

Pay for Performance (P4P) in Medical Imaging: The Time Has (Finally) Come

Bruce I. Reiner and Eliot L. Siegel

INTRODUCTION

In life, there tends to be two (somewhat simplistic) views on payment. The first (and more common) is that services rendered generate a fixed payment. It does not matter whether you take the perspective of the employer, employee, or customer. All three parties know up-front what they are expecting and willing to pay (and receive), based upon a mutually agreeable compensation schedule. In the workplace, employers and employees will periodically review this compensation schedule and make systematic adjustments based on market conditions. Although this commonly follows an upwards trajectory (i.e., progressively increasing salaries), this is not always the case as evidenced by the recent decline in compensation benefits within the airline industry. When salary increases do occur, they are frequently passed on to the customers, often creating inflationary pressures within the marketplace. This has certainly been the case throughout the health care industry, where prices for goods and services have continuously trended upwards, often irrespective of the quality and timeliness of these services. Recent estimates by the Centers for Medicare and Medicaid Services (CMS) predict national health care expenditures to grow 2.5% faster than the gross domestic product (GDP), with the percentage of GDP spent on health services predicted to grow from its current rate of 15% to 17% in 2011.¹ This figure is almost twice as high as comparable health care expenditures in the United Kingdom and is projected to exhaust the Medicare hospital trust fund by 2019.²

The second (and less common) school of thought on payment is one which is incentive-

based, tying payment to performance. This “sliding scale” approach is often met with skepticism, largely out of concerns over how (and by whom) the payment will be calculated. If a supplier of goods and services was to defer to the customer on determining “fair and equitable” compensation, the recommended figure would in all likelihood be less than the “standard market rate”. If, on the other hand, the consumer of these goods and services was to appropriate pricing to the provider, an equal degree of disparity will emerge. In the end, determining “fair and equitable” pricing becomes highly subjective, especially when objective comparison data is in such short supply. On face value, this pay-for-performance (P4P) pricing model becomes too subjective and uncertain to gain wide acceptance. But does that necessarily have to be the case?

PRESSURES MANDATING P4P IN MEDICINE

Recent high visibility publications³⁻⁵ from the Institute of Medicine (IOM) have explicitly pointed out deficiencies in the existing health care system, specifically as it relates to quality and safety. As a result, a renewed emphasis on quality and patient safety has been advocated by multiple stakeholders including the general public, policy

From the Baltimore Veterans Affairs Medical Center, 10 N. Greene St., Baltimore, MD 21201, USA.

Correspondence to: Bruce I. Reiner, 6 Greenleaf Lane, Seaford, DE 19973, USA; tel: +1-302-629-6611 ; fax: +1-302-628-3460; e-mail: breiner1@comcast.net

Copyright © 2006 by SCAR (Society for Computer Applications in Radiology)

Online publication 05 December 2006

doi: 10.1007/s10278-006-1053-2

makers, employers, and insurers, and the medical profession. According to the IOM,⁴ current deficiencies in health care quality are primarily “systems” problems rather than “poor provider” problems, and therefore call for a fundamental and sweeping reform of the entire health care system, based on six arms:

1. Safety (avoiding injury and illness)
2. Effectiveness (medical service predicated on scientific knowledge)
3. Patient-centeredness (care focused on patient values and preferences)
4. Timeliness (reducing delays)
5. Efficiency (avoiding waste)
6. Equity (uniform care across geographic, socioeconomic, gender, and ethnic boundaries)

In *Crossing the Quality Chasm*,⁴ the IOM specifically recommended the establishment of financial incentives to health care providers for the practice of evidence-based medicine and improved patient outcomes.⁴ In *Leadership by Example*,⁵ the IOM challenged the government to take the lead in health care quality improvement by instituting P4P initiatives in federally funded health care programs. These should have the dual purpose of offering financial rewards and public recognition to practitioners who achieve superior levels of quality. This financial commitment to quality should be supplemented by the public dissemination of quality information to consumers (through the World Wide Web).

TURNING P4P INTO A REALITY

What if (and this is a big “if”) all parties can agree on the parameters and metrics with which the P4P model is calculated? If a radiology or hospital administrator can produce a payment schedule for technologists that can prospectively tie specific productivity and quality measures to performance, will that be perceived as acceptable? For this setup to be acceptable, the following conditions must first be met:

1. The payment schedule in question must be agreed upon by both parties.
2. The quantitative and qualitative metrics used must be unambiguously defined and reproducible.
3. Any subjective measures must be subject to validation by a “neutral” third party.

Although the employer (hospital/radiology administrator) and employees (technologists) agree in principle that P4P is conceptually viable, will the customers receiving such services also agree on this model? The answer partly depends on how one defines the customer and what eventual effect this “new” model will have on cumulative pricing for these services. One customer group would of course be the individual patient, who is the subject of the imaging services being rendered. Another customer group would be that of a large employer, who contracts out imaging services for a large commercial employee pool. A third customer group would be that of a governmental agency (e.g., CMS), which serves both as a comprehensive payer and barometer of imaging services over a large distributed network. Although the individual payer (i.e., patient) may be amenable to the P4P model, the larger group payer may be less inclined, particularly if the cumulative costs for this model exceed that of the traditional “fixed” payment model. A relevant example can be seen in the payment of prescription drugs, where more affluent individual payers may opt for the more expensive “name brand” pharmaceutical, if they believe there is an overall health benefit derived from it, as compared to a less expensive counterpart. Group payers (and less affluent individual patients), instead commonly opt for the less expensive generic equivalents, even though some studies have reported decreased drug efficacy of generics.^{6,7}

In the future, health care decisions will be increasingly consumer-driven, as employers switch from defined benefit to defined contribution health plans. This will likely result in an increasingly consumer-driven market, where individuals use quality data to assist in the selection of health care providers and services.⁸

The challenge to this economic conundrum for both parties (service deliverers and customers) is to devise a P4P model that can balance the often competing demands of higher quality with cost efficacy. If higher quality ultimately costs more, can we cost-justify this added expense, and if so to what degree? Are there objective measures and/or scientific data to justify higher costs, and how do we ultimately determine “fair value”? The answer to this challenge comes in the form of outcomes research, which to date has several limitations, particularly as it relates to medical

Table 1. Standardization and Quality Assurance Factors in Medical Imaging

1.	Equipment
2.	Imaging exam acquisition protocols
3.	Image processing techniques
4.	Image display parameters
5.	Utilization review
6.	Decision support tools
7.	Reporting/Communication

imaging. This is attributable to a number of factors including the large number of confounding variables in the analysis, difficulty in aggregating multi-institutional databases, and the relative lack of standardization in radiology today (Table 1). Before we can contemplate large-scale radiology outcomes research, we must first find a reproducible and efficient means of introducing quality assurance and standardization in the manner in which images are acquired, processed, displayed, interpreted, requested, and reported. In particular, outcomes research assessing radiologist performance requires for large numbers of radiology reports to be entered into minable databases. Before that can practically occur, however, reports must have a common terminology (lexicon) and organization (structured reporting).

In the end, the one reliable means to prevent radiology from becoming a commodity (contracting radiology services to the lowest bidder) is to develop quantifiable quality performance indicators throughout the radiology continuum, that are clearly tied to improved patient outcomes.

HOW TO START

If the medical imaging community wants to create a P4P reimbursement model, then its constituents must begin to take a proactive approach. To do this, we must redirect our focus on development and testing of quality performance measures and invest heavily in large-scale outcomes research. Here are a few ideas and recommendations to facilitate this:

1. *Create quality performance metrics that transcend the entire radiology process.* This should include the individual steps and functions of exam ordering, image acquisition, image storage and distribution, image display, image processing, interpretation, and reporting.

2. *Enlist the services and support of medical equipment providers to tie quality performance measures (and reimbursement) to the imaging modalities themselves.* Vendors should prefer to have objective performance measures introduced into the reimbursement model, specifically rewarding innovations and products that enhance medical care delivery. In its current fixed form, payment has no direct ties to equipment performance. A magnetic resonance imaging (MRI) exam performed on an antiquated, mobile 0.5-T MRI scanner is reimbursed the same as the same exam performed on a state-of-the art 3-T MRI scanner. Does that make any sense? (If you think it does, maybe you should have your next cardiac MRI on the mobile 0.5-T unit.)

3. *Lobby for state and federal legislation that ties reimbursement rates to the education and training level of the service providers.* Although it may seem convenient to have your x-ray taken by an aide and read in the office of your family practitioner, it is less likely to be of the same quality as that taken by a trained registered technologist and interpreted by a board-certified radiologist. One must remember that providing imaging services involves more than just image acquisition and interpretation, and requires specialized knowledge of image processing, quality assurance, and patient safety. Education is a lifelong process and all providers must continuously update their knowledge through continuing medical education and testing, *with no exceptions!*

4. *Create or identify a neutral third party organization that establishes, oversees, and tests universal quality assurance (QA) standards and metrics.* With the advent of teleradiology, this takes on even greater importance as many service providers work out of other countries, with different malpractice laws and professional standards. QA standards need to become far more rigorous and follow several precepts:

- a. QA should be prospective, not retrospective. (Why not provide QA services in real-time, before the image exam and report has been completed?)
- b. QA should be taken out of the hands of individuals with an inherent conflict of interest.

- c. QA should be a billable service, or else it remains somewhat marginalized.
 - d. QA results should be public information and be accessible to all customers.
5. *Support the standardization and integration of referenceable databases.* For the imaging community, this requires a number of parallel initiatives including the development of universal structured reporting and a modality/anatomy independent lexicon. Information technology vendors [picture archiving and communication system (PACS), IS, and EMR] need to integrate supporting technologies such as natural language processing (NLP) and intelligent agents to facilitate “on the fly” data collection and analysis.

Through the creation of structured text reports (using a standardized lexicon), radiologist performance can be analyzed in an automated fashion. This allows one to objectively review a number of performance measures including the frequency with which additional imaging studies are recommended, the use of qualifying terms (i.e., uncertainty), and diagnostic accuracy. This does not need to be perceived as a punitive process, but can instead be used for educational purposes and ultimately raise the standard of radiology practice.

6. *Track QA performance metrics from the patient perspective and incorporate them into the comprehensive QA analysis.* Imaging professionals tend to focus on the image as the end product and ignore the patient and their experience. Although timeliness is important, it is certainly not the sole criterion on which patient satisfaction should be evaluated. Patients should be routinely surveyed and their subjective perceptions incorporated into the QA analysis. At the same time, QA data should be made public knowledge to allow for discriminating customers to make educated decisions as to choosing providers. In the end, an educated customer is the quality providers’ best friend (and advocate).
7. *Tie medical malpractice rates directly to individual QA performance metrics, as opposed to global rates devoid of intrinsic QA measures (i.e., P4P malpractice).* An individual physician should be rewarded (or penalized) by their own performance record, which should be unequiv-

ocally tied to QA reference data. Occupation or subspecialty alone should not be the primary determinant of malpractice rates. If a consumer with a stellar driving record can get a performance-based discount, should not the same apply to a health care professional? This becomes ever more practical as universal QA standards and metrics are developed, measured, and analyzed on a regular basis.

8. *Make utilization review (UR) an integral part of the P4P reimbursement model, rewarding those individuals and institutions that proactively work to improve utilization of medical services.* P4P models must work in parallel to improve the quality, timeliness, and economics of service deliverables. More efficient UR has the synergistic effect of improving operational efficiency, eliminating unnecessary and/or redundant exams, and reducing overall health care costs. Just like any other “value-added” service (including QA), this needs to be directly compensated for by rewarding providers to invest the time, technology, and personnel required. Radiologists should not become gatekeepers, but instead take an active role as imaging consultants. Although technology [such as computerized physician order entry (CPOE) systems] can streamline the UR process, it is not a replacement for physician-to-physician communication. By collectively mining the clinical, administrative, and economic databases, physician and institutional profiles can be collected and analyzed for the purposes of education and utilization optimization. A proactive, technology-driven UR program has the added advantage of providing more accurate and detailed clinical information for the radiologist at the time of image interpretation, which can theoretically improve diagnosis.
9. *Invest (on multiple levels) in the value of outcomes research.* This will provide a means with which to cost-justify additional reimbursements for new technologies, services, and applications that lead to documented improvements in patient safety, diagnostic accuracy, and improved patient management. To facilitate this research, a number of efforts are required from a number of participants including:
- a. Create greater financial incentives for academicians and researchers, who are current-

- ly reimbursed far less than their clinical colleagues.
- b. Encourage greater involvement from all segments of the imaging community (no free rides). Although private practice radiologists may not have great interest or aptitude in outcomes research, they can offer valuable clinical services and insight, which is an undervalued resource.
 - c. Require imaging and IT vendors to actively participate in outcomes-based research. Vendors' participation is essential in the development and integration of industry-wide standards and they obviously benefit by becoming more competitive in the marketplace.
 - d. Create an incentive (e.g., increased reimbursements) and mechanism for community-based hospitals and practitioners to actively participate in multicenter trials. This would in effect eliminate the "academic bias" that tends to plague most research efforts and improve dialogue and collegiality throughout the professional community.
10. *Tie reimbursement directly to the practice of evidence-based radiology (EBR).* Although medicine remains an art form, the more science (and data) introduced into the process, the better for all parties. With the ability to mine large imaging databases and incorporate decision support tools into the interpretation process, EBR is more practical (and relevant) than ever before.
 11. *Reward innovation.* Most successful companies with a long-term approach allocate substantial percentages of their revenue to continued R&D efforts. Companies that tend to limit their focus to the present and near-term only, tend to eventually become marginalized by their myopic approach. Medicine is no different and the various stakeholders involved in the delivery of imaging services (radiologists, industry, administrators, and technologists) must maintain a long-term perspective on quality improvement and safety. To accomplish this, P4P programs must financially reward and encourage innovation, to ensure that practitioners do not focus solely on the proverbial "lowest hanging fruit". Technical innovations (PACS, EMR, speech recognition) and new applications (virtual colonography, functional MRI) require significant resource expenditures before they become ready for "prime time". In the end, innovation is what drives the advancement of products and services and requires financial incentives for continuation.
 12. *Eliminate selection bias in establishing metrics.* A number of potential sources of bias can be introduced into the metrics being collected and analyzed, including the patient population being served, institutional demographics, and available technology. Clearly, a rural radiology provider in a 50-bed hospital will not have the same available resources as a 500-bed tertiary care facility. At the same time, the patient population being served in a VA hospital is less compliant and frequently has more complex medical disease than their ambulatory counterparts in a suburban outpatient imaging center. One must incorporate "peer" modifiers into the analysis of metrics to allow for these differences. At the same time, one must realize that "poor" data tends to be underreported and often gets "lost". All data must be collected and analyzed to ensure that the data being analyzed is fair, reproducible, and accurate.

CURRENT P4P MARKETPLACE

Although P4P is currently a grass roots effort, there is a growing momentum underway. Price-WaterhouseCoopers consultants estimate that approximately 1/3 of conventional medical insurance plans now implement some pay-for-performance models.⁹ A study by Med-Vantage reported a threefold increase in P4P provider programs from 2003 to 2005, from 35 to 104.¹⁰ In June of 2005, the Medicare Value Purchasing (MVP) Act of 2005¹¹ was introduced in Congress that would require Medicare to implement a P4P program to cover a portion of payments made.

A recent study in JAMA¹² reported the preliminary results of a prototype P4P program in cervical cancer screening, mammography, and hemoglobin A1c testing. Although quality improvement scores were observed in the bonus incentive group (compared with the control group) for cervical cancer and mammography screening, the physician groups with baseline performance above the performance threshold improved the

least over the study period, yet garnered the largest share of the bonus payment. This suggests that existing P4P programs have the potential to selectively reward “high performance” baseline physician groups (who merely maintain the status quo in order to receive bonus payments), with little overall gain in quality.

This same study also showed that the “lower performance” baseline physician groups accounted for the majority of quality improvement over the study period, even though they did not realize the same degree of financial reward. This suggests that P4P programs offer the potential benefit of “refocusing” the collective medical community’s efforts on quality, with the long-term result being a higher standard of patient care. Although additional longitudinal research is required to validate these preliminary observations, the potentially derived benefit is substantial.

At present, P4P programs and data tracking are largely limited to general medical practitioners, and radiology has been largely ignored. Quality measures currently collected are limited and in the very early stages of development. Data currently being tracked come from a few organizations, including the American Medical Association (AMA) and National Committee on Quality Assurance (NCQA). Guidelines and measures come from the American Academy of Family Physicians (AAFP), Centers for Medicaid and Medicare Services (CMS), and National Quality Forum (NQF).¹³

The time is right for radiology activists and professional societies to heed the call and take a proactive role in leading the medical imaging community into a new era of quality-oriented,

performance-based reimbursement. Like the famous advertisement used to say, “You can pay me now, or pay me later”. Paying now provides a seat at the table and allows for radiologists to have greater control over their long-term destiny, while enhancing patient care in the process.

REFERENCES

1. <http://www.cms.hhs.gov/charts/default.asp>
2. <http://www.cbo.gov/showdoc.cfm>
3. Institute of Medicine. To Err is Human: Building a Safer Health System. Washington, DC: National Academy Press, 2000
4. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press, 2001
5. Institute of Medicine. Leadership by Example: Coordinating Government Roles in Improving Health Care quality. Washington, DC: National Academy Press, 2002
6. Borgheini G: The bioequivalence and therapeutic efficacy of generic *versus* brand-name psychoactive drugs. Clin Ther 26:801–802, 2004
7. Meredith P: Bioequivalence and other unresolved issues in generic drug substitution. Clin Ther 25:2875–2890, 2003
8. Seidel RL, Nash DB: Paying for performance in diagnostic imaging: current challenges and future prospects. J Am Coll Radiol 1:952–956, 2004
9. <http://www.pwchealth.com/cgi-local/hcregister.cgi?link=pdf/p4p.pdf>
10. http://www.medvantageinc.com/Pdf/MV_2004_P4P_National_Study_Results-Exec_Summary.pdf
11. http://www.ncqa.org/communications/news/Medicare_legislation_statement_05.htm
12. Rosenthal MB, Frank RG, Li Z, Epstein AM: Early experience with pay-for-performance: from concept to practice. JAMA 293:1788–1793, 2005
13. DiIulio: P4P primer. Medical Imaging 18–19, 2005, September