Redefining the Practice of Peer Review Through Intelligent Automation Part 1: Creation of a Standardized Methodology and Referenceable Database

Bruce I. Reiner¹

Published online: 25 July 2017
© Society for Imaging Informatics in Medicine 2017

Abstract Conventional peer review practice is compromised by a number of well-documented biases, which in turn limit standard of care analysis, which is fundamental to determination of medical malpractice. In addition to these intrinsic biases, other existing deficiencies exist in current peer review including the lack of standardization, objectivity, retrospective practice, and automation. An alternative model to address these deficiencies would be one which is completely blinded to the peer reviewer, requires independent reporting from both parties, utilizes automated data mining techniques for neutral and objective report analysis, and provides data reconciliation for resolution of finding-specific report differences. If properly implemented, this peer review model could result in creation of a standardized referenceable peer review database which could further assist in customizable education, technology refinement, and implementation of real-time context and user-specific decision support.

Keywords Peer review · Data mining · Report analysis

Introduction: Peer Review and the Medical Standard of Care

In law, the term “standard of care” is defined as “the caution that a reasonable person in similar circumstances would exer-

cise in providing care to a patient.” In medical practice this is routinely translated as the duty to exercise the degree of care expected of a minimally competent physician in the same specialty and under the same circumstances [1]. An important point to be made is that the standard of care does not equate to perfection in practice. In the course of actual delivery, the actions of a physician can result in a negative clinical outcome without violating the standard of care [2]. “A competent physician is not liable per se for a mere error in judgment, mistaken diagnosis, or the occurrence of an undesirable result” [3].

For radiologists, the standard of care is largely defined by peer review, which is an evaluation by a professional colleague with comparable experience or training, and routinely performed through retrospective analysis of the radiology report [4]. A key concept is that of *comparability*, which should relate not only to the practitioner (i.e., peer reviewer) but also the task being performed, available data, and technology in use.

The challenge therefore is to define how comparable peers would act under the same circumstances. For a radiologist tasked with interpretation of an imaging exam, this equates to providing a fellow radiologist with the same data and technical resources and asking them to render professional service in their “usual and customary fashion.” But what exactly is usual and customary? Do radiologists perform their duties in a consistent and reproducible fashion? If the answer is no, then what is a reasonable expectation of inter- and intra-radiologist variability? Without objective and verifiable data, it is difficult to reliably determine minimal competence and similarity in practice. The net result is that peer review and the standard of care are inextricably tied to one another, and any bias or deficiency in one will adversely affect the other. It is therefore incumbent upon both the medical and legal communities to create well-defined and standardized methodologies which

✉ Bruce I. Reiner
breiner1@comcast.net

¹ Department of Radiology, Veterans Affairs Maryland Healthcare System, 10 North Greene Street, Baltimore, MD 21201, USA

remove conjecture, bias, and inference and are instead defined by objective and reproducible data.

Current Practice: the Biases and Prejudices of Conventional Peer Review and Standard of Care Analysis

Determining the standard of care in the case of alleged medical malpractice is often left to the subjectivity of the jury. In this analysis, jury members often rely on perceptions related to the credibility of the defendant physician, sympathy for the plaintiff, and the comparative believability of opposing expert witnesses [5]. Physician credibility may be determined by a wide array of subjective perceptions including physical appearance, emotional demeanor, perceived intellect, and “likability.” Plaintiff sympathy is in large part defined by the severity of the clinical outcome; the individual jury member’s feelings of empathy, sympathy, and compassion; and their own personal experiences. The comparative perception of expert witness believability may be affected by the same subjective traits used in analyzing the defendant physician, which may eventually lead to a decision of guilt or innocence on the basis of the “battle of opposing experts” [6].

While the American College of Radiology (ACR) has written a wide array of standards of radiology practice to serve as *general guidelines*, they may be modified in accordance with the specifics of each individual case and as a result provide a great deal of variability in defining the standard of care for each individual circumstance. The challenge therefore in attempting to objectify the standard of care is to create a scientific methodology which can take into account both generalizable (i.e., community wide) standards and case-specific (i.e., individual) variables.

The single most common cause for medical malpractice in radiology is the “missed” finding [4], with “missed” error rates reported to be as high as 30% [7]. Given the common occurrence of missed findings, one must seek to determine when a missed finding is the result of negligence and when it is justifiable. The traditional method of addressing this question is peer review, where professional peers (i.e., fellow radiologists) are asked to independently review the same imaging dataset and provide their own unbiased analysis. In the setting of a documented missed finding, a number of biases can be introduced into the peer review process, which can negatively impact the validity of the peer review process [8]. One such bias is *contextual bias*, which occurs when the reviewer has some degree of suspicion that the case in question has medicolegal consequences, and as a result, he or she changes their attention level and diagnostic threshold in the interpretation process, resulting in increased observer sensitivity and decreased specificity [9, 10]. A second form of observer bias in peer review is *hindsight bias* [11], which refers to the

tendency to overestimate the conspicuity of a given finding once it has been established, despite conscious attempts to review the case objectively [12]. The third form of bias is *outcome bias* [13], which refers to the impact of pre-existing knowledge of a missed finding on clinical outcome, which tends to increase the perception that negligence has occurred [14]. The net effect is that even when attempting to remain objective, traditional peer review often introduces bias, which can unfairly assign negligence to the physician of record.

Innovation Opportunity: Automated, Blinded, Continuous, and Data-Driven Peer Review

With the exception of breast imaging, most hospital-based radiology peer review is limited to the Joint Commission requirement of 5% of all cases performed [15]. This results in an exceedingly large number of cases (up to 95%) being excluded from formal peer review, with the potential for clinically significant diagnostic errors going undetected. In addition to this small sample size, conventional peer review is further limited by the customary practice of randomized selection, which is theoretically designed to represent the full spectrum of radiology practice [16]. One could reasonably argue that a preferable approach to peer review selection would be a combination of random and targeted case selection, the latter of which is intended to identify *high-risk* cases a priori and automatically incorporate them into the peer review queue with the goal of improving clinical outcomes by prospectively identifying cases more prone to diagnostic error.

If one was to create technology which could automate this case selection process on a continuous basis (through combined randomized and targeted case selection) and increase the overall percentage of cases beyond 5% of case volume, peer review would theoretically be improved both in qualitative and quantitative terms. A greater challenge and opportunity for improvement lies in addressing the existing biases and prejudices inherent in conventional peer review practice. In order to accomplish this, one must create a process which is completely blinded to peer review participants. In short, the peer reviewer would be unaware that he or she was participating in peer review and instead believe they were interpreting an exam in their routine and customary fashion.

To illustrate how such a system would work, we take an example of a CT exam which has been designated as “high risk,” based on computerized analysis of peer review risk factors (which will be discussed in detail in part 2 of this series). Upon selection as an “active peer review” case, the exam and accompanying data would be automatically transferred to a peer review database which would be tasked with selecting a qualified and comparable peer reviewer and transmitting the exam data to the institutional archive of the designated peer reviewer. Once received, the case would be automatically

routed and assigned to the reading queue of the designated peer review reader, without any designation of its peer review status. The peer reviewer would open the case and be presented with data customary to the normal workflow and exam-related data. In certain instances, masking may be required to provide ancillary data (e.g., ordering physician name, acquisition technology identifying data) in order to hide the peer review status of the case. Upon completion of the case, the resulting report would be automatically transferred to the peer review database for comparative analysis with the comparable report of the primary reader. (The methodology for this comparative report analysis is described in detail in article 3 of this series).

The net result is that the targeted peer review case is independently interpreted by multiple readers in a truly blinded fashion, with comparable data availability and removal of any underlying bias. Upon completion of inter-report analysis, the data is recorded into a referenceable peer review database where it can undergo standardized peer review analysis. In the case of alleged medical malpractice, the same peer review process can be used with multiple peer reviewers, in order to better define the standard of care over a larger size reader group. The ultimate goal would be to standardize the peer review process though objective, reproducible, and unbiased data analysis.

Peer Review Database and Analytics

At the core of the proposed innovation strategy is the referenceable peer review database, which is designed to contain standardized peer review data which can be comingled among large numbers of institutional and individual user groups to create large sample size peer review statistics. Table 1 lists a number of finding-specific analytics which can be derived from the peer review database.

While the radiologist community remains an important focus of analysis, other individual participants playing a role in the imaging event can also be analyzed, with the goal of gaining insight as to their respective roles in diagnostic

accuracy and clinical outcomes. Technologists play an important role since they are directly involved in protocol selection, exam acquisition, image processing, and retrieval of pertinent clinical and imaging data. While not routinely included in conventional peer review, referring clinicians also play an important role in peer review analysis since they are directly involved in exam ordering and providing relevant clinical exam, history, and test data. In some circumstances, referring clinicians may also have access to historical imaging report data performed outside of the host institution which may be inaccessible to the radiologist and prove useful in determination of temporal change, diagnosis, and clinical significance. Another participant for inclusion in peer review analysis is the patient, who plays a critical yet frequently overlooked role in diagnostic accuracy. A variety of patient attributes of consideration include body habitus, ambulatory status, morbidity, compliance, and venous accessibility. Any one of these variables may adversely affect imaging data quality, and in turn diagnostic accuracy. In addition, patients often serve as important resources for clinical data.

Since report accuracy is directly related to the quality and quantity of available data, technology (i.e., both hardware and software) also plays a critical role in peer review analysis. This technology dependence takes on even greater importance in the current practice environment, in which aggressive radiation dose reduction is used to promote patient safety, at the potential expense of image quality and diagnostic accuracy. Only through standardized data analysis, can the potential interaction effect of technology and report accuracy be properly elucidated.

When available, clinical outcome data (e.g., pathology reports, clinical test data, discharge summary) can be incorporated into the peer review database to provide valuable knowledge in determining ground truth. The ultimate goal of this collective data is to provide customizable peer review analytics which can be used for targeted user and context-specific education, workflow distribution, deployment of decision support tools, and technology refinement.

References

1. Strauss DC, Thomas. What does the medical profession mean by “standard of care”? *J Clin Oncology* 2009; 27: 192–193
2. Moffett P, Moore G: The standard of care: legal history and definitions: the good news and the bad news. *West J Emerg Med* 12:109–112, 2011
3. Hall v Hillburn, 466 So.2d 856 (Miss. 1985)
4. Kaewlai R, Abujudeh H: Peer review in clinical radiology practice. *AJR* 199:158–162, 2012
5. Vidmar N: Juries and medical malpractice claims: empirical facts versus myths. *Clin Orthop Relat Res* 2:367–375, 2009
6. Weintraub MI: Expert witness testimony: an update. In: Weintraub MI Ed.. *Neurologic Clinics of North America*. Philadelphia: W.B.Saunders, 1999, pp. 363–370

Table 1 Finding-specific analyses derived from the peer review database

1.	Technology
2.	Radiologist
3.	Technologist
4.	Referring clinician
5.	Exam type
6.	Modality
7.	Anatomy/organ system
8.	Clinical context
9.	Patient
10.	Disease state

7. Krupinski EA: Current perspectives in medical image perception. *Alteration, Perception & Psychophysics* 5:1205–1217, 2010
8. Durand DJ, Robertson CT, Agarwal G et al.: Expert witness blinding strategies to mitigate bias in radiology malpractice cases: a comprehensive review of the literature. *J Am Coll Radiol* 11:868–873, 2014
9. Harvey JA, Fajardo LL, Innis CA: Previous mammograms in patients with impalpable breast carcinoma: retrospective vs blinded interpretation. *AJR* 161:1167–1172, 1993
10. Egglin TK, Feinstein AR: Context bias. A problem in diagnostic radiology. *JAMA* 276:1752–1755, 1996
11. Berlin L: Hindsight bias. *AJR* 175:597–601, 2000
12. Erly WK, Tran M, Dillon RC et al.: Impact of hindsight bias on interpretation of nonenhanced CT head scans for acute stroke. *J Comput Assist Tomogr* 34:229–232, 2010
13. Berlin L: Outcome bias. *AJR* 183:557–560, 2004
14. LaBine SJLG: Determinations of negligence and the hindsight bias. *Law Hum Behav* 21:501–516, 1996
15. McEnery KW, Sutor CT, Hildebrand S et al.: Integration of radiologist peer review into clinical review workstation. *J Digit Imaging* 13:101–104, 2000
16. Mahgerefteh S, Kruskal JB, Yam CS et al.: Peer review in diagnostic radiology: current state and a vision for the future. *RadioGraphics* 29:1221–1231, 2009