42 ^{***}	345.91
Complex MM	14,070.46 ^{***b}	1,380.63	1,957.92	17,291.43 ^{***}	1,110.73 ^{**}	1,798.86
Chiro	− 1,745.82 ^{*b}	− 1,117.59	− 1,328.16	− 641.32	− 1,033.81 ^{**}	− 191.69
PT	5,058.98 ^{***b}	− 239.90	788.21	3,638.84 ^{***}	292.55 [*]	648.21
LBP guideline aspect						
Had first x-ray within 28 days of episode start	2,236.54 ^{***c}	− 725.44	336.44	1,823.07 ^{***}	− 189.67 ^{***}	18.83 [*]
Had first MRI/CT within 28 days of episode start	5,045.12 ^{***c}	− 233.50	990.05	5,525.16 ^{***}	896.14	1,296.43 ^{***}
Had first MRI/CT within 84 days of episode start	5,345.62 ^{***c}	− 509.56	730.79	5,980.82 ^{***}	944.05	1,260.81 ^{**}
Had First Chiro/PT Session within 14 days of episode start	− 2,205.94 ^{***c}	− 1,793.78	− 1,169.52	− 1,057.67 ^{**}	− 1,165.81 ^{**}	− 217.75
Had major back surgery within 42 days of episode start	22,678.16 ^{***c}	598.80	− 167.17	30,770.33 ^{***}	973.22 ^{***}	− 539.58 ^{***}
Had major back surgery within 365 days of episode start	28,077.78 ^{***c}	3,766.25 [*]	219.19	31,802.48 ^{***}	3,633.80 ^{**}	1,099.14 ^{**}
Filled opioids Rx before and after 84-day mark	7,822.78 ^{***c}	4,250.08 ^{***}	3,225.00 ^{***}	4,477.50 ^{***}	1,598.16 ^{***}	1,666.96 ^{***}
Filled NSAIDs Rx before and after 84-day mark	3,784.56 ^{***c}	3,366.08 ^{***}	1,701.83 [*]	2,724.26 ^{***}	2,190.90 ^{***}	1,649.15 ^{***}
Filled other pain Rx before and after 84-day mark	4,746.85 ^{***c}	5,324.70 [*]	4,476.35 ^{**}	6,762.18 ^{***}	1,472.72	2,312.24 [*]
Filled muscle relaxant Rx before and after 28-day mark	2,450.18 ^{***c}	712.17	− 819.79	831.81 ^{***}	− 134.21	387.98
Filled benzodiazepine Rx before and after 28-day mark	5,086.42 ^{***c}	2,936.96 [*]	821.44	3,001.57 ^{***}	2,190.26 ^{***}	729.69 ^{**}
Had 1+ guideline-incongruent imaging or surgical procedure OR Rx regimen	5,478.46 ^{***c}	1,586.99	1,601.88 [*]	3,220.07 ^{***}	1,223.86 ^{**}	928.55 ^{**}

P* < 0.05; *P* < 0.01; ****P* < 0.001.

^aLow back/neuro and low back/nonneuro groups only.

^bRow entries are average cost differences for group minus the TalkInfo group, expressed in 2009 dollar and adjusted for age, sex, LBP episode type and start date year, comorbidities, count of claims lines, had prior episode, and left-censored episode status.

^cRow entries are average cost differences for episodes with aspect minus episodes without aspect, expressed in 2009 dollar and adjusted for age, sex, LBP episode type and start date year, comorbidities, count of claims lines, had prior episode, and left-censored episode status.

CT, computed tomography.

each contributed significantly to these totals in the first year across both types of episodes and in many cases across years 2 and 3 as well. The “PT/Chiro within 2 weeks” aspect again proved to be the lone exception, with differentials linking strikingly lower total costs to incongruent use being recorded on both LBP/neuro and LBP/nonneuro episodes across all 3 years. Rerunning these analyses to test 2 refinements—1) examining first episodes only and 2) lengthening the “before” and “after” intervals from 6 to 12 months—did not alter these conclusions (results available on request).

The other set of clues for the initial treatment pattern cost differences came from the comparisons above linking initial treatment patterns with guideline aspect incongruence rates (Table 8). As noted previously, the Complex MM approach was the most likely to record incongruent imaging, surgery and medication episodes, followed by the PT approach. Across guideline aspects, the lone exception came on the “Chiro/PT visits within 2 weeks” measure, which had the Chiro approach and then the PT approach recording the highest rates.

Combined, these results strongly affirmed both hypotheses. Not only did incongruence on 10 of 11 guideline aspects each show appreciable levels of prevalence and strikingly positive, substantial and lingering relationships with total costs, as per hypothesis #1 (summarized in Fig. 1); the treatment approaches most likely to exhibit incongruent use of imaging, surgery, and medications—Complex MM, then PT—were, in turn, associated with the greatest total costs per hypothesis #2. As summarized in Figs. 2 and 3, these treatment approach effects stayed intact across breakouts for acute, subacute, and chronic episodes. The lack of sensitivity to propensity controls underscored the robustness of these findings.

DISCUSSION

This study—in effect, an exploration of the actuarial consequences of clinical decision making—suggests several conclusions for those proactive employers and occupational health personnel that are turning to LBP to improve its care:

1. Back pain is costly, regardless of the type of back pain EEs have or whether they seek treatment for it. It matters little whether the

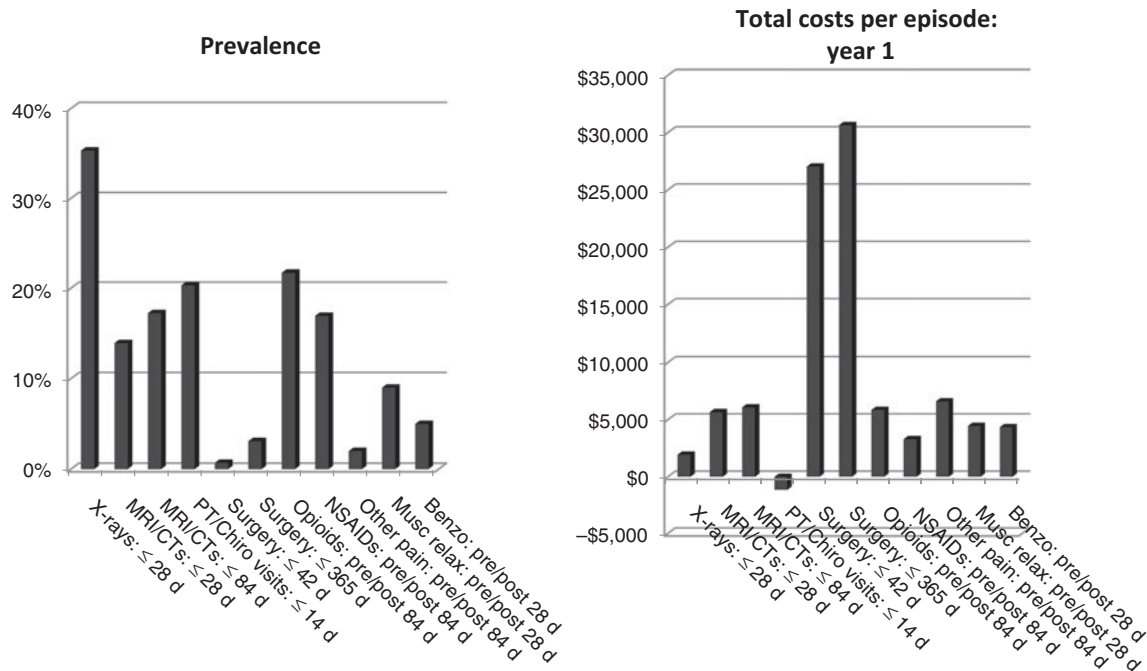


FIGURE 1. Guideline aspect incongruence: prevalence and total costs per episode at year 1.

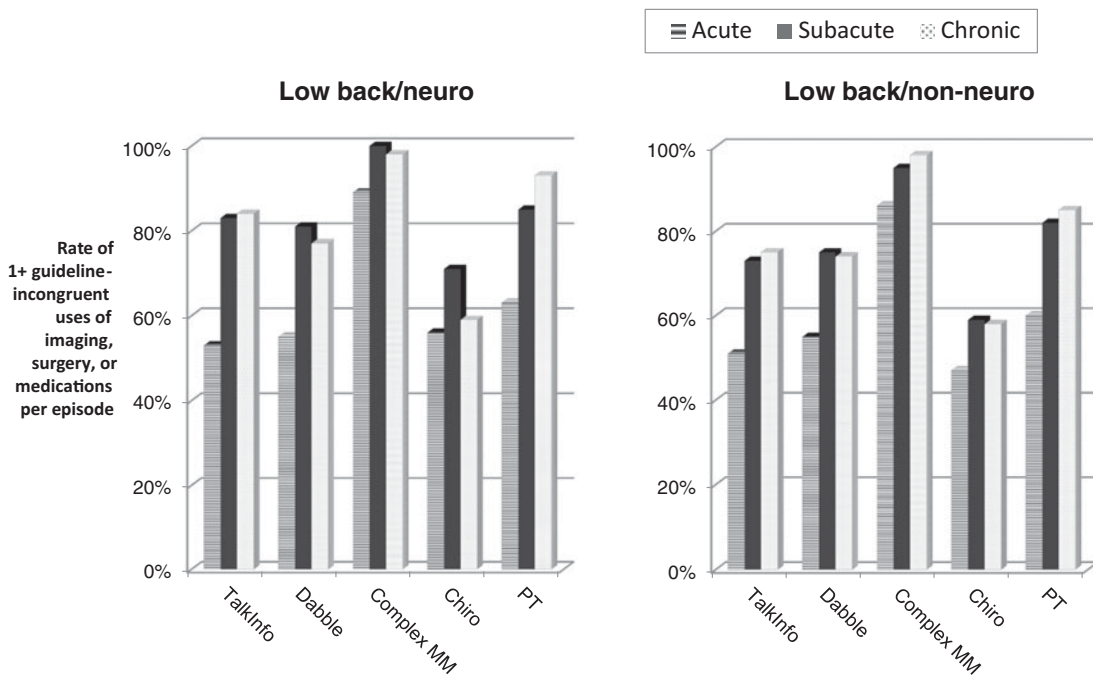


FIGURE 2. Initial treatment pattern by episode type, duration, and guideline-incongruent use of imaging, surgeries, and medications.

problem with which EEs present is in the low back or cervical or thoracic spine or of neurological or nonneurological LBP origin. And, it matters little whether their care starts out by simply seeking information and advice, or by launching into a sustained course of treatment from physicians, chiropractors, or physical therapists, or by exploring two or more of these types of providers. On average, their direct health care and indirect lost productivity costs will total in excess of \$4000 by the end of the episode's first

year and continue—albeit at levels that taper off—across years 2 and 3.

2. The initial treatment choices that back pain EEs/patients make in concert with their providers matter. Previous treatment choices they have made will influence their current treatment choice, above and beyond the type of back pain diagnosis they receive or the length of the episode that ensues. These choices, in turn, can have a major differential effect on the total cost of their episode,

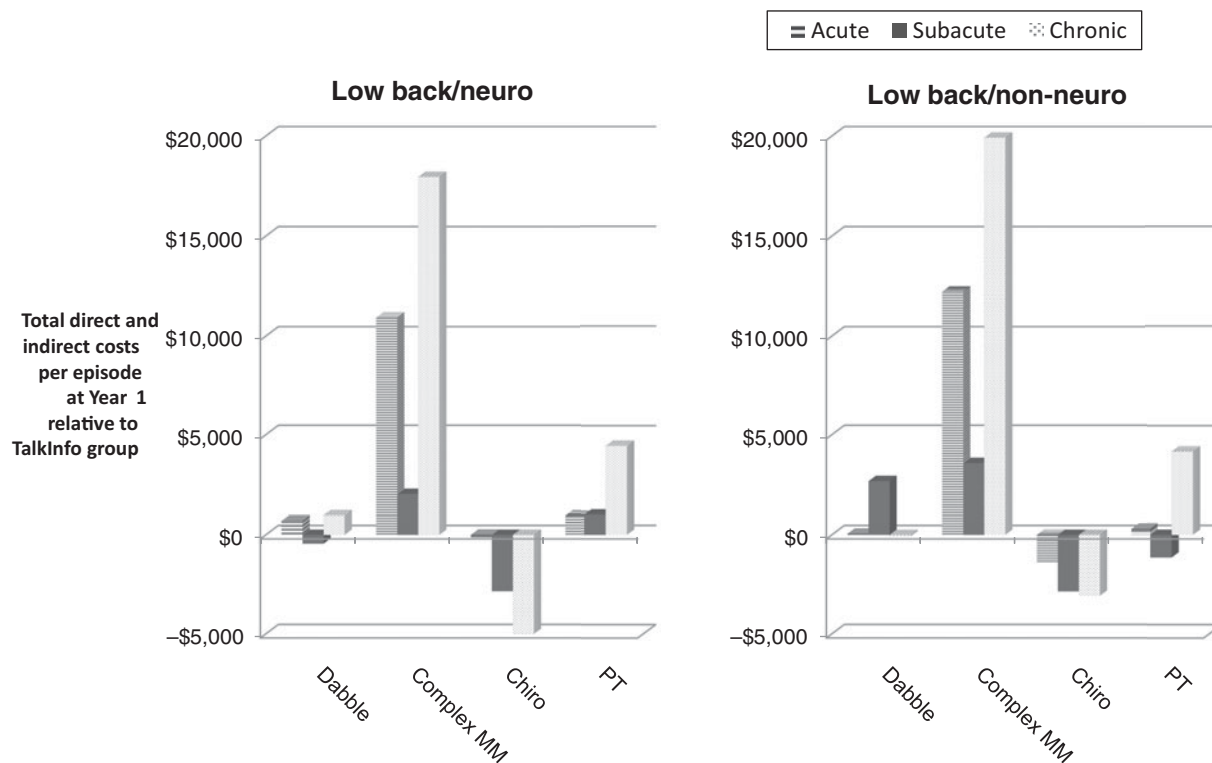


FIGURE 3. Initial treatment patterns by episode type, duration, and total costs at year 1. All cost entries computed relative to the TalkInfo Initial Treatment group.

particularly when these treatment choices are not congruent with certain timing-related aspects of current guidelines for the use of imaging, surgeries, and medications. This finding aligns with the Choosing Wisely initiative of the ABIM Foundation, which encourages physicians, patients, and other health care stakeholders to discuss and utilize the available evidence when determining an individual patient’s treatment choices.⁴⁹

- Of the initial treatment patterns, the Complex MM approach will, on average, be associated with the greatest total cost—not only in the first year of episodes but also in years 2 and 3—in significant part because the providers practicing the approach are more likely to deploy the expensive medical options of the clinical repertoire in ways that are guideline-incongruent. Those EEs who access the PT approach will likewise be associated with higher costs if they persist with treatment although a first visit to a PT within the initial 14 days may have cost-reducing effects. The EEs accessing the Chiro approach will tend to be the least expensive because they are less likely to be prescribed medications or end up with complex medical procedures and because they are less likely to record guideline-incongruent use of imaging, procedures, and medications when the latter are delivered.
- Of all treatment components for LBP, surgery exerts the most potent cost effects. These costs not only are concentrated in year 1 but also can linger into years 2 and 3. Surgeries stand out, especially in this regard when conducted prior to the end of the waiting periods recommended by the pertinent guidelines. Of all the procedures assessed in relation to guideline aspects pertaining to LBP/neuro episodes that were examined in this study, surgeries prior to the 6-week mark accumulated the greatest total cost, averaging \$33,519 across years 1, 2, and 3 combined. Similarly, of all the procedures examined in relation to guidelines pertaining to

LBP/nonneuro episodes, surgeries on or before 365 days accumulated the greatest total cost, averaging \$41,580 across the 3 years.

- Individuals-prescribed pain-related medications are more costly, and these expenses can linger into the future if prescription filling is not congruent recommended guidelines for short-term use. Although much has been written recently about opioids in this regard, this pattern holds for all five Rx classes assessed in relation to guidelines in this study. The potential for improved management of this pattern and its costs holds for all three types of back pain examined here.
- Although guidelines are structured to reflect the latest thinking and evidence on which processes are most efficient and effective at delivering on average the best patient outcomes at the lowest cost, it pays to subject their implementation to empirical scrutiny. Not all the expectations generated may prove accurate. In 1 of the 11 instances examined here, where the recommendation is that the first visit to PT or chiropractor should be made at least 2 weeks after the episode start date, incongruent care actually led to significantly less total cost. Although this result may not necessarily warrant a revisiting of this particular guideline, it merits the attention of all stakeholder groups with an interest in optimizing its usefulness.

Study Limitations

The process used to generate the evidence for these conclusions requires some caveats. Consider, for example, the breakouts by episode duration reported in Table 6, which found notable percentages of guideline-incongruent medication use reported beyond the timeframes specified for acute and subacute episodes. For instance, opioid Rx filling both before and after the 84-day mark was linked to just more than 14% of acute LBP/neuro cases, which by

definition last up to 28 days, and just more than 23% of subacute LBP/nonneuro, which by definition last up 84 days.

Such results point to the need to strengthen the longitudinal paradigm used for this study. Take the issue of the same individual incurring more than one episode of back pain. While bringing simplicity to the current work that helped reduce the complexity of the results, the approach used here to treat episodes as independent units of analysis did not delve into the effects associated with serial episodes by the same patient. Or, take the classification of treatment strategies based on just the first 6 weeks of claims activity alone. While also helping reduce the complexity of results, this second feature precluded delving into what happens when initial diagnoses are subsequently modified for the same patient either within or across episodes. Or, take the absence of any controls for changes in clinical context over time—new procedures, new drugs (and marketing), new modalities, etc. Although the variations in rates of use from the first to the second year of episodes reported in Table 2 hint at these contextual changes, they do not reflect the evolution in practice patterns and modalities that other exploratory analyses indicate occurred across the 9-year study period (results available on request). Any of these adjustments in longitudinal approach would likely shed new light on the results reported here.

A second caveat stems from a timing issue. The bulk of the work that led to release of the guidelines selected for this study occurred during the 2007 to 2012 period, whereas the claims activity examined in this study spanned 2001 to 2009. For many of the episodes selected for this analysis, it is reasonable as such to assume that the providers involved either were not aware of the guidelines or probably had no possibility of even knowing that they were under development. For this reason, the characterizations of congruence deployed here have been carefully crafted to refrain from the use of any standard harboring the explicit expectation that the guideline aspects should somehow have been “followed” per se. Future work is needed that establishes context where provider behavior can be fairly assessed as such.

A third caveat concerns the sorting of EEs/patients into different treatment strategies and time frames for the delivery and filling of procedures and medications. It is likely that this sorting is not random. Those with more severe back pain could self-select into more medically intensive strategies. More severe back pain can also have prompt earlier imaging and surgeries or other guideline-incongruent choices. The propensity adjustments used in this study each provided statistical control for an extraneous factor arguably directly or indirectly related to disease severity. Yet, in the tests here self-reported symptom severity was not one such factor. As shown elsewhere,¹⁶ future work would benefit from inclusion of such self-reports as a control.

A fourth concern focuses on the likely underestimates of the true health burden of back pain that stem from the use of the study's total cost measure. This measure did not include presenteeism or job performance impaired by back pain. Much previous work has found that presenteeism is, in fact, the largest component of total health burden, not only on the indirect side but also on the direct side.^{15,50,51} The mix of complexities due to data availability and analytic requirements that led to presenteeism not being included here is a priority issue for future research in this area.

A final concern is less a limitation and more a direction for future research. The comparisons underlying the results reported here pitted episodes in which care was congruent with selected aspects of guidelines versus episodes in which care was not congruent. The comparisons did not take the additional step of directly pitting incongruent episodes versus congruent episodes per se. The trade-off rendered by this decision meant that analytic complexities deemed better left for future exploration in this line of research could be bypassed, but at the expense of curtailing the capacity to report not just losses averted when care was not congruent but also savings achieved

when care was congruent. Future work making this distinction will provide a fuller picture of guideline impact.

Implications

This said, this study not only replicated the considerable burden of back pain in this workforce in a framework conducive to upgrading the management of care for LBP but also juxtaposed current LBP care with aspects of guidelines for this care and found that care not congruent was associated with much added cost burden. The evidence here would seem to merit a new outreach to company health plans and providers. The focus of this outreach would be the applicability of these findings for improving the health care and productivity loss outcomes of LBP EEs/patients.

To be effective, the process by which guidelines are implemented is complex and requires the input of many parties.⁵² The employer perspective merits being considered as one such input. Echoing what has been observed elsewhere in contexts for guideline development and adherence (eg, the National Committee for Quality Assurance's HEDIS rules for LBP imaging for health plan performance assessment⁵³ and the Choosing Wisely initiative⁴⁹), noteworthy rates of x-rays, MRIs/CTs, surgeries, and medications are being provided for both types of LBP during timeframes that are guideline-incongruent. Much added questionable expense to the employer and ultimately its EEs seems to be occurring as a result. In each instance, the implication is that closer adherence to the guideline will lead to considerable cost savings. On the contrary, incongruent behavior relative to the “first PT/Chiropractor visit within 2 weeks” guideline is also occurring at an appreciable rate but seems instead to be a significant deterrent for long-term total costs. All else held equal, it seems that the less that care is congruent with this latter guideline, the greater the cost savings that will result.

These considerations are not necessarily more important than the improved clinical efficiencies, coordination of care, or other non-fiscal factors that might result from the incorporation of these guidelines. But, the priorities that employers as health care purchasers are increasingly placing on long-term H&P outcomes are arguably no less substantive or relevant than the priorities of providers, patients, health plans, and other affected stakeholders. Evidence like that produced here, which bears on any group's capacity to manage its value/sustainability challenge in the marketplace, belongs in the discussion.

Coaxing Change in Context

Efforts to promote guideline-induced change to incorporate these implications do not, of course, take place in a vacuum. Any such effort in fact would be well served by taking into account the larger context where possible. At the employer for this study, the aggregate-level reductions in total costs per active employee from 2001–2002 to 2008–2009 noted previously encompassed widespread drops across healthy and disease groups. One such area was back problems, for which a drop of just more than \$3500 per EE/patient (excluding presenteeism and absenteeism) was recorded.

Among the broad array of disease management, benefit, and prevention initiatives that contributed to these cost reductions for back problems specifically was a staged musculoskeletal educational intervention.⁵⁴ During 2002 to 2004, the company targeted physicians providing musculoskeletal care at a major company facility with this program. Its features included the assignment of a carefully selected physician/nurse team on-site to each new disability case, usually within 1 to 3 days of the injury. Tests of program effect at the intervention site found that mean days lost per work-related injury dropped from 35.1 to 27.6 and that mean annual indemnity per work-related injury decreased from \$9327 to \$4493. During the next 5 years, much effort was made to broaden the program to other sites throughout the company. The sharp drops in disability incidents

and costs from 2001–2002 to 2008–2009 subsequently found for the aggregate workforce across the company's US sites¹⁵ reflect its apparent success in this regard.

Yet, systematic efforts to implement and refine guidelines for LBP care were not a particular focus for this educational intervention or, for that matter, the broad array of management, benefit, and preventive initiatives undertaken by the company from 2001 through 2009. A trend that would seem to register this lack of focus was the little movement shown by the incongruent “hit” rates on the guideline aspects. From the 18-month period from July 1, 2001, to December 12, 2002, to the 18-month period beginning January 1, 2007 (both selected to maximize the observations for pre-post comparisons in this data set), the only aspect to post a significant drop in incongruent rate was the time-excessive NSAID Rx measure (for LBP/neuro: 28.0% [2001–2002] vs 13.6% [2008–2009]; $t = 4.1$, $P = 0.00$; for LBP/nonneuro: 23.0% [2001–2002] vs 12.4% [2008–2009]; $t = 6.9$, $P = 0.00$). All of the 16 other one-sided t tests for reductions across the set of guidelines applicable to each LBP group found either no change or a modest increase of three percentage points or less. Much potential for promoting well-informed guideline adherence would seem possible in this setting.

CONCLUSIONS

Although employers and their occupational health personnel are not necessarily looking to test the accuracy or appropriateness of guidelines for LBP care per se, they are seeking to better manage the long-term total cost outcomes of their LBP EE/patients. Do guidelines hold much promise for advancing this objective? The bottom line answer from this study is an empirically based yes.

REFERENCES

- Deyo R, Tsui-Wu Y. Descriptive epidemiology of low-back pain and its related medical care in the United States. *Spine (Phila Pa 1976)*. 1987;12:264–268.
- Wheeler S, Wipf J, Staiger T, Deyo R. Approach to the diagnosis and evaluation of low back pain in adults. Available at: http://www.uptodate.com/contents/approach-to-the-diagnosis-and-evaluation-of-low-back-pain-in-adults?detectedLanguage=en&source=search_result&search=approach+to+the+diagnosis+and+evaluation+of+low+back+pain+in+adults&selectedTitle=1~150&provider=noProvider. Accessed October 2, 2013.
- Knight C, Deyo R, Staiger T, Wipf J. Treatment of acute low back pain. Available at: <http://www.uptodate.com/contents/treatment-of-acute-low-back-pain>. Accessed October 2, 2013.
- Allen H, Burton W. On value, sustainability and the rollout of health care reform: a call to the implementers. *Popul Health Manag*. 2013;16:288–290.
- Ivanova J, Birnbaum H, Schiller M, Kantor E, Johnstone BM, Swindle RW. Real-world practice patterns, health care utilization, and costs in patients with low back pain: the long road to guideline-congruent care. *Spine J*. 2011;11:622–632.
- Goetzl R, Hawkins K, Ozminkowski R, Wang S. The health and productivity burden of the “top ten” physical and mental health conditions affecting six large U.S. employers in 1999. *J Occup Environ Med*. 2003;45:5–14.
- Bunn W, Stave G, Allen H, Naim A. Evidence-based benefit design: toward a sustainable healthcare future for employers. *J Occup Environ Med*. 2010;52:951–955.
- Bunn W, Pikelny D, Paralkar S, Slavin T, Borden SIV, Allen HM Jr. The burden of allergies—and the capacity of medications to reduce this burden—in a heavy manufacturing environment. *J Occup Environ Med*. 2003;45:941–955.
- Allen H, Bunn W. Using self-report and adverse event measures to track health's impact on productivity in known groups. *J Occup Environ Med*. 2003;45:973–983.
- Allen H, Borden S IV, Pikelny D, Paralkar S, Slavin T, Bunn WB III. An intervention to promote appropriate management of allergies in a heavy manufacturing setting: an evaluation of health and productivity outcomes. *J Occup Environ Med*. 2003;45:956–972.
- Allen H, Bunn W. Validating self-reported measures of productivity at work: a case for their credibility in a heavy manufacturing setting. *J Occup Environ Med*. 2003;45:926–940.
- Allen H, Slavin T, Bunn W. Do long workhours impact health, safety and productivity at a heavy manufacturer? *J Occup Environ Med*. 2007;49:148–171.
- Allen H, Woock C, Barrington L, Bunn W. Age, overtime and employee health, safety and productivity outcomes: a case study. *J Occup Environ Med*. 2008;50:873–894.
- Bunn W, Allen H, Stave G, Naim A. How to align evidence-based benefit design with the employer bottom-line: a case study. *J Occup Environ Med*. 2010;52:956–963.
- Allen H, Rogers W, Bunn W, Pikelny D, Naim A. Reducing total health burden from 2001 to 2009: a major self-insured employer's counter-trend success story and implications for health care reform. *J Occup Environ Med*. 2012;54:904–916.
- Allen H, Bunn W, Naim A. The self-reported health and productivity burden of auto immune inflammatory disorders at a major self-insured employer. *J Occup Environ Med*. 2012;54:1049–1063.
- Allen H, Rogers W, Bunn W. Managing the burden of chronic obstructive pulmonary disease on workforce health and productivity: upping a leading employer's game. *J Occup Environ Med*. 2012;54:1064–1077.
- van Tulder M, Assendelft W, Koes B, Bouter L. Spinal radiographic findings and non-specific low back pain: a systematic review of observational studies. *Spine (Phila Pa 1976)*. 1997;22:427–434.
- Croft P, Macfarlane G, Papageorgiou A, Thomas E, Silman AJ. Outcome of low back pain in general practice: a prospective study. *BMJ*. 1998;316:1356–1359.
- Cherkin D, Deyo R, Street J, Barlow W. Predicting poor outcomes for back pain seen in primary care using patients' own criteria. *Spine (Phila Pa 1976)*. 1996;21:2900–2907.
- Webster B, Cifuentes M. Relationship of early magnetic resonance imaging for work-related acute low back pain with disability and medical utilization outcomes. *J Occup Environ Med*. 2010;52:900–907.
- Mehling W, Gopisetty V, Bartmess E, et al. The prognosis of acute back pain in primary care in the United States: a two-year prospective cohort study. *Spine (Phila Pa 1976)*. 2012;37:678–684.
- Carey TS, Freburger JK, Holmes GM, et al. A long way to go: practice patterns and evidence in chronic low back pain care. *Spine (Phila Pa 1976)*. 2009;34:718–724.
- Deyo R, Weinstein J. Low back pain. *N Engl J Med*. 2001;344:363–370.
- Webster B, Bauer A, Choi Y, Cifuentes M, Pransky G. Iatrogenic consequences of early magnetic resonance imaging in acute, work-related, disabling low back pain. *Spine (Phila Pa 1976)*. 2013;38:1939–1946.
- Chou R, Fu R, Carrino J, Deyo RA. Imaging strategies for low-back pain: systematic review and meta-analysis. *Lancet*. 2009;373:463–472.
- Cifuentes M, Willets J, Wasiak R. Health maintenance care in work-related low back pain and its association with disability recurrence. *J Occup Environ Med*. 2011;53:393–404.
- Wasiak R, Kim J, Pransky G. The association between timing and duration of chiropractic care in work-related low back pain and work-disability outcomes. *J Occup Environ Med*. 2007;49:1124–1134.
- Tao X, Lavin R, Yuspen L, Bernacki E. Natural history of opioid dosage escalation post-injury. *J Occup Environ Med*. 2012;54:439–444.
- Tao X, Lavin R, Yuspen L, Bernacki E. Impact of combined use of opioids and surgical procedures on WC cost among a cohort of injured workers in the state of Louisiana. *J Occup Environ Med*. 2012;54:1513–1519.
- Cifuentes M, Powell R, Webster B. Shorter time between opioid prescriptions associated with reduced work disability among acute low back pain opioid users. *J Occup Environ Med*. 2012;54:491–496.
- Chou R, Qaseem A, Snow V, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann Intern Med*. 2007;147:478–491.
- van Tulder MW, Becker A, Bekkering T, et al. European guidelines for the management of acute low back pain in primary care. *Eur Spine J*. 2006;15(suppl 2):S169–S191.
- Koes BW, van Tulder M, Lin CW, Macedo LG, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J*. 2010;19:2075–2094.
- Lin C, Maher C, Machado L, van Tulder M. Cost-effectiveness of guideline-endorsed treatments for low back pain: a systematic review. *Eur Spine J*. 2011;20:1024–1038.
- Pillastrini P, Gardenghina I, Bonetta F, et al. An updated overview of clinical guidelines for chronic low back pain management in primary care. *Joint Bone Spine*. 2012;79:176–185.
- Krawciw D, Atlas S. Occupational low back pain: evaluation. Available at: <http://www.uptodate.com/contents/occupational-low-back-pain-evaluation?>

- detectedLanguage=en&source=search_result&search=occupational±low±back±pain&selectedTitle=2~3&provider=noProvider. Accessed October 2, 2013.
38. Krawciw D, Atlas S. Occupational low back pain: treatment. Available at: http://www.uptodate.com/contents/occupational-low-back-pain-treatment?detectedLanguage=en&source=search_result&search=chou±subacute±or±chronic±low±back±pain%3A±nonsurgical±interventional±treatment&selectedTitle=4~150&provider=noProvider. Accessed October 2, 2013.
 39. Staiger T, Medley G, Wipf J, Deyo R. Diagnostic testing for low back pain. Available at: http://www.uptodate.com/contents/diagnostic-testing-for-low-back-pain?detectedLanguage=en&source=search_result&search=diagnostic±testing±for±low±back±pain&selectedTitle=1~150&provider=noProvider. Accessed October 2, 2013.
 40. Chou R. Subacute and chronic low back pain: surgical treatment. Available at: http://www.uptodate.com/contents/subacute-and-chronic-low-back-pain-surgical-treatment?detectedLanguage=en&source=search_result&search=chou±subacute±or±chronic±low±back±pain%3A±nonsurgical±interventional±treatment&selectedTitle=2~150&provider=noProvider. Accessed October 2, 2013.
 41. Chou R. Subacute and chronic low back pain: pharmacologic and noninterventional treatment. Available at: http://www.uptodate.com/contents/subacute-and-chronic-low-back-pain-pharmacologic-and-noninterventional-treatment?detectedLanguage=en&source=search_result&search=chou±subacute±or±chronic±low±back±pain%3A±nonsurgical±interventional±treatment&selectedTitle=3~150&provider=noProvider. Accessed October 2, 2013.
 42. World Health Organization. International Classification of Diseases, 9th Revision. Available at: <http://www.who.int/classifications/icd/en/>. Accessed January 15, 2013.
 43. Agency for Healthcare Research and Quality. Clinical classifications software (CCS) for ICD-9-CM. Available at: <http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp>. Accessed January 15, 2013.
 44. Underwood M. Diagnosing acute non-specific low back pain: time to lower the red flags? *Arthritis Rheum*. 2009;60:2855–2857.
 45. American Society of Health-System Pharmacists. The AHFS pharmacologic-therapeutic classification system. Available at: <http://www.ahfsdruginformation.com/class>. Accessed November 15, 2011.
 46. Imbens G. The role of the propensity score in estimating dose–response functions. *Biometrika*. 2000;87:706–710.
 47. Lechner M. Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In: Lechner M, Pfeiffer F, eds. *Econometric Evaluation of Labour Market Policies*. Heidelberg: Springer-Verlag; 2001:1–18.
 48. Bryson A., Dorsett R, Purdon S. The use of propensity score matching in the evaluation of labour market policies. Working paper No. 4, Department for Work and Pensions, UK.
 49. Choosing Wisely®. An initiative of the ABIM Foundation. Available at: <http://www.choosingwisely.org/>. Accessed March 15, 2014.
 50. Loeppke R, Taitel M, Hauffe V, Parry T, Kessler RC, Jinnett K. Health and productivity as a business strategy: a multiemployer study. *J Occup Environ Med*. 2009;51:411–428.
 51. Goetzl RZ, Long SR, Ozminkowski RJ, Hawkins K, Wang S, Lynch W. Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting US employers. *J Occup Environ Med*. 2004;46:398–412.
 52. Provonost P. Enhancing physicians' use clinical guidelines. *JAMA*. 2013;310:2501–2502.
 53. National Committee for Quality Assurance. *State of Health Care Quality: 2012*. Washington, DC: National Committee for Quality Assurance; 2012.
 54. Bunn W, Bayer R, Ehni T, Stowers A, et al. Impact of a musculoskeletal disability management program on medical costs and productivity in a large manufacturing company. *Am J Manag Care*. 2006;12:SP27–SP33.