Neurosurgery Quality: Pay-for-Performance, Guidelines, and Outcome Measures

Mark E. Linskey, M.D.

The healthcare quality movement can trace its origins back to the man considered to be the architect of healthcare quality, Avedis Donabedian (1919–2000), who did his major work on defining quality in healthcare from the 1960s through 1985, and who published three seminal volumes, entitled *Explorations in Healthcare Quality and Monitoring* between 1980 and 1985 (21–23). Donabedian first classified and characterized healthcare quality characteristics into *Structural, Process,* and *Outcomes* measures.

The work of Donabedian launched initial forays into the healthcare "quality assurance" (QA) movement. Eventually, the healthcare QA movement became strongly influenced by the "continuous quality improvement" (CQI) aspects of "total quality management" (TQM),[18,20,41] and the result was a shift to more of a systems-based "quality improvement" (QI) approach.[6,8,25,44–47]

"Pay-for-performance" (P4P) is the latest quality initiative to come along in the continuing evolution of healthcare QI. Although the general origins of healthcare quality have already been outlined above, P4P arose from two very specific root sources. The Institute of Medicine (IOM) Quality Initiative that began in 1996 collided with the Center for Medicare and Medicaid Services (CMS) strategic planning initiative stemming from the Medicare prescription drug improvement and Modernization Act of 2003 (MMA 2003), to produce the CMS QI Roadmap of 2005, which included P4P as one of its five system strategies.

P4P affects both hospitals and physicians, but this chapter deals only with physicians. The interplay of elected versus government officials and agencies as well as private institutes and organizations in the P4P development, implementation, and oversight processes can be dizzyingly complex and very confusing, as can the new array of organization abbreviations involved. The schematic outlined in *Figure 27.1* as well as the abbreviation list in *Table 27.1* are provided to assist the reader in navigating this interactive maze. Elected government officials are indicated in royal blue, appointed government officials and agencies in light blue, and private agencies and organizations in green. Formal lines of influence are indicated with solid arrows, whereas less formalized relationships are indicated with dashed arrows. Areas with lobbying potential and important impact on neurosurgical practice are indicated with asterisks.

Figure 27.1 is not intended to represent an organizational chart. Rather, it represents one person's perception at one snapshot in time of perceived lines and directions of influence between the various agencies, organizations, and officials involved. These relationships have shown themselves to be very liquid and dynamic during the last 12 months, and a similar chart created 6 to 12 months from now might differ in certain details, particularly as they relate to private organizations and agencies. It is hoped that *Figure 27.1* will assist the reader in maintaining overall orientation and perspective as we proceed through a reductionist analysis of the individual agencies involved.

MEASURING AND INFLUENCING QUALITY IN HEALTHCARE

The three volumes Donabedian produced in his *Explorations in Healthcare Quality and Monitoring* series, published between 1980 and 1985, included *Definition of Quality and Approaches to its Assessment* (Volume 1), *The Criteria and Standards of Quality* (Volume 2), and *Methods and Findings of Quality* (Volume 3).[21–23] In his work, Donabedian divided quality measures into "Structural," "Process," and "Outcomes" measures. Examples of structural measures include such things as certifications (e.g., did you complete an Accreditation Council for Graduate Medical Education [ACGME]-accredited neurosurgery residency training program or are you certified by the American Board of Neurological Surgery [ABNS]), proof of case volume for different procedures, or does your practice have the health information technology (HIT) necessary for documenting, reporting, and monitoring quality measures. Unfortunately, when studied in terms of either clinical outcomes or cost effectiveness, structural measures of healthcare quality are only poorly correlated with the quality of healthcare delivered.

Process measures can be very broad in scope and include things that are now referred to as "quality measures"

0148-703/07/5401-0157

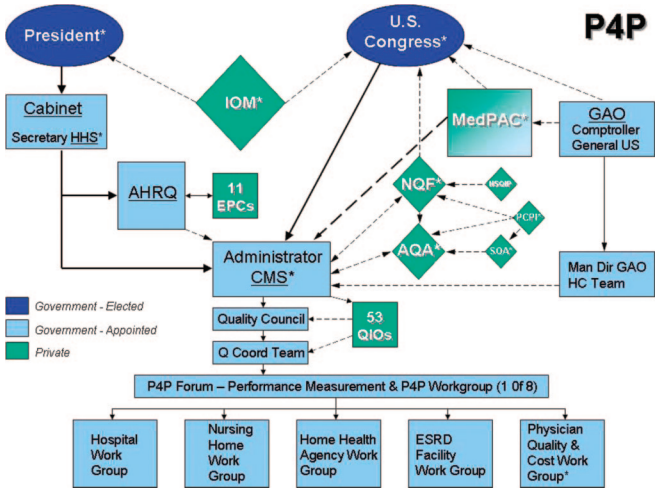

FIGURE 27.1. A hypothetical schematic diagram depicting perceived connections of influence in the national P4P movement. Elected government officials are depicted in dark blue. Agencies with appointed government officials are depicted in light blue. Private entities or organizations are depicted in green (those who are funded by government agencies are depicted as squares and those who are financially independent are depicted as triangles). Solid lines depict formal, direct, chain-of-command influence. Dashed lines depict informal lines of influence. The bolder and thicker the arrowed line, the more influential the connection. This figure is not intended to depict an organizational chart. It is an interpretation of one observer only, and many other depictions are possible. It should serve to assist the reader in navigating the enumeration of organizations and their relations contained in the text of the manuscript. Asterisks mark potential sites for influence by organized neurosurgery educational, advocacy, and lobbying efforts. Man Dir, managing director; HC, healthcare; Q coord team, quality coordinating team.

as well as "efficiency measures." Examples of quality process measures include such things as whether or not your patient received the appropriate antibiotics before surgical incision or whether or not your surgical patient received antithromboembolism prophylaxis. Examples of efficiency process measures include such things as whether or not the procedure performed was justifiable under evidence-based care (EBC) criteria, your average length of inpatient stay after a particular procedure, or your cost-of-care for performing a specific procedure. Although process measures generally perform better than structural measures in predicting quality of healthcare delivered, in their current form, their correlation with crude measures of clinical outcome remains disappointingly poor.[9,60] Individuals and institutions can become very proficient and compliant with documenting and reporting process measures without necessarily improving the care of individual patients.

Outcome measures are the most difficult and expensive measures to collect and analyze. Crude measures of outcome, such as mortality rates or "rate of discharge to other than home" require proper and careful risk adjustment to ensure fairness in assessment and avoid being misleading. More detailed outcomes measures tend to be very diagnosis and procedure specific, which requires multiple and differing reporting tracks, systems, and metrics. Outcome measures remain our best means of measuring healthcare quality for a particular procedure. Unfortunately, they are unable to assess whether the procedure should have been performed in the first place, and, thus, do little to assess healthcare cost effectiveness or efficiency.

Data Sources and Measurement Instruments

The source of data used to measure healthcare quality is another critical issue. Physicians are skeptical of data produced by outside stakeholders, such as government agencies or employer coalitions, because of concerns regarding the quality of the data and the validity of measures created using the data.[7,46,56] In general, physicians do not trust groups that are nonclinical to develop valid metrics that truly focus on quality (rather than cost).[40,52] Performance measures that lack clinical face validity or sufficient scope and sophistication tend to be poorly received and actively resisted by physicians.[7,52] Physicians generally prefer process measures that assess the correct clinical decisions and that appropriate diagnostic test or treatments are chosen, rather than those that assess outcomes, which are strongly influenced by patient factors outside a provider's control.[28,52,56]

In the absence of detailed electronic medical records (EMR) that are consistent and compatible across health systems, and in the presence of significant time pressure to "get started," the CMS and private third-party payers have focused on existing medical claims data because of its ready availability in electronic format. Unfortunately, medical claims data is structured for billing and reimbursement purposes, and it is coded to maximize reimbursement rather than accurately reflect the complexity of patient care and risk acuity, or the true relationship of a crude outcome to the performance of a specific medical or surgical intervention.

Drilling into the actual medical record of an individual or small group of patients is the most reliable means to assure proper linking of measured outcomes to specific medical or surgical interventions, as well as allow a fair risk adjustment for the complexity of patient care and risk acuity. Unfortunately, this approach is impractical from a logistical standpoint on an ongoing and mass scale, and would be cost prohibitive in the absence of universal EMRs. Risk adjustment is generally less effective and less accurate when administrative databases are used, because detailed and specific clinical context information is typically unavailable. Although some have asserted that the addition of a few simple clinical variables to existing administrative data would be sufficient to make risk adjustment comparable to that

which can be achieved with sophisticated, disease-specific databases developed by professional societies,[40,43,51] this assertion is far from proven, and certainly not accepted by most clinicians. Two recent neurosurgery examples will serve to illustrate this point.

An academic medical center with a multidisciplinary stroke service in Boston, Massachusetts was rated as the best hospital for stroke care in Massachusetts for 2003 based on the lowest reported mortality rate for the stroke diagnosis-related group (DRG) as calculated by Agency for Healthcare Research and Quality (AHRQ) risk-adjusted methodology and hospital report card reporting system. The next year, the same medical center was rated the second worst for stroke care in Massachusetts based on stroke DRG mortality rates calculated and reported using the same methodology. There had been no change in inpatient clinical care pathways or protocols in the intervening period and no change in medical staff or treatment philosophy. The mortified dean of the medical school ordered an immediate investigation. An in depth review of all 31 stroke DRG deaths for 2004 revealed only one instance of potentially avoidable care delivery morbidity.[29] All remaining deaths were expected based on the severity of patient's initial clinical condition and nonintervention choices appropriately made by family. The AHRQ reporting methodology was insensitive to these issues. The report card was public, with patients and referring physicians left to draw their own conclusions.

"Risk-adjusted" outcomes for key surgical procedures for California hospitals are routinely reported to the Office of Statewide Health Planning and Development (OSHPD). OSHPD then posts this data on their website for use by the general public and all interested parties. Recently, the state began tracking ventriculostomy mortality rates as a neurosurgical indicator of quality. In 2005, our institution's mortality rate for ventriculostomy was reported to be 55%, which was not only shockingly high (the mortality rate should be $<1\%$), but was far in excess of similar data reported for other state university hospitals (although one was listed with a 37% mortality rate). The figure raised immediate concerns from the offices of both the Dean and the Chancellor of the university, and an immediate and very time-consuming investigation was initiated.

In reality, the actual ventriculostomy-related mortality rate at our institution for 2005 was confirmed to be zero. It turns out that placing a ventriculostomy allows a patient to move out of the stupor and coma DRGs into DRG 1 (craniotomy age 17 yr, with complications and comorbidities), which carries much greater reimbursement for the hospital doing the coding. In patients too sick to stabilize and either recover spontaneously or recover after a real craniotomy intervention, who subsequently die, the ventriculostomy is listed as the major procedure accounting for their DRG status, and the mortality is attributed to the ventriculostomy rather than their underlying condition for which the ventriculostomy is inserted. Our coders were doing their best to help the hospital recoup revenue for caring for very sick patients. In fact, they were doing a better reimbursement coding job than at the comparison hospitals. However, the providers were paying the price in the form of negative value judgment and expenditure of significant time, effort, and perception damage control for the hospital coder's efficiency coupled with the use of a database to measure quality that was not designed or intended for that purpose.

The latest iteration of DRG risk adjustment is the "All Patient Refined" (APR) DRG system, which is currently being implemented across the country with the endorsement of the AHRQ, usually with the aid of software designed by 3M (DRG Assurance Program through 3M Consulting Services and 3M APR-DRG Software; 3M Health Information Systems, Salt Lake City, UT). This program purports to classify patients into clinically meaningful groups and then divide them into four severity of illness and four risk of mortality subclasses within each APR-DRG, for subsequent analysis. We had an opportunity to assess this system for several of our typical neurosurgical diagnosis, and our assessment led to serious concerns.

For example, most neurosurgeons will tell you, and a wealth of peer-reviewed clinical evidence supports the assertion, that the most likely significant variables for predicting severity of subarachnoid hemorrhage (SAH) and likely mortality from SAH include:

1) Patient clinical grade (e.g., Hunt-Hess classification,[30] Glasgow Coma Score,[58] etc.).
2) Computed tomographic (CT) scan Fisher grade.[27]
3) Patient age.
4) Presence of hydrocephalus.
5) Development of rehemorrhage.
6) Development of vasospasm.
7) Development of intracranial hypertension.
8) Requirement for intubation and mechanical ventilation.
9) The development of cerebral salt wasting or syndrome of inappropriate antidiuretic hormone release.
10) The development of status epilepticus, among others.

These factors are far more likely to be significant for predicting outcome in the specific diagnosis of SAH than such medical comorbidities as diabetes, hypertension, hyperlipidemia, coronary artery disease, etc. However, even with the APR-DRG system, the actual documented comorbidities that do count for risk adjustment for the principle diagnosis subclass of SAH are outlined in *Table 27.2*. You will notice that, with the possible exception of ventriculostomy (which could potentially be a secondary marker of presence of hydrocephalus or development of intracranial hypertension), none of the obvious clinically significant variables are taken into account, and none of the variables are diagnosis specific. We clearly have a long way to go with administrative claims

TABLE 27.1. Abbreviations used throughout the manuscript

AAFP	American Academy of Family Practitioners
AANS	Association of Neurological Surgeons
AARP	American Association of Retired Persons
ABNS	American Board of Neurological Surgery
ACGME	Accreditation Council on Graduate Medical Education
ACP	American College of Physicians
ACS	American College of Surgeons
AHIP	America's Health Insurance Plans
AHRQ	Agency for Healthcare Research and Quality (formerly AHCPR—Agency for Health Care Policy and Research)
AMA	American Medical Association
AOA	American Osteopathic Association
AOB	Annual operating budget
APR-DRG	All patient refined diagnosis-related group
AQA	Ambulatory Quality Alliance (formerly ACQA—Ambulatory Care Quality Alliance)
CABG	Coronary artery bypass graft
CAHPS	Consumer assessment of healthcare providers and systems
CEO	Chief Executive Officer
CMS	Center for Medicare and Medicaid Services (formerly HCFA—Health Care Finance Administration)
CNS	Congress of Neurological Surgeons
CQI	Continuous quality improvement
CSNS	Council of State Neurosurgical Societies
DRG	Diagnosis-related group
EMR	Electronic medical record
EBC	Evidence-based care
EBM	Evidence-based medicine
EMTALA	Emergency Medical Treatment and Active Labor Act
HER	Electronic health record
EPC	Evidence Practice Center
ESRD	End-stage renal disease
GAO	General Accountability Office (formerly the General Accounting Office)
HHS	Department of Health and Human Services
HIT	Health information technology
HMO	Health maintenance organization
HQA	Hospital Quality Alliance
HQID	Hospital Quality Incentive Demonstration
IOM	Institute of Medicine
MedPAC	Medicare Payment Advisory Committee
MMA	Medicare Prescription Drug Improvement and Modernization Act
NAS	National Academy of Science
NGC	National Guidelines Clearinghouse
NHQR	National Healthcare Quality Report
NQF	National Quality Forum
NSQIP	National Surgical Quality Improvement Program (of the ACS and the VA)
NTTA	National Technology Transfer and Advancement Act
OMB	Office of Management and Budget
OSHPD	Office of Statewide Health Planning and Development (CA)
P4P	Pay-for-performance

(Continued)

TABLE 27.1. *(Continued)*

PAC	Political action committee
PACCPQHI	President's Advisory Commission on Consumer Protection and Quality in the Healthcare Industry
PCPI	Physician Consortium for Performance Improvement (of the AMA)
PHCD	Physician-Hospital Collaboration Demonstration
PHQID	Premier Hospital Quality Incentive Demonstration
PRO	Peer Review Organization
PVRP	Physician Voluntary Reporting Program
QA	Quality assurance
QI	Quality improvement
QIO	Quality Improvement Organizations
QIW	Quality Improvement Workgroup (of the AANS/CNS Washington Committee)
SAH	Subarachnoid hemorrhage
SGR	Sustainable growth rate
SQA	Surgical Quality Alliance (of the ACS)
SSA	Social security act
STS	Society of Thoracic Surgeons
TQM	Total quality management
VA	Veteran's Administration

data before it becomes convincing, reliable, and fair for assessing quality when reporting and studying neurosurgical pathologies and procedures.

IOM AS A PRIMARY DRIVER

The IOM is one of the newer divisions of the National Academy of Science (NAS), which was founded as an honorific society by then President, Abraham Lincoln, March 3, 1863. Since its inception, the NAS has not only served as an honorific society, but has served to periodically advise the Office of the President as well as the United States Congress on issues related to the intersection of science and public policy. Members of the NAS had to be nominated by existing members of the NAS, and, as such, the society has functioned very much like a fraternity since its inception (? ΝΑΣ). The IOM (? ΙΩΜ) was established within the NAS in 1970. It currently has approximately 1440 members.

The IOM currently has significant issues with legitimacy when it comes to representing the interests and experience of practicing physicians to the federal government and planning for their future. This remains particularly true for practicing surgeons. Approximately 9.9% of IOM members are from Harvard (Harvard University, Harvard Medical School, and Harvard School of Public Health). This represents a 12.4-fold academic overrepresentation, given Harvard's position as only 1 of 125 United States accredited medical schools. Academicians who now occupy predominantly administrative and/or university leadership positions; nonclinical faculty, such as public health educators and epidemiologists; as well as academic clinicians who are not solely reliant on clinical revenue for their living are overwhelmingly represented. Those members that do actually see and directly take care of patients may only do so on an occasional weekly or monthly basis, and then usually in the protected environment of house officer insulation and amplification. Private practice clinicians are exceedingly rare and usually represent academicians who happened to return to a private practice environment, rather than physicians who have always made their living in private practice.

Surgeons are also scarce in the IOM. There are only 56 members in section 6a (the surgery subsection). Even given that other surgical subspecialists (e.g., ophthalmology, orthopedic surgery, otorhinolaryngology, etc.) would need to be accounted for, it is likely that the total surgical membership in the IOM is less than 144, which would be less than 10% of all IOM members. Unfortunately, neurosurgeons seem to be the scarcest of all. Currently, there are only seven neurosurgeons (0.5%) in the IOM (*Table 27.3*), and they are often excluded from IOM studies that directly involve and effect neurosurgery, such as the recently completed study on emergency medical services in the United States.[34]

The IOM cannot be considered to adequately or proportionately represent practicing physicians in the United States, let alone surgeons or neurosurgeons engaged in the care of United States citizens. The cross-sectional make up suggests a strong academic, public health, preventative medicine, and primary care bias, as well as a certain understandable practical naiveté regarding the realities and practical complexity of healthcare delivery in our country.

TABLE 27.2. All patient refined (APR) diagnosis-related group (DRG)

		Principle diagnosis of subarachnoid hemorrhage with secondary diagnosis of:	
		Severity of illness	**Expected mortality**
1	Minor	Hyperlipidemia	Dehydration
		Coronary artery disease	Malnutrition
			Chronic obstructive pulmonary disease (COPD)
2	Moderate	Dehydration	Dehydration, malnutrition
		Malnutrition	Malnutrition, COPD
		COPD	Dehydration, COPD
		Decubitus ulcer	Decubitus ulcer
		Congestive heart failure (CHF)	CHF
			Hypotension
		Hypotension	
3	Major	CHF, Hypotension, Dehydration	CHF, hypotension, dehydration
			Dehydration, decubitus, CHF
		Acute respiratory failure	
		Decubitus ulcer, Malnutrition	CHF, (PPx)—ventriculostomy
4	Extreme	One set from each of 1–3 above	One set from each of 1–3 above plus acute respiratory failure

In 1996, the IOM ceased to exist solely as an honorific society when it launched its own internal United States healthcare quality initiative. This extremely influential initiative has gone through three distinct phases. Phase One, from 1996 to 1999, dealt with documenting the seriousness and pervasiveness of the United States healthcare quality problem and culminated in the publication of *To Err is Human*.[38] Phase Two, from 1999 to 2001, dealt with defining the nature of the problem in terms of overuse, misuse, and underuse of healthcare services, and laid out the IOM's vision for how the healthcare system and related policy environment must be radically transformed in the publication, *Crossing the Quality Chasm*.[32] Phase Three, which began in 2002 and is still ongoing, tries to operationalize the IOM quality vision through multiple efforts focusing on reform in three overlapping levels of the system: the environmental level, the level of the healthcare organization, and the interface between clinicians and patients. Examples of Phase Three publications include *Health Professions Education: A Bridge to Quality*[33] and *Awarding Provider Performance: Aligning Incentives in Medicine*.[37]

In addition to its own internal agenda and efforts, this highly respected and influential private organization is regularly commissioned by government offices and agencies to perform specific healthcare studies and provide formal reports of their findings. Government agencies or offices involved include the Office of the President, the United States Congress, the Department of Health and Human Services (HHS), and the AHRQ. It may safely be asserted that there is no private organization that has had more influence on governmental healthcare agencies and on United States healthcare policy than the IOM.

As a result of both Phase One and Two of their healthcare quality initiative, the IOM concluded that healthcare today harms too frequently and routinely fails to deliver its potential benefits. In their *Crossing the Quality Chasm* publication, they made 13 specific recommendations for improving healthcare in the United States.[32] The IOM defined healthcare quality as the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge. They specifically enumerated that quality healthcare must fulfill six characteristics. It must be safe, effective, efficient, patient centered, timely, and equitable. They noted unexplainable variation in use of healthcare resources and performance of procedures geographically and physician-to-physician across the United States. Most importantly, they noted that quality healthcare should not just be measured by the quality of a service delivered, but is also reflected in the misuse, underuse, and overuse of healthcare services.

CMS AS A PRIMARY DRIVER

Formerly known as the Health Care Finance Administration (HCFA), CMS is a major division of the Department

TABLE 27.3. Neurosurgeons who are members of the Institute of Medicine[a]

Henry Brem
Mahlon DeLong
Julian Hoff
Edward Laws
Robert Martuza
Bryce Weir
Charles Wilson

[a]n = 7; 0.5% of Institute of Medicine (IOM) membership.

of HHS. Since the Medicare and Medicaid programs were signed into law under Title XIX of the Social Security Act (SSA) in 1965, CMS has been the primary agency tasked with their administration. As such, CMS is the primary force involved with P4P. The major statutory supports and basis for CMS legally implementing P4P are outlined in *Table 27.4*. Just as the definition and measurement of quality in healthcare has undergone evolution since the 1960s, the interpretation and implementation of healthcare quality by CMS has undergone a similar parallel evolution.

CMS QI Roadmap

MMA 2003 was probably the most important and significant piece of legislation regarding improving Medicare since the SSA of 1985, which first called for quality control through peer review. Although many of us listening to the politicians talk about the MMA on television were led to think that this legislation was predominantly focused on a Medicare drug prescription reform, the legislation actually had far more reaching requirements and implications. It was the requirements of the MMA that led CMS to proceed through a strategic planning process that ultimately led to the CMS P4P program as part of the CMS QI Roadmap.[17] The MMA also serves as the legislative authority authorizing CMS to proceed with P4P initiatives and demonstrations.

TABLE 27.4. Legal sources of authority for Center for Medicare and Medicaid Services (CMS) pay-for-performance

SSA 1982—Social Security Act
- ○ Medicare use and quality control peer review program

NTTA 1995, OMB Circular A-119, revision 1998—National Technology Transfer and Advancement Act, Office of Management and Budget Circular (Executive Office of the President of the United States)
- ○ If medical quality indicators are endorsed by voluntary consensus standard bodies (e.g., National Quality Forum, and Ambulatory Quality Alliance), the government is obligated to adopt them

BIPA 2000—Medicare, Medicaid, and State Childrens Health Insurance Program Benefits Improvement and Protection Act

MMA 2003—Medicare Prescription Drug Improvement and Modernization Act
- ○ Pay-for-performance initiatives and demonstrations

The CMS strategic planning process took place from 2003 to 2005. The strategic plan identified four strategies that, if adopted by *providers*, could lead to high healthcare performance. These four key strategies were:

1) Measurement and reporting of quality.
2) Adoption and use of HIT.
3) Redesign of care processes.
4) Change in organizational culture and management.

The final CMS strategic plan, which was issued in July 2005 as the QI Roadmap,[17] was very strongly influenced by the results of the first two phases of the IOM, which were completed in 1999 and 2001, respectively.[32,38]

The CMS QI Roadmap articulated a vision for United States Healthcare that was "*the right care for every person every time*." It articulated six goals for improving United States healthcare that were lifted directly from the IOM's *Crossing the Quality Chasm*,[32] namely, that care funded by CMS needed to be:

1) Safe.
2) Effective.
3) Efficient.
4) Patient centered.
5) Timely.
6) Equitable.

The CMS QI Roadmap identified five system strategies for improving healthcare funded by CMS. These system strategies are:

1) Work through partnerships (within CMS, with Federal and State agencies, and with nongovernmental partners).
2) Publish quality measurement information (including both the beneficiary audience and the professional/provider/purchaser audience).
3) Pay in a way that expresses commitment to quality and rewards, rather than inadvertently punishing providers and practitioners for doing the right thing.
4) Promote HIT.
5) Become an active partner in creating and using information regarding the effectiveness of healthcare technologies to bring effective innovations to patients more rapidly and to monitor the effectiveness of technologies for which they are paying.

Implementation of the CMS QI Roadmap required a restructuring of CMS in a way that reflected the strategic initiative and mission and facilitated roadmap focus and implementation. The basic outline of this new structure as it relates to physician P4P is represented diagrammatically in *Figure 27.1*. A CMS Quality Council was formed to meet biweekly and report directly to the CMS Director. This council would be supported by a CMS Quality Coordination Team that would directly manage, track, and plan the road-

map implementation process, report regularly to the Quality Council, and provide technical support to the two levels of working groups assigned to specific implementation tasks. The first level of working groups would each focus on one of eight main CMS QI Roadmap tasks or projects. The first-level working group concerned with physician P4P is called the P4P Forum and is charged with performance measurement as it relates to P4P. There are five second-level working groups within the P4P Forum, and the one specifically tasked with physician P4P measurement is the Physician Quality and Cost workgroup. Each of these layers of new CMS structure are populated and led by government appointees within CMS that can change relatively frequently, and the CMS bureaucracy can be difficult to penetrate through routine search strategies to identify all of the individuals involved.

AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

The AHRQ was established in December 1989 as the Agency for Health Care Policy and Research (AHCPR) under Public Law PL 101–239. Just like CMS, it is one of the major sister government agencies within the Department of HHS. After the collapse of the healthcare reform debate in 1993 to 1994, intramural analysis of national outcomes data and creation and dissemination of national clinical guidelines was effectively removed from the scope of its mission. In 1999, its name was changed to AHRQ, in part to reflect elimination of direct influence on United States healthcare policies. Currently, approximately 80% of its budget goes to funding extramural Evidence-Based Practice Center (EPC) grants (originally 13 in 2002, now down to 11; see *Table 27.5*). The extramural EPCs are now tasked with studying EBC and producing Evidence-Based Clinical Practice Parameter Guidelines. The AHRQ supports the internet-based National Guidelines Clearinghouse (NGC; http://www.guideline.gov) in partnership with the American Medical Association (AMA) and the America's Health Insurance Plans (AHIP) institute.

Of note, the AHRQ produces the annual National Healthcare Quality Report (NHQR), which has been produced every December since 2003. The NHQR reports United States hospital healthcare performance on 176 core quality measures. Currently, there are no neurosurgery core measures among the 176 measures analyzed.

The AHRQ is also the government agency that has tasked and granted the Ambulatory Quality Alliance (AQA) with implementing authority for quality and efficiency measures that relate to P4P. In addition to influencing and assisting CMS with their intramural P4P efforts, AHRQ is actively promoting and assisting with P4P programs for nongovernment, private third-party payers.[24] The AHRQ is also in the process of developing a national uniform patient satisfaction measurement instrument, which they have titled the Consumer Assessment of Healthcare Providers and Systems (CAHPS).[1]

GOVERNMENT ACCOUNTABILITY OFFICE

Formerly known as the General Accounting Office, the Government Accountability Office (GAO) was founded in 1921 as an independent, nonpartisan legislative branch agency that works as an investigative arm for Congress and also advises heads of executive agencies (e.g., HHS and AHRQ, CMS). The GAO essentially serves as a "congressional watchdog," tasked with studying and financially auditing any area, program, or initiative that relies on tax expenditures. The GAO has extensive independent discretionary authority, power, and influence. The current Comptroller General of the GAO is David M. Walker, who is the seventh comptroller general in the agency's history. The GAO has a formal Health Care division, which is empowered and authorized to audit the Medicare and Medicaid programs and provides direct advise and testimony to Congress as well as the Secretary of the HHS and the Directors of CMS and AHRQ. The current Managing Director for Health Care within the GAO is Marjorie E. Kanof. One of the special functions of the Comptroller General of the GAO is that he/she is the government official authorized to appoint new members and member 3-year term renewals, as well as the Chairman of the Medicare Payment Advisory Committee (MedPAC).

MEDICARE PAYMENT ADVISORY COMMITTEE

MedPAC is an independent federal body established by the Balanced Budget Act of 1997 (PL 105–33) and tasked with advising Congress on issues effecting Medicare. MedPAC reports to Congress twice every year. MedPAC has 17 members appointed by the Comptroller General of the United States GAO in 3-year, renewable, staggered terms. The Chairman of MedPAC is also appointed by the Comptroller General of the United States GAO. Although appointed by the GAO, MedPAC is independent of the GAO chain of command. Although still government appointees, members are private citizens who serve part time. MedPAC is, thus, shaded both light blue and green in *Figure 27.1*. As a small independent quasi-private agency, MedPAC has extensive influence on Congress, CMS, and AHRQ regarding Medicare issues.

Table 27.6 lists the current members of MedPAC. Although the list includes five physicians (29.4%), it should be noted that only two of these are practicing physicians who rely on Medicare for practice income (one academic general surgeon, one private practice urologist), which only make up 11.8% of MedPAC. The remaining three physicians are either health maintenance organization (HMO) physicians (independent of Medicare), a health systems chief executive officer (CEO), and a CEO of a healthcare purchasing consulting firm. As is readily apparent on perusal of *Table 27.6*, Med-

TABLE 27.5. Evidence-based practice centers (EPCs) receiving federal grants from the Agency for Health Care Research and Quality to produce evidence-based clinical guidelines[a]

Blue Cross and Blue Shield Association, Technology Evaluation Center (TEC), (in collaboration with Kaiser Permanente)	Chicago, IL http://www.bcbs.com/tec/index.html
• Naomi Aronson, Ph.D., Executive Director	
• David M. Eddy, M.D., Ph.D., Scientific Advisor	
Duke University, Center for Clinical Health Policy Research (CCHPR)	Durham, NC http://www.clinpol.mc.duke.edu/
• David B. Matchar, M.D., Co-Director	
• Douglas McCrory, M.D., Co-Director	
ECRI—Emergency Care Research Institute	Plymouth Meeting, PA http://www.ecri.org/
• Charles Turkelson, Ph.D., Project Manager	
Johns Hopkins EPC	Baltimore, MD http://www.jhsph.edu/epc
• Eric B. Bass, M.D., M.P.H., Director	
McMasters University EPC	Hamilton, Ontario, Canada http://hiru.mcmaster.ca/epc/
• Parminder Raina Ph.D., Director, EPC	
Metaworks, Inc.	Boston, MA
• 1997–2001, now defunct	
Oregon, EPC (OHSU, Portland VAMC, and Kaiser Permanente collaboration)	Portland, OR http://www.ohsu.edu/epc/
• Mark Helfand, M.D., M.S., M.P.H., Director, EPC	
RTI-UNC EPC (Research Triangle Institute and University of North Carolina, Chapel Hill collaboration)	Chapel Hill, NC http://www.rti.org/epc/home.html
• Kathleen Lohr, Ph.D., Co-Director, RTI	
• Timothy Carey, M.D., M.P.H., Co-Director, UNC	
Southern California—RAND, EPC (RAND, UCLA, UCSD, USC, Cedars-Sinai Medical Center/ZYNX Health, Children's Hospital Los Angeles collaboration)	Santa Monica, CA http://www.rand.org/health/epc/
•• Paul G. Shekelle, M.D., Ph.D., Director	
• Sally C. Morton, Ph.D., Co-Director	
Stanford-UCSF, EPC (Stanford-UCSF collaboration)	Stanford, CA http://healthpolicy.stanford.edu/stanford-ucsf-epc/
• Douglas K. Owens, M.D., M.S., Director	
• A. Eugene Washington, M.D., M.Sc., Co-Director	
Tufts-New England Medical Center, EPC	Boston, MA http://www.nemc.org/dccr/Evidence-based%20Practice.htm
• Joseph Lau, M.D., Director	
University of Alberta, EPC (University of Alberta and Capital Health Authority in Edmonton collaboration)	Edmonton, Alberta, Canada http://www.epc.ualberta.ca/index.htm
• Terry Klassen, M.D., M.Sc., Director	
University of Minnesota, EPC	Minneapolis, MN http://evidence.ahc.umn.edu
• Robert Klane, M.D., Director	
University of Ottawa, EPC	Ottawa, Canada http://www.uo-epc.org/index.html
• Howard Schachter, Ph.D., Co-Director	
• David Moher, M.Sc., Co-Director	
University of TX HSC, San Antonio, EPC	San Antonio, TX
• 1997–2001; now defunct	

[a]As of June 2002.

PAC is dominated by health policy consultants (29.4%), health system CEOs (17.6%), healthcare economists and business administrators (11.8%), and healthcare purchasing consultants (11.8%). Practicing physicians, let alone surgeons, who have direct personal experience with Medicare reimbursement and are at least partially dependent on Medicare reimbursement for their livelihood, are in a distinct minority.

TABLE 27.6. Current members of Medicare Payment Advisory Committee (MedPAC) with terms of appointment

2006–2009

Glenn Hackbarth, J.D.—Chairman

- Lawyer; former chief executive officer (CEO) of a healthcare group; former government official with Department of Health and Human Services

Robert Reischauer, Ph.D.—Vice Chairman

- Public health consultant; president of the Urban Institute; member of the Institute of Medicine (IOM)

Mitra Behroozi, J.D.

- Lawyer; executive director of a labor union pension fund

Karen Borman, M.D.

- Physician; Professor of Surgery, vice president for Education at the University of Mississippi; American College of Surgeons (ACS) C & R Committee; American Medical Association (AMA) procedural terminology panel

Ronald Castellanos, M.D.

- Physician; practicing urologist in Florida; Chair of Professional Physicians Advisory Council (PPAC)

Douglas Holtz-Eakin, Ph.D.

- Economist; former director of the Congressional Budget Office (CBO)

2005–2008

Nancy-Ann DeParle, J.D.

- Lawyer; senior advisor to JP Morgan Partners; board member National Quality Forum (NQF); formerly with Health Care Finance Administration (HCFA); formerly director for health and personnel at the White House Office of Management and Budget (OMB)

David Durenberger, J.D.

- Lawyer; senator, chairman of the National Institute of Health Policy; Board member NCQA

Jennie Chin Hansen, R.N., M.S.N.

- Nurse; board member AARP, Member of Lumestra (CA QIO)

Nancy Kane, D.B.A.

- Business administration; directs HSPH Healthcare management program, former physical therapist

Nicholas Wolter, M.D.

- Physician; pulmonary and critical care, Healthcare Administrator, CEO of Billings Clinic

2004–2007

John Bertko, F.S.A.

- Actuary; vice president and chief actuary, Humana, Inc.

Shiela Burke, R.N., M.P.A.

- Nurse; deputy secretary and chief operating officer, Smithsonian Institution; former Executive Dean of Public Policy, Harvard University; former Chief of Staff of Senate Majority Leader; Member IOM

Francis Crosson, M.D.

- Physician; executive director, Permanente Medical Group

Arnold Milstein, M.D., M.P.H.

- Physician; health care purchasing consultant, medical director, Pacific Business Group on Health; co-founder of Leapfrog Group

Ralph Muller, M.A.

- Hospital administrator; CEO University of Pennsylvania Health System; board member, NCQA

William Scanlon, Ph.D.

- Health policy expert/advisor; former managing director of healthcare issues United States Government Accountability Office (GAO)

The influence of MedPAC on Congress and on CMS through its testimony to Congress should not be underestimated. It was MedPAC that recommended, during Congressional testimony and in their subsequent report to Congress, that CMS use claims data to measure physician's resource use and educate them regarding their performance relative to their peers, that initially 2% and eventually up to 50% of physician Medicare reimbursement eventually be linked through P4P, and that P4P for physicians should be budget neutral through reductions in the physicians fee schedule (March 2005).[49] It was also MedPAC that reported to Congress in March of 2006 that P4P was ready for implementation and that recommended moving forward.[50]

QI ORGANIZATIONS

In accordance with Title XI of the SSA of 1982, which was the first Congressional attempt to improve Medicare through institution of some form of quality control, CMS established the Medicare Utilization and Quality Control Peer Review Program. Its goal was to improve the efficiency,

effectiveness, economy, and quality of the services delivered to Medicare beneficiaries. The first iteration of this program consisted of the formation of physician Peer Review Organizations (PROs) for each state to respond to beneficiary complaints of poor care. PROs led to individual cases of care improvement related to individual providers, and quality of physician care was assessed by physicians. However, the changes achieved were not systematic, system wide, or quantifiable, and the process was primarily complaint and appeal driven.

In response to an IOM study and report in the early 1990s, CMS changed the name of PROs to QI Organizations (QIOs). Coincident with the name change was an expanded mission that supplemented peer individual case review with the collection of limited crude quality measures data for State Medicaid participants, along with the offer of technical support assistance and advice to providers. The goal was to include the beginnings of quality measurement and improvement along with physician peer review by identifying and disseminating best-practice information and offering training and assistance to providers on a voluntary basis.

QIOs are independent organizations contracted with CMS. There are currently 53 CMS-contracted QIOs (one for each state, the two territories, and the District of Columbia). Contracts are bid with the following requirements: 1) all work for the full range of QIO activities must go to one contractor per state, 2) the QIO must meet physician sponsorship and/or physician access criteria. The QIO for California is Lumetra, based in San Francisco.

The MMA of 2003 tasked the IOM with reviewing the CMS QIO program, and this report was received in March 2006.[36] The IOM recommended: 1) divesting QIOs of their beneficiary complaint and appeal management role, 2) making QIOs regional or national to eliminate duplication and increase efficiency, 3) eliminating the QIO physician sponsorship and/or access criteria requirements, and 4) narrowing the focus of QIOs to that of technical assistance to providers for performance measurement and QI. The IOM did not recommend doing away with QIOs and thought that the program should be redesigned to support the processes of national reporting of performance measures. As a result of the IOM report, it seems that CMS intends to reengineer QIOs in a manner that better fits with their QI Roadmap and, in particular, the P4P initiative. The exact role that QIOs will take in advising CMS on matters related to P4P as well as administering and monitoring the program is not yet clear, but is likely to be significant.

It should be noted that recommendations 2 and 3 above would require statutory change because these requirements are specifically spelled out in the SSA of 1982. Recommendation 3 is particularly worrisome from the standpoint of practicing physicians, because it would further disconnect them and further mute their voice and diminish their leadership role in measuring and assessing the quality of healthcare delivered by physicians. The QIO program website is www.medqic.org.

ALLIED PRIVATE ORGANIZATIONS AND THE DEVELOPMENT, ENDORSEMENT, AND IMPLEMENTATION OF PHYSICIAN QUALITY MEASURES

The Department of HHS, including CMS and AHRQ, has supported the conclusions of the CMS QI Roadmap. The first key system strategy in the QI Roadmap is working through partnerships that include private nongovernmental organizations. The area in which this strategy has moved forward fastest is with the development of quality measures for measuring and reporting the P4P initiative. Regarding physician quality measures, a series of Executive Office Circulars and CMS and AHRQ contracts have led to the general schema outlined in *Figure 27.2*. In essence, the original plan was for physician quality measures to be developed by the Physician Consortium for Performance Improvement (PCPI) of the AMA, to be endorsed by the National Quality Forum (NQF), and to be implemented by the AQA.

AMA Physician Consortium for Performance Improvement

The PCPI was convened by the AMA in 2000. It consists of more than 100 national medical specialties and state medical societies, the Council of Medical Specialty Societies, the American Board of Medical Specialties (and two of its member Boards), AHRQ, and CMS. The American Association of Neurological Surgeons (AANS) and the Congress of Neurological Surgeons (CNS) are both represented through a Washington Committee, QI Workgroup (QIW) representative, who is currently Dan Resnick. The ABNS is

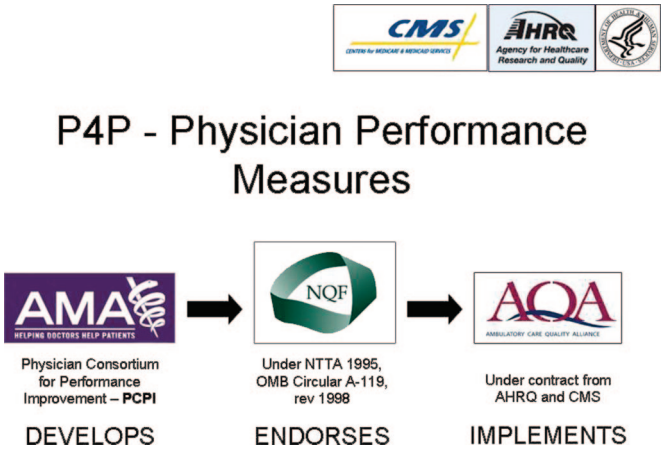

FIGURE 27.2. A diagrammatic representation of the physician P4P quality and efficiency measure development, endorsement, and implementation process.

not represented. The AMA-PCPI is tasked with developing evidence-based physician performance measures. Measure development can proceed through intramural initiative or in response to requests from other agencies, such as the NQF or the AQA. Measure development proceeds through a focused working group mechanism.

The National Quality Forum

The President's Advisory Commission on Consumer Protection and Quality in the Healthcare Industry (PACCPQHI) was established by Executive Order 13017 on September 5, 1996, by then President, William Clinton.[26] The first report of the PACCPQHI, November 20, 1997, called for establishment and endorsement of a Consumer Bill of Rights and Responsibilities (Patient's Bill of Rights).[54] Their final report entitled, *Quality First: Better Healthcare for All Americans*, called for formation of a single forum of providers, business, labor, consumers, insurers, and government to set healthcare quality standards for measurement and reporting.[55] The goal was to identify and endorse "valid," consensus-based, quality measures. The NQF was incorporated as a private organization in May 1999.

In 1998, the Office of Management and Budget (OMB) of the Executive Office of the President of the United States issued a revision of their OMB Circular A-119 relating to the National Technology Transfer and Advancement Act (NTTA) of 1995.[53] In this circular revision, it is directed that if medical quality indicators are endorsed by voluntary consensus standard bodies (e.g., the NQF), the government is obligated to adopt them.

As of September 2006, the NQF had 335 members. As can be seen in *Table 27.7*, the NQF is dominated by hospitals, hospital associations, and integrated health delivery networks (36.1%), as well as quality certification bodies, QI associations, and healthcare management and consulting groups (23%). Physician associations, coalitions, or group practices currently make up only 11.9% of members. Broad representation is one of the strengths of the NQF for consensus input, particularly because it not only endorses physician quality measures, but must also do so for hospitals and health plans. However, the representation is skewed in a way that leads to bias based on its current financial membership admission structure, as outlined below. Currently, neither the AANS, CNS, Washington Committee, nor the ABNS are members, and our only voice is indirect, through the AMA and the American College of Surgeons (ACS).

The NQF must carry on its mission without a stable source of funding. Indeed, its sole source of funding seems to be the dues fees that NQF members are assessed annually. You have to pay to be part of the NQF. In this setting, you might suspect that dues fees would be proportional to annual budget across special interests to not skew membership based on cost concerns as well as to maximize revenue from larger, more wealthy organizations. However, the reality is the opposite. Currently, to be a member of the NQF, a physician organization or coalition must pay $15,750 if their annual operating budget (AOB) is higher than $10,000,000 (0.16% AOB). Whereas, to be a member of the NQF, a health plan, healthcare providing institution, corporate employer, or other healthcare purchaser only has to pay $15,750 if their AOB is between $1,000,000,000 and 1,900,000,000 (0.0008–0.0016% AOB). The membership playing field for physician organizations and healthcare corporate entities is dramatically uneven, by more than 100-fold (2 log difference in percent of AOB requirement for membership). Given this disparity in proportional cost of membership, it should be no surprise that national neurosurgery organizations and certification boards are not members whereas the individual community hospital down my street in Newport Beach, CA is a full member of the NQF.

Other than a restricted and skewed membership as outlined above, the other major problem with the NQF is its approach to endorsing "valid," consensus-based, quality measures. To most physicians and scientists, validity has a very specific meaning and must be both internal and external. In other words, quality measures must be accurate, nonconfounded, reproducible, and have low interobserver variability (internal validity), but must also apply across environments and circumstances and be linked to the desired outcome(s) in question by strong empiric evidence (external validity). For the NQF, however, measure "validity" is assessed only on

TABLE 27.7. Membership composition of the National Quality Forum (NQF) (335 members)

121 (36.1%)	Hospitals, hospital associations, integrated health delivery networks
77 (23.0%)	Certification bodies, quality improvement associations, healthcare management and consulting groups
40 (11.9%)	Physician associations, coalitions, or group practices
24 (7.2%)	Patient advocacy or watchdog groups and employee unions
23 (6.9%)	Federal, state, or city agencies (including Center for Medicare and Medicaid Services [CMS] and Agency for Healthcare Research and Quality [AHRQ])
20 (6.0%)	Insurance companies, 3rd-party payers, health maintenance organizations (HMOs)
13 (3.9%)	Large employers or employer healthcare purchasing consortiums
10 (3.0%)	Drug, implant, or medical supply companies
5 (1.5%)	Pharmacist or pharmacy associations
5 (1.5%)	Nursing associations
2 (0.6%)	Optometry associations

data fidelity: namely can it be measured? Whether there is an evidence-based proven rationale for measuring it is not a requirement, and the decision regarding whether or not the measure should be included is left to expert consensus rather than an evidence-based medicine (EBM) approach that its fulfillment improves quality (clinical outcomes or public health).

The NQF is organized into a steering committee as well as technical advisory panels composed of experts in the field in question. The NQF is perceived by P4P government agencies to possess five key attributes. Namely, a standardized body, an openness of process, a spectrum of interests across all stakeholders, due process for decision making, and a consensus approach.

Ambulatory Quality Alliance

The AQA was originally known as the Ambulatory Care Quality Alliance (ACQA). It was formed as a collaborative effort initiated September 2004 by the American Academy of Family Physicians (AAFP), the American College of Physicians (ACP), AHIP, and the AHRQ. As such, nonsurgical primary care societies are the lead physician organizations in the AQA.

Currently, the AQA has more than 125 members representing physicians, consumers, employers, government, health insurance plans, and accreditation/QI programs. The steering group is composed of the American Association of Retired Persons (AARP), the AHRQ, the AAFP, the ACP, the ACS, the AMA, the American Osteopathic Association (AOA), the AHIP, the National Partnership for Women and Families, the Pacific Business Group on Health, and the Society of Thoracic Surgeons (STS). Physician societies make up 54.5% of the membership (45.5% insurance, government, and patient advocates), however, there is very poor surgical representation. Only 18% of steering group members are surgical societies (ACS and STS). Neurosurgery is even more poorly represented. Currently neither the AANS, CNS, Washington Committee, nor the ABNS are members, and our only voice is indirect, through the AMA, the ACS, and the Surgical Quality Alliance (SQA; which represents only 0.8% of the general AQA membership).

The AQA is formally contracted with both the AHRQ and CMS to design and perform pilot studies for public reporting on quality measures. In July 2006, the AQA joined with the Hospital Quality Alliance (HQA) into a new National Quality Alliance Steering Committee. The AQA is charged with transparent reporting for both public and private entities of both quality and cost-of-care measures. The AQA initially operated under an informal general principle that only NQF-approved measures will be implemented through their pilot programs. However, it is clear that, similar to the NQF, the AQA meets all of the requirements specified by the NTTA OMB Circular A-119 revision 1998 for a voluntary consensus standard body, and it seems probable that the AQA will soon proceed with endorsing as well as implementing quality measures for P4P independent of the NQF and/or before NQF endorsement.

Surgical Quality Alliance

The SQA is a coalition of 13 surgical societies (see *Table 27.8*) functioning under the rubric of the ACS. Both the AANS and the CNS joined the SQA and are represented by a representative from the QIW of the Washington Committee. Those representatives are currently Robert Harbaugh and Gary Bloomgarden. The hope was that the SQA would be considered to be for surgery, whereas the AQA is considered to be for primary care and internal medicine, and would, thus, provide an equally strong voice for surgery with CMS and the AHRQ.

Unfortunately, the SQA has not achieved that status. Unlike the situation with the AQA; insurance, government, and patient advocate groups have not become members of the SQA, and, thus, it does not meet all of the requirements specified by the NTTA OMB Circular A-119 revision 1998 for a voluntary consensus standard body. The SQA has no seat independent of the ACS at the NQF and is not contracted with AHRQ and CMS for surgical quality measure implementation. It is part of the AQA, but instead of being a parallel counter-balancing organization, constitutes only 0.8% of AQA membership.

National Surgical QI Program

The ACS became interested in developing a nongovernmental form of the National Surgical QI Program (NSQIP) after reviewing the Veterans Administration's (VA) experi-

TABLE 27.8. Membership of the Surgical Quality Alliance (13 surgical societies)

American Academy of Ophthalmology
American Academy of Otolaryngology, Head and Neck Surgery
American Association of Neurological Surgeons
American Association of Orthopedic Surgeons
American College of Osteopathic Surgeons
American College of Surgeons
American College of Anesthesiologists
American Society of Cataract and Refractory Surgery
American Society of Colon and Rectal Surgeons
American Society of General Surgeons
American Society of Plastic Surgeons
American Urological Association
Congress of Neurological Surgeons
Society for Vascular Surgery
Society of American Gastrointestinal Endoscopic Surgeons
Society of Thoracic Surgeons

ence with their own national surgical risk study. The VA National Surgical Risk Study was performed from 1991 to 1993, and the results were presented in 1994, resulting in the VA-NSQIP.[42] The initiative proved so effective that, in 2002, the IOM sanctioned the VA-NSQIP as the "best in the nation" for measuring and reporting surgical quality and outcomes at a hospital level.[35] The ACS initiative began in 2001 with a grant from the AHRQ to fund the initial study for non-VA hospitals.[4] Beginning in 2004, the ACS began offering NSQIP as a subscription service for hospitals for a fee of $35,000/yr.

NSQIP requires formal training of dedicated nursing personnel to record and track key data on the preoperative, intraoperative, and postoperative course of every surgical patient, which is then entered into the NSQIP software model to calculate expected rates of outcomes, which can be later compared against actual measured data in a surgical quality report card fashion.

NSQIP is currently designed and intended for measuring surgical quality at an inpatient and hospital level. However, few adjustments in software design and reporting format would be necessary to apply the program to an outpatient surgery setting or to allow results to be analyzed and scrutinized by provider. NSQIP is already considered a major quality measure input into the NQF for hospital quality measures, and it would not be surprising if it became a source for physician quality measures for P4P as well. It would have the distinct advantage of being applicable to all neurosurgical procedures at once, rather than validating measures for many different procedures and diagnoses one-at-a-time through expensive and time-consuming individual outcome studies. It would also have the distinct advantage of allowing for stringent empirical clinical validation through iterations of mathematical modeling and comparison with actual outcomes. As such, it would free evidence-based validation requirements from only relying on peer-reviewed literature studies, without sacrificing verification through objective measurable evidence. It is sole reliance on peer-reviewed literature and study methodology stringency that slows and limits EBM clinical practice guideline development.

The major issue regarding NSQIP for neurosurgery is that the calculation model was generated predominantly using general and vascular surgery cases. Very little neurosurgery operative case data was involved. As a result, it is not clear that its validity for general and vascular surgery case assessment applies equally well to neurosurgery cases. This area remains to be explored. Efforts are currently underway at the University of Michigan to begin exploring the applicability and usefulness of NSQIP for assessing and reporting quality for neurosurgical operative procedures on a provider level for departmental performance improvement purposes (John A Cowan, Jr., personal communication, October 11, 2006).

CURRENT CMS P4P DEMONSTRATION PROJECTS AND STATUS

There are several principles that seem to be operative in the CMS implementation approach to P4P. The first would seem to be pilot in demonstration projects first and then generalize to system-wide. The second would seem to be pilot first in the hospital setting in which data is more readily available and buy-in is greatest, and then begin piloting with physicians. The third would seem to be begin pilot studies with less objectionable quality measures (process and outcomes measures—how well did you do what you did) and then progress to efficiency measures (cost measures and procedure appropriateness assessment—should you have done what you did and did you do it cheaply enough). The fourth seems to be to start with a low amount of reimbursement at risk (e.g., 2%) and then progressively ratchet that amount up to the MedPAC-recommended penetration level of 50%.

There also seems to be an overwhelming rush to implementation within CMS, even at the potential expense of initiative optimization and fairness, as well as potential unintended negative downstream consequences. To paraphrase a senior CMS representative who met with the AANS/CNS Washington Committee July 7, 2006: "There is a sense of urgency at CMS regarding P4P. CMS has too little time to consider other models. CMS does not want to hear, and is not receptive to, resistance. CMS is only open to practical solutions that can be applied or implemented within the perceived Congressional timeline." —Tom Valuck M.D., Medical Officer and Senior Advisor, Center for Medicare management, CMS

HOSPITAL P4P DEMONSTRATION INITIATIVES

Understanding and briefly reviewing CMS demonstration initiatives for hospital P4P is important for gaining insight into the strengths and weaknesses of the initial CMS approaches as well as potential design features that will also likely be applied to initiatives for physicians. Two CMS hospital P4P initiatives for Medicare are currently ongoing and a third is in place for Medicaid.

In 2003, CMS began the Hospital Quality Incentive Demonstration (HQID).[11] This demonstration initially involved 10 core quality measures, which was later expanded to 17. Hospital participation was voluntary; the goal was to asses how many hospitals would be interested in participating in similar programs in the future. *Hospitals that voluntarily participated and performed well would receive an increase in DRG payments of 0.4% above the standard DRG reimbursement*. This represented additional reimbursement and was not budget-neutral regarding all hospital DRG reimbursement system wide. With additional money available to claim, it should be no surprise that, by 2005, 98.6% of nonfederal United States hospitals were participating.

Also in 2003, CMS began the Premier HQID (PHQID).[16] This demonstration included 300 hospitals that were assessed on 34 quality measures related to five clinical conditions—heart attack, heart failure, pneumonia, coronary artery bypass grafting (CABG), and hip and knee replacements. Demonstration design dictated that, at the end of the first year, the top 10% performing hospitals would receive a 2% bonus, the next 10% would receive a 1% bonus, and the next 30% would not receive additional reimbursement, but would benefit from published public recognition of quality.

Once again, *hospitals that performed well received additional reimbursement rather than return of a budget-neutral withhold.* This demonstration project is notable for establishing public reporting of quality without additional reimbursement as a potential benefit and positive reinforcement lever. The progression design for the second year of the demonstration was also notable. Based on first-year demonstration data, the project called for setting performance data baselines for the lowest performing 10% and 20% of hospitals. During the second year of the project, the rewards realizable during the first year would remain unchanged, but now there would be a 2% reimbursement cut for the hospitals performing in the lowest 10%, and a 1% cut for the hospitals performing in the next-lowest 10%. There would now be "a stick along with the carrot." It is very probable that we will see some form of both of these features in future physician P4P demonstration projects.

In 2005, CMS began hospital P4P demonstration projects for Medicaid in at least 12 states (CA, IA, MD, MI, NE, NJ, NM, NY, NC, PA, RI, and DC).[12] With these projects, the hospitals would receive *additional incentive reimbursement payments for performing well* on the 17 quality measures in the Medicare HQID demonstration. Nonremunerative incentives in the form of public reporting as one of the highest quality hospitals in the program, similar to that designed into the PHQID Medicare demonstration, were also included. However, an additional potential nonremunerative incentive in the form of eligibility for preferential auto-enrollment for the top-performing health systems was included for the first time.

PHYSICIAN P4P DEMONSTRATION INITIATIVES

Physician-Hospital Collaboration Demonstration

In 2006, CMS began a P4P demonstration project called the Physician-Hospital Collaboration Demonstration (PHCD).[14] In essence, it allows hospitals under a demonstration initiative to legally transfer Medicare Part A funds that come to hospitals under the DRG system to physicians, based on mutually achieved cost savings through inpatient care pathway development (savings from decreased services, discounted surgical implants from volume contracting, reduced average length of stay [ALOS], etc.). For CMS, this is attractive because it allows hospitals and physicians to work together to lower costs (presumably without quality reduction) by aligning incentives across both Part A and B reimbursement. To this point, physicians have already been under tremendous pressure to improve hospital margin by reducing ALOS without getting anything for it, and, in fact, we hurt ourselves by lowered work relative value units when current procedural terminology (CPT) codes are revalued in the setting of lowered ALOS. The interesting part of this initiative is that the additional reimbursement for physicians would not come from an initial "below the line" withhold, but through transfer of a portion of Medicare Part A funds from the institutions affiliated with the groups, based on savings realized in the delivery of care. The concept is very similar to that of "gainsharing" demonstrations that have been attempted in the past but were thwarted by legal challenges regarding the mixing of Part A and Part B funds.

Physician Voluntary Reporting Program

Also in 2006, CMS initiated the Physician Voluntary Reporting Program (PVRP).[15] This demonstration initiative initially involved 36 core quality measures (process measures) approved by NQF, which was subsequently reduced to a 16-measure starter set. Physician participation is voluntary and currently does not involve any effect of reimbursement. Measures are reporting using G codes and specified CPT Category II codes effective April 1, 2006. Participating physicians report data and receive confidential feedback on performance versus their peers. The first 3-month report period is for April 1, 2006 to June 30, 2006 data (3 months), and reports for this period should be available for feedback by December 2006. It is likely that this initiative will eventually expand to the full set of 36 measures in the future, which includes a few outcomes measures along with the purely process measures included in the starter set of 16 measures. Reimbursement under this or similar programs will only effect Medicare Part B funds. Once advanced to the reimbursement phase, *it is expected that these initiatives will involve an initial withhold of reimbursement, which will later be reclaimed by the highest performing physicians as a "below the line" incentive.*

Eight of the 16 starter set quality measures involved in the PVRP involve surgery. Five apply to CABG only, one applies to surgical fistulas for end-stage renal disease (ESRD), and only two are general surgical measures (antibiotic prophylaxis and thromboembolism prophylaxis). Each is assigned three possible G codes ("Yes," "No," and "Patient Ineligible"). Each measure is only reported for a closed set of requested surgical procedures identified by CPT code. For antibiotic prophylaxis, the patient must be documented *to receive* their antibiotics 1-hour before incision (2 h for vancomycin or fluoroquinolone). Currently, none of the 283

tracked CPT procedure codes are for neurosurgical procedures. For thromboembolic prophylaxis, the patient must be documented *to receive* thromboembolic prophylaxis before incision. Currently, 109 of the 538 tracked CPT procedure codes (20%) are for neurosurgical procedures, and this is the one area that currently has neurosurgical impact.

Participation in the PVRP involves three activities. The first is data reporting either via claims submission or electronic health record (EHR). The second involves declaring your intent to participate by going to the web site, *www.qualitynet.org/pvrpintent,* and filling out a survey that takes less than 5 minutes. The third involves a formal registration process to receive your feedback report. You must establish a QualityNet account at the same web site to access these reports. This service was scheduled to be available in June 2006 and reportedly takes approximately 15 minutes. The degree to which neurosurgeons are currently participating in the PVRP is currently unknown.

Other Physician P4P Projects Under Development

CMS is currently developing a Medicare Management Performance Demonstration initiative, which will be a 3-year demonstration involving four states (AR, CA, MA, and VT). The goal is to promote adoption and use of HIT to improve the care of chronically ill patients.[13] Implementation anticipates *secondary return of an initial reimbursement withhold.*

CMS is also currently developing a Medicare Health Care Quality Demonstration, which will be a 5-year demonstration (personal communication; Tom Valuck, M.D., Medical Officer and Senior Advisor, Center for Medicare management, CMS, July 7, 2006, AANS/CNS Washington Committee). The goal will be to increase quality by increasing patient safety and reducing variation in use of health services by encouraging the use of *EBC* and *best-practice guidelines.* Implementation anticipates *secondary return of an initial reimbursement withhold*.

Finally, CMS is also currently working on a Physician Resource Use Reports for Highly Utilized Imaging Services project that will involve analyzing claims data in Ohio and Wisconsin. Phase 1 is designed to look at ordering echocardiograms for heart failure, and Phase 2 is designed to look at the ordering of magnetic resonance imaging (MRI) and CT scans for diagnosing chronic neck pain (personal communication; Tom Valuck, M.D., Medical Officer and Senior Advisor, Center for Medicare management, CMS, July 7, 2006 AANS/CNS Washington Committee). Both phases involve assessing and reimbursing the performance of diagnostic studies based on clinical practice guideline assessment of appropriateness (efficiency measure). The latter phase is of particular importance to neurosurgery from a practice standpoint.

The project is also important from a philosophical and EBM perspective. Not all guidelines are equivalent in quality. According to Woolf, there are three main methods of guideline development—informal consensus, formal consensus, and evidence-linked development.[61] From the standpoint of EBM, only the latter have evidentiary status for EBM decision making. Indeed, the IOM hopes to eventually restrict the use of the term "guideline" to systematically developed advisory statements created according to validated methodology.[31] Some consider consensus guidelines as intellectually suspect by reflecting expert opinion, which, when promulgated as a "guideline," can formalize unsound practice.[19] Without strict adherence to systematic and validated methodology, panelists may be pooling ignorance as much as distilling wisdom.[57] Some guidelines are of questionable quality and there have been calls for guidelines regarding how to devise guidelines.[39]

However, although the IOM supports restricting guidelines to those with valid methodology leading to an evidentiary status, CMS has chosen the American College Radiology (ACR) methodology (ACR Appropriateness Criteria) for Phase 2 of this initiative. The ACR Appropriateness Criteria methodology starts with EBM evidence tables for assessing published literature evidence, but instead of solely linking the level of recommendation to the level of evidence, provides a numeric "Delphi" voting mechanism for a restricted panel of experts to judge the level of recommendation to publish.[2] Thus, ACR Appropriateness Criteri are guidelines developed by formal consensus regarding an evidence review, rather than complete evidence-linked development.

According to ACR Appropriateness Criteria, when it comes to evaluating patients with neck pain, plain films should always be obtained before an MRI (regardless of additional clinical symptoms or findings), a CT scan should be obtained only when an MRI scan is contraindicated, and no patient should be evaluated with both an MRI and a CT scan.[3] However, most medical and surgical spine specialists realize that there are cervical spine conditions in which the information from both CT and MRI scans is desirable for optimal patient counseling and surgical planning because, in certain cases, the studies are complimentary rather than redundant. Perhaps these forms of unqualified restriction are to be expected from guidelines developed by physicians who are not directly responsible for the clinical care of patients rather than a multidisciplinary EBM development effort produced by, and endorsed by, neurosurgeons, orthopedic spine surgeons, and physiatrists. The problem is that, in the absence of EBM clinical practice parameter guidelines on the subject either produced by, or endorsed by, organized neurosurgery, CMS is left with the ACR as their benchmark "partner." The onus is on us to provide and defend a better alternative.

P4P BEYOND CMS

CMS, Medicare, and Medicaid are not the only arenas in which P4P will play out. Third-party payers have expressed wide interest in pursuing P4P as a condition of third-party reimbursement and Department of HHS government agencies are in full support. Indeed, in April 2006, the AHRQ published a monograph entitled, *Pay for Performance: A Decision Guide for Purchasers*, intended to assist third-party payers with P4P planning and implementation.[24] Several private insurers have already begun P4P pilot programs, including the PacifiCare of CA Quality Incentive Program, the Integrated Healthcare Association of CA, Premier Blue Cross of WA State, the Alliance of WI—Bridges to Excellence Program, Anthem Blue Cross Blue Shield of NH, and Wellpoint Quality Incentive Program NY. HMOs have been even quicker on the uptake, and penetration among HMOs may already be as high as 50%. Clearly, simply dropping Medicare and Medicaid from a neurosurgical practice portfolio will not insulate neurosurgeons to the reality of P4P.

PROBLEMS

There is not sufficient space in this one chapter to enumerate or discuss in depth all of the potential problems inherent in current P4P plans. However, a few key points need to be made.

Improvement in Quality and Efficiency Measures May Not Lead to Improved Clinical Outcomes

As outlined in the section, "Measuring and Influencing Quality in Healthcare," at the beginning of this manuscript, although process measures generally perform better than structural measures in predicting quality of healthcare delivered, in their current form, their correlation with crude measures of clinical outcome remains disappointingly poor.[9,60] As Williams et al.[60] discovered, improved hospital compliance with 15 of 17 process measures on acute myocardial infarction, heart failure, and pneumonia were possible without any improvement of in-hospital mortality after myocardial infarction. Individuals and institutions can become very proficient and compliant with documenting and reporting process measures without necessarily improving the care of individual patients. If this observation also holds up for neurosurgery, it is not clear that demanding adherence to process measures will improve the quality of neurosurgical care unless we tautologically and trivially define improvement in quality solely as extent of adherence to process measures.

Potential Negative Effects of P4P on Patient Care and the United States Healthcare System

Even with the best of intentions, there are several a priori, predictable, negative effects of a P4P program as currently envisioned that will need to be effectively managed. The first deals with physician patient selection. In a P4P reimbursement environment, particularly one in which up to 50% of reimbursement is determined by P4P criteria, it is highly likely that physicians will avoid sicker patients, given the perception that risk adjustment from the databases used is inadequate.[59] Indeed, after public reporting of CABG data was instituted in New York state, two-thirds of cardiothoracic surgeons admitted that they avoided most severely ill patients.[10] To limit this predictable effect, CMS would do well to focus on structural and process measures until EBM-validated outcome measures for specific diagnoses and procedures are more mature and generally available.

P4P, as currently envisioned, also has the potential to direct attention away from other aspects of care that are not examined under the initial set of quality measures used. It is likely that clinical attention will be preferentially focused on conditions for which there is measurement and augmented payment to the detriment of other, potentially equally important, clinical areas.

The P4P program as currently envisioned has the potential to widen existing performance gaps in United States healthcare. This potential widening is more likely if programs provide only rewards for top performance and fail to penalize lower performers who may give up in their perceived futile attempt to rise within the pack. In this area of concern, safety net healthcare systems with limited means and resources to invest in and augment their current performance are most at risk.

It is possible that if P4P targets are relative and move each year, rather than being fixed, that enthusiasm for continued effort will be dampened with each subsequent passing year of the program.

In the current P4P program, demonstration projects, individual practitioners, and small groups are preferentially penalized for their inability in make the capital investments in HIT necessary to participate and compete with larger groups and integrated healthcare delivery organizations. New CMS demonstration programs for HIT assistance/support are being developed nationally as well as at the state level via the 53 QIOs to address these concerns, and new Stark and anti-kickback safe harbors are being developed, but many of these items will require statutory change.

In a projected budget-neutral P4P environment for Medicare Part B reimbursement, there is no avoiding the fact that 100% of the Medicare Part B withhold from baseline physician reimbursement will never be secondarily returned to physicians via a P4P program because program overhead must also be accounted for in a budget-neutral environment. It is hard to imagine a government-administered program that will not exceed at least 30 to 40% overhead costs. If the initial withhold is 50% of baseline reimbursement, than this would ultimately lead to a 15 to 20% reimbursement reduction even

to the top performing 10% of Medicare providers. We have severe difficulty facing the possibility of potential annual 5 to 8% reimbursement cuts dealing with the existing onerous and patently unfair sustainable growth rate (SGR) formula, let alone facing a 15 to 20% reimbursement cut, even if we are the best healthcare providers in the nation.

Current P4P program considerations take no account of reimbursing clinicians for the increases in time devoted to increasing complexity of practice. They are also insensitive to the increasing overhead costs involved with HIT investment and increased coding and claims reporting complexity.

There are also significant concerns that there is a rush to proceed to efficiency measures, particularly cost measures, before there has been any demonstration of actual improvement in healthcare delivery quality related to the institution of quality measures. If cost measures come to dominate policy and reimbursement decisions, then the overall stated goal of improving the quality of healthcare delivered to United States citizens is at risk.

Budget Neutrality, Fairness, and Realism

Both the hospital P4P demonstration projects and the physician P4P demonstration projects outlined above reveal that CMS is implementing P4P under a double standard when it comes to Medicare Part A versus Part B funds. Hospitals are being approached with an "above the line incentive," namely additional reimbursement above standard baseline for participating and performing well; whereas physicians are being approached with a budget neutral, "below the line" "incentive" in the form of an initial withhold of standard reimbursement that must be reclaimed (minus overhead costs of administering the program and with a potentially significant time delay from time of delivery of service).

The concept of budget neutrality of physician P4P is a major problem for fairness (especially when compared against CMS proposals for hospital P4P), ultimate physician acceptance and buy-in, as well as maintenance of access to care in the setting of ever-increasing difficulty of practice viability in a continually shrinking physician reimbursement environment. At the urging of the Council of State Neurosurgical Societies (CSNS), the AANS and CNS have taken a strong position and a lead role on this issue. Largely on our initiative and with our urging and support, in 2005, the AMA issued a new AMA policy—H-450.947 *Pay-for-Performance Principles and Guidelines*.[5] This policy requires new funds for P4P payments as part of "fair and equitable program incentives." It is not yet clear whether the AMA will stick to their own policy regarding P4P in terms of steadfastly opposing any budget-neutral government-mandated program.

PELL MELL

The headlong, pell mell manner with which CMS is pursuing P4P implementation seems to be precipitous and rash. This time urgency seems to be in response to a perceived Congressional timeline pressure that will hold CMS government appointees accountable. There seems to be far more interest in getting to the point of public reporting and the public perception of responding to calls to improve healthcare quality than in actually improving it. The operative assumption seems to be that the very act of measuring and reporting on quality, no matter how imperfect or unintentionally damaging, will actually lead to improvements in healthcare delivery quality, and that the process can always go on to be corrected and refined "on the fly." The fact that we are still struggling to deal with the unfunded mandate and unintended consequences of another precipitously implemented federal program, namely the Emergency Medical Treatment and Labor Act (EMTALA), has had little impact on thinking regarding P4P.

Clearly, a national CMS P4P program is not yet ready for implementation. The existing HIT databases currently available for mining in terms of claims data are fundamentally, and perhaps fatally, flawed for assessing quality of clinical care delivery, and they must await development and implementation of a universal and reliable EMR. The issue of fairness and desirability of an "above the line" reimbursement incentive for hospitals versus a "below the line" reimbursement plan for physicians does not pass muster for either physician buy-in or any conception of fairness. The development of truly valid quality and efficiency measures must await further development of disease- and procedure-specific outcomes studies and truly evidence-linked multidisciplinary clinical practice parameter guidelines development. Holding physicians hostage for redress of another major wrong to fair physician compensation (namely, fixing the SGR formula) with the price of acceptance of a precipitously initiated and fundamentally flawed P4P program does little to sway our perception or position.

PHYSICIAN VOICE, REPRESENTATION, AND LEADERSHIP

Ultimately, no one knows more regarding the quality of actual patient care delivery or cares more about the quality of care delivered to our patients than the physicians providing that care. However, the organizational structure for CMS QI and the private partnerships and lines of influence outlined in *Figure 27.1* have served to distance, insulate, diminish, and mute the voice of the physician involved in direct, day-to-day, patient care when it comes to the national dialog as well as the design, approval, and implementation of national policy regarding healthcare QI in general and P4P in particular.

The dark blue boxes of elected officials contain no practicing physicians, although a select few are physicians who gave up clinical practice for the responsibilities of a full-time politician. The light blue boxes are government

appointees. Although a few are physicians by degree (especially in CMS), there are very few who continue to see and care for patients on a regular and ongoing basis. Practicing surgeons are almost unheard of, and I know of no neurosurgeon holding an appointed post in any of the agencies listed. For the most part, they have chosen a career in healthcare administration, healthcare policy, and appointed politics, and have largely left their patient care responsibilities behind. The green boxes representing the private partnerships sought out as part of the CMS QI Roadmap are supposed to provide the consensus voice that adequately represents the practicing clinician. Unfortunately, as has been repeatedly pointed out in this manuscript, this is "the great fallacy in assumption" on which we must publicly shine the bright light of truth and exposure.

The IOM cannot be considered to adequately or proportionately represent the average practicing physician in the United States, let alone surgeons or neurosurgeons engaged in the care of United States citizens. Their cross-sectional make up suggests a strong academic, public health, preventative medicine, and primary care bias, as well as a certain understandable practical naiveté regarding the realities and practical complexity of healthcare delivery in our country. Surgeons represent less than 10% of membership and neurosurgeons currently represent less than 0.5% of IOM membership.

Only 11.8% of the MedPAC membership are practicing physicians who rely on Medicare for practice income (one academic general surgeon, and one private practice urologist). MedPAC is dominated by health policy consultants (29.4%), health system CEOs (17.6%), healthcare economists and business administrators (11.8%), and healthcare purchasing consultants (11.8%).

Physician associations, coalitions, or group practices currently make up only 11.9% of NQF members. Surgical representation (let alone neurosurgical representation) is an even smaller minority. For the most part, the NQF is dominated by hospitals, hospital associations, and integrated health delivery networks (36.1%), as well as quality certification bodies, QI associations, and healthcare management and consulting groups (23%).

Currently, physician societies make up 54.5% of AQA membership (45.5% insurance, government, and patient advocates), however, there is very poor surgical representation. Only 18% of the steering group members are surgical societies (ACS and STS). Neurosurgery, via the Washington Committee QIW, is currently directly represented by David McKalip and indirectly represented through the SQA, however, these two positions only account for 1.6% of the general AQA membership.

In the best of all possible worlds, the physician component of the United States healthcare QI initiative should be led by clinicians. In the current situation, we struggle mightily to even realize a modicum of reasonable representation, let alone our rightful position as key stakeholders and leaders of the process.

NEUROSURGERY POSITION AND STRATEGY

United States neurosurgery is the strongest possible supporter of improved healthcare delivery to our patients and improved access to optimal neurosurgical services for all United States patients. However, we have some serious concerns regarding the well intentioned, but potentially flawed current P4P efforts sweeping the healthcare landscape. The CMS QI Roadmap produced in response to MMA 2003 and strongly influenced by Phases 1 and 2 of the IOM quality initiatives is strongly and broadly supported by Congress, the Department of HHS, its sister agency the AHRQ, the GAO, and MedPAC. At this juncture, it is unlikely that any attempts by organized medicine, let alone organized neurosurgery, will be able to halt or block inevitable implementation. Realistically speaking, at best, neurosurgeons can position ourselves nationally to be prepared for the predictable phases of development and implementation, and can try to refine, better focus, and redirect the initiative(s) into a healthier and more worthwhile focus and direction. The AANS/CNS Washington Office, the AANS/CNS Washington Committee, and the Neurosurgery Political Action Committee (PAC), are all organizing to address this issue

Neurosurgery PAC

Each of the areas indicated with an asterisk in *Figure 27.1* is a potential area of influence that requires education and input from organized neurosurgery, and is a potential area for decision making and policy modification influence. The primary function of the neurosurgery PAC is gaining access to policy makers to present our case and issues and positively influence changes in United States healthcare. The message needs to include: 1) our commitment to contribution and leadership in the healthcare QI arena; 2) our commitment to developing evidence-based quality and outcome measures from national neurosurgery outcomes studies and a vigorous program of high-quality EBM clinical practice parameter guidelines development, 3) our concerns regarding the fairness, adequacy, and unintended negative downstream consequences of using claims data and current crude risk adjustment methods for public reporting of "quality data"; 4) a continued call to fairness in physician P4P to only proceed under "above the line" financial incentives and to scrap the current budget-neutral requirements; 5) a debunking of the prevailing assumption that the NQF and the AQA adequately represent the voice of practicing physicians and, particularly, practicing surgeons; 6) a debunking of uncritical acceptance of the results of studies performed by the IOM as adequately representing the experience and concerns of actual practicing physicians; and 7) a debunking of the perspective and legit-

imacy of MedPAC, as currently comprised, for understanding, or being concordant with, the interests and needs of clinicians who actually care for patients.

We need to approach these key individuals with positive suggestions. These might include leveling the NQF annual dues schema across interest groups; considering broadening the SQA mission and membership to better balance the AQA as well as establish AHRQ and CNS contracts to legitimize and validate that role; improving surgical and neurosurgical representation on the NQF and the AQA (elevating the SQA to AQA steering committee membership?); improving MedPAC through more representative membership; and lobbying for increased practicing physician, surgeon, and neurosurgical membership in the IOM. We might also lobby for funding through the AHRQ for grants for multidisciplinary professional society EBM clinical practice parameter guidelines development performed in collaboration with existing AHRQ-funded EPCs. The Neurosurgery PAC is currently headed by Gary Bloomgarden.

Washington Office and Washington Committee Structure and Approach

The AANS/CNS approach to dealing with the current issue can be summarized as follows. We need to aggressively fight on every front to limit the negative effects and promote positive embellishments of current CMS and AHRQ P4P proposals and demonstration projects. To the best extent possible, we need to ensure that we are not subjugated to poor or inappropriate quality and efficiency measures. An example of a neurosurgery move along these lines is our initiative through the AQA via the SQA to change the antibiotic prophylaxis and thromboembolism prophylaxis G-code wording to read that prophylaxis was "ordered" by the surgeon rather than "received" by the patient, given the factors outside of a surgeon's controlling effecting whether or not the order is performed by hospital, surgicenter, and/or anesthesia personnel. Ultimately, we need to play for time to allow our own neurosurgery-specific outcome measures and EBM clinical practice parameter guidelines efforts to be realized and come to full maturity. We will then need to work very hard to get our measures accepted over alternative competing measures from other potential sources. To realize this strategy, the Washington Office and Washington Committee have modified their organization, as follows.

Quality Improvement Workgroup

The QIW of the Washington Committee is currently chaired by Robert Harbaugh. Originally tasked with developing disease-specific outcomes studies for neurosurgery, it is now also tasked with interfacing with the AMA-PCPI and the SQA for participating in the development and assessment of quality measures (process and outcome measures). Currently, Dan Resnick is the QIW representative to the AMA-PCPI, Robert Harbaugh is our representative to SQA, and David McKalip is our representative to the AQA. The disease-specific outcome study, which has come on line from the QIW, is an outcome study related to the surgical treatment of lumbar stenosis.

Outcome measures are the most difficult and expensive measures to collect and analyze. Outcome measures remain our best means of measuring healthcare quality for a particular procedure. As such, they are crucial for development of adequate quality measures (process and outcome measures). Unfortunately, they are unable to assess whether the procedure should have been performed in the first place, and, thus, do little to assess healthcare cost effectiveness or efficiency. They will be of little use in the arena of efficiency measures (cost and appropriateness of procedure selection).

Guidelines Committee

Less than 1 year old, the Washington Committee Guidelines Committee is Co-Chaired by Mark Linskey and David Adelson. The committee is tasked with developing a national prioritization agenda for developing multidisciplinary EBM clinical practice parameter guidelines for the areas of neurosurgery most likely to be targeted by P4P. Development can be neurosurgery initiated, or can involve neurosurgery in a multidisciplinary initiative from another medical society in which we participate in development and/or approval. Current ongoing and developing efforts include adult severe head injury revision, cervical spondylosis, lumbar radiculopathy, and metastatic brain tumor. We are hoping to initiate new efforts focused on thoracolumbar fracture and congenital pediatric communicating hydrocephalus in the near future.

EBM clinical practice parameter guidelines are one of our best sources of potential process measures (one type of quality measure), but also are our best source of efficiency measures. Only through EBM clinical practice parameter guidelines development can we arrive at reasonable and defensible measures assessing whether or not a procedure should be performed. Although an outcome measure can tell how well a procedure is executed, it is silent regarding whether the procedure should have been performed in the first place. Only efficiency measures will allow us to address the concerns of the IOM and CMS regarding unexplainable variation in procedural frequency between regions and between clinicians. These concerns involve overuse, misuse, and underuse of interventions. Only through our own rigorous methodological efforts can we preserve procedural options as equivalent Level 3 recommendations where appropriate, and prevent inappropriate consensus-driven restrictions at the Level 2 or Level 1 recommendation level (guidelines or standards).

Washington Office Quality Division

The Washington Office provides technical and administrative support to the neurosurgery PAC as well as the AANS/CNS Washington Committee. Regarding the latter, the AANS/CNS Washington Office has been reorganized to better serve and support the mission of rising to the challenge of P4P. A new Quality Division of the Washington Office has been formed, and a new full-time personnel (Rachel Groman) hired to provide ongoing technical and administrative support to both the QIW and the Guidelines Committee of the Washington Committee.

Washington Committee AMA and ACS Representatives

In the absence of an AANS, a CNS, a Washington Committee, or an ABNS representative at the NQF table, it becomes crucial that our Washington Committee AMA and ACS representatives effectively and vociferously represent us to these organizations that do have positions on the NQF. Currently, our AMA representatives are Mark Kubala, Monica Wehby, and Phil Tally, and our ACS representative is Clarence Watridge. All of our AMA, ACS, PCPI, SQA, and AQA representatives need to coordinate their approach and activities for maximal effectiveness.

ADDITIONAL STRATEGIES TO CONSIDER

Private Third-Party Payer P4P Legal Challenge

Government agencies are unique in their special protected positions insulating them from collusion and antitrust litigation. On the other hand, third-party payers are private, independent agencies and are subject to scrutiny and oversight. As such, there may be legal avenues open for addressing expansion of a uniform, coordinated P4P reimbursement program across multiple third-party payers, particularly if they dominate a regional market either alone or in combination.

Specifically, there may be avenues open to lobby via the Federal Trade Commission of the Department of Justice, based on an antitrust argument. One might contend that joint agreement regarding "Best Practice Standards" among the majority of third-party payers within a region and compelling their application among practitioners, and/or jointly agreeing to a uniform bonus payment schedule or schema for practitioners might arguably be interpreted as collusion. One might also make an unfair competition argument, in which individual practitioners and small groups may be able to argue that requirements for HIT investment to participate in commercial P4P, if it becomes dominant in a region, represent an unfair competition advantage for larger groups, Health Systems, or integrated health networks.

Although these legal challenges and Department of Justice/Federal Trade Commission legal approaches may delay or even prevent private third-party payer P4P implementation, they have certain inherent drawbacks. First of all, they would be extremely expensive to prosecute and sustain and, thus, would be unlikely to be sustainable unless a consortium of physician organizations agreed to a long-term funding commitment along these lines. Second, they would serve to place neurosurgery on the wrong side of the perceived healthcare quality initiative, where, instead of leading and setting the corrected direction, we would likely be perceived as resisting QIs and quality healthcare for patients.

Join the NQF

Currently, organized neurosurgery has limited their P4P input and influence to the AMA, the ACS, the AMA-PCPI, the ACS-SQA, and the AQA. Although these are important areas for liaison and contribution, we currently do not have direct input into a major organization dealing with physician P4P quality and efficiency measure endorsement and implementation, namely the NQF (see *Fig. 27.2*). Although joining this organization would be expensive, with ongoing and recurring annual costs, P4P is such an important issue for the future of neurosurgical patient care and neurosurgical professional practice that it is perhaps time to reconsider whether a seat at this table might be a worthwhile investment.

Exploring NSQIP

Outcomes studies and EBM clinical practice parameter guidelines development are important areas for ongoing neurosurgery quality efforts that are already addressed in via the Washington Committee QIW and Guidelines Committee. However, both are limited by the very significant time and financial investment necessary to bring multiple disease- and procedure-specific outcomes studies and guidelines production initiatives to completion. Guidelines production is further limited by limitations inherent in using the published peer-reviewed literature as the only source of evidence.

As discussed in the section on NSQIP, the ACS NQIP program, although originally designed for hospital reporting, is potentially modifiable for surgeon quality measurement. It would have the distinct advantage of being applicable to all neurosurgical procedures at once, rather than validating measures for many different procedures and diagnoses one-at-a-time through expensive and time-consuming individual outcome studies. It would also have the distinct advantage of allowing for stringent empirical clinical validation through iterations of mathematical modeling and comparison with actual outcomes. As such, it would free evidence-based validation requirements from only relying on peer-reviewed literature reports, without sacrificing verification through objective measurable evidence. It is perhaps time for the Washington Committee to fully explore the potential of the NSQIP program for meeting many of our neurosurgery P4P process measure and outcome measure needs.

CONCLUSION

It is an old adage in the Marine Corps that, when surprised, or when encountering a new effective tactic in battle, that the best strategy is to 1) survive, 2) adapt, and 3) overcome. At the present time, when it comes to P4P, we are trying to affect the first strategy while we play for time to realize the second and third strategies. This chapter provides the background information and influence structure needed to understand the issue, outlines our current organizational response to the issue, defines some of the challenges and obstacles we have yet to overcome, and outlines some suggested tactics and strategies for moving ourselves forward in a positive and effective manner.

REFERENCES

1. Agency for Healthcare Research and Quality, Consumer Assessment of Healthcare Providers and Systems (CAHPS). https://www.cahps.ahrq.gov/default.asp. Accessed 10/13/06.
2. American College of Radiology. ACR Appropriateness Criteria. Background and Development. http://www.acr.org/s_acr/bin.asp?CID=1847&DID=16124&DOC=FILE.pdf. Accessed 10/13/06.
3. American College of Radiology. ACR Appropriateness Criteria. Chronic Neck Pain. Last reviewed 2005. https://webapps.acr.org/ACRAC/VariantList.aspx?topicid=30223. Accessed 10/13/06.
4. American College of Surgeons, National Surgical Quality Improvement Program. https://acsnsqip.org/login/default.aspx. Accessed 10/13/06.
5. American Medical Association. Official AMA Policy—Principles for Pay-for-Performance and Guidelines for Pay-for-Performance. http://www.ama-assn.org/apps/pf_new/pf_online?f_n=resultLink&doc=policyfiles/HnE/H-450.947.HTM&s_t=pay+for+performance&catg=AMA/HnE&catg=AMA/BnGnC&catg=AMA/DIR&&nth=1&&st_p=0&nth=2&. Accessed 10/13/06.
6. Berwick DM: Continuous quality improvement as an ideal in health care. **N Engl J Med** 320:53–56, 1989.
7. Berwick DM, Wald DL: Hospital leaders' opinions of the HCFA mortality data. **JAMA** 263:247–240, 1990.
8. Blumenthal D: The future of quality measurement and management in transforming healthcare system. **JAMA** 278:1622–1625, 1997.
9. Bradley EH, Herrin J, Elbel B, McNamara RL, Magid DJ, Nallamothu BK, Wang Y, Normand SL, Spertus JA, Krumholz HM: Hospital quality for acute myocardial infarction: Correlation among process measures and relationship with short-term mortality. **JAMA** 296:72–78, 2006.
10. Burack JH, Impellizzeri P, Homel P, Cunningham JN Jr: Public reporting of surgical mortality: A survey of New York State cardiothoracic surgeons. **Ann Thoracic Surgery** 68:1195–1200, 1999.
11. Centers for Medicare and Medicaid Services. Hospital Quality Initiative. http://www.cms.hhs.gov/HospitalQualityInits/. Accessed 10/13/06.
12. Centers for Medicare and Medicaid Services. Medicaid and SCHIP Quality Practices. Pay for Performance—Payment Aligned with Quality. http://www.cms.hhs.gov/MedicaidSCHIPQualPrac/03_P4P.asp#TopOfPage. Accessed 10/13/06.
13. Centers for Medicare and Medicaid Services. Medicare Care Management Performance Demonstration. http://www.cms.hhs.gov/DemoProjectsEvalRpts/MD/itemdetail.asp?filterType=none&filterByDID=-99&sortByDID=3&sortorder=ascending&itemID=CMS057286. Accessed 10/13/06.
14. Centers for Medicare and Medicaid Services. Physician Hospital Collaboration Demonstration. http://www.cms.hhs.gov/DemoProjectsEvalRpts/MD/itemdetail.asp?filterType=none&filterByDID=-99&sortByDID=3&sortorder=ascending&itemID=CMS1186653. Accessed 10/13/06.
15. Centers for Medicare and Medicaid Services. Physician Voluntary Reporting Program. http://www.cms.hhs.gov/PVRP/. Accessed 10/13/06.
16. Centers for Medicare and Medicaid Services. Premier Hospital Quality Incentive Demonstration. http://www.cms.hhs.gov/HospitalQualityInits/35_HospitalPremier.asp. Accessed 10/13/06.
17. Centers for Medicare and Medicaid Services. Quality Improvement Roadmap. http://www.medicaldevices.org/public/issues/documents/CMSMedicareroadmap.pdf. Accessed 10/13/06.
18. Crosby PB: *Let's Talk Quality*. New York, McGraw-Hill, 1989.
19. Delamothe T: Wanted: Guidelines that doctors will follow. Implementation is the problem. **BMJ** 307:218, 1993.
20. Deming WE: *Out of the Crisis*. Cambridge, MA, MIT Center for Advanced Engineering Studies 1986.
21. Donabedian A. *Definitions of Quality and Approaches to its Assessment. Explorations in Quality Assessment*. Chicago, Health Administrations Press, vol 1, 1980.
22. Donabedian A. *The Criteria and Standards of Quality. Explorations in Quality Assessment*. Chicago, Health Administrations Press, vol 2, 1981.
23. Donabedian A. *Methods and Findings of Quality Assessment and Monitoring: An Illustrated Analysis. Explorations in Quality Assessment*. Chicago, Health Administrations Press, vol 3, 1985.
24. Dudley RA, Rosenthal MB: *Final Contract Report—Pay for Performance: A Decision Guide for Purchasers*. Washington DC: AHRQ Publication No. 06–0047, April 2006.
25. Eddy DM: Performance measurement: Problems and solutions. **Health Affairs** 17:7–25, 1998.
26. Executive Order 13017—Advisory Commission on Consumer Protection and Quality in the Healthcare Industry. September 5, 1996. http://presidency.ucsb.edu/ws/index.php?pid=53292. Accessed 10/13/06.
27. Fisher CM, Kistler JP, Davis JM: Relation of cerebral vasospasm to SAH visualized by CT scanning. **Neurosurgery** 6:1–9, 1980.
28. Gupta N, Kotler PK, Dudley RA: A report card approach to ICUs. **J Intensive Care Med** 17:211–218, 2002.
29. Heilman CB, Hwang S. First to worst. Does DRG assessment of hospital mortality indicate quality? 56th Annual Meeting of the Congress of Neurological Surgeons, Chicago, IL, October, 2006 [Abstract #857].
30. Hunt WE, Hess RM: Surgical risk as related to time of intervention in the repair of intracranial aneurysms. **J Neurosurg** 28:14–20, 1968.
31. Institute of Medicine, Committee to Advise the Public Health Service on Clinical Practice Guidelines: *Clinical practice guidelines: Directions for a new program*. Field MJ, Lohr KN (eds). Washington, DC, National Academy Press 1990.
32. Institute of Medicine: *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC, National Academy Press 2001.
33. Institute of Medicine: *Health Professions Education: A Bridge to Quality*. Washington, DC, National Academy Press 2003.
34. Institute of Medicine: *Hospital-Based Emergency Care: At the Breaking Point*. Washington, DC, National Academy Press 2006.
35. Institute of Medicine: *Leadership by Example: Coordinating Government Roles in Improving Health Care Quality*. Washington, DC, National Academy Press 2002.
36. Institute of Medicine: *Medicare's Quality Improvement Organization Program: Maximizing Potential*. Washington, DC, National Academy Press 2006.
37. Institute of Medicine: *Rewarding Provider Performance: Aligning Incentives in Medicine*. Washington, DC, National Academy Press 2006.
38. Institute of Medicine: *To Err is Human: Building a Safer Health System*. Washington, DC, National Academy Press 1999.
39. Jackson R, Feder G: Guidelines for clinical guidelines: A simple, pragmatic strategy for guideline development. **BMJ** 317:427–428, 1998.
40. Jones RH, Hannan EL, Hammermeister KE, Delong ER, O'Connor GT, Luepker RV, Parsonnet V, Pryor DB: Identification of preoperative variables needed for risk adjustment of short-term mortality after coronary artery bypass graft surgery. The working group panel on the cooperative CABG database project. **J Am Coll Cardiology** 28:1478–1487, 1996.
41. Juran JM: *Juran on Planning for Quality*. New York, Free Press 1988.
42. Khuri SF, Daley J, Henderson W, Hur H, Demakis J, AUst JB, Chong V, Fabri PJ, Gibbs JO, Grover F, Hammermeister K, Irvin G 3rd, McDOnald G, Passero E Jr, Phillips L, Scamman F, Spencer J, Stremple JF: The Department of Veterans Affairs' NSQIP: The first national

validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. **Ann Surg** 228: 491–507, 1998.
43. Krumholz HM, Chen J, Wang Y, Radford MJ, Chen YT, Marciniak TA: Comparing AMI mortality among hospitals inpatients 65 years of age and older: Evaluating methods of risk adjustment. **Circulation** 99: 2986–2992, 1999.
44. Linskey ME: The total quality movement: A paradigm shift in leadership and management philosophy for neurological surgery and healthcare in general, Chapter 1, in Linskey ME, Rutigliano MJ (eds): *Quality and Cost in Neurological Surgery. Concepts in Neurosurgery,* Philadelphia, Lippincott Williams and Wilkins, Volume 10:pp 3–31, 2001.
45. Linskey ME: Obstacles to implementing a quality philosophy in neurological surgery, Chapter 14, in Linskey ME, Rutigliano MJ (eds): *Quality and Cost in Neurological Surgery. Concepts in Neurosurgery*. Philadelphia, Lippincott Williams and Wilkins, vol 10:pp 260–284, 2001.
46. Linskey ME, Rutigliano MJ (eds): *Quality and Cost in Neurological Surgery. Concepts in Neurosurgery*. Philadelphia, Lippincott Williams and Wilkins, vol 10, 2001.
47. Lohr KN: How do we measure quality? **Health Affairs** 16:22–25, 1997.
48. Luce JM, Thiel GD, Holland MR, Swig L, Currin SA, Luft HS: Use of risk-adjusted outcome data for quality improvement by public hospitals. **Western J Med** 164:410–414, 1996.
49. Medicare Payment Advisory Commission: Report to Congress: Medicare Payment Policy. March 2005. http://www.medpac.gov/publications/congressional_reports/Mar05_TOC.pdf. Accessed 10/13/06.
50. Medicare Payment Advisory Commission: Report to Congress: Medicare Payment Policy. March 2006. http://www.medpac.gov/publications/congressional_reports/Mar06_TOC.pdf. Accessed 10/13/06
51. Medicare Payment Advisory Commission: Strategies to improve care: Pay for performance and information technology. Chapter 4, in: Report to Congress: Medicare Payment Policy. March 2005. http://www.medpac.gov/publications/congressional_reports/Mar05_TOC.pdf. Accessed 10/13/06.
52. Mehrota A, Bodenheiner T, Dudley RA: Employers' efforts to measure and improve hospital quality: Determinants of success. **Health Affairs (Millwood**) 22:60–71, 2003.
53. Office of Management and Budget, Executive Office of the President of the United States, Circular No. A-119, revised February 10, 1998. Federal Participation in the Development and Use of Voluntary Consensus Standards and in Conformity Assessment Activities. http://www.whitehouse.gov/omb/circulars/a119/a119.html. Accessed 10/13/06.
54. President's Advisory Commission on Consumer Protection and Quality in the Healthcare Industry: Consumer Bill of Rights and Responsibilities. Report to the President of the United States, November 1997. http://govinfo.library.unt.edu/hcquality/cborr/index.htm. Accessed 10/13/06.
55. President's Advisory Commission on Consumer Protection and Quality in the Healthcare Industry: Quality First: Better Health Care for all Americans. Final Report to the President of the United States, March 1998. http://www.hcqualitycommission.gov/final/. Accessed 10/13/06.
56. Romano PS, Rainwater JA, Antonius D: Grading the graders: How hospitals in California and New York perceive and interpret their report cards. **Medical Care** 37:295–305, 1999.
57. Scott EA, Black N: When does consensus exist in expert panels? **J Public Health Med** 13:35–39, 1991.
58. Teasdale G, Jeannett B: Assessment of coma and impaired consciousness: A practical scale. **Lancet** 2:81–4, 1974.
59. Werner RM, Asch DA: The unintended consequences of publicly reporting quality information. **JAMA** 293:1239–1244, 2005.
60. Williams SC, Schmaltz SP, Morton DJ, Koss RG, Loeb JM: Quality of care in U.S. hospitals as reflected by standardized measures, 2002–2004. **N Engl J Med** 353:255–264, 2005.
61. Woolf, SH: Practice guidelines, a new reality in medicine. II. Methods of developing guidelines. **Arch Intern Med** 152:946–952, 1992.

