RadioGraphics

Quality Initiatives

Peer Review in Diagnostic Radiology: Current State and a Vision for the Future[1]

Shmuel Mahgerefteh, MD • Jonathan B. Kruskal, MD, PhD • Chun S. Yam, PhD • Arye Blachar, MD • Jacob Sosna, MD

ONLINE-ONLY CME

See www.rsna.org/education/rg_cme.html

LEARNING OBJECTIVES

After reading this article and taking the test, the reader will be able to:

- Identify the key aspects of performance evaluation, including peer review, as defined by the ABMS and JCAHO.
- Describe the advantages and disadvantages of blinded double reading, interdisciplinary evaluation, and workstation-integrated systems for peer review in diagnostic radiology.
- Discuss common challenges and solutions when implementing peer review in a radiology department.

TEACHING POINTS

See last page

Over the past decade, the level of interest in improving the quality of healthcare in the United States has increased. New requirements established by regulatory organizations require the ongoing practice-based evaluation of physician performance. Peer review, a key process in physician performance evaluation, is geared primarily toward measuring diagnostic accuracy. Accuracy may be measured in terms of interpretive agreement or disagreement during a blinded double reading or in workstation-integrated evaluations. Each method of assessing diagnostic accuracy has strengths and weaknesses that should be carefully considered before it is implemented in a particular departmental or institutional setting.

Abbreviations: ABMS = American Board of Medical Specialties, ACGME = Accreditation Council for Graduate Medical Education, ACR = American College of Radiology, JCAHO = Joint Commission on Accreditation of Healthcare Organizations

RadioGraphics 2009; 29:1221–1231 • **Published online** 10.1148/rg.295095086 • **Content Codes:** PR QA

[1]From the Department of Radiology, Hadassah Hebrew University Medical Center, PO Box 12000, Jerusalem, Israel 91120 (S.M., J.S.); Department of Radiology, Beth Israel Deaconess Medical Center, Boston, Mass (J.B.K., C.S.Y., J.S.); and Department of Radiology, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel (A.B.). Received April 7, 2009; revision requested April 27 and received April 29; accepted May 6. J.S. is a consultant with ActiViews and has received research funding from Koninklijke Philips Electronics; all other authors have no financial relationships to disclose. **Address correspondence to** J.S. (e-mail: *jacobs@hadassah.org.il*).

Introduction

Various regulatory agencies are responsible for assessing hospitals and doctors for purposes of accreditation, certification, licensing, credentialing, and privileging. These agencies include state and specialty boards, the American College of Radiology (ACR), the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), and the Accreditation Council for Graduate Medical Education (ACGME). Regulatory agencies base their assessments of medical staff partly on ongoing performance-based evaluations that include peer review. Detailed definitions of the five functions performed by agencies that regulate healthcare delivery are important for understanding the processes involved (Table 1).

Teaching Point

The article surveys the key aspects of medical staff performance evaluation, compares the evaluation methods that are currently used in radiology, and describes a setting-specific evaluation system that includes confidential peer review and that can be applied in a radiology department.

Healthcare in the United States

In the decade since the Institute of Medicine reported that tens of thousands of deaths in the United States were linked to medical errors, the quality of medical care has evoked increasing concern (3–5). In 2001, the Institute of Medicine stated that "the U.S. health care delivery system does not provide consistent, high-quality medical care to all people," that healthcare "harms patients too frequently," and that "between the health care that we now have and the health care that we could have lies not just a gap, but a chasm." Quality improvement requires growth and adaptability, and these in turn require knowledge and communication of goals and obstacles in a changing environment. As W. Edwards Deming said, "It is not enough to do your best; you must know what to do, and then do your best" (6). The performance of individual physicians is a significant factor in the overall quality of healthcare. Healthcare providers whose loved ones have needed medical care may have witnessed the unfortunate reality firsthand: Doctors are fallible, and the care they provide is not always optimal. In fact, much of medical practice is not evidence-based (5). This is true also of radiology (7). Moreover, the quality of a physician's performance tends to decrease with an increasing number of years in practice (8).

Quality Improvement Efforts and Physician Performance Evaluation

The American Board of Medical Specialties (ABMS) took up the call for quality improvement, focusing on individual physician performance. In 2000, the ABMS instituted a program for maintenance of certification, stipulating that the renewal of certification should be contingent on continuous evaluation and ongoing demonstration of quality in four areas: professional standing (eg, licensure status); lifelong learning; cognitive expertise (eg, evidenced by performance on standardized tests); and performance in practice (9,10). Since performance in practice is the principal parameter of interest in most peer-review efforts in radiology, our article focuses on physician clinical performance assessment. Landon et al (11), writing for the ABMS, defined physician clinical performance assessment as "the quantitative assessment of physician performance based on the rates at which their patients experience certain outcomes, and/or the rates at which physicians adhere to evidence-based processes of care during their actual practice of medicine."

The JCAHO, aligning its efforts with those of the ACGME and the ABMS, issued new guidelines for medical credentialing and privileging that require ongoing, practitioner-specific evaluation of six general competencies. These include patient care, technical skills, professionalism and communication skills, systems-based practice, practice-based learning and improvement, and medical knowledge and clinical judgment. Of these six general competencies, the last—medical knowledge and clinical judgment—is most closely related to performance in practice (12).

Peer Review in Diagnostic Radiology

Of the factors relating to overall physician competence, performance in practice is the most difficult to assess. Limited procedure numbers and patient populations for individual physicians, as well as confounding variables such as variations between patient populations, are typical obstacles to a statistically meaningful assessment of individual physician performance in practice. Since clinical outcomes data are difficult to collect and analyze, few specialties have developed a comprehensive, evidence-based system for physician clinical performance assessment (11).

Table 1
Regulatory Functions of Healthcare-related Agencies in the United States

Function	Definition of Function	Agency Performing the Function
Accreditation	Granting of formal approval to graduate medical education programs; designed to ensure and improve the quality of physician training	ACGME, ACR
Licensure	Granting of legal permission to individual physicians to practice within their area of certification (certification is issued by a noncivil testing or training authority)	State (specialty board)
Credentialing	Assessment and confirmation of the qualifications of healthcare practitioners; based on their education, training, certification, licensure, and practice history	Medical center in conjunction with JCAHO
Privileging	Authorization of healthcare practitioners to provide specific services to patients; based on existing certification or a direct assessment of skills	Medical center

Sources.—References 1 and 2.

Teaching Point

In radiology, evaluation of performance in practice—or what could be called performance at the workstation—may be uniquely feasible, since the average case volume per physician in radiology exceeds that in many other specialties. In addition, adherence to standards of care in radiology may be defined by the degree of interpretive agreement between readers, which is believed to be useful for gauging diagnostic accuracy. Hence, in radiology, the evaluation of physician performance may be reduced to an assessment of diagnostic accuracy—whether or not a certain diagnostic feature was perceived, interpreted correctly, and reported.

Teaching Point

Peer review is the most commonly used method for assessing performance in terms of medical and clinical knowledge among radiologists (12). Ideally, it involves the fair and transparent evaluation of performance by a physician's peers to identify opportunities for additional education, error reduction, and self-improvement. If quantitative data about clinical outcomes were collected at the regional or national level, peer review might also facilitate objective benchmarking of radiologists and facilities according to their overall diagnostic performance.

The ACR currently requires that medical centers participate in physician peer review to obtain and maintain accreditation; only institutions and programs that specialize in mammography and stereotactic and ultrasonography-guided breast biopsy procedures are exempted (13). However, the ACR requirement is voluntary, and noncompliance incurs no penalty other than refusal of accreditation. The JCAHO similarly requires that all staff participate in peer review, with a continuous random review of 5% of cases for ongoing credentialing (14).

Because considerable variation exists in the identification and interpretation of abnormal imaging findings, the identification of a particular reading as "accurate" is subjective and possibly controversial. For this reason, we have focused on the measurement of diagnostic performance in the terms mandated by an increasing number of regulatory groups (15).

Attributes of Effective Peer-Review Systems

The key attributes of an effective method of clinical performance evaluation have been defined by the ABMS and JCAHO. Evaluation of performance in practice should reveal opportunities for quality improvement, help ensure competence, help improve individual outcomes, and provide evidence in cases of adverse and sentinel events (11,16). In addition, the peer-review process must be unbiased, fair, and balanced. Cases ideally should be selected at random to broadly represent the work performed in the radiology department. The views of the physician whose work is undergoing review, and any minority opinions, should be recorded. The evaluation process should be consistent, with all staff being aware of and adhering to established rules and procedures. The process also should be timely so as to represent the current state of performance; interpretations should be evaluated within a reasonable interval after the initial report. Moreover, peer review should be ongoing so that data can be tracked over time and analyzed to reveal trends. As with any part of the performance evaluation process, peer review should be nonpunitive, have a minimal effect on work flow, and allow easy participation.

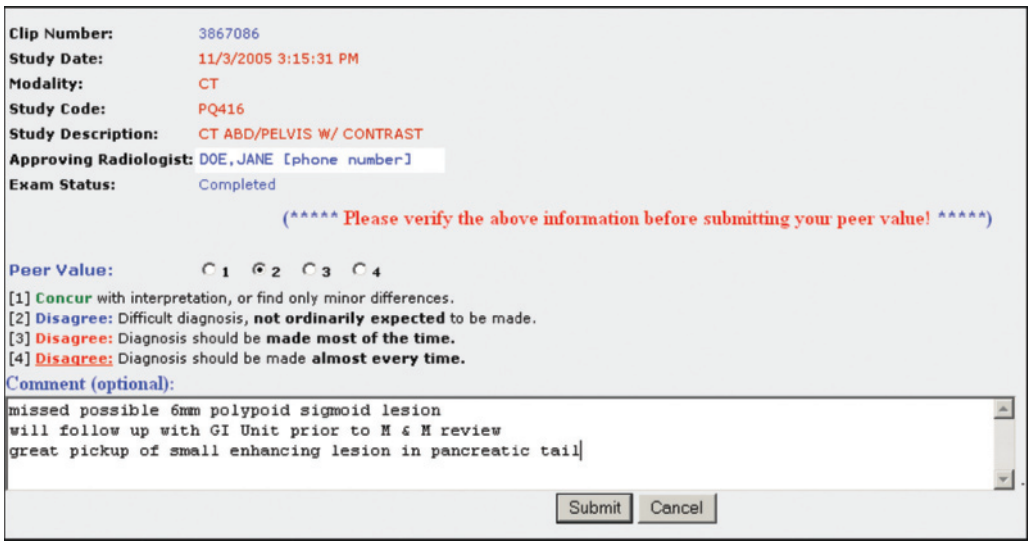
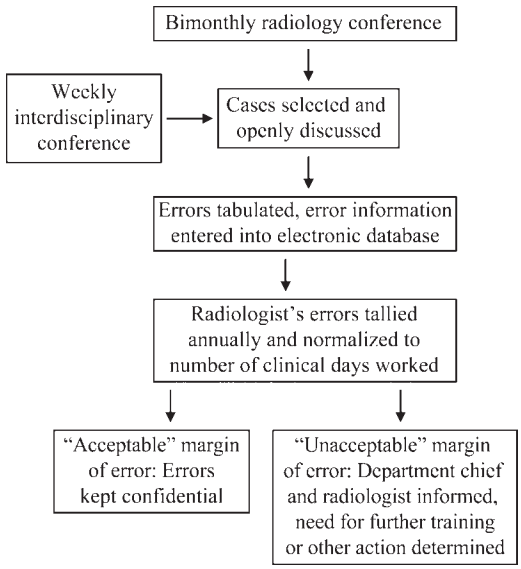


Figure 1. Screen capture shows the case review submission page in the RadReview workstation-integrated system.

Figure 2. Flowchart shows the process for performance-based assessment of radiology faculty as described in Jackson et al (22) and Accreditation Council for Graduate Medical Education (1). At bimonthly radiology conferences, discrepant cases may be submitted by either the initial reader or the reviewing radiologist. At weekly interdisciplinary conferences, clinicians from any specialty may submit review cases and contribute information. Halsted (23) believes that this system "empowers staff members to handle discrepant cases as they deem appropriate" because "they will be given the first opportunity to deal with errors themselves" and to notify patients and clinicians or dictate addenda to reports. Tracking of errors made by individual staff members, although it is required by the JCAHO, may be considered a drawback.



It is therefore best to incorporate the peer-review process into a preexisting evaluation system, such as a department's quality assurance program. Such systems are already used by attending radiologists to evaluate residents, fellows, and other staff. However, a simple, minimally intrusive, objective, and fair system for the routine clinical peer review of attending radiologists' diagnostic performance has yet to be developed at many institutions.

Existing peer-review systems have been classified as either proactive or reactive (17). One example of a proactive peer-review system would be a blinded, randomly assigned, routine double reading by separate radiologists (18). A proactive peer review might easily be incorporated into the protocol for performing and interpreting standardized radiologic studies; however, to our knowledge, no attempt has been made to establish such a peer-review system. In a reactive peer-review system, performance is assessed and documented only when discrepant reports are retroactively compared during routine clinical work (12,19). This type of peer review may be biased and subjective.

Workstation-integrated Peer Review

Workstation-integrated peer review is a reactive evaluation performed at the workstation. It requires no additional reading by the reviewing

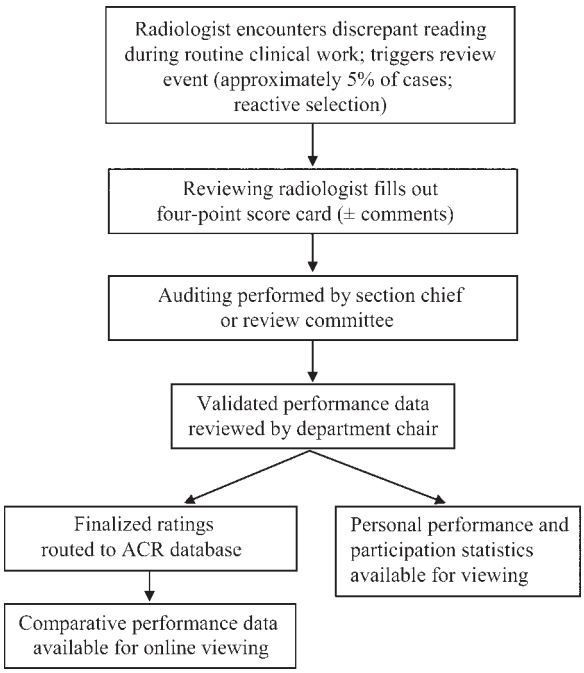


Figure 3. Flowchart shows the peer-review process in the RADPEER and RadReview workstation-integrated systems. A four-point scoring system is used to assess interpretive accuracy. Overlooked or misinterpreted findings in a routine case are grounds for a score of 3 or 4, which mandates auditing by the section chief or a peer-review committee. Validated performance data are stored in departmental and external databases for subsequent viewing and analysis. The online accessibility of performance data may provide motivation for improvements in individual and departmental performance and peer-review participation. However, interdepartmental comparisons are tenuous because of the absence of validated benchmarks.

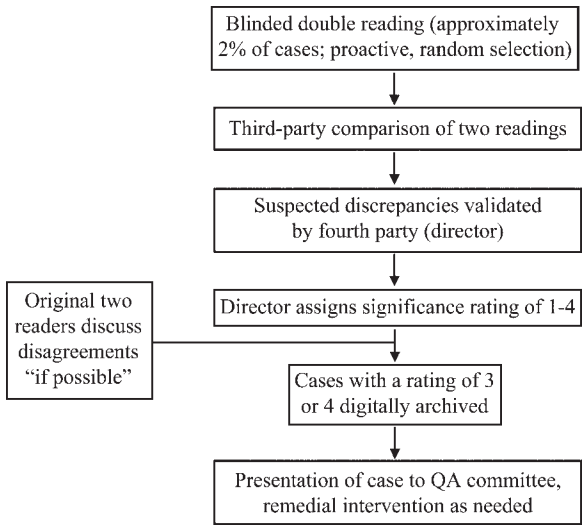


Figure 4. Flowchart shows the quality assessment process in a group radiology practice described by Soffa et al (18). The rating system used is similar to the four-point scoring system in RADPEER and RadReview, with a score of 3 or 4 indicating interpretive disagreement or outright misinterpretation of findings. *QA* = quality assurance.

radiologist and, hence, causes minimal interference with the regular work flow. It involves the review of findings previously reported by another radiologist.

One example of a workstation-integrated peer-review system is RADPEER, which was developed by the ACR. When a radiologist encounters a previously read study during routine interpretation of a current study, he or she may choose to use RADPEER to evaluate the previous radiologist's interpretation.

More recently, two other online peer-review systems were developed by using the same principles as RADPEER for scoring and reporting clinical performance data. These are eRADPEER, also developed by the ACR, and RadReview (Fig 1), a noncommercial program developed at the Beth Israel Deaconess Medical Center. The benefits of online peer-review systems include secure, personalized Web-based log-in, electronic data entry, and automated routing of performance data to the appropriate databases and quality assurance officials. No significant differences exist between these systems in terms of scoring or data entry.

In the RadReview system, data undergo a complex process before submission to the ACR database. The first step after scoring is data verification, wherein the section chief reviews rating statistics and selects cases with a score of 3 or 4 for auditing. Next, the section chief audits all reviews with these scores and may exercise the option of adjusting a score if he or she disagrees with the reviewer's assessment. The final ratings are submitted to the department chair for review. RadReview also includes designated screens for submission statistics and performance statistics of all participating staff, two powerful ways of encouraging improvement in performance and ongoing active participation in the peer-review process. eRADPEER allows online review of comparative performance data from other centers, although no benchmarking standards have yet been established.

In addition to these workstation-integrated systems, many other methods may be used for peer review of diagnostic performance in radiology (Figs 2–4; Tables 2, 3).

Table 2
Description of Processes for Peer Review of Diagnostic Accuracy in Radiology

Source and Reference Number	Peer-Review Metric	Performance Data Collection Method*		
		Description	Advantages	Disadvantages
Soffa et al (18)	Disagreement rate between two radiologists in a group practice	Blinded double reading of 2% of all cases	Not vulnerable to selection bias	Resource intensive: two parties initially read the study; third and fourth parties compare discrepant reports to render a final decision
Donnelly (12), Donnelly and Strife (20)	Medical and clinical knowledge; frequency of faculty errors	Incidental observation of discrepant interpretation of a study in any clinical scenario†	Meets guidelines set by JCAHO; is easily integrated into existing interdisciplinary clinical system; maximizes educational opportunities for whole department	Vulnerable to selection bias; unable to control for practitioner-specific case type or volume
McEnery et al (14), Borgstede et al (19), Strife et al (21)	Accuracy of interpretation	Workstation-integrated peer review (RADPEER, RadReview)	Meets guidelines set by ABMS MOC program Part IV; integrated into regular work flow, minimally disruptive‡	Selection bias; unable to control for practitioner-specific case type or volume

Note.—MOC = maintenance of certification.
*Regardless of the method of data collection used, data may be shared with all staff members at discrepancy review meetings to allow learning from previous mistakes.
†For example, discrepancies might be noticed during comparison of a current study with a previous one, during a clinical consultation, or during a clinical case-review conference.
‡Part IV of the ABMS MOC program involves an evaluation of practice performance that is linked to an ongoing practice quality improvement process for diagnostic radiology. Radiologists participate in one personal quality initiative project selected from a list of five targeted areas, including accuracy of interpretation.

Table 3
Timeline for Completion of Tasks in the Practice Quality Improvement Program

Year of Cycle	Task to Be Completed by Radiology Diplomate
1	Learn about PQI process, select project and metric(s)
2	Collect baseline data
3	Analyze data, work on improvement plan
4	Collect data, compare with initial data, summarize results
5	Modify improvement plan
6	Analyze data; if goals were achieved, select another PQI project
7	Summarize data, refine improvement plan
8	Continue collecting data
9	Complete collection of improvement plan data, analyze data, summarize data
10	Prepare final report of results and conclusions, maintain gain of first cycle, select topic for next cycle

Source.—Reference 21.
Note.—PQI = Practice Quality Improvement.

Table 4
Questions That Affect the Collection of Performance Data for Peer Review

What guidelines should be set for assuring adequate representation of all study types?
To what extent should complex studies (eg, multisequence MR imaging examinations) be reviewed retrospectively?
Who should do the initial review or double readings, and who should validate the results?
How many reports must be reviewed in order to ensure statistically meaningful representation of a radiologist's performance?*
How should cases be selected for review?
How should case review quotas be calculated—on a monthly or a yearly basis?
How can the anonymity of initial readers and reviewers be protected in a strictly reactive peer-review system?
How can a balanced review (eg, with random selection of cases) be achieved in a strictly reactive system?
How can reviews consistently be performed in a timely manner in a strictly reactive system?
Who should bear the financial burden of peer review?†

*The card-based review system offered by the ACR (RADPEER) allots as many as 50 cards per physician per month (19). If a radiologist's average annual case load is 10,000 studies, this translates into review of approximately 5% of the radiologist's annual case load. Soffa et al (18) applied blinded double reading to 2% of the annual case load. However, to our knowledge, the validity of these approaches has never been demonstrated.
†The annual fee for use of RADPEER, which is approximately $100–$400, decreases with an increasing number of participants. More objective and proactive systems of peer review can be expected to cost more than less objective ones. The cost of peer review likely will be absorbed by each facility, but if it is attributed to the bottom line, it could lead to an increase of 1%–2% in the average cost of radiologic studies.

Controversies, Challenges, and Future Possibilities

Several aspects of diagnostic radiology peer review remain controversial. For example, how should previously reported studies be selected for review, and to what extent should complex studies (eg, multisequence magnetic resonance [MR] imaging evaluations) be reviewed? Routine peer review of complex studies would clearly be onerous. Moreover, there is no consensus about what percentage of studies should be reviewed. Some institutions and organizations randomly select 5% of the annual radiology case load for peer review, but this proportion is not, to our knowledge, evidence-based. Other challenges that may arise include biases for or against the initial reader or modality used for a specific application, and collective targeting of individual physicians; however, such occurrences can be detected with regular reviews of scored cases and data. In addition, controversy may surround decisions about how to manage abnormal findings when these are detected retrospectively and are likely to affect patient management. Departmental or institutional policy should be used as a guide in these situations, which may include decision making about disclosures to the ordering physician and the patient.

In keeping with Deming's "plan-do-study-act" model for continuous quality improvement, efforts to optimize the quality of performance in clinical practice must proceed in conjunction with evidence about the validity of those efforts. Although existing peer-review systems have some of the attributes prescribed by the JCAHO and ABMS, rigorous evidence of the validity, reliability, and reproducibility of these systems is lacking; there is as yet no proof that they will yield useful information about performance and lead to quality improvement in future. Challenges to the collection of useful performance data and controversies surrounding their interpretation and use are summarized in Table 4 and discussed in the next section.

Teaching Point

Assessing the Validity and Reliability of Current Peer-Review Measures

Physicians' clinical performance may be assessed in terms of either clinical outcomes or rates of adherence to standards of care (11). As noted by Jackson et al (22), the relative feasibility of assessing adherence to standards of care by using interpretive agreement as the measure has led to a principal emphasis on interpretive agreement in peer review. Additional documentation of the clinical significance of potential errors has been accomplished by some (17,18). The second ACR task force on patient safety recently added this option to the existing dimensions of evaluation in RADPEER (22). Statements concerning the clinical significance of errors, while important, may not always be objective, and the task force has suggested that such qualifications be based on "'gut' assessments of the likelihood of impact … on patient care" (22).

Diagnostic accuracy has a great impact on the overall effectiveness of radiology practice—not only in terms of clinical outcomes but also in terms of resource utilization (7). For this reason, it may be worthwhile to expand the scope of performance evaluations beyond the number of false-positive findings, currently the single parameter used in some peer-review models. As part of a quality assurance effort, Lee (17) evaluated breast imaging reports according to the following three parameters: positive predictive value, disease detection rate, and abnormal interpretation rate (ie, the frequency of reports that led to further testing or procedures). Of course, that depth of analysis is feasible only in the most objective clinical settings—for example, when a study is performed to rule out a definitive disease entity or in cases where ample follow-up information exists. Still, some assessment of false-positive rates may be appropriate in a substantial number of cases, and the addition of that parameter to the scoring systems described earlier may be beneficial.

Reactive forms of peer review are plagued by yet another problem: selection bias. At an institution where an interdisciplinary model of peer review is applied, it has been asserted that "'no news is good news,' because any significant errors will be reported" (23). Contrary to this notion, our experience shows that bias may in fact be in play and result in the underreporting of errors. In addition, poor general compliance with peer review has been reported even with the application of a workstation-integrated model, which may represent the method of review that is least disruptive to work flow (19). Even if all discrepancies discovered during routine clinical work were openly reported, reactive peer review might still fail to reliably indicate performance quality. Some errors might be lost in the sea of clinical information, never to be followed up, reviewed, or reported. A partial solution to this problem might be to mandate peer review of a specified number of cases of each type read by each radiologist. Another solution that might be appropriate for larger academic departments is for all retrospective reviews to be performed by imaging subspecialists.

Related to the issue of selection bias is the question of whether physicians should be reviewed anonymously. Ideally, a radiologist's identity should remain confidential during the generation, auditing, and validation of performance data. While some measure of anonymity may be possible in proactive peer review, it is impossible in reactive peer review. Failure to ensure anonymity could adversely affect the objectivity of evaluation and infringe on the privacy of radiologists whose performance is being reviewed. Greater automation of the performance evaluation process and allotment of a certain percentage of cases for mandatory peer review may help safeguard anonymity. Systems could also be established for monitoring scored cases to detect any trend suggestive of bias.

Interpretive Disagreement: Analysis and Standardization of Performance Data

All of the currently applied peer-review methods assess interpretive disagreement between readers. Such assessments are based on the premise that variability implies potential oversight or error. However, since variability is not always the consequence of error (4,17,24), a third mediating party, such as the section chief or a review panel, must validate reviewers' scores. In future, the auditing of initial scores by a review committee may prove unnecessary: In a preliminary study, we found a high statistical correlation in ratings between radiologists and their section chiefs (unpublished data presented at the RSNA Annual Meeting, November 30–December 5, 2008).

Broad faculty involvement in performance evaluations may benefit even those physicians who are not directly involved in a peer-review event; however, mandatory participation in all such events may unnecessarily constrain resources and limit the scope of the peer-review process. Ideally, quality assurance should be as automated as possible, requiring faculty oversight only in special instances. With the advance of systems-based medicine and online patient medical records, we expect much greater automation also in the performance evaluation system.

Increasing automation should allow the electronic tracking of radiology reports over time to detect matching or mismatching of key-words with clinical outcomes.

Another problem created by reliance on measurement of interobserver discrepancy is that the data extracted can be used only to set internal standards, not to establish interfacility benchmarks. Borgstede et al (19) showed that intradepartmental committee-validated disagreement rates obtained with the use of RADPEER varied significantly between facilities. Even if intradepartmental discrepancy rates are consistent in many medical centers, as was reported by Jackson et al (22) and confirmed by our preliminary assessments performed with the use of RadReview, statistical averages of such data would be useless

for deriving absolute benchmarks of performance that would be valid for all facilities; interfacility disagreement rates may well exceed intradepartmental rates. To our knowledge, no study of interpretive agreement between members of different facilities has been attempted.

Effect of Peer Review on Individual Radiologists

Practitioner-specific clinical performance data have various uses and ramifications for individual radiologists. Collected on an ongoing basis, such data are necessary for credentialing and certification. As for licensing, the House of Delegates of the Federation of State Medical Boards of the United States declared in 2004 that "state medical boards have a responsibility to the public to ensure the ongoing competence of physicians seeking relicensure," although legislation to that effect will require state-by-state adoption of this policy (25,26). Performance data also may guide continuous self-improvement, and qualifications based on these data can be expected to affect healthcare choices of patients and providers. Moreover, pay-for-performance reimbursement, whereby a portion of the payment to a physician is based on his or her adherence to quality and efficiency standards, is increasingly common among social and private insurers. Federal and state laws afford some legal immunity in situations relating to peer review for the purposes of improving the quality of healthcare services. The legal and malpractice implications of peer-review systems are discussed below.

Managing Underperformance and Retrospectively Addressing Errors

If peer review is to have any beneficial effect, procedures must be put in place to correct or ameliorate individual radiologists' mistakes. For institutional accreditation, the ACR requires that all RADPEER scores of 3 or 4 be internally validated and acted on appropriately. Specific recommendations for action are seldom published because generalizations are difficult and may be counterproductive: If significant errors occur in routine cases, they are likely to be radiologist specific, and their correction in any case will be at the discretion of the institutional faculty.

In difficult cases, by contrast, interpretive discrepancy is more common, and it is there that quality improvement efforts may be most fruitful (19). Noting that diagnostic errors tend to occur in patterns, Halsted (23) argued that efficient quality improvement must involve open disclosure and discussion of discrepancies in interpretation at regular conferences, enabling all members of a department to learn from common pitfalls. This approach has led to quality improvement at various levels in a department where an interdisciplinary model of peer review was applied (Fig 2) (12,23). Department-wide conferences may be incorporated into any peer-review system.

The management of retrospectively detected errors that affect patient care should be guided by departmental and institutional policy.

Legal Considerations

In order to encourage free participation in quality improvement measures, federal and state laws afford certain legal privileges and immunities to performance-related data extracted or documented in the course of peer review and related proceedings. Immunity from discovery by the patient is contingent on state-specific laws. For instance, data initially generated by a peer-review committee specifically for the purpose of furthering quality improvement are legally privileged; however, documents produced by other entities, such as the hospital administration, or for purposes other than quality assessment and improvement are not privileged (27,28). The Health Care Quality Improvement Act of 1986, a federal law, affords legal immunity, "in suits brought by disciplined physicians, from liability for money damages to those who participate in professional peer review activities" (29). Actions undertaken by a peer-review committee must meet certain standards in order to qualify for immunity from legal claims brought by physicians who have been corrected or disciplined after making a gross error. The peer-review committee must have taken action *(a)* with the reasonable belief that the purpose of such action was to further healthcare quality; *(b)* after a reasonable effort to discover all the facts; *(c)* after providing adequate notice and a fair hearing to the physician; and *(d)* with the reasonable belief that disciplinary measures were warranted (30). The U.S. Supreme Court accordingly has ruled that actions by a peer-review committee are not protected by the Health Care Quality Improvement Act in antitrust or anticompetition proceedings (31).

In short, federal and state laws afford some privileges or immunities to peer-review bodies because legislators recognize the need to encourage candor and objectivity in peer investigations in order to improve healthcare quality. However, the need to prevent a chilling effect on

the accurate evaluation of healthcare facilities is balanced against the rights of individuals who are subject to performance review and litigation. The legal immunities and privileges conferred on peer-review committees therefore are restricted to legitimate actions taken by such committees, actions that are genuinely intended to improve healthcare.

Because legal protections for documents and information emanating from peer review are not uniform, radiologists are best advised to seek local counsel. Berlin (32) offers fascinating insights into these issues as they relate to the practice of radiology.

Implementation and Maintenance of a Peer-Review System

Whatever the method used for peer evaluation, the staff may be reluctant to participate. (Reasons for resistance to the implementation of peer-review procedures, as voiced by members of our staff, are shown in Table 5.) If peer review was only recently incorporated into the daily routine, frequent friendly reminders may be needed. Such reminders might take the form of pop-up windows on workstation monitors, e-mails, or text messages. It is important that peer-review systems be simple, have minimal effects on regular work flow, and demonstrate immediate as well as long-term benefits of participation. The latter objective can be accomplished in the short term with a performance summary like that provided by RadReview—for example, a breakdown of performance statistics by study type for all standardized studies, or by case type for all cases in specific categories. Such disclosures may reveal shortcomings that can be targeted for further focused training. Long-term demonstration of the benefits of any method of physician clinical performance assessment requires a continuous accumulation of data.

Teaching Point

Having considered these perspectives, it is worth noting the following statement made by Borgstede et al (19): "The philosophy of current quality improvement approaches … is to employ practical, easily obtained quality measures, although they may not be altogether unimpeachable, and to concentrate resources and energy on the work of improving quality, not on measurement." Indeed, although work to define the parameters of quality of care in radiology is ongoing (33), efforts to construct strict and standardizable methods for performance evaluation have been limited. Time will tell if more robust evaluation tools are necessary.

Table 5
Reasons for Resistance to Implementation of Peer Review in Our Department

"Peer review is time consuming." (Reviewing cases disrupts work flow; submitted data must be analyzed further.)
"Show me how the system really helps; I want to see the benefits of my efforts."
"Data can be used against us; I don't want to get my colleague into trouble."
"Peer review is too subjective."
"Some staff members don't seem to play fair."

Conclusions

Quality management in medicine is a growing endeavor. Recently updated criteria for certification, credentialing, and privileging place increased emphasis on the collection of practitioner-specific performance data and ongoing professional evaluation. Attributes of physician clinical performance assessment have been defined, and specific metrics have been recommended by the ABMS, JCAHO, and ACGME. In radiology, the assessment of reading accuracy, which allows quantitation of at least some performance parameters, has been implemented with minimal interference in the regular work flow. However, the validity and reliability of this method have not yet been rigorously demonstrated. Expansion of the parameters of peer review to include the number of false-positive findings, along with mandated review of specific case types and volumes, may solve some logistical problems but will likely face implementation challenges. For now, the primary challenge is to attain enthusiastic participation in peer review. Toward that end, we recommend frequent friendly reminders and ongoing confidential disclosures of performance data to participating radiologists.

Acknowledgments: The authors thank Rana S. Ziaee, an attorney in private practice, for her review of state and federal law and legal implications of peer-review programs; and Shifra Fraifeld, a research associate in the Department of Radiology at the Hadassah Medical Center, for her editorial contribution to the manuscript.

References

1. Accreditation Council for Graduate Medical Education. Understanding the difference between accreditation, licensure and certification. http://www.acgme.org/acWebsite/RRC_140/140_20UnderstandingtheDifference.pdf. Accessed February 23, 2009.
2. U.S. Department of Health and Human Services. Health Resources and Services Administration. Center for Risk Management. Policy Information Notice 01-16: Credentialing and Privileging of Health Center Practitioners. http://bphc.hrsa.gov/policy/pin0116.htm. Published July 17, 2001. Revised November 20, 2006. Accessed February 15, 2009.
3. Committee on Quality of Health Care in America, Institute of Medicine, National Academy of Sciences. Crossing the quality chasm: a new health system for the 21st century. Washington, DC: National Academies Press, 2001. http://books.nap.edu/catalog.php?record_id=10027#toc. Accessed February 12, 2009.
4. Board of the Faculty of Clinical Radiology, Royal College of Radiologists. Standards for radiology discrepancy meetings. http://www.rcr.ac.uk/docs/radiology/pdf/Stand_radiol_discrepancy.pdf. Published August 2007. Accessed June 1, 2009.
5. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348:2635–2645.
6. Deming WE. Elementary principles of the statistical control of quality. Tokyo, Japan: Nippon Kagaku Gijutsu Renmei, 1950.
7. Swensen SJ, Johnson CD. Radiologic quality and safety: mapping value into radiology. J Am Coll Radiol 2005;2:992–1000.
8. Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. Ann Intern Med 2005;142:260–273.
9. Assembly of the American Board of Medical Specialties. Description of the competent physician. Evanston, Ill: American Board of Medical Specialties, 1999.
10. Strife JL, Kun LE, Becker GJ, Dunnick NR, Bosma J, Hattery RR. The American Board of Radiology perspective on maintenance of certification: part IV—practice quality improvement in diagnostic radiology. Acad Radiol 2007;14:631–636.
11. Landon BE, Normand SL, Blumenthal D, Daley J. Physician clinical performance assessment: prospects and barriers. JAMA 2003;290:1183–1189.
12. Donnelly LF. Performance-based assessment of radiology practitioners: promoting improvement in accordance with the 2007 Joint Commission standards. J Am Coll Radiol 2007;4:699–703.
13. New accreditation physician peer-review requirements effective April 1, 2007. American College of Radiology Web site. http://www.acr.org/SecondaryMainMenuCategories/quality_safety/radpeer/new_requirements.aspx. Accessed February 23, 2009.
14. McEnery KW, Suitor CT, Hildebrand S, Downs RL. Integration of radiologist peer review into clinical review workstation. J Digit Imaging 2000;13:101–104.
15. Lockyer JM, Violato C, Fidler HM. Assessment of radiology physicians by a regulatory authority. Radiology 2008;247:771–778.
16. Joint Commission on the Accreditation of Healthcare Organizations. Comprehensive accreditation manual for hospitals (CAMH): the official handbook. Oakbrook Terrace, Ill: Joint Commission on the Accreditation of Healthcare Organizations, 2007.
17. Lee JK. Quality—a radiology imperative: interpretation accuracy and pertinence. J Am Coll Radiol 2007;4:162–165.
18. Soffa DJ, Lewis RS, Sunshine JH, Bhargavan M. Disagreement in interpretation: a method for the development of benchmarks for quality assurance in imaging. J Am Coll Radiol 2004;1:212–217.
19. Borgstede JP, Lewis RS, Bhargavan M, Sunshine JH. RADPEER quality assurance program: a multifacility study of interpretive disagreement rates. J Am Coll Radiol 2004;1:59–65.
20. Donnelly LF, Strife JL. Performance-based assessment of radiology faculty: a practical plan to promote improvement and meet JCAHO standards. AJR Am J Roentgenol 2005;184:1398–1401.
21. Strife JL, Kun LE, Becker GJ, Dunnick NR, Bosma J, Hattery RR. American Board of Radiology Perspective on Maintenance of Certification: Part IV—Practice Quality Improvement for Diagnostic Radiology. RadioGraphics 2007;27:769–774.
22. Jackson VP, Cushing T, Abujudeh HH, et al. RADPEER scoring white paper. J Am Coll Radiol 2009;6:21–25.
23. Halsted MJ. Radiology peer review as an opportunity to reduce errors and improve patient care. J Am Coll Radiol 2004;1:984–987.
24. Alpert HR, Hillman BJ. Quality and variability in diagnostic radiology. J Am Coll Radiol 2004;1:127–132.
25. Special report on the maintenance of licensure: interim report to the Board of Directors—February 2005. Web site of the Federation of State Medical Boards of the United States. http://www.fsmb.org/pdf/Special_Committee_MOL_Interim_Report_BOD_February2005.pdf. Accessed February 26, 2009.
26. Steinbrook R. Renewing board certification. N Engl J Med 2005;353:1994–1997.
27. Chicago Trust Company v Cook County Hospital. 698 NE2d 641 (Ill App 1998).
28. Green v Lake Forest Hospital. 781 NE2d 658 (Ill App 2002).
29. Bryan v James E. Holmes. 33 F3d 1318, 1321 (11th Cir 1994).
30. Mathews v Lancaster General Hospital. 87 F3d 624 (US App 1996).
31. Patrick v Burget. 486 US94 (1998).
32. Berlin L. Performance improvement and peer-review activities: are they immune from legal discovery? AJR Am J Roentgenol 2003;181:649–653.
33. Siewert B, Sosna J, McNamara A, Raptopoulos V, Kruskal JB. Missed lesions at abdominal oncologic CT: lessons learned from quality assurance. RadioGraphics 2008;28:623–638.

This article meets the criteria for 1.0 AMA PRA Category 1 Credit™. To obtain credit, see www.rsna.org/education/rg_cme.html.

RadioGraphics

Peer Review in Diagnostic Radiology: Current State and a Vision for the Future

by Shmuel Mahgerefteh et al

RadioGraphics 2009; 29:1221–1231 • **Published online** 10.1148/rg.295095086 • **Content Codes:** PR QA

Page 1222
Regulatory agencies base their assessments of medical staff partly on ongoing performance-based evaluations that include peer review.

Pages 1223
In radiology, evaluation of performance in practice—or what could be called performance at the workstation-—may be uniquely feasible, since the average case volume per physician in radiology exceeds that in many other specialties. In addition, adherence to standards of care in radiology may be defined by the degree of interpretive agreement between readers, which is believed to be useful for gauging diagnostic accuracy.

Page 1223
Peer review is the most commonly used method for assessing performance in terms of medical and clinical knowledge among radiologists. Ideally, it involves the fair and transparent evaluation of performance by a physician's peers to identify opportunities for additional education, error reduction, and self-improvement.

Page 1227
Although existing peer-review systems have some of the attributes prescribed by the JCAHO and ABMS, rigorous evidence of the validity, reliability, and reproducibility of these systems is lacking; there is as yet no proof that they will yield useful information about performance and lead to quality improvement in future.

Page 1230
It is important that peer-review systems be simple, have minimal effects on regular work flow, and demonstrate immediate as well as long-term benefits of participation.