

adapted: some have become students of the coding procedures, but many have outsourced billing to professional coders trained to search for keywords. Although many ED physicians don't know exactly what is billed in their name, physicians commonly receive regular feedback on their average billing performance through automated reports.

Early adoption of electronic records by the ED may in part explain the sharper billing increases in emergency medicine than in other clinical specialties. The EHR facilitates billing by presenting clickable check-boxes that easily satisfy coding-complexity criteria, and some EHRs even issue notifications when documentation needed for certain billing levels has not been achieved. These changes ensure that no billable action goes unnoticed and have reduced undercoding. In fact, EHR vendors tout this effect to justify the cost of their products. In other ways, however, the EHR has become a double-edged sword, potentially undermining its intended goal of reducing medical errors. Throughput suffers when time that could be better spent with patients is wasted on elaborate documentation. The EHR may also facilitate improper behavior, such as clicking multiple items in the "review of systems" that patients were not directly asked about. Of even

greater concern is the possibility of deliberate, systematic use of easily selected templates designed to ensure billing at the highest possible level, rather than promoting validated clinical decision rules and protocols designed to improve efficiency and quality. Although ED physicians are increasingly employed by hospitals, hospital chains, or contract groups with productivity-based compensation,<sup>5</sup> the OIG holds individual physicians accountable for billing done in their name, regardless of who directly manages the billing operations.

What should be done about the trend in billing? A first step is to do what the OIG report proposes: educate physicians about the importance of proper billing, review billing records to ensure that results match performance, and scrutinize physicians who consistently bill at higher levels than their peers.<sup>1</sup> From a broader perspective, the science of ED operations should be advanced to facilitate timely care. These advances should include the development of a more effective business model for the digital era that allows ED practitioners to get away from the computer and back to the bedside of sick and injured patients.

The EHR is one reason behind increased ED billing, and fraud may be facilitated by these new systems. However, this simple ex-

planation does not capture the broader story of what happened in U.S. EDs during the decade the OIG examined. While the ED has remained the social safety net, it has also gradually inherited roles previously handled by office-based physicians. EDs have become a central staging area for acutely ill patients, for the use of diagnostic technology, and for decisions about hospital admission, all of which makes ED care increasingly complex.

Disclosure forms provided by the author are available with the full text of this article at [www.nejm.org](http://www.nejm.org).

From the Department of Emergency Medicine, Emory University School of Medicine, Atlanta.

1. Department of Health and Human Services, Office of Inspector General. Coding trends of Medicare Evaluation and Management Services. May 2012 (<http://oig.hhs.gov/oei/reports/oei-04-10-00180.pdf>).
2. Pitts SR, Pines JM, Handrigan MT, Kellermann AL. National trends in emergency department occupancy, 2001 to 2008: effect of inpatient admissions versus emergency department practice intensity. *Ann Emerg Med* 2012;60:679-86.
3. Cubanski J, Huang J, D'Amico A, Jacobsen G, Neuman T. Medicare chartbook. 4th ed. Washington, DC: Kaiser Family Foundation, 2010 (<http://www.kff.org/medicare/8103.cfm>).
4. Pitts SR, Carrier ER, Rich EC, Kellermann AL. Where Americans get acute care: increasingly, it's not at their doctor's office. *Health Aff (Millwood)* 2010;29:1620-9.
5. O'Malley AS, Bond AM, Berenson RA. Issue brief no. 136: rising hospital employment of physicians: better quality, higher costs? Washington, DC: Center for Studying Health System Change, August 2011:1-4 (<http://www.hschange.com/CONTENT/1230>).

DOI: 10.1056/NEJMp1211315

Copyright © 2012 Massachusetts Medical Society.

## Appropriateness Criteria and Elective Procedures — Total Joint Arthroplasty

Hassan M.K. Ghomrawi, Ph.D., M.P.H., Bruce R. Schackman, Ph.D., and Alvin I. Mushlin, M.D.

Many of the most common inpatient surgeries in the United States are elective procedures. With health insurance cov-

erage expanding under the Affordable Care Act, utilization of elective surgery is likely to increase — with implications for

costs and the expansion of capacity required to meet the new demand and achieve good outcomes.

Cost-containment proposals have focused primarily on payment reforms, with approaches such as pay for performance and bundled payment generating great interest. But the nonemergency nature of elective procedures provides another opportunity for reducing costs. Using clinical appropriateness criteria to determine priorities for care can reduce or slow the growth in the number of procedures performed. Appropriateness criteria have not yet been developed for most common elective procedures. In instances in which they have been established, however, studies applying these criteria have generally revealed overutilization.<sup>1</sup> The development and implementation of evidence-supported appropriateness criteria that help to identify the subgroups of patients likely to benefit the most from a given procedure (thereby creating benchmarks for reimbursement) could help to combat increasing health care costs while enhancing access and quality. We believe that the case of total joint arthroplasty offers a prime example of the opportunities and challenges involved in creating and implementing appropriateness criteria.

Elective total hip and knee arthroplasties for the treatment of advanced osteoarthritis are among the most common inpatient surgeries in the United States; more than 1 million such procedures were performed in 2009.<sup>1</sup> They are performed in a wide variety of patients, ranging from those requesting surgery to facilitate their highly active lifestyle to those who require surgery in order to perform routine activities of daily living. The growing obesity epidemic coupled with aging of the population will almost certainly accelerate the demand

for these procedures. Estimates indicate that demand will quadruple by 2030, exceeding 4 million operations, and that more than 50% of patients will be younger than 65 years of age.<sup>2</sup> It has been proposed that episode-based bundled payment be used to cover the inpatient and postacute care of patients undergoing total joint arthroplasty, a procedure that seems to be an ideal candidate for achieving the efficiencies of care and reductions of costs associated with that mechanism of payment. Still, although bundling might reduce the cost per case, it won't provide incentives for reducing the number of procedures performed — and hence will not solve the utilization problem.

Using appropriateness criteria could slow the increase in utilization. Studies in other countries have found that 60 to 80% of total joint arthroplasties were considered appropriate according to the evidence-based criteria established by those countries.<sup>3,4</sup> There are two potential sources of reductions in use. First, a one-time decrease could occur when the appropriateness criteria are implemented; for instance, if procedures deemed to be inappropriate are not reimbursed by health insurers, only a minority of patients not meeting the criteria will be willing or able to pay out of pocket for them. Second, if reimbursement going forward is contingent on meeting the criteria, the upward trend can be expected to slow down from that point on. Basing reimbursement on appropriateness criteria also has the potential to enhance the overall quality of care by preventing complications that might have occurred in operations that were inappropriate to begin with. For example, quality improvement has

been observed with appropriate use of carotid endarterectomy, which has reduced the rate of strokes in patients who require and undergo the procedure but also reduced surgery-induced strokes by preventing the use of carotid endarterectomy for inappropriate reasons.<sup>5</sup>

Although implementing appropriateness criteria for total joint arthroplasty has not succeeded in the past, there are a few reasons it's likely to work now. First, opinion leaders in the U.S. orthopedics community, primarily at the American Academy of Orthopedic Surgeons, have recognized the importance of such criteria and have already started developing them as guidelines for other orthopedic procedures. Second, accountable care organizations and other institutions pursuing similar health care delivery models are becoming influential, and as they move away from procedure-based payments, they may well need to use such criteria to limit overall costs. Primary care trusts in England, for example, have already adopted the Oxford Knee Score, developed to assess outcomes of total knee arthroplasty, to determine eligibility for coverage. Third, recent developments in health information technology allow very complex appropriateness criteria (the Spanish criteria for total knee arthroplasty include 624 different potential combinations of factors<sup>3</sup>) to be readily integrated into decision-support tools for timely evaluation of appropriateness.<sup>1</sup>

Significant challenges to implementing appropriateness criteria must be overcome. First, we need to achieve consensus about the criteria themselves. There are currently no appropriateness criteria for total joint arthroplasty

in the United States beyond general statements from the National Institutes of Health about the use of these procedures.<sup>1</sup> Developing consensus will be challenging because the criteria for the procedures depend in part on patients' symptoms, activities, and preferences. This clinical situation thus differs from that of carotid endarterectomy, for example, in which the appropriateness determination is based primarily on measurements of blood flow. Most criteria developed in other countries for total joint arthroplasty have included both clinician-graded radiographic findings and patients' perceptions of pain and function. Moreover, to simply adopt criteria developed elsewhere would be to risk incorrect classification of patients, because physicians and patients in different countries have different expectations and preferences. For example, in Spain, previous surgical and nonsurgical management figures prominently in the criteria, whereas in Canada it does not.<sup>3</sup> In addition, although both these sets of criteria include pain and functional disability, each country defines and weights these elements differently.

Clinical opinion leaders and patient representatives must be involved in developing appropriateness criteria so that they are credible to physicians and patients and don't limit necessary care. Clinical leaders should also recognize that the quality of the criteria will be enhanced if representatives of multiple clinical disciplines are included in the development process — not only orthopedic surgery, but also gen-

eral internal medicine, family medicine, rheumatology, radiology, and rehabilitation medicine.

Another challenge is that accountable care organizations and third-party payers may apply appropriateness criteria variably, in part because of differences in their risk pools. Some payers may use the criteria as a benchmark for the level of reimbursement or to determine whether to reimburse at all for total joint arthroplasty. Others may use them as a basis for requiring prior authorization for referral to an orthopedic surgeon but not as a basis for reimbursing the surgeon. Such policy differences will create challenges for physicians and patients in making decisions about surgery and, if the criteria are not correctly applied, may limit necessary care.

Additional considerations may limit the extent to which implementing these criteria reduces the number of procedures performed. Retrospective studies may overestimate the inappropriate use of the procedure simply because patients' charts lack sufficient evidence about the relevant factors. Patients who can choose among health plans may eschew plans that apply appropriateness criteria in determining whether to issue prior authorization or to reimburse providers, even if those plans are less expensive than the alternatives. Finally, partial reliance on patients' self-reported subjective symptoms would permit surgeons and patients to override the criteria in order to justify the procedure.

Ultimately, payment reforms alone will probably be insuffi-

cient to restrain costs, especially for procedures whose use is expanding rapidly. Integrating appropriateness criteria into the reimbursement and care delivery systems could help bend the cost curve, although the achievement of savings will depend on the criteria's acceptance by physicians and patients. The challenges in developing and implementing appropriateness criteria for total joint arthroplasty probably apply to other elective procedures as well. But such evidence-based criteria, if applied wisely and fairly, may be the most powerful tool for controlling the cost and enhancing the quality of elective procedures.

Disclosure forms provided by the authors are available with the full text of this article at [NEJM.org](http://NEJM.org).

From the Department of Public Health, Weill Cornell Medical College (H.M.K.G., B.R.S., A.I.M.); and the Hospital for Special Surgery (H.M.K.G.) — both in New York.

1. Lawson EH, Gibbons MM, Ingraham AM, Shekelle PG, Ko CY. Appropriateness criteria to assess variations in surgical procedure use in the United States. *Arch Surg* 2011; 146:1433-40.
2. Kurtz SM, Lau E, Ong K, Zhao K, Kelly M, Bozic KJ. Future young patient demand for primary and revision joint replacement: national projections from 2010 to 2030. *Clin Orthop Relat Res* 2009;467:2606-12.
3. Quintana JM, Arostegui I, Escobar A, Azkarate J, Goenaga JJ, Lafuente I. Prevalence of knee and hip osteoarthritis and the appropriateness of joint replacement in an older population. *Arch Intern Med* 2008;168:1576-84.
4. van Walraven CV, Paterson JM, Kapral M, et al. Appropriateness of primary total hip and knee replacements in regions of Ontario with high and low utilization rates. *CMAJ* 1996;155:697-706.
5. Halm EA, Tuhim S, Wang JJ, Rojas M, Hannan EL, Chassin MR. Has evidence changed practice? Appropriateness of carotid endarterectomy after the clinical trials. *Neurology* 2007;68:187-94.

DOI: 10.1056/NEJMp1209998

Copyright © 2012 Massachusetts Medical Society.