Cost-effectiveness Analysis of an Established, Effective Procedure

THE NUMBER OF TOTAL KNEE ARTHROPLASTY (TKA) procedures performed in the United States has been rising rapidly. In 2006, approximately 500 000 TKAs were performed, incurring direct medical costs of roughly $11 billion (our unpublished estimate). Use of this procedure is expected to continue to rise due to both the obesity epidemic and the aging of the population. One study estimates that 3.5 million TKAs will be performed annually by the year 2030.[1] The increasing use of this procedure has prompted an increased interest in its evaluation. For instance, national TKA registries now exist in Australia, Denmark, Norway, and Sweden. The US Food and Drug Administration (FDA) has begun exploring the possibility of a national TKA registry to help to evaluate the potential strengths and weaknesses of different implant designs.[2]

See also page 1113

Although TKA is a safe and effective treatment for advanced knee osteoarthritis,[2-4] lingering questions remain regarding variations in patient outcomes due to differences among patients undergoing the procedure and among the hospitals where it is performed. Elderly patients, minorities, those with more comorbid conditions, and those who undergo TKA later in their functional decline or at hospitals where only a few TKAs are performed each year (ie, "low-volume" hospitals) are more likely to have worse outcomes than patients who do not fit these criteria.[5-10] What remains in question is whether strategies exist that can further improve outcomes in these "at-risk" groups and whether all centers performing the procedures should be doing so.

In this issue of the *Archives*, Losina et al[11] examine these questions from the perspective of cost-effectiveness, with a focus on Medicare enrollees who were 65 years or older. The overall findings were favorable to TKA, which had an incremental cost-effectiveness ratio of $18 300 per quality-adjusted life year (QALY) gained compared with medical treatment alone. This figure falls below the cost-effectiveness thresholds often mentioned as appropriate, such as the £20 000 to £30 000 (approximately $29 000 to $44 000) per QALY threshold used by the British National Health Service's National Institute for Health and Clinical Excellence (www.nice.org.uk). The cost-effectiveness ratio of TKA compared with no TKA varied from $9200 (low-risk patients in high-volume hospitals) to $29 800 (high-risk patients in low-volume hospitals). Even among high-risk patients, the authors found TKA to be relatively cost-effective ($29 000-$30 000/QALY) compared with no TKA. In other words, although the ratio varied somewhat by both patient and hospital characteristics, TKA appeared to be a cost-effective strategy compared with medical management.

The study by Losina et al[11] has other important features worth highlighting. The volume categories (<25, 25-200, or >200 TKAs per year) represent only the indemnity Medicare-covered portion of a hospital's overall TKA volume. Nearly 60% of the TKAs performed in the United States are in Medicare beneficiaries (our unpublished analysis of Healthcare Utilization Project data [http://hcupnet.ahrq.gov/HCUPnet.jsp]), but this proportion varies by region of the country due to variations in TKA utilization, obesity prevalence, and underlying age distribution. So hospitals classified by their "Medicare volume" might be classified differently by their "total volume."

The value of moving patients undergoing TKA from local hospitals to higher-volume regional centers remains to be determined because "value" involves more than just cost-effectiveness. A regionalization approach to TKA could widen health disparities. Younger, wealthier, nonminority patients might be more likely to travel for care rather than undergo TKA at a local, lower-volume community hospital, further lowering volume and experience at the local hospitals that serve older, poorer, minority patients. An alternative strategy might be to transfer models of nursing and rehabilitation care from high-volume centers to lower-volume hospitals: differences in outcomes between these hospital types might have more to do with perioperative care than with surgical proficiency.

Analyses such as the one conducted by Losina et al,[11] carefully conducted and wholly transparent, highlight several of the dilemmas policy makers face in evaluating widely used medical technologies. At least in the United States, even well-performed cost-effectiveness analyses do not influence either payers or physicians directly. Payers do not use the results to make coverage determinations nor do physicians use them to make treatment decisions. How we move from this current state to a system in which cost-effectiveness of procedures affects medical practice is unclear.

Losina et al[11] also highlight a challenge ahead for comparative effectiveness research. This year, Congress included $1.1 billion in the American Recovery and Reinvestment Act ("the stimulus package") to support comparative effectiveness research. The research relies on having high-quality outcomes data so that outcomes

achieved through alternative treatment strategies can be validly compared.

The overall estimate of cost-effectiveness reported by Losina et al[11] was relatively sensitive to their estimate of the QALY gain that resulted from TKA. If the quality of life gain was 15% too optimistic in their model, then the actual incremental cost-effectiveness ratio was $106 700 per QALY, as summarized in their Table 2.[11] This number is higher than most policy makers would consider acceptable, so the accuracy of the QALY gain is critical to the interpretation of TKA's cost-effectiveness.

As with many other procedures, the estimate of effectiveness for TKA is not based on the findings of randomized controlled studies in which the outcomes of similar patients undergoing continued medical management (the "base case" in the study by Losina et al[11]) vs TKA could be compared. Instead, the only relevant comparative trials use a before/after design, and Losina et al[11] cite 2 of these. There are many more. In a before/after design, a patient's status is assessed before and after treatment.

Results from such studies have been misleading in some cases, making the intervention appear more effective than it may actually be. For instance, the findings of a before/after study assessing lung volume reduction surgery for the treatment of emphysema suggested that most patients had large quality of life improvements, reduced symptoms, and improved physical function.[12] However, the results of a randomized trial conducted later suggested that the benefits were actually far more modest.[13]

This creates an intriguing dilemma. The sensitivity of the cost-effectiveness estimate reported by Losina et al[11] might suggest the need for large randomized controlled trials of TKA to generate unbiased estimates of utility, which would enable more certain estimates of cost-effectiveness. However, TKA is already recognized as a highly effective procedure. The numerous existing before/after studies and decades of experience with TKA confirm that. Over 90% of patients undergoing TKA experience significantly higher functional ability, less pain, and higher mobility levels than prior to TKA.[14] So a randomized study cannot be launched because it would be unethical to deny an effective procedure to patients merely to generate a more precise estimate of the utility gain conferred.

Alternatively, one could argue that the FDA should mandate that all new technologies be subjected to a superiority standard at the time of approval, which would in most cases necessitate the conduct of randomized trials. Currently, most medical devices need only be shown to be safe as long as there is a similar previously approved device on the market. One problem with changing the FDA approval standard is that the impact would be asymmetric: the new standard could not be applied retroactively. Another problem is that slowing the path to approval for equivalent substitutable devices could impede competition between manufacturers, allowing those with devices already on the market to increase prices. In other words, the FDA cannot easily solve the "imperfect data" problem without potentially causing other problems.

Policy makers are aware that the imperfect data problem is not unique to orthopedic procedures. Alternative treatments for prostate cancer, including multiple different radiation treatment strategies, also see wide use despite a paucity of high-quality effectiveness data. Whether these treatments are effective is not at issue—how much better a newer (and typically more expensive) technology is than the one it replaces is far harder to know.

These important follow-up issues do not detract from the quality or importance of the study by Losina et al.[11] It provides insight into the volume-outcome question in TKA while also illustrating one of the key challenges intrinsic to technology assessment of widely used procedures.

Stephen Lyman, PhD
Robert G. Marx, MD
Peter B. Bach, MD, MAPP

Correspondence: Dr Bach, Memorial Sloan-Kettering Cancer Center, New York, NY.
Financial Disclosure: None reported.

REFERENCES

1. Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am.* 2007;89(4):780-785.
2. Department of Health and Human Services, US Food and Drug Administration. Evaluation of long-term safety and effectiveness of joint replacement devices. https://www.fbo.gov/index?s=opportunity&mode=form&tab=core&id=9105136558343 0bcedb3b2dc6383a8fe&_cview=0&cck=1&au=&ck=. Accessed April 16, 2009.
3. Ethgen O, Bruyere O, Richy F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty: a qualitative and systematic review of the literature. *J Bone Joint Surg Am.* 2004;86(5):963-974.
4. Jones CA, Beaupre LA, Johnston DW, Suarez-Almazor ME. Total joint arthroplasties: current concepts of patient outcomes after surgery. *Clin Geriatr Med.* 2005;21(3):527-541.
5. Dunlop DD, Song J, Manheim LM, Chang RW. Racial disparities in joint replacement use among older adults. *Med Care.* 2003;41(2):288-298.
6. Escalante A, Barrett J, del Rincon I, Cornell JE, Phillips CB, Katz JN. Disparity in total hip replacement affecting Hispanic Medicare beneficiaries. *Med Care.* 2002; 40(6):451-460.
7. Hawker GA, Wright JG, Coyte PC, et al. Differences between men and women in the rate of use of hip and knee arthroplasty. *N Engl J Med.* 2000;342(14):1016-1022.
8. Holtzman J, Saleh K, Kane R. Gender differences in functional status and pain in a Medicare population undergoing elective total hip arthroplasty. *Med Care.* 2002; 40(6):461-470.
9. Ibrahim SA, Stone RA, Han X, et al. Racial/ethnic differences in surgical outcomes in veterans following knee or hip arthroplasty. *Arthritis Rheum.* 2005; 52(10):3143-3151.
10. Losina E, Barrett J, Baron JA, Levy M, Phillips CB, Katz JN. Utilization of low-volume hospitals for total hip replacement. *Arthritis Rheum.* 2004;51(5):836-842.
11. Losina E, Walensky RP, Kessler CL, et al. Cost-effectiveness of total knee arthroplasty in the United States: patient risk and hospital volume. *Arch Intern Med.* 2009;169(12):1113-1121.
12. Yusen RD, Lefrak SS, Gierada DS, et al. A prospective evaluation of lung volume reduction surgery in 200 consecutive patients. *Chest.* 2003;123(4):1026-1037.
13. Fishman A, Martinez F, Naunheim K, et al. A randomized trial comparing lung-volume-reduction surgery with medical therapy for severe emphysema. *N Engl J Med.* 2003;348(21):2059-2073.
14. Franklin PD, Li W, Ayers DC. The Chitranjan Ranawat Award: functional outcome after total knee replacement varies with patient attributes. *Clin Orthop Relat Res.* 2008;466(11):2597-2604.