ORIGINAL ARTICLE

# Radiologist Peer Review by Group Consensus

*H. Benjamin Harvey, MD, JD[a,b,c], Tarik K. Alkasab, MD, PhD[a,b], Anand M. Prabhakar, MD[a,b], Elkan F. Halpern, PhD[b,c], Daniel I. Rosenthal, MD[a,b], Pari V. Pandharipande, MD, MPH[a,b,c], G. Scott Gazelle, MD, PhD, MPH[a,b,c]*

## Abstract

**Purpose:** The objective of this study was to evaluate the feasibility of the consensus-oriented group review (COGR) method of radiologist peer review within a large subspecialty imaging department.

**Methods:** This study was institutional review board approved and HIPAA compliant. Radiologist interpretations of CT, MRI, and ultrasound examinations at a large academic radiology department were subject to peer review using the COGR method from October 2011 through September 2013. Discordance rates and sources of discordance were evaluated on the basis of modality and division, with group differences compared using a $\chi^2$ test. Potential associations between peer review outcomes and the time after the initiation of peer review or the number of radiologists participating in peer review were tested by linear regression analysis and the *t* test, respectively.

**Results:** A total of 11,222 studies reported by 83 radiologists were peer reviewed using COGR during the two-year study period. The average radiologist participated in 112 peer review conferences and had 3.3% of his or her available CT, MRI and ultrasound studies peer reviewed. The rate of discordance was 2.7% (95% confidence interval [CI], 2.4%-3.0%), with significant differences in discordance rates on the basis of division and modality. Discordance rates were highest for MR (3.4%; 95% CI, 2.8%-4.1%), followed by ultrasound (2.7%; 95% CI, 2.0%-3.4%) and CT (2.4%; 95% CI, 2.0%-2.8%). Missed findings were the most common overall cause for discordance (43.8%; 95% CI, 38.2%-49.4%), followed by interpretive errors (23.5%; 95% CI, 18.8%-28.3%), dictation errors (19.0%; 95% CI, 14.6%-23.4%), and recommendation (10.8%; 95% CI, 7.3%-14.3%). Discordant cases, compared with concordant cases, were associated with a significantly greater number of radiologists participating in the peer review process (5.9 vs 4.7 participating radiologists, $P < .001$) and were significantly more likely to lead to an addendum (62.9% vs 2.7%, $P < .0001$).

**Conclusions:** COGR permits departments to collect highly contextualized peer review data to better elucidate sources of error in diagnostic imaging reports, while reviewing a sufficient case volume to comply with external standards for ongoing performance review.

**Key Words:** Radiology, peer review, RADPEER, quality, safety, error, quality assurance, quality improvement, regulation, consensus, COGR

*J Am Coll Radiol 2016;■:■-■. Copyright © 2016 American College of Radiology*

## INTRODUCTION

Physician peer review is widely recognized as a fundamental component of health care quality assurance [1]. Experts believe that physician peer review will result in better clinical outcomes by monitoring the quality of care, increasing adherence to recognized standards, and creating a culture of transparency around issues of patient safety [2-4]. Trying to measure the impact of peer review programs on clinical practice and patient outcomes is fraught with difficulties, and studies to date have been limited in scope with mixed findings [5-9]. Nonetheless, a large Cochrane review found that audit and feedback interventions, including peer review, can drive quality improvement if physician feedback remains a primary focus [5].

The radiology community was an early adopter of physician peer review, with workflow-integrated peer review systems in widespread use as early as 2006 [10,11].

[a]Department of Radiology, Massachusetts General Hospital, Boston, Massachusetts.

[b]Harvard Medical School, Boston, Massachusetts.

[c]Institute for Technology Assessment, Massachusetts General Hospital, Boston, Massachusetts.

Corresponding author and reprints: H. Benjamin Harvey, MD, JD, Massachusetts General Hospital, Department of Radiology, 175 Cambridge Street, Suite 200, Boston, MA 02114; e-mail: hbharvey@partners.org.

Outside the submitted work, Dr Gazelle has received personal fees from GE Healthcare. Dr Pandharipande has received research funding from the Medical Imaging and Technology Alliance, for unrelated research.


1546-1440/16/$36.00 ■ http://dx.doi.org/10.1016/j.jacr.2015.11.013

Since then, many accrediting bodies and third-party payers have mandated the adoption of radiologist peer review by diagnostic imaging groups [12]. RADPEER is a workstation-integrated peer review system developed by the ACR and represents the earliest and the most widely used peer review system in diagnostic imaging [10]. Modeled from the traditional process of double reading, the ease and convenience of RADPEER-style peer review has driven its widespread adoption. Yet critics of RADPEER feel that it is too limited in its focus and fails to address many important aspects of quality in diagnostic imaging, such as report clarity and length and adherence of the interpretation to national guidelines and standards [13]. RADPEER also has relatively weak feedback mechanisms, an essential aspect for effective peer review [5].

Attempting to harness the strengths of RADPEER while increasing the robustness of peer review, our department developed a novel peer review process for radiologists, known as consensus-oriented group review (COGR) [13]. The COGR process has been previously described in detail, but in brief, it is a software-enabled peer review process in which groups of radiologists meet regularly to review randomly selected cases and record consensus on the acceptability of the issued reports [13]. The purpose of this study is to evaluate the feasibility of the COGR method of radiologist peer review within a large subspecialty radiology department.

## MATERIALS AND METHODS

### Human Subjects Compliance

This retrospective, HIPAA-compliant study was approved by our institutional review board. The need to obtain patient consent was waived.

### Peer Review Data Collection

The study was performed in a radiology department at a 950-bed tertiary care academic center. The radiology department has more than 100 staff radiologists, greater than 85% of whom are subspecialized by organ system, and more than 500,000 diagnostic imaging studies are performed and interpreted in the radiology department annually. All data were collected during the study period from October 1, 2011, to September 30, 2013, the initial two years of COGR.

Peer review data were prospectively recorded during the COGR process of radiologist peer review. COGR consisted of regular division-specific meetings of a minimum of three radiologists tasked to peer-review a randomly selected sample of recently interpreted cases in a conference setting. To be available for review, the case must have been read by one of the radiologists participating in the COGR conference. The participating radiologists were able to select the time window from which cases are selected, but it is generally 3 to 7 days. On the basis of the participating radiologists and the specifications chosen for the peer review conference (ie, time window, modality type, number of cases per participant), the COGR software tool randomly selects the cases to be reviewed. For each case, the group of radiologists reviewed the images and the report (deidentified with respect to the radiologist) on a PACS-enabled workstation or projected onto a larger screen. The group then attempted to arrive at a consensus as to whether the issued report was adequate or needed to be changed. Consensus required unanimous agreement of all the radiologists participating in the conference. The outcome for each case could be consensus that the report as issued was acceptable, consensus that the report should be changed, or an inability to reach consensus.

If the outcome of the case was consensus that the report should change or inability to reach consensus, the software provided a free text box for the group to record relevant case information, in particular the rationale for the decision reached. The time required for a group of radiologists to review a case and record consensus was generally less than 1 to 2 minutes, depending on the complexity of the case. The radiologist who originally interpreted the case participates in the peer review discussion and consensus decision and may choose to remain anonymous or not.

Radiologists chose which COGR conferences to attend within their division on the basis of their availability but were required to participate enough such that at least 2% of their advanced imaging cases underwent peer review.

### Data Analysis

The COGR software tool was queried for the following data for all cases peer reviewed during the study period: subspecialty division, interpreting radiologist, examination type, modality, date of peer review, peer review conference participants, consensus decision, and, if applicable, a corresponding free text entry and addendum date. Additionally, the COGR software tool was queried for the following radiologist participation metrics for the study period: number of conferences attended, number of studies peer reviewed, and total number of studies available for review (ie, all CT, MRI, or ultrasound studies interpreted by the radiologist during the study period). All data were recorded and analyzed in Microsoft Excel 2010 (Microsoft Corporation, Redmond Washington).