

D. Seidenwurm  
P. Turski  
J. Barr  
J. Connors  
M. Lev  
S. Mukherji  
E. Russell

## Performance Measures in Neuroradiology

**SUMMARY:** Performance measurement has been added to the Medicare payment scheme as of July 2007. Two performance measures are applicable to neuroradiology, pertaining to brain and vascular imaging in stroke. These measures are early attempts to rigorously define the meaning of effective performance of neuroradiology.

The cost of health care in the United States is high. The American health care system has been unable to deliver universal coverage or aggregate health care outcomes commensurate with either our expenditure level or our economic standing.<sup>1</sup> In addition, barely half of the care recommended by practice guidelines promulgated by specialty societies based upon the best available outcome data is actually delivered.<sup>2</sup> For these and other reasons, attention has recently focused on metrics for evaluating the performance of physicians in providing health care paid for by the government, other third parties, and patients themselves. Compensation of physicians in part related to quality metrics has been proposed as a method for improving performance of the health care system and is now part of the Medicare payment scheme.<sup>3</sup>

Whether one believes that market or social models of health care delivery are best, determination of the quality of the product, however it is paid for, seems crucial. In the social model, care is paid for either directly by the government or indirectly via intermediaries contracted with or regulated by the government payer. These organizations, as stewards of the taxpayers' resources and advocates for citizens, desire information on which to base compensation and coverage decisions. Quantitative data will allow the documentation and, subsequently, improvement of health care quality.

Market models of health care delivery rely on patients themselves, their families, or other entities with which patients have associated voluntarily to make health care decisions. Because market models rely on competition among providers to guide informed choices by consumers, data on health care quality seem necessary for efficient markets to function.

In the United States, a hybrid market-social health care system is in place. The system is about 60% government funded, if one takes into account direct government expenditures and indirect expenditures resulting from tax code provisions favoring certain health care spending by others.<sup>4</sup> If one includes the cost of all regulation pertaining to health care, the proportion is even higher. The reach of government influence on health care is yet greater than these calculations suggest, because many private entities look to the government for

guidance, and local, state, and federal government regulations determine medical policies. In addition, government-funded research sets the stage for the commercial developments that follow clinical research, much of which is itself funded directly by governmental agencies or structured according to the dictates of drug or medical-device approval mechanisms.

Performance measures in health care delivery have traditionally focused on system-level variables or primary care and have not specifically involved radiology at all, much less neuroradiology. Mammography screening has been subject to the greatest degree of scrutiny, with federally mandated training and experience to qualify for reimbursement and mandated audits of reader performance, in addition to technical standards. Mammography referral rate has been included in numerous quality datasets as well. Even these standards are almost solely related to inputs rather than outcomes of patient care. In general, performance-based compensation has been linked to performance at the system level but not at the level of the individual physician. Even when indices of individual performance seem reasonable intuitively, it is quite rare for individual physician volume to exceed a sample-size threshold necessary for statistical significance.<sup>5,6</sup>

Recently, budgetary and political considerations led Congress and the Executive Branch to adopt a revised Medicare compensation scheme linked in part to performance measures. Whereas recent changes in Congressional leadership have produced some uncertainties, pay for performance became part of the Medicare payment system beginning in July. This system has been designed to act as a carrot rather than a stick and to function as a bit of sugar to help ease the bitterness associated with the accompanying cuts in the fee schedule.<sup>1</sup>

The American Medical Association Physician Consortium for Performance Improvement under a contract with Mathematica, a consulting firm engaged by the Centers for Medicare & Medicaid Services (CMS), undertook to develop performance measures for a variety of specialties, including radiology. The first of the radiology performance measures fell under the Stroke and Stroke Rehabilitation project. The American College of Radiology and the American Academy of Neurology were selected as the lead specialty organizations for the project. A neurologist and a neuroradiologist were selected as clinical co-chairs along with a third co-chair with methodologic expertise.<sup>7</sup>

The task was to develop evidence-based physician performance measures relevant to stroke care that were reasonably expected to help improve patient outcome by addressing dif-

From the Radiological Associates of Sacramento (D.S.), Sacramento, Calif.; University of Wisconsin (P.T.), Madison, Wis; Baptist Hospital (J.B.), Memphis, Tenn; the Miami Vascular Institute (J.C.), Miami, Fla; Massachusetts General Hospital (M.L.), Boston, Mass; the University of Michigan (S.M.), Ann Arbor, Mich; and Northwestern University (E.R.), Chicago, Ill.

Please address correspondence to David Seidenwurm, MD, Radiological Associates of Sacramento, 1500 Expo Pkwy, Sacramento, CA 95816; e-mail: dseidenwurm@comcast.net

DOI 10.3174/ajnr.A0672

ferences between care actually delivered as compared with that recommended by guidelines of major medical societies. The standards were quite strict in that the measured elements of care must be under the direct control of individual physicians and attributable to them, and compliance, including documentation, must be simple and not unduly costly in time, effort, or other resources. The suggested action must be feasible for most intended physicians and, ideally, promote communication and coordination of care among all physicians involved in caring for a particular patient. The measures must be rigorously defined in terms of easily available clinical data and specifiable in terms of existing data elements like the International Classification of Diseases-9 and Current Procedural Terminology (CPT) codes or other items generally found in the paper or electronic medical record. Perhaps most important, there must be no identifiable unintended consequence or perverse incentives within the measure that would be counterproductive, especially with respect to access to care of underserved populations in rural or inner city urban areas. In general, guidelines and standards promulgated by major specialty societies formed the body of relevant evidence available to the panel.

Evidence-based measures in stroke imaging are hard to construct because the direct evidence of improved outcome due to imaging is sparse. The deficiencies in the evidence base are easily understandable due to the obvious benefits of imaging in stroke. Consider that the advantages of parachute use in falls from great heights have not been subjected to randomized placebo-controlled trials.<sup>8</sup> The benefits of universal imaging in stroke care are implicit in the design of each of the randomized trials on which the construction of numerous practice guidelines has relied. Furthermore, the methods of the trials generally specified the relevant information sought from the imaging studies. The performance measures were developed from guideline statements promulgated by relevant specialty organizations. Individual studies, even of the highest quality, were not, by themselves, sufficient to support performance measures in the absence of such consensus support. These guidelines and randomized trials formed the basis for the stroke imaging performance measures for cross-sectional brain imaging studies and imaging studies of the carotid arteries.<sup>9,10</sup>

Before describing the radiology performance measures adopted by the panel, we will consider some of the ideas rejected and the reasons they failed to meet the standards of the panel. One proposed measure concerned time to brain imaging in stroke. Despite the importance of efficient delivery of care in the acute stroke setting, this measure was rejected because the data collection would have been arduous and inaccurate, the attribution of responsibility to an individual radiologist would have been impossible, and the burdens unmanageable considering the relatively small number of patients with stroke who actually present within the time window that requires counting minutes. Another measure rejected by the panel set a standard requiring brain and vascular imaging in transient ischemic attack and stroke that might have encouraged overuse or delayed care in some cases, interfered with effective patterns of care in others, and was not supported directly by previously published guideline statements.

The panel eventually agreed on 2 imaging measures for stroke, and these are the first of the measures to apply directly to radiologists in the CMS dataset. The successful develop-

ment of these measures will allow radiologists to participate in the voluntary program allowing Medicare physicians to recoup 1.5% of the 5% fee cut that was scheduled to take effect in the summer of 2007.<sup>2</sup> Some may consider the resulting measures to be so rudimentary as to be useless, whereas others may decry them as radical departures from physician autonomy. It is crucial to understand that these measures are the first steps in a long process, so it is important that they represent a consensus rather than break new ground.

One measure concerned the interpretation of cross-sectional imaging in acute stroke. This measure is satisfied when the interpreting radiologist evaluates the presence or absence of hemorrhage, mass, and acute infarct when reading a CT or MR imaging performed within 24 hours of hospital admission of a patient with acute stroke. The measure is agnostic with respect to the precise technique used. For the purposes of this program, CT and MR imaging are deemed equivalent overall, though each radiologist has an opinion as to which is the best procedure in a specific clinical setting for any one of a number of valid reasons. It is left to the individual radiologist to determine that the MR imaging technique used is sufficiently sensitive to hemorrhage or that the CT technique is sufficiently sensitive to acute infarct to be clinically useful in risk stratification in stroke care.

The second measure concerns the interpretation of carotid artery imaging studies. The measure simply requires that the percentage of stenosis based on the angiographic North American Symptomatic Carotid Endarterectomy Trial (NASCET) method of comparing the narrowest segment with the distal luminal diameter be related to the reported measure of arterial narrowing. The measure does not privilege one imaging technique over any other, recognizing that many reasons may exist for preferring ultrasound (US), CT angiography (CTA), MR angiography (MRA), or conventional angiography in a given situation. It simply requires that radiologists use a standardized method for stenosis quantification validated in clinical intervention outcome trials. There is some room for discussion, perhaps because different studies use different physical principles to arrive at estimates of narrowing, but there is little room for debate about the need to convert the inferences drawn from these disparate techniques into a common metric applicable across all of them. Physicians need imaging-derived stenosis measurements translatable directly into widely accepted practice guidelines that rely on imaging to direct patient management. This seems beyond responsible controversy.<sup>11,12</sup>

Each of these measures satisfies the strict standards set for face validity in physician performance measures; they are directly related to the care of patients with stroke. One cannot deny the need to identify hemorrhage, mass, and acute infarct when studying the brain of a patient with stroke, nor can one deny the need to state the degree of stenosis of the carotid arteries with respect to the most reproducible, generally accepted, and rigorously outcome-validated method. The measures are attributable to individual physicians, in this case the radiologists interpreting the study, and they are actionable, under their control. The measures specify the pathologic conditions and interpretation methodology that a radiologist should use to receive credit for satisfying the requirements of quality performance.

Assessing and documenting the presence or absence of

mass effect, intracranial hemorrhage, and acute infarct signs at CT or MR imaging ought to be feasible, and interpreting carotid imaging studies with reference to the NASCET methodology seems readily achievable, as well.

To the extent that they are not feasible at present in some centers, MR imaging techniques in acute stroke may need to be modified so as to maximize the likelihood of detecting acute blood. Acute stroke signs on CT ought to be learned, and CT techniques refined. Reference standards for US, CTA, MRA, and angiographic interpretation that coordinate with the NASCET methodology of angiographic interpretation ought to be selected. Simply stating that such a reference standard has been used over the appropriate range of stenoses satisfies the measure.

Deficiencies or gaps in care in stroke imaging were inferred from the geographic variation in carotid endarterectomy not obviously related to regional differences in risk factors or outcomes and differences in outcomes between Medicare populations and the randomized trial cohorts. The low rate of tissue plasminogen activator (tPA) use among eligible patients nationwide also represents a deficiency in care that needs to be addressed. Clearly, these gaps in care are neither solely, nor perhaps even to a great extent, based upon inadequate radiologic evaluation, but to the extent that variations in the rate of endarterectomy are due to heterogeneity of radiologic practice, standardization in methods of interpretation ought to minimize them. Standardization of measurement methods for carotid stenosis ought to at least establish a common vocabulary through which the conversation can continue. To the extent that lack of confidence in imaging findings results in a failure to treat with tPA, knowledge that exclusion criteria are explicitly sought may be of help.<sup>13-15</sup>

The measures are easily specifiable and should not be unduly costly to audit because the data elements on which they are based are readily available. The final radiologic report need only contain mention of mass, bleed, and acute infarct in the case of brain CT and MR imaging; and the final carotid US, CTA, MRA, or angiographic report need only mention that the stenosis is estimated using standards related to the NASCET method. By their very nature, these measures facilitate coordination of care and communication among members of the patient care team. Because they specify the precise nature of the information in the radiology report to be communicated in order that other physicians may approach the patient in a manner consistent with the best available evidence, they promote coordinated care.

One challenge will be the development of guidelines for stenosis characterization that produce similar results across all techniques. Ideally, the methodology should be easy to perform, with high levels of accuracy and reproducibility. The organization that takes on this task will be making a significant contribution to the consistency of stenosis reporting across techniques. It is clear, however, that adequate data are already available to permit implementation of the performance measure.<sup>16-20</sup> Additional work might focus on whether noninvasive techniques offer additional benefits in predicting patient outcome.

It is difficult, though not impossible, to imagine how anodyne measures like these can result in unintended or adverse consequences, but unintended consequences are almost by definition unanticipated, so some caution is indicated. It is also hard to think of how compliance with the quality stan-

dards for stroke imaging could limit access to care among underserved populations, but these are areas that might benefit from further study, as well.<sup>21,22</sup>

It is to be expected that performance measures will evolve with the state of medical knowledge and change as social values change. It must also be understood that as the quality and consistency of medical care improve, the standards for quality will rise. Industrial engineering teaches us that first we make each of the widgets the same as all the others, and then we make better widgets.<sup>23</sup> Quality improvement is best imagined as a ratchet rather than a treadmill.

The process that led to the production and promulgation of these performance measures was internally consistent within a conceptual framework that accepts the validity of evidence-based physician performance measures. However, it is necessary to state that the evidence to support this approach is spotty. Some have suggested that the acceptance of this methodology is risible, because the trials performed to date have generally failed to provide decisive support for the underlying hypothesis. The studies demonstrating that paying physicians for satisfying predetermined performance metrics has a beneficial effect on patient outcomes are often of poor quality, limited in their generalizability, or otherwise flawed. The evidence for centralized implementation of evidence-based medicine is slim.<sup>24,25</sup>

Neuroradiologists make arguments like these from a weak position, because the evidence for much of what we do every day is slimmer still. If the parachute analogy applies to imaging, it is equally applicable to performance measures. There is certainly a wide body of scientific data, as well as experience in everyday life supporting the ideas that incentives influence behavior and that people tend to do what they are paid to do. As we begin to pay for performance in medicine nationally, we ought to do so with our eyes open to the potential pitfalls as well as the potential benefits.<sup>26</sup> We must make sure that there are no adverse outcomes to the widespread application of outcome data. Also, we must demonstrate rigorously that we image patients with specific techniques to improve their lives, not just to produce pretty pictures.<sup>27</sup>

## References

1. Tanne JH. US gets mediocre results despite high spending on health care. *BMJ* 2006;333:672
2. McGlynn EA, Asch SM, Adams J, et al. The quality of healthcare delivered in the United States. *N Engl J Med* 2003;348:2635-45
3. Glendenning D. Last-minute Medicare pay package freezes rates, adds reporting bonus. *American Medical News* December 25, 2006:1
4. Bitton A, Kahn JB. Government share of health care expenditures. *JAMA* 2003;289:1165
5. Rosenthal MB, Frank RG, Li Z, et al. Early experience with pay-for-performance. *JAMA* 2005;294:1788-793
6. Christianson JB, Knutson DJ, Mazze RS. Physician pay-for-performance implementation and research issues. *J Gen Intern Med* 2006;21:S9-13
7. Drozda J, Holloway R, Seidenwurm D, et al. Stroke and stroke rehabilitation physician performance measure set. *AMA Physician Consortium for Quality Improvement*, 2006. Available at: <http://www.ama-assn.org/ama1/pub/upload/mm/370/strokekw120606.pdf>. Accessed January 7, 2007
8. Potts M, Prata N, Walsh J, et al. Parachute approach to evidence-based medicine. *BMJ* 2006;333:701-03
9. Kwiatkowski TG, Libman RB, Frankel M, et al. Effects of tissue plasminogen activator for acute ischemic stroke at one year. *N Engl J Med* 1999;340:1781-87
10. Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis: North American Symptomatic Carotid Endarterectomy Trial Collaborators. *N Engl J Med* 1991;325:445-453
11. Sacco RL, Adams R, Albers G, et al. Guidelines for prevention of stroke in patients with ischemic stroke or transient ischemic attack: a statement for healthcare professionals from the American Heart Association/American

- Stroke Association Council on Stroke—co-sponsored by the Council on Cardiovascular Radiology and Intervention: the American Academy of Neurology affirms the value of this guideline. *Stroke* 2006;37:577–617
12. Albers GW, Hart RG, Lutsep HL, et al. **AHA scientific statement: supplement to the guidelines for the management of transient ischemic attacks—a statement from the Ad Hoc Committee on Guidelines for the Management of Transient Ischemic Attacks, Stroke Council, American Heart Association.** *Stroke* 1999;30:2501–11
  13. Saleh SS, Hannan EL. **Carotid endarterectomy utilization and mortality in 10 states.** *Am J Surg* 2004;187:14–19
  14. Connors JJ, Seidenwurm D, Wojak JC. **Treatment of atherosclerotic disease at the cervical carotid bifurcation: current status and review of the literature.** *AJNR Am J Neuroradiol* 2000;21:444–50
  15. Katzan IL, Hammer MD, Hixson ED. **Utilization of intravenous tissue plasminogen activator for acute ischemic stroke.** *Arch Neurol* 2004;61:346–50
  16. Bartlett ES, Walters TD, Symons SP, et al. **Quantification of carotid stenosis on CT angiography.** *AJNR Am J Neuroradiol* 2006;1:13–19
  17. Nederkoorn PJ, van der Graaf Y, Hunink MGM. **Duplex ultrasound and magnetic resonance angiography compared with digital subtraction angiography in carotid artery stenosis: a systematic review.** *Stroke* 2003;34:1324–31
  18. Grant EG, Benson CB, Moneta GL, et al. **Carotid artery stenosis: gray-scale and Doppler US diagnosis—Society of Radiologists on Ultrasound Consensus Conference.** *Radiology* 2003;229:340–46
  19. Gaitini D, Soudack M. **Diagnosing carotid stenosis by Doppler sonography: state of the art.** *J Ultrasound Med* 2005;24:1127–36
  20. Sabeti S, Schillinger M, Mlekusch W, et al. **Quantification of internal carotid artery stenosis with duplex US: comparative analysis of different flow velocity criteria.** *Radiology* 2004;232:431–39
  21. Fisher ES. **Paying for performance: risks and recommendations.** *N Engl J Med* 2006;355:1845–47
  22. Werner RM, Asch DA. **The unintended consequences of publicly reporting quality information.** *JAMA* 2005;293:1239–44
  23. Maani KE, Putterill MS, Sluti DE. **Empirical analysis of quality improvement in manufacturing.** *International Journal of Quality & Reliability Management* 1994;11:19–37
  24. Rosenthal MB, Frank RG. **What is the empirical basis for paying for quality in health care?** *Med Care Res Rev* 2006;63:135–57
  25. Werner RM, Bradlow ET. **Relationship between Medicare's hospital compare performance measures and mortality rates.** *JAMA* 2006;296:2694–702
  26. Rosenthal MB, Dudley RA. **Pay-for-performance: will the latest payment trend improve care?** *JAMA* 2007;297:740–44
  27. Bush RW. **Reducing waste in the US healthcare systems.** *JAMA* 297:871–74