

# Bias in Neuroradiology Peer Review: Impact of a “Ding” on “Dinging” Others

 P. Charkhchi,  B. Wang,  B. Caffo, and  D.M. Yousem

## ABSTRACT

**BACKGROUND AND PURPOSE:** The validity of radiology peer review requires an unbiased assessment of studies in an environment that values the process. We assessed radiologists' behavior reviewing colleagues' reports. We hypothesized that when a radiologist receives a discrepant peer review, he is more likely to submit a discrepant review about another radiologist.

**MATERIALS AND METHODS:** We analyzed the anonymous peer review submissions of 13 neuroradiologists in semimonthly blocks of time from 2016 to 2018. We defined a discrepant review as any one of the following: 1) detection miss, clinically significant; 2) detection miss, clinically not significant; 3) interpretation miss, clinically significant; or 4) interpretation miss, clinically not significant. We used random-effects Poisson regression analysis to determine whether a neuroradiologist was more likely to submit a discrepant report during the semimonthly block in which he or she received one versus the semimonthly block thereafter.

**RESULTS:** Four hundred sixty-eight discrepant peer review reports were submitted; 161 were submitted in the same semimonthly block of receipt of a discrepant report and 325 were not. Receiving a discrepant report had a positive effect on submitting discrepant reports: an expected relative increase of 14% (95% CI, 8%–21%). Notably, receiving a clinically not significant discrepant report (coefficient = 0.13; 95% CI, 0.05–0.22) significantly and positively correlated with submitting a discrepant report within the same time block, but this was not true of clinically significant reports.

**CONCLUSIONS:** The receipt of a clinically not significant discrepant report leads to a greater likelihood of submitting a discrepant report. The motivation for such an increase should be explored for potential bias.

Peer review is one form of evaluation of a radiologist's performance, mostly targeting the diagnostic accuracy of interpretation.<sup>1</sup> The 2007 medical staff standards of The Joint Commission (<https://www.jointcommission.org/>) have strengthened the peer review process by explicitly requiring focused and ongoing professional practice evaluations. These standards evaluate a practitioner's knowledge, skill, and behavior. Focused professional practice evaluations involve an intense assessment of a practitioner's credentials and current competence at the initial appointment in a practice. Ongoing professional practice evaluations are the routine monitoring of current physician competency, which includes but is not limited to assessment of a

practitioner's ongoing interpersonal and communication skills, professional behavior, practice competency, and behavior as a team member.<sup>2</sup> To address these standards, most radiology practices use some form of peer review to assess radiologists' accuracy and performance.<sup>3,4</sup>

The primary goal of radiology peer review is to reduce diagnostic errors, educate radiologists to their blind spots and areas for improvement, and improve patient safety. In addition to evaluating the radiologist's technical performance, peer review can evaluate communication skills, interpersonal relationships, team cooperation, and responsiveness.<sup>5</sup>

The American College of Radiology currently recommends that medical centers participate in physician peer review to obtain and maintain accreditation. Many radiology groups have committed to using a peer review system due to hospital requirements, The Joint Commission standards, or recommendations from specialty societies.<sup>6</sup> There are different types of peer review systems, including RADPEER, implemented by American College of Radiology (<https://www.acr.org/Clinical-Resources/RADPEER>). Our department is currently using a home-grown peer review system with the advantage of an integrated information technol-

Received July 10, 2018; accepted after revision October 24.

From the Department of Radiology and Radiological Science (P.C., D.M.Y.), Division of Neuroradiology, Johns Hopkins Medical Institutions, Baltimore, Maryland; and Department of Biostatistics (B.W., B.C.), Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland.

Please address correspondence to David M. Yousem, MD, MBA, Johns Hopkins Bloomberg School of Public Health, 600 N Wolfe St, Baltimore, MD 21287; e-mail: [dyousem1@jhu.edu](mailto:dyousem1@jhu.edu); [@dyousem1](mailto:@dyousem1)

<http://dx.doi.org/10.3174/ajnr.A5908>

ogy solution that allows the review of cases within 24 hours of completion, thereby catching errors early, rather than several months later. The process is anonymized so that neither the reviewer nor the original reader knows the author of the reports or the peer reviews. Both RADPEER and our internal system are scoring-based peer review systems.

As Kaewlai and Abujudeh<sup>5</sup> indicated, 2 critical areas for success in peer review are a positive peer review culture and a committed team. Larson et al<sup>7</sup> indicated that scoring-based systems tend to drive radiologists inward, against each other and against practice leaders. Our aim in this communication was to critically assess the radiologists' behavior in the setting of reviewing colleagues' reports. We hypothesized that when a radiologist receives a discrepant review, he or she would be more likely to report a discrepant review (colloquially referred to as a "ding") on another person's report within 2–4 weeks.

## MATERIALS AND METHODS

Because this project dealt with quality-improvement processes, it was deemed by the Johns Hopkins institutional review board to be exempt from review. The data collected were independent of protected health information; therefore, the study was Health Insurance Portability and Accountability Act–compliant.

### Data Source

This is a retrospective study. We used peer review data from the division of neuroradiology because of the early implementation of the internal system by this team. The peer review system is completely anonymous: The radiologists are aware of neither who has reviewed their reports nor whose report they are reviewing. The program randomly selects colleagues' reports from the previous 24 hours and assigns them to the peer reviewer, providing the report without the author of the report being identified. It opens the images on the case. The radiologist reviewing the case must choose whether he or she concurs with the interpretation, identifies a detection miss, or identifies an interpretation miss. If a miss is identified, it is scored as clinically significant or clinically not significant. If such a miss is identified, whether significant or not, the peer reviewer fills in a text box identifying the miss. The radiologist submits the case; and if a miss has been identified, the original reader receives an e-mail immediately thereafter notifying him or her that there is a discrepancy and to review the case.

We collected data from January 1, 2016, to January 30, 2018, in increments from the first of the month to the 15th day of the month and from the 16th day of the month to the last day of the month (ie, twice a month for 25 months) for all 13 neuroradiologists who were practicing in our facility for the full duration of the study. The system provides data on the number of cases that are read by the radiologist, the number of peer reviews that the radiologist completed, the number of discrepant reviews that the radiologist submitted, the number of the radiologist's cases that were reviewed by neuroradiology colleagues, the number of discrepant cases in which a lesion was not detected and whether it was clinically significant, and the number of cases in which a lesion was appropriately detected but its etiology was misinterpreted and whether that misinterpretation was clinically significant. We defined a discrepant review in our study as any one of the

following: 1) detection miss clinically significant, 2) detection miss clinically not significant, 3) interpretation miss clinically significant, and 4) interpretation miss clinically not significant.

The members of the neuroradiology division are encouraged to perform peer review each day that a neuroradiologist has clinical duties, and all members must review cases at a rate equaling at least 3% of the total number of cases they read each month (ie, if they read 600 cases, they have to peer review at least 18 cases from colleagues). Fulfilling this participation rate is part of their end of year bonus "quality and safety" calculation.

### Study Variables

We defined the independent variable as the receipt of a discrepant report. We defined 1 dependent variable as submitting a discrepant review within the semimonthly time block of receiving a discrepancy and a second dependent variable as submitting a discrepant review in the semimonthly block after receiving the discrepant review. As an example, if someone received a discrepant report on day 7 of the month, we surveyed for discrepant reports within the block that included the first 15 days of the month (first dependent variable) and the last half of the month (second dependent variable) for a discrepant submission by that person. If the discrepancy was received, for example, on the 18th day of the month, then the second half of the month becomes the first dependent variable and the first 15 days of the next month were the second dependent variable. We did not include more than 4 weeks because we assumed that the likelihood of a reflexive response diminished after a 2- to 4-week interval.

By virtue of collecting the data twice a month for 2 years of practice from January 1, 2016, to December 31, 2017, and including the 4 weeks of follow-up extending to January 31, 2018, we had 50 data points. We also assessed the association between the type of discrepant report (clinically significant versus clinically not significant) and submitting the discrepant report on others.

### Data Analysis

We used random-effects Poisson regression models to assess the effect of receiving a discrepant review on submitting a discrepant report within the received block and the next block (with doctor as the random effect). Also, we included a multivariate regression model in our analysis using clinically significant and non-clinically significant reports as covariates to assess the association of receiving different types of discrepant reviews (clinically significant versus clinically not significant) and submitting one. All analyses were performed with R statistical and computing software (Version 3.4.3; [www.r-project.org](http://www.r-project.org)).

We ran 2 sensitivity analyses by including and excluding outliers. In the first analysis, we excluded 1 radiologist with the largest number of submitted discrepant reports. This radiologist reported, on average, 3.26 discrepant reports, while the mean of all radiologists was  $0.75 \pm 1.6$ . We thought this radiologist may influence the results because this radiologist submitted the highest number of discrepant reviews in the study time period. In the second sensitivity analysis, we checked for any extreme observations and excluded greater than 5 discrepancy reports received or submitted in a block and repeated the analysis.

## RESULTS

The overall distribution of submitted reports for each neuroradiologist is presented in Fig 1.

In the 2-year period, 486 discrepant peer review reports were submitted, of which 161 were submitted by individuals in the same 2-week block in which they received notice of a discrepant report; 325 were not. There was a positive effect (coefficient = 0.13; 95% confidence interval, 0.08–0.19) between submitting a discrepant report within the block of receiving one. The relative rate was 14% (95% CI, 8%–21%). In other words, according to the model, for every 5 discrepancy reports received, the number of

discrepancies submitted will be doubled (times 1.93). If one does not receive any discrepant reports, then he or she will submit 0.47 discrepancy reports (on average) in a 2-week block.

There was no statistically significant effect between receiving a discrepant report in one block and then submitting one in the following 2-week time block (coefficient,  $-0.09$ ; 95% CI,  $-0.19$ – $0.02$ ) (Table and Fig 2).

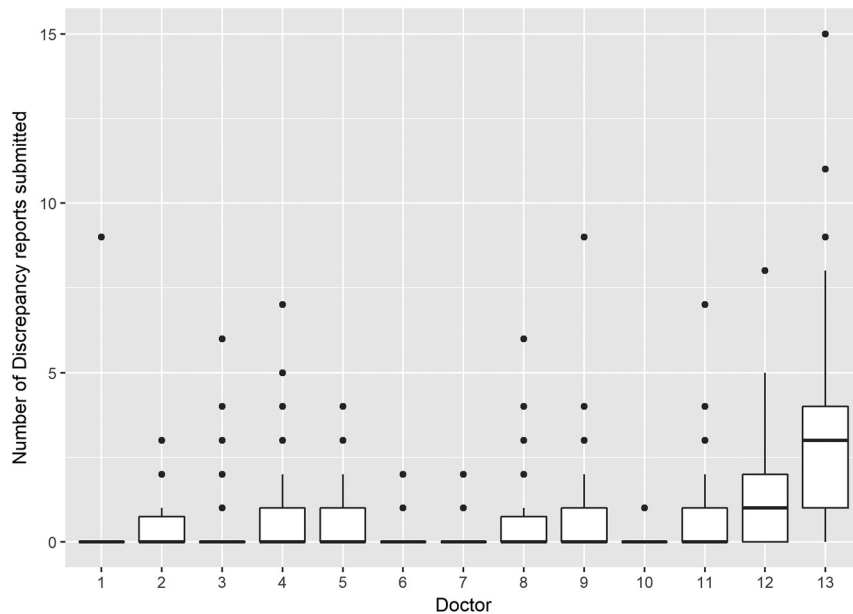
If one ran a multivariate regression analysis to assess the effect of different discrepant reports on submitting a ding on others, there was a significant association between receiving a not clinically significant (coefficient = 0.13; 95% CI, 0.05–0.22) report

and submitting a discrepant report, while there was no statistically significant association between receiving a clinically significant (coefficient = 0.26; 95% CI,  $-0.04$ – $0.55$ ) discrepant report and submitting a discrepancy in the same time block. There was no significant association between receiving clinically significant or not clinically significant discrepant reports and submitting a ding on others in the next following 2-week time block (Table).

After excluding 1 outlier radiologist who submitted the most discrepancies, there were no changes in our results: For the same block analysis, the coefficient changed from 0.13 to 0.12 (95% CI, 0.03–0.21) and remained statistically significant; for the next time block, the impact remained nonsignificant (coefficient =  $-0.05$ ; 95% CI,  $-0.18$ – $0.07$ ).

The coefficient for each radiologist is reported in Fig 3. If one compared the 2 time blocks, there was no significant association between receiving and submitting a discrepant review for any of the radiologists in the next time block (Fig 3).

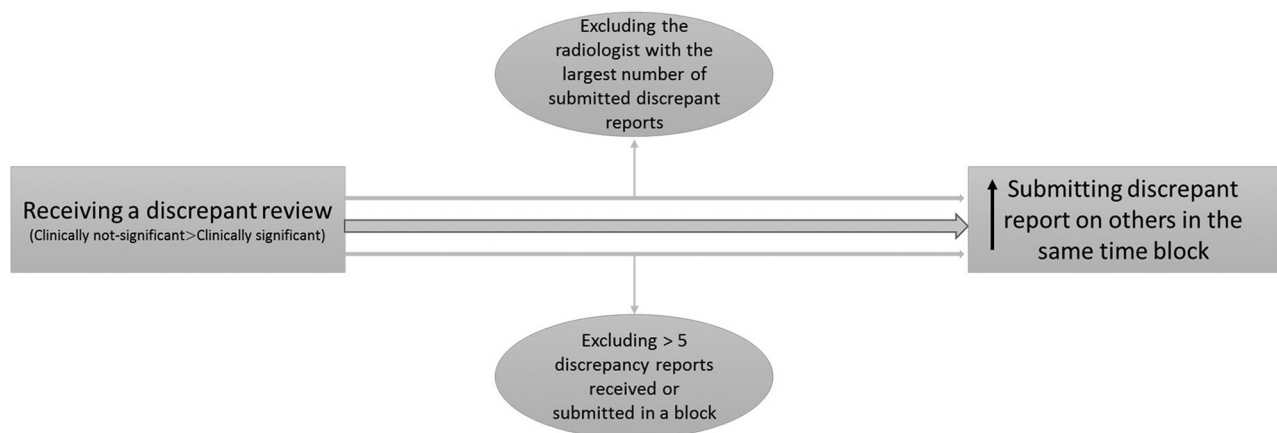
Removing outliers (more than 5 discrepant or submitted reports) kept the



**FIG 1.** Mean number of discrepancy reports submitted for each neuroradiologist.

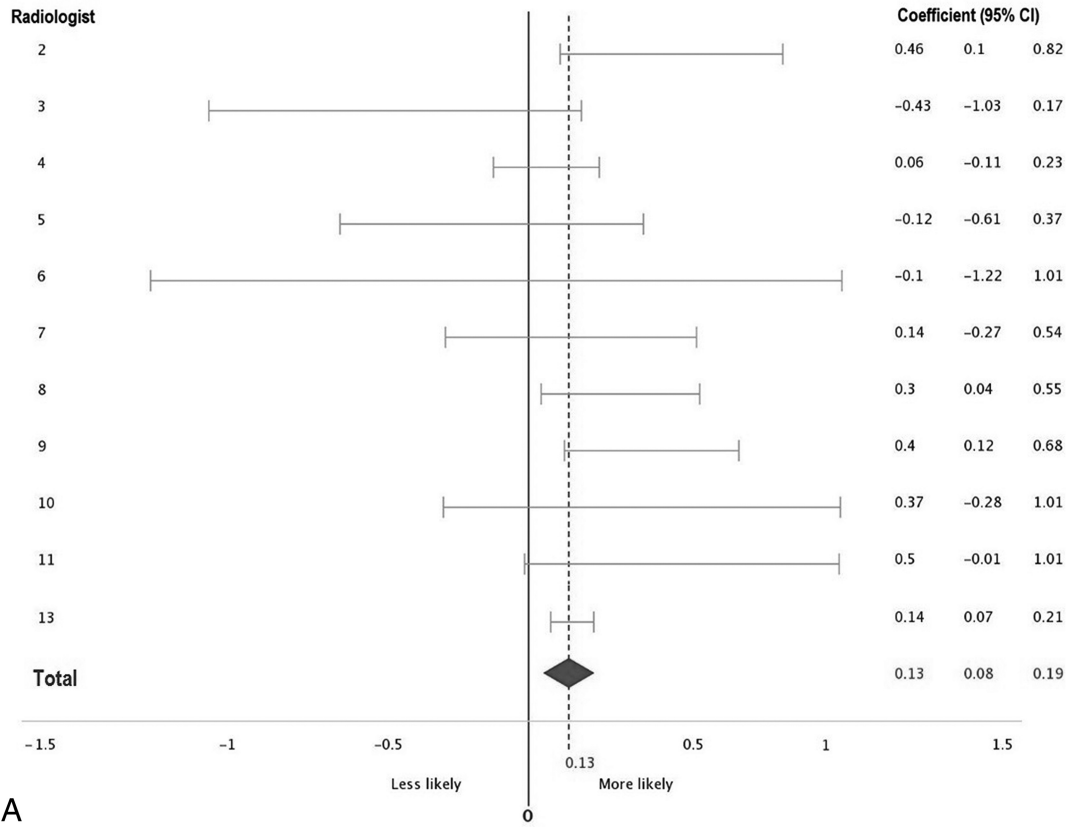
### The association between receiving and submitting discrepant reports in the same and next time block in a clinical radiology peer review system

Received a Discrepant Report	Submitting a Discrepant Report within Same Time Block (Coefficient) (95% CI)	Submitting a Discrepant Report in the Next Time Block (Coefficient) (95% CI)
Any kind of discrepant report	0.13 (0.08–0.19)	$-0.09$ ( $-0.19$ – $0.02$ )
Clinically significant discrepant report	0.26 ( $-0.04$ – $0.55$ )	$-0.30$ ( $-0.81$ – $0.20$ )
Clinically not significant discrepant report	0.13 (0.05–0.22)	$-0.05$ ( $-0.18$ – $0.07$ )



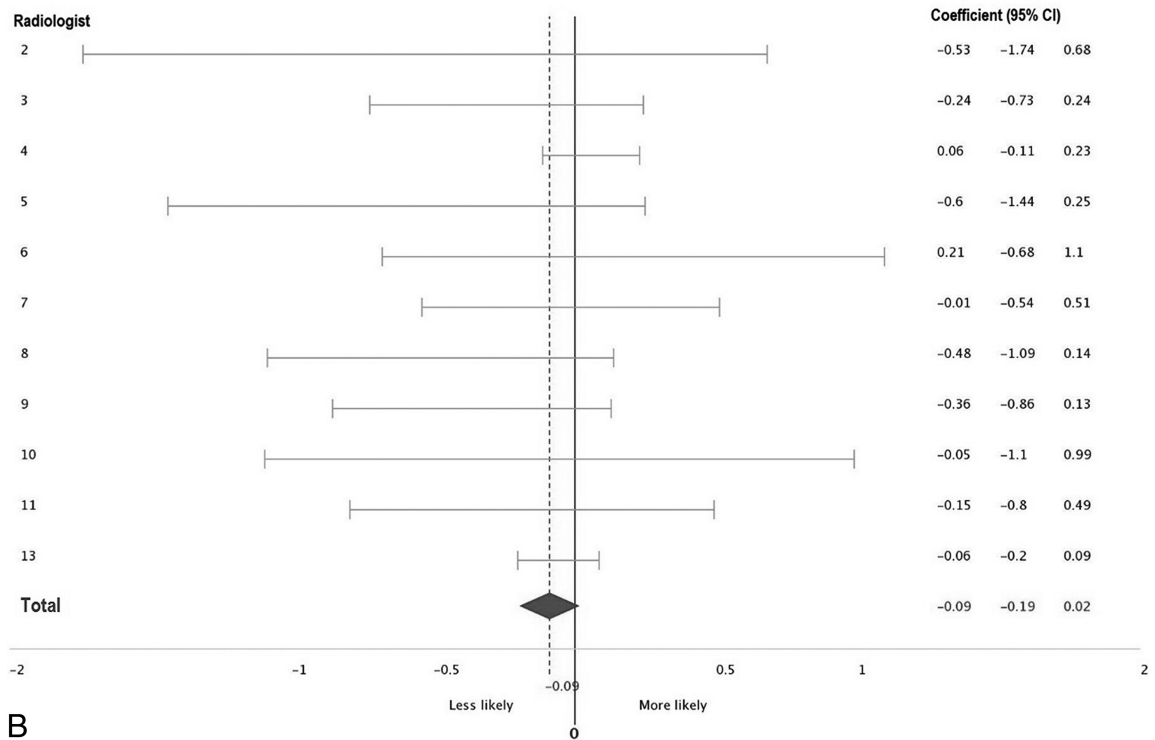
**FIG 2.** The association between receiving and submitting a discrepant report and 2 sensitivity analyses.

Association of submitting a discrepant report with receiving one.



A

Association of submitting a discrepant report with receiving one.



B

**FIG 3.** Association of receiving a discrepant review with submitting one for each radiologist. Plot A (top) indicates the current time block; plot B (bottom), the next time block.

current time block significant (coefficient = 0.32; 95% CI, 0.12–0.34) and the next time block nonsignificant (coefficient = 0.06; 95% CI, –0.08–0.19).

## DISCUSSION

We found that when a radiologist in our study received a discrepant report, he or she was more likely to submit a discrepant peer review report within the 2-week time block of receiving it. The observed effect was not seen in the following 2-week block of time, suggesting an immediate reaction to the ding rather than a delayed or sustained effect. Receiving a clinically not significant report and submitting a discrepant report on others are significantly positively correlated compared with receiving a “clinically significant” report.

To the best of our knowledge, this is the first article studying physicians’ reactions to a discrepant report in the clinical setting. Data are well-published on causes of discrepancies in radiology and also strategies to prevent them.<sup>8–10</sup> However, none of these articles qualitatively or quantitatively studied the radiologists’ behavior on receiving a discrepant report.<sup>7,11</sup>

Generally, there is a negative attitude toward the peer review system among radiologists. In an American College of Radiology survey assessing the RADPEER program, most radiologists opined that the peer review system is only performed to meet accreditation and hospital credentialing requirements.<sup>12</sup> Nearly half believed that their practice patterns had not changed as a result of peer review. One-third of respondents admitted that there was underreporting of disagreements in the peer review process at their practice.<sup>12</sup> This underreporting highlights the current peer review systems deficits. Peer review may elicit anxiety, shame, humiliation, and fear, leading to a reluctance to report disagreements.<sup>7</sup> These factors may lead to the behavior demonstrated in our study. If the peer review system is converted into a retributive instrument among colleagues, it becomes worse than meaningless; it becomes destructive.

On the other hand, the positive effect of receiving a discrepant report on submitting discrepant reports may illustrate a positive bias rather than a negative reaction. While previous studies have shown that radiologists tend to underreport discrepancies on peer review,<sup>12,13</sup> in contrast, our data suggest that receiving a discrepant report may motivate the radiologist to review their colleagues’ reports more diligently and potentially identify errors that might otherwise be overlooked. However according to our findings, participants tend to submit more discrepancy reports on their colleagues when they receive a not clinically significant report compared with a clinically significant one. We posit that this result may be in favor of a motivation for a retributive reaction rather than motivating the reviewer to be more conscientious. When a radiologist receives a discrepant report that is clinically significant, she or he may react with gratitude and not negatively react to it, but if she or he receives a clinically not significant (“nuisance”) discrepant report, the radiologist may be more likely to respond by submitting a reciprocal discrepant report on a colleague.

There are a few limitations associated with this study. First, the peer review system we use is unlike most peer review programs that use historical studies for review. In other words, most peer review systems require the radiologist to review a comparison study from months to years earlier. In that gap, the diagnosis may

become clear and, for example, growth of a missed cancer can be readily detected. By limiting our peer review system to reviews within the previous 24 hours, we identify discrepancies earlier, but a final diagnosis may not be clear at that point. Second, if the reviewing radiologist wanted to enter the Radiological Information System (RIS) or Electronic Medical Record (EMR), he or she could break the anonymity of the self-contained peer review program and identify who read each study. Third, we used semi-monthly time intervals because our peer review system data are collected this way. We cannot determine whether the radiologist immediately submitted a discrepant report the same hour or day that he or she received a discrepant report because we do not have the data on the exact time of receiving and submitting reports. We do not monitor the peer review system at such a granular level. On a similar note, the reporting function of the program is able to document that a dinged physician submitted a discrepant report but not the type (detection versus interpretation/significant or not) of report. Finally, if a discrepancy is challenged by the receiver, the division chief then adjudicates the 2 reviews, which could change the initial discrepant designation by the dinged physician.

How can we address this potential bias in the peer review system? We could write code to the program that after receiving notice of a discrepant report and reviewing it, that individual is “frozen” from submitting any peer review reports for 7 days. Thus, the more immediate “gut” reaction could be assuaged. Education and re-education continuously on the purpose of peer review may also be helpful. Providing data showing the overall results and how well individuals perform may decrease the psychological impact of a solitary discrepant review. In addition, department leadership support to keep peer review results completely anonymous, blinded to leadership, and accessible only by individual physicians can improve rate of participation in the peer review system.

## CONCLUSIONS

When a radiologist in our study received a discrepant report, he or she was more likely to submit a discrepant report within the semi-monthly block of time of receiving it. The observed effect was not seen in the following block of time, suggesting an immediate reaction to the ding rather than a delayed or sustained effect. The impact was maximal after receiving a clinically Radiological Information System (RIS) discrepant peer review.

Disclosures: Brian Caffo—UNRELATED: Consultancy: for d8lab. Comments: I do personal consulting from time to time; Grants/Grants Pending: National Institutes of Health\*; Royalties: Leanpub Publishing. Comments: book royalties; Payment for Development of Educational Presentations: Becton Dickinson. Comments: on-line courses. David M. Yousem—UNRELATED: Expert Testimony: medicolegal work; Payment for Lectures Including Service on Speakers Bureaus: American College of Radiology Education Center speaker; Royalties: Elsevier for 5 books; Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed: Radiological Society of North America 2018 as an awardee. \*Money paid to the institution.

## REFERENCES

1. Alpert HR, Hillman BJ. **Quality and variability in diagnostic radiology.** *J Am Coll Radiol* 2004;1:127–32 [CrossRef Medline](#)
2. Freedman S. **How 2007 joint commission standards expand hospital peer review.** 2007. <https://www.psqh.com/analysis/peer-review-how-2007-joint-commission-standards-expand-hospital-peer-review/>. Accessed November 17, 2018



3. 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;184:1398–401 CrossRef Medline
4. 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 CrossRef Medline
5. Kaewlai R, Abujudeh H. **Peer review in clinical radiology practice.** *AJR Am J Roentgenol* 2012;199:W158–62 CrossRef Medline
6. Kruskal J, Eisenberg R. **Focused professional performance evaluation of a radiologist: a Centers for Medicare and Medicaid Services and Joint Commission requirement.** *Curr Probl Diagn Radiol* 2016;45:87–93 CrossRef Medline
7. Larson DB, Donnelly LF, Podberesky DJ, et al. **Peer feedback, learning, and improvement: answering the call of the Institute of Medicine Report on Diagnostic Error.** *Radiology* 2017;283:231–41 CrossRef Medline
8. Brady AP. **Error and discrepancy in radiology: inevitable or avoidable?** *Insights Imaging* 2017;8:171–82 CrossRef Medline
9. Bruno MA, Walker EA, Abujudeh HH. **Understanding and confronting our mistakes: the epidemiology of error in radiology and strategies for error reduction.** *Radiographics* 2015;35:1668–76 CrossRef Medline
10. Brady A, Laoide RÓ, McCarthy P, et al. **Discrepancy and error in radiology: concepts, causes and consequences.** *Ulster Med J* 2012;81:3–9 Medline
11. FitzGerald R. **Radiological error: analysis, standard setting, targeted instruction and teamworking.** *Eur Radiol* 2005;15:1760–67 CrossRef Medline
12. Abujudeh H, Pyatt RS, Bruno MA, et al. **RADPEER peer review: relevance, use, concerns, challenges, and direction forward.** *J Am Coll Radiol* 2014;11:899–904 CrossRef Medline
13. Moriarity AK, Hawkins CM, Geis JR, et al. **Meaningful peer review in radiology: a review of current practices and potential future directions.** *J Am Coll Radiol* 2016;13:1519–24 CrossRef Medline