Integrating the ACR Appropriateness Criteria Into the Radiology Clerkship: Comparison of Didactic Format and Group-Based Learning

Marjorie W. Stein, MD*, Susan J. Frank, MD*, Jeffrey H. Roberts, MD*, Malka Finkelstein, MD*, Moonseong Heo, PhD*

Abstract

Purpose: The aim of this study was to determine whether group-based or didactic teaching is more effective to teach ACR Appropriateness Criteria to medical students.

Methods: An identical pretest, posttest, and delayed multiple-choice test was used to evaluate the efficacy of the two teaching methods. Descriptive statistics comparing test scores were obtained.

Results: On the posttest, the didactic group gained 12.5 points ($P < .0001$), and the group-based learning students gained 16.3 points ($P < .0001$). On the delayed test, the didactic group gained 14.4 points ($P < .0001$), and the group-based learning students gained 11.8 points ($P < .001$). The gains in scores on both tests were statistically significant for both groups. However, the differences in scores were not statistically significant comparing the two educational methods.

Conclusions: Compared with didactic lectures, group-based learning is more enjoyable, time efficient, and equally efficacious. The choice of educational method can be individualized for each institution on the basis of group size, time constraints, and faculty availability.

Key Words: American College of Radiology, ACR Appropriateness Criteria, didactic lectures, group-based learning, educational, imaging, radiology

INTRODUCTION

The skyrocketing cost of medicine in the United States cannot be sustained in the face of limited economic resources. A portion of this increase is related to the use of inappropriate imaging [1]. Evidence-based guidelines have been developed by the ACR, called the ACR Appropriateness Criteria®, to help physicians choose the most appropriate imaging test for more than 200 clinical scenarios [2]. The criteria offer information on alternative imaging tests and relative radiation exposure. Although most radiologists are familiar with the ACR Appropriateness Criteria, the vast majority of clinicians, in our experience, are unaware of the existence of this online resource [3]. Many practicing physicians decide which imaging test to order on the basis of past experience and intuition, rather than on evidence-based guidelines. They then transmit this knowledge to their trainees. Medical students are our future physicians, and it is our responsibility as educators to integrate the ACR Appropriateness Criteria into their curricula.

There are various educational methods available to teach medical students how to use the ACR Appropriateness Criteria. Traditionally, didactic lecturing was the standard format with which to transmit information to students. Recently, there has been a shift to more active educational techniques, such as group-based learning, which appeals to technology-oriented millennial learners [4]. Passively listening to a lecture and recording...
information may not be fully engaging to millennials, whereas a group discussion while using a flexible search-based computer resource more closely mirrors how millennials process information and make decisions in their daily lives. The purpose of our study was to determine which of these instructional techniques is most effective to teach specific ACR Appropriateness Criteria to third-year medical students on their 2-week radiology clerkship.

METHODS
At our institution, medical students have a 2-week required radiology clerkship during their third year of medical school, composed of groups of four or five students. Since July 2012, the ACR Appropriateness Criteria have been incorporated into our radiology clerkship using two different educational methods, which were alternated monthly. The two educational methods were didactic lectures and group-based learning; the content of each session was identical. The methods were compared to determine which is most efficacious in teaching selected ACR Appropriateness Criteria to our students. This prospective study was approved by our institutional review board.

During the 1-hour didactic lecture, one of two faculty members who were clerkship site directors described the study of choice and alternative imaging examinations for specific clinical scenarios involving gastrointestinal and genitourinary pathology, as outlined in the ACR Appropriateness Criteria. The relative radiation exposure for each imaging test was also discussed. Pertinent cross-sectional imaging was presented.

At the start of the 1-hour group-based learning activity, the students received a packet containing 10 clinical vignettes focused on the same ACR Appropriateness Criteria as the didactic group. The students were provided the ACR Appropriateness Criteria and then collaborated to decide the best imaging test to order for each scenario. After 35 min, the same two faculty members joined the group to facilitate group discussion. They reviewed the students' responses and guided them when necessary to the most appropriate clinical decision.

Three tests were administered to evaluate the efficacy of the teaching methods. Students were given identifying numbers, and thus their scores were anonymous. Participation was voluntary, as explained to the students at the outset of the clerkship. The students took the same 10-question multiple-choice pretest on the first day of the clerkship and then 2 days after the intervention. Three months after completion of the radiology clerkship, a 10-question multiple-choice delayed test was submitted to the students using SurveyMonkey (www.surveymonkey.com), a web-based survey and questionnaire tool. The content of the three tests were identical, but patient names and question order were varied. These three 10-question multiple-choice tests were developed to test basic knowledge and recall. The score range of the 10-question test was 0 to 100. In the last year of the study, four higher level questions with scores ranging from 0 to 100, testing analytic skills and image recognition, were added and separately graded (Appendix 1).

Statistical Analysis
Descriptive statistics are provided in terms of mean and SD. Because both didactic lectures and group-based learning sessions were carried out in a class format, we took into account for statistical inference the potential correlations of scores among students within the same classes by applying a mixed-effects linear model with class-level intercepts as random. This approach was applied to both within- and between-group comparisons as well as to testing the significance of slopes. The reported P values are therefore based on the mixed-effects linear models. Two-sided P values less than .05 were considered to indicate statistical significance. All statistical analyses were conducted with SAS version 9.3 (SAS Institute, Cary, North Carolina).

RESULTS
From July 2012 to July 2015, 221 students took the pretest. Of these, 215 students (97.3%) took the posttest, and 78 of the initial participants (36.3%) went on to take the three-month delayed test. A total of 53 students took the four higher level questions pretest, of whom 51 (96.2%) took the posttest and 14 (26.4%) the delayed test.

Both groups learned and retained knowledge, as evidenced by gains in test scores on the 10-question multiple choice test. When the pretest and posttest scores were compared, the didactic group gained 12.5 points ($P < .0001$), and the group-based learning students gained 16.3 points ($P < .0001$). When the delayed and pretest scores were compared, the didactic group gained 14.4 points ($P < .0001$), and the group-based learning group gained 11.8 points ($P < .001$). The incremental gains in scores for both the posttest and three-month delayed test compared with pretest scores were statistically significant for both groups. However, when comparing the two
educational methods, the differences in scores between the two groups were not statistically significant (Table 1).

The additional test containing four higher level learning questions had mixed results. When the pre- and posttest scores were compared, the didactic group gained 2.1 points \((P = .022)\), and the group-based learning group gained 1.5 points \((P = .099)\). When the delayed pretest and posttest scores were compared, the didactic group gained 0.9 points \((P = .451)\), and the group-based learning group gained 0.0 points \((P = 1.00)\). Again, the gains in scores between the two learning groups were not statistically significant (Table 2).

There was a significant association between absolute values of pretest and posttest scores and clinical experience on the 10-question test. Increasing clinical experience was defined as the number of rotations the student completed before the radiology clerkship. The change in score per unit change in rotation was defined by the \(\beta\) score, which was 1.91 \((P < .0001)\) for the pretest and 1.14 \((P = .006)\) for the posttest, both statistically significant (Fig. 1). However, the slope difference was not significant. The association between absolute test scores and clinical experience was not statistically significant for the delayed test and the four-question pretest, posttest, and delayed tests.

**DISCUSSION**

In this era of excessive and often inappropriate imaging utilization, soaring health care costs, decreasing reimbursements, and limited monetary resources, it is essential that medical student educators not only expose their students to the existence of the ACR Appropriateness Criteria but integrate these guidelines into the curriculum. Faculty with this online resource will help these doctors in training be conscientious about radiation exposure and order the most appropriate imaging test, enabling them to become wise guardians of limited health care dollars.

Our study shows that our students' test scores, which evaluated their knowledge, higher level analytic skills, and retention, were not statistically different whether the students were taught the ACR Appropriateness Criteria using a didactic lecture or group problem-based learning setting. Once we determined the preferred teaching method for our study group, our objective was to integrate this method into teaching sessions covering Appropriateness Criteria for conditions involving all organ systems at all teaching sites at our institution. However, we were left with two equivalent educational methods and needed to decide which method to use going forward.

<table>
<thead>
<tr>
<th>Table 1. Comparison of test results between two modalities of teaching: 10-question test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variable</strong></td>
</tr>
<tr>
<td><strong>n</strong></td>
</tr>
<tr>
<td>Pretest</td>
</tr>
<tr>
<td>Posttest</td>
</tr>
<tr>
<td>Delayed test</td>
</tr>
<tr>
<td>Posttest—pretest</td>
</tr>
<tr>
<td>Delayed test—pretest</td>
</tr>
<tr>
<td>Delayed test—posttest</td>
</tr>
</tbody>
</table>

Note: \(P\) values are based on mixed-effects linear models.

*Within-group tests.
<table>
<thead>
<tr>
<th>Table 2. Comparison of test results between two modalities of teaching: four higher level learning question test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variable</strong></td>
</tr>
<tr>
<td><strong>N</strong></td>
</tr>
<tr>
<td>Pretest</td>
</tr>
<tr>
<td>Posttest</td>
</tr>
<tr>
<td>Delayed test</td>
</tr>
<tr>
<td>Posttest—pretest</td>
</tr>
<tr>
<td>Delayed test—pretest</td>
</tr>
<tr>
<td>Delayed test—posttest</td>
</tr>
</tbody>
</table>

Note: \(P\) values are based on mixed-effects linear models.

*Within-group tests.

*Between-group tests.
We have decided to use the group-based learning method for future medical student groups because it is educationally equivalent, in our experience more enjoyable for both teacher and learner, and more effective in terms of time management. We spend 1 hour per session directly teaching the students when we deliver the didactic lecture. In comparison, with the small-group problem-based approach, we spend 25 min per session, as the students independently research and solve problems before we facilitate the discussion. The initial time creating the didactic lecture and problem-based sessions is similar.

Our required radiology clerkship occurs in 2-week blocks over the course of the medical students’ third year. As expected, pretest scores are lower earlier in the year, when students have been exposed to fewer clinical rotations and thus have less clinical experience. Pretest scores are higher toward the end of the year when students have been exposed to more rotations and have gained more clinical experience. Therefore, students who rotate in radiology earlier have a greater incremental gain on the posttest, as they start out with lower scores, and students who rotate later have less of an incremental gain, as they have higher pretest scores.

Review of the educational literature reveals higher learner satisfaction with more active problem-based learning [5-10]. However, there are conflicting data regarding the benefits of didactic versus small-group problem-based learning in terms of acquisition and retention of knowledge. Doucet et al [11] compared the efficacy of the two educational methods on teaching headache diagnosis and management to primary care physicians and found equivalent results on the pretests but higher scores on the posttest and three-month delayed test after problem-based learning. They concluded that new information is better retained and applied if taught in an environment similar to the actual work environment. In addition to imparting information, group-based learning enhances teamwork and analytic problem-solving skills, as well as retrieval of information. In a study conducted by Yoo and Park [12], a lecture-based curriculum was compared with a case-based format for nursing students, and they concluded that the latter method improved not only problem-solving abilities but also enhanced communication skills and the ability to perform self-directed learning. White et al [13] conducted a study comparing the efficacy of the two educational methods in teaching management of asthma using evidence-based medical guidelines to family physicians. They found no significant difference in test scores on pretests, posttests, and three-month delayed tests, similar to our results. The group-based learning group rated the subjective value of their educational experience higher than the didactic group [13]. Similarly, Fischer et al [5] conducted a study comparing lecture and small-group discussion in a third-year obstetrics and gynecology course and found that the students preferred the more active small-group discussion, although there was no difference in test scores between the groups.

A review conducted by Thurley and Dennick [6] highlighted the benefits of group-based learning: teaching problem-solving skills useful to the future physician-educator, as well as expanding on prior knowledge, in an enjoyable, collaborative atmosphere. However, group-based learning has certain drawbacks. Group-based learning is ideal in the setting of small, but not large, groups. For large groups, multiple sessions would be needed, which is more time consuming and faculty intensive compared with the traditional lecture format. Computers are often needed to promote self-directed learning in group-based learning sessions, which increases departmental costs [6].

Limitations of our study include the presence of only two instructors conducting both the didactic and group-based learning sessions, thus limiting the ability to generalize to other instructors. Also, our subjects were from a single medical school. Because there is no established approach for group-based learning, our exercise may be more or less complex than that used at other institutions. This difference might affect the time spent creating and facilitating the group-based session. Furthermore, there was attrition of subjects taking the delayed test.

It is imperative that we expose future physicians to the ACR Appropriateness Criteria. Our experience, as well as the educational literature, demonstrates that this generation of learners prefers a more active, self-directed education.
learning environment. Group-based learning is enjoyable, time efficient, and comparably effective to didactic lectures. The choice of educational method to use can be individualized for each institution on the basis of group size, time constraints, and faculty availability.

TAKE-HOME POINTS
- Many practicing physicians decide which imaging test to order on the basis of past experience and intuition rather than evidence-based guidelines. They then transmit this knowledge to their trainees.
- Medical students are our future physicians, and it is our responsibility as educators to integrate the ACR Appropriateness Criteria into their curriculum.
- Our study shows that our students’ test scores, which evaluated their knowledge, higher level analytic skills, and retention, were not statistically different whether the students were taught the ACR Appropriateness Criteria using a didactic lecture or group problem-based learning setting.
- The choice of educational method to use can be individualized for each institution on the basis of group size, time constraints, and faculty availability.

ADDITIONAL RESOURCES
Additional resources can be found online at: http://dx.doi.org/10.1016/j.jacr.2015.12.020.

REFERENCES