

Orion: A Web-Based Application Designed to Monitor Resident and Fellow Performance On-Call

Jason N. Itri · Woojin Kim · Mary H. Scanlon

Published online: 20 January 2011
© Society for Imaging Informatics in Medicine 2011

Abstract Radiology residency and fellowship training provides a unique opportunity to evaluate trainee performance and determine the impact of various educational interventions. We have developed a simple software application (Orion) using open-source tools to facilitate the identification and monitoring of resident and fellow discrepancies in on-call preliminary reports. Over a 6-month period, 19,200 on-call studies were interpreted by 20 radiology residents, and 13,953 on-call studies were interpreted by 25 board-certified radiology fellows representing eight subspecialties. Using standard review macros during faculty interpretation, each of these reports was classified as “agreement”, “minor discrepancy”, and “major discrepancy” based on the potential to impact patient management or outcome. Major discrepancy rates were used to establish benchmarks for resident and fellow performance by year of training, modality, and subspecialty, and to identify residents and fellows demonstrating a significantly higher major discrepancy rate compared with their classmates. Trends in discrepancies were used to identify subspecialty-specific areas of increased major discrepancy rates in an effort to tailor the didactic and case-based curriculum. A series of missed-case conferences were developed based on trends in discrepancies, and the impact of these conferences is currently being evaluated. Orion is a powerful information technology tool that can be used by residency program directors, fellowship programs directors, residents, and fellows to improve radiology education and training.

Keywords Software design · Quality improvement · Residency · Medical informatics applications · Natural language processing · Performance measurement · Discrepancies · Performance · Web-based · Residents

Background

The Institute of Medicine’s (IOM) report “To Err Is Human” brought the quality and safety of patient care delivered in this country to the forefront of recent health care reform efforts by the government and medical community. More recently, the IOM cited a lack of adequate resident supervision and excessive fatigue as significant contributors to medical errors. The IOM’s recommendations to restrict resident responsibilities and reduce work hours place a clear emphasis on immediate patient safety over education and training. With any proposed changes in physician training, it is critically important to recognize that independent call is one of the most valuable and significant learning experiences of residency, providing the opportunity for residents to make independent decisions and assume graded levels of responsibility that are an integral part of training.

Academic medical centers are tasked with the responsibility of maintaining the highest quality of patient care while training the next generation of physicians. Providing residents and fellows opportunities to make independent decisions is an essential part of medical training. However, this aspect of training should not negatively impact patient care. To ensure that both of these activities are maintained without compromising patient safety, it is important to develop systems to evaluate resident and fellow performance on-call.

J. N. Itri (✉) · W. Kim · M. H. Scanlon
Hospital of the University of Pennsylvania,
3400 Spruce Street,
Philadelphia, PA 19104, USA
e-mail: Jason.iritri@uphs.upenn.edu

Radiology residency and fellowship training provides a unique opportunity to evaluate trainee performance. Every radiological study interpreted by a radiology resident or fellow must be reviewed by a faculty radiologist. In addition, there are usually electronic records of both trainee and faculty interpretations, allowing direct comparison and archiving of study information for subsequent analysis. The process of radiology residents and fellows providing preliminary interpretations in electronic format with subsequent faculty review provides a wealth of information that can be used to evaluate performance.

The Orion software application described in this manuscript was designed specifically to take advantage of these unique features of radiology resident and fellow training, allowing the development of a system to evaluate and track resident and fellow performance on-call.

Methods

Institutional review board exempt status was obtained for this study.

Dictating Preliminary Interpretations On-Call

From January 1 to June 15, 2010, radiology residents and fellows dictated preliminary reports on-call into the Radiology Information System (RIS; GE Centricity® RIS-IC, GE Medical Systems, Waukesha, WI) using two different speech recognition (SR) systems, Centricity Precision Reporting (GE Medical Systems) and RadWhere (Nuance Healthcare Solutions, Burlington, MA). After a resident or fellow signs a

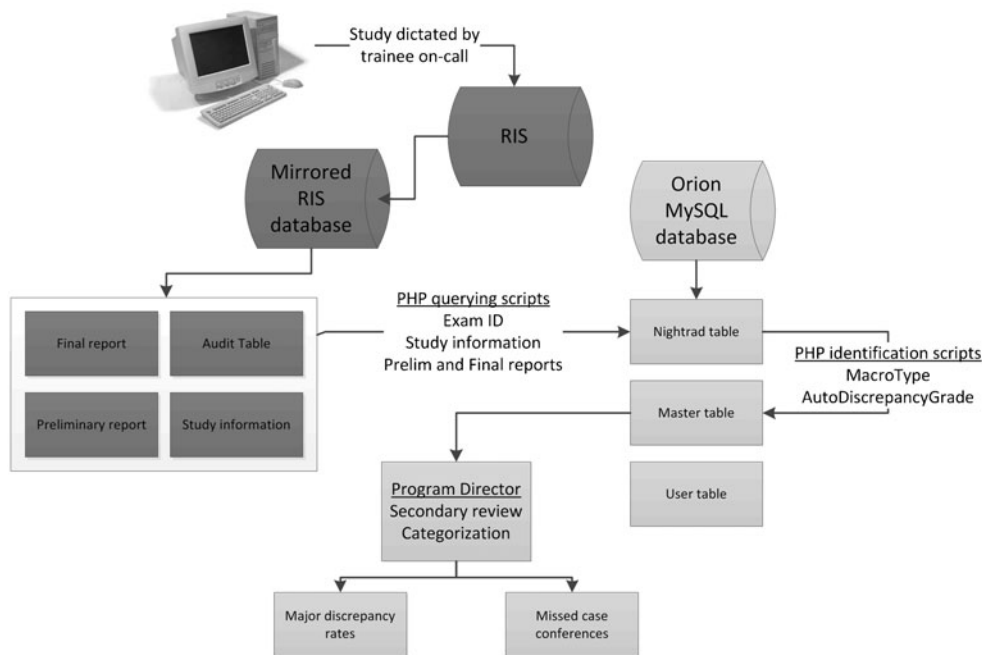
report into preliminary status, the report is saved into the RIS and sent to the electronic medical record for referring clinicians to review. Preliminary reports can be changed or updated depending on additional clinical information or images. All versions of draft and preliminary reports stored in the RIS can be accessed using Microsoft® (Redmond, VA) Structured Query Language (MS SQL) queries. At our hospital system, there are approximately eight different types of call shifts ranging from 4 to 12 h in length during which residents and fellows can cover up to three different hospitals: the Hospital of the University of Pennsylvania, Penn-Presbyterian Medical Center, and the Philadelphia VA Medical Center.

Radiology residents and fellows interpreting radiological studies on-call must assign a provider prior to signing the report into preliminary status, which allows the referring physician or service to view the preliminary report in the electronic medical record system. On-call studies are assigned to a generic provider called “Night Radiologist”, which has a corresponding unique provider identification (ID) number defined in the user table within the RIS. We use the provider ID number for “Night Radiologist” to identify all reports interpreted by residents and fellows on-call.

Development of the Core Orion Software Application

The Orion software application was developed using the open-source tools PHP: Hypertext Preprocessor (PHP) and MySQL, a relational database management system (Fig. 1). An initial PHP querying script accesses a mirrored RIS database every 90 min and identifies all reports signed out to “Night Radiologist” using the corresponding provider ID

Fig. 1 Diagrammatic representation of the relational databases, tables, and queries used to monitor resident and fellow discrepancies



number in the report audit table. When a report signed out to “Night Radiologist” is identified, the PHP script uses SQL to query the unique study ID number, accession number, patient information, exam type and code, modality, subspecialty, patient location, date and time that the study is completed in the RIS, date and time of the most recent preliminary interpretation, and all other relevant study information from four tables in the RIS database. The study information is linked across the four tables by a unique exam ID.

All study information queried from the mirrored RIS database is stored in a MySQL table (called “nightrad”) on a separate dedicated server. Additional PHP querying scripts use SQL to query the most recent preliminary and final reports from separate tables in the RIS, which are then stored in the “nightrad” table. A final PHP querying script uses SQL to search the audit table in the RIS and identify all studies in which our automatic electronic notification system to the emergency department was used (Emtrac notification system) and subsequently updates the “nightrad” table to indicate that the Emtrac notification system was used for a particular study. The Emtrac notification system is used for both minor and major discrepancies in reports for studies performed on patients in the emergency department who were discharged prior to the final report being issued. This notification system ensures 100% follow-up of all modified reports, including those with new recommendations.

Once the preliminary and final reports are in the “nightrad” table with the patient and interpreting physician information, two types of PHP identification scripts are used to identify discrepancies. The first PHP identification script (called “MacroType”) identifies unique text strings that are part of four standard report macros that faculty are required to use when reviewing on-call reports. We have simplified the RadPeer scoring system and use three general grades: agreement, minor discrepancy, and major discrepancy. The difference between minor and major discrepancy is that a major discrepancy has the potential to impact patient management or outcome. Each macro has slightly different introductory statement, which allows for detection and categorization using the PHP scripts. The macro introductory statements corresponding to agreement, minor discrepancy, and major discrepancy are listed below:

- “ATTENDING RADIOLOGIST AGREEMENT” indicates agreement without modification
- “ATTENDING RADIOLOGIST ADDITION” indicates agreement with additional text
- “*ATTENDING RADIOLOGIST CHANGE” indicates a minor discrepancy that has no clinical impact
- “**ATTENDING RADIOLOGIST CHANGE” indicates a major discrepancy that may have clinical impact

When the script identifies one of these text strings in the final report, it updates the “MacroType” field in the

MySQL table “master” to indicate agreement, minor discrepancy, or major discrepancy.

Automatic Detection of Discrepancies

The second PHP identification script (“AutoDiscrepancy”) evaluates entries in which a standard review macro text string is not identified in the final report, indicating that the faculty member reviewing the report did not use one of the aforementioned report review macros. The “AutoDiscrepancy” script first performs a simple comparison between the preliminary and final reports to determine if there has been a change. If there is no change in text between the preliminary and final reports, the script updates the “AutoDiscrepancyGrade” field of the “master” MySQL table to indicate that the report is unchanged. If there is a change between the preliminary and final report, the “AutoDiscrepancyGrade” script calculates a score based on the percentage of text change and presence of specific words and phrases that indicate discrepancy. Based on a pre-determined threshold, reports are then graded as “agreement” or “discrepancy” in the “AutoDiscrepancyGrade” field of the “master” table. The details of the formula used to determine the likelihood of discrepancy and its development are further discussed in the [Results](#) section.

User Login and Interface

User information is stored in a user table (“rftable”) in the MySQL database on a secure server accessible only via the hospital intranet or virtual private network. Each user is assigned a username and password based on their unique RIS identification number to login to the application, an access setting to determine which modules the user will be able to access, a user designation (resident, fellow, fellowship program director, residency program director, or administrator), specialty (only for fellows), and year of graduation (only for residents). The residency program director has administrator access and can use all Orion modules (Fig. 2a). Fellowship program directors have mid-level access, which allows them to review specialty-specific on-call studies and performance for fellows within their specialty. Residents and fellows have individual user access and can only review their own reports (Fig. 2b).

The report review screen (Fig. 3) has a navigation bar at the top with all relevant study information immediately below. Grading options and a comment field are located below the study information section. The residency program director and fellowship program directors can grade reports as agreement, minor discrepancy, or major discrepancy based on the preliminary and final reports, which are

Fig. 2 Administrator (a) and resident/fellow screens (b) demonstrating the various functions of the Orion application.

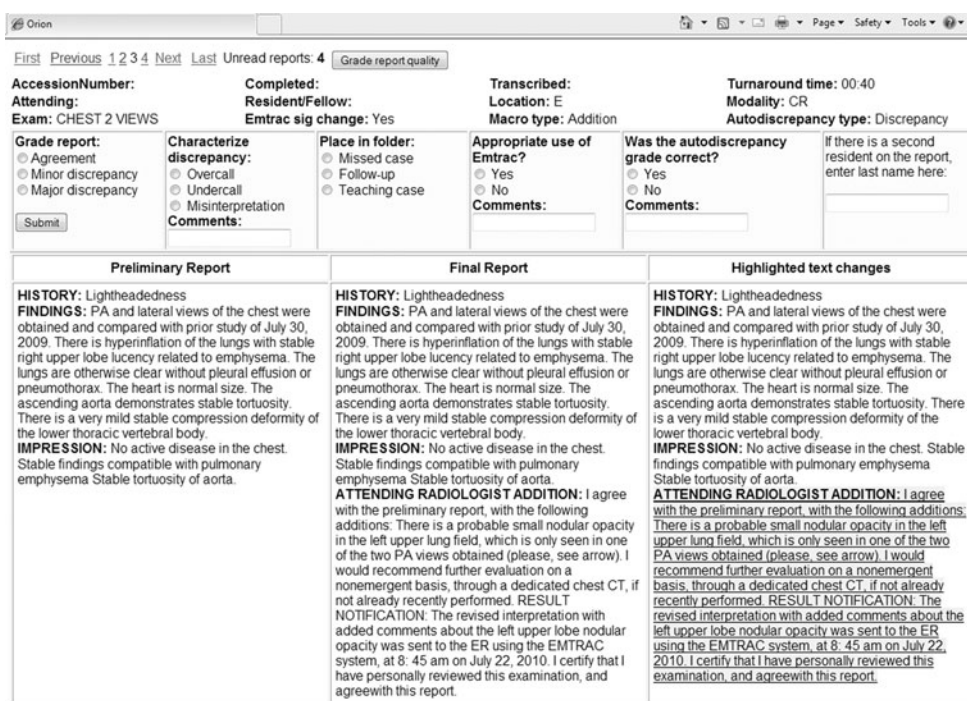
Administrators can review on-call reports by discrepancy type, resident/fellow or accession number, generate summary reports, review report quality, evaluate resident and fellow performance on-call, and monitor compliance. Residents and fellows can review their own reports by discrepancy type, review daytime reports, review on-call performance, and generate summary reports



displayed in tandem as part of a three-column table. The third column is the text differential of the preliminary and final reports with highlighting and strikethrough to more clearly display the changes between the preliminary and final reports. In addition to grading, the residency program director can provide a comment, save the case to a folder (teaching, missed, follow-up), classify the discrepancy (overcall, undercall, misinterpretation), and indicate if the Emtrac notification system and/or required faculty review macros were used appropriately.

For the residency program director, there is an additional “Grade report quality” button adjacent to the navigation car that links the active study to the report quality module described below under “Evaluating report quality”. Individual residents and fellows can only mark studies as reviewed, provide comments, and save cases in folders. Discrepancy grades, comments, and folder designations for each study are recorded separately for the residency program director, fellowship program directors, and individual residents and fellows in the “master” table. When a report is graded or

Fig. 3 Report review screen demonstrating the navigation bar with option to grade report quality (*top*), study information, grading/folder options, and comment fields. The preliminary report, final report, and highlighted text differential are displayed in tandem for easy comparison. Note the standard review macro used to identify this report as “agreement”



marked reviewed, it is immediately dropped from the review list, and the next study is loaded and displayed.

Secondary Review of Discrepancies

All studies in which faculty members reviewing an on-call preliminary report used the minor change macro, major change macro, or Emtrac notification system are reviewed by the residency program director (M.H.S.). In addition, all studies identified as “discrepancy” by the “AutoDiscrepancy” script are also reviewed by the residency program director. This ensures that nearly 100% of all resident and fellow discrepancies are reviewed and categorized. The final determination of the type of discrepancy is made by the residency program director. The discrepancy grade is stored in the “master” table and linked to the “nightrad” table using the unique exam ID for the study.

Monitoring Faculty Compliance with Use of Required Macros

The purpose of monitoring compliance is to ensure that faculty members use one of the required macros when they modify a preliminary report. Faculty compliance with using the required macros is determined by using the “MacroType” and “AutoDiscrepancy” script results. Effective compliance is calculate by dividing the number of modified reports without a required macro in the final report text by the total number of modified reports and reported real-time for the last 30, 60, and 90 days. The 30-day report is converted to portable document

format (pdf) format and emailed with specific comments to faculty members with less than 90% compliance.

Monitoring Resident Compliance with Reviewing Studies

Radiology residents at our institution do not review on-call cases with faculty at the end of the call shift. There is a resident version of Orion that allows residents to review all minor and major discrepancies and mark them as reviewed. The resident can provide comments about the case, save the case in a folder (missed, teaching, follow-up), and generate a summary report of all cases by type of discrepancy or folder. The residency program director monitors resident compliance with reviewing all minor and major discrepancies through Orion and reviews both resident performance and compliance at the semi-annual review. The resident discrepancy grade is stored in the “master” table and linked to the “nightrad” table using the unique exam ID for the study.

Analyzing Resident and Fellow Performance

The analytics module allows both the residency program director and fellowship program directors to evaluate minor and major discrepancy rates for all residents and fellows in real-time with multiple filter options. Total volume of on-call studies with overall minor and major discrepancy rates can be generated for all residents and fellows for a specified time range and is used for the quarterly clinical quality report for the Department of Radiology. Individual resident

minor and major discrepancy rates are used to establish overall performance benchmarks and identify residents who demonstrate a significantly higher major discrepancy rate compared with classmates. Individual resident and fellow minor and major discrepancy rates are generated by modality and subspecialty to target specific areas for improvement. Minor and major discrepancy rates by modality and subspecialty are filtered by residency class to establish performance benchmarks by year of training.

A separate analytics module allows evaluation of major discrepancy rate by time of day and shift length. A schedule table (“nftable”) containing the resident nightfloat call schedule from January 1 to June 15, 2010, was added to the MySQL database along with a separate table (“calltype”) to define the type of nightfloat call (1-week versus 2-week assignments) and shift length (12 h). Modality and subspecialty filters can be employed, and date ranges can be specified.

Generating Summary Reports

Comprehensive summary reports can be generated by the residency program director for a specified time period by discrepancy type, specialty, or resident/fellow. Individual resident summary reports are generated every 6 months, reviewed with residents during the semi-annual review, and placed in the resident’s learning portfolio. Summary reports by specialty are generated every 6 months and forwarded to each subspecialty section chief and fellowship director in an effort to tailor the didactic and case-based curriculum for that specialty to the specific needs of the residents and fellows.

A complete summary report of all major discrepancies is reviewed by the resident missed-case committee every 6 months to identify trends in major discrepancies. Regular resident missed-case conferences are designed to target specific types of misses identified as major discrepancies over the previous 6–12 months. In addition, every PGY-3 resident is required to give a missed-case conference to all residents after taking independent call based on their own discrepancies.

System Maintenance

The system maintenance script reports if there are any entries in the MySQL database still signed out to nightrad (not reviewed by a faculty member), as well as entries without a preliminary report, “MacroType”, or “AutoDiscrepancyGrade”. It also reports when the query scripts were last performed and if any errors were encountered. This maintenance report ensures that the query scripts run every day without error and that no issues are encountered with the PHP scripts identifying review macros. It can also identify reports that have not been signed by a faculty member within 24 h, but there are other processes in place at our institution to identify these reports.

Results

Accuracy and Reliability of the MS SQL Queries

From January 1 2010 to June 15 2010, 33,153 reports signed out to nightrad were identified and entered into the Orion MySQL database. Residents were responsible for interpreting 19,200, whereas fellows interpreted 13,953 reports. This amounts to an average of 200 radiological studies per day being interpreted on-call by residents and fellows at our institution. Nearly 100% of on-call reports are identified by the MS SQL queries and entered into the MySQL database, which was confirmed by manual review. Less than 0.1% of on-call reports are missed by the core MS SQL queries.

Faculty Compliance with Required Review Macros

The four standard review macros described in the [Methods](#) section were implemented in December 2009, with Orion being formally implemented on January 1, 2010. Initial faculty compliance using the review macros was approximately 30% on February 1, 2010, 60 days after implementing the macros. A summary report of effective compliance was sent via email to all faculty members with an effective compliance less than 90%, along with a more detailed summary of the modified reports in which a standard review macro was not used (if requested). The Orion system and standard review macros were presented during a departmental section chief meeting in an effort to improve faculty compliance. From May 15 to June 15, 2010, effective compliance was 93% with 2,270 out of 2,434 modified reports containing one of the standard review macros.

Development and Accuracy of the AutoDiscrepancy PHP Script

Prior to implementation of Orion, a separate software application (Minerva) was used to review resident and fellow on-call reports from March 2008 to December 2009. During this time period, 96,036 on-call reports were added to the Minerva database, of which 23,228 reports were reviewed and graded by the residency program director as agreement, minor discrepancy, or major discrepancy. The four standard review macros had not yet been instituted while Minerva was used by the residency program director to review resident and fellow on-call performance.

Initially, 1,260 reports graded as agreement (201), minor discrepancy (684), and major discrepancy (368) were used to develop the weighted formula employed in the “AutoDiscrepancy” script. A development script was used to compare the preliminary and final reports and

calculate a percent text change. The text added by faculty reviewing preliminary reports was then separated from the final report and analyzed independently in an excel spreadsheet. Eighteen unique words or phrases indicating discrepancy (e.g., “disagree” or “ignore”) or communication of results (e.g., “communicated to”) were identified. The frequency with which these words and phrases appeared in final reports with minor and major discrepancies was used to generate a weighted score for each word or phrase, with weighted scores ranging from 0.6 (“reviewed with”) to 32.4 (“disagree”).

The percent text change for graded reports was independently scored. A percent text change of less than 5% indicated agreement with only two exceptions (0.2% of all discrepancies). More than 30% text change between the preliminary and final reports indicated a high probability of discrepancy in this limited subset of reports. The percent text change was multiplied by a conversion factor with 30% text change resulting in a score of 3.

The cumulative scores for percent text change and weighted words or phrases in the final report were summed with a threshold of 3 to indicate discrepancy. When this formula was incorporated into the “AutoDiscrepancy” script and used on the original 1,260 reports, it correctly identified 97 out the 201 agreements as “agreement” (48.3%), 534 out of 684 minor discrepancies as “discrepancy” (78.1%), and 333 out of 368 major discrepancies as “discrepancy” (90.5%).

The “AutoDiscrepancy” script was then tested on the 9,721 graded reports from June 2009 to December 2009 (Table 1). For 2,428 studies performed in the Emergency Department, the script correctly identified 186 out of 228 total discrepancies (81.6%), 64 out of 78 major discrepancies (82.1%), and 1,764 out of 2,200 agreements (80.2%). The script incorrectly identified 436 out of

2,200 agreements as “discrepancy” (19.8%) and 42 out of 228 discrepancies as “agreement” (18.4%). For 7,293 studies performed on inpatients, the script correctly identified 347 out of 467 total discrepancies (74.3%), 106 out of 124 major discrepancies (85.5%), and 5,917 out of 6,826 agreements (86.7%). The script incorrectly identified 909 out of 6,826 agreements as “discrepancy” (13.3%) and 120 out of 467 discrepancies as “agreement” (25.7%).

From January 1 to June 15, 2010, there were 15,362 final reports without a review macro. The “AutoDiscrepancy” script identified 12,670 final reports as having less than 5% text change between the preliminary and final reports. The script then used the weighted formula to identify 1,145 final reports as “discrepancy”. Secondary review of these final reports demonstrated that 310 out of the 1,145 reports identified as “discrepancy” were actual minor or major discrepancies (27.1%).

Baseline Resident and Fellow Performance Data

From January 1 2010 to June 15 2010, 19,200 on-call studies were interpreted by 20 residents and 13,953 on-call studies were interpreted by 25 fellows representing eight subspecialties. The overall major discrepancy rate for residents was 0.85% with a standard deviation (SD) of 0.42%. There were two residents with a major discrepancy rate more than 1SD above the class average and one resident with a major discrepancy rate more than 2SD above the class average. The overall major discrepancy rate for fellows was 1.47% with a SD of 0.93%. There were five fellows with a major discrepancy rate more than 1SD above the fellow average and no fellows with a major discrepancy rate more than 2SD above the fellow average (Fig. 4).

Table 1 Performance of the “AutoDiscrepancy” script in discriminating “agreement” reports from “discrepancy” reports

	Emergency department	Inpatient
Total number of studies	2,428	7,293
Overall discrepancies correctly identified by the Autodiscrepancy code	186/228 (81.6%)	347/467 (74.3%)
Major discrepancies correctly identified by the Autodiscrepancy code	64/78 (82.1%)	106/124 (85.5%)
Minor discrepancies correctly identified by the Autodiscrepancy code	122/150 (81.3%)	241/343 (70.3%)
Agreements correctly identified by the Autodiscrepancy code	1,764/2,200 (80.2%)	5,917/6,826 (86.7%)
Agreements INCORRECTLY identified as discrepancies by the Autodiscrepancy code	436/2,200 (19.8%)	909/6,826 (13.3%)
Discrepancies INCORRECTLY identified as agreement by the Autodiscrepancy code	42/228 (18.4%)	120/467 (25.7%)
Overall true-negative rate	80.20%	86.70%
Overall true-positive rate	81.60%	74.30%
Major discrepancy true-positive rate	82.10%	85.50%
Minor discrepancy true-positive rate	81.30%	70.30%
Overall false-negative rate	18.40%	25.70%

Resident	Total volume	Macro Minor %	Reviewed Major %
	544	7.2%	0.9%
1		0%	0%
1509		5.7%	1.1%
76		5.3%	0%
386		7.8%	1%
4		0%	0%
3		0%	0%
504		13.7%	1.8%
1032		8.4%	0.8%
694		8.1%	0.4%
575		9.6%	0.2%
1		0%	0%
777		6.2%	1%
470		9.6%	0.6%
3337		9.1%	1.5%
1360		15.3%	1.3%
1792		5.2%	0.7%
44		0%	0%

Fig. 4 Resident performance page listing each resident with total volume of on-call studies, percentage of total studies in which the interpreting radiologist used the minor discrepancy macro, and percentage of total studies in which the interpreting radiologist and secondary reviewer (residency program director) designated the study as a major discrepancy. There are filters for PGY year of training and modality/subspecialty for more detailed analysis. A similar performance page is available for fellows

Discussion

Development and Implementation

The design of Orion was based on an earlier software program (Minerva) developed by programmers from the information technology (IT) department in conjunction with the residency program director (M.H.S.) at the University of Pennsylvania Hospital [1]. The original motivation behind Minerva was occasional telephone calls and/or email messages from faculty to the residency program director questioning the ability of specific residents to perform adequately on-call based on limited experience reviewing that resident's overnight reports. In response to the need for a formal process to evaluate resident and fellow performance, Minerva was conceived. Briefly, the original software application used similar SQL queries to identify all reports signed out to nightrad and entered relevant study information into a Microsoft Access database. Visual Basic for Windows was used for the graphical user interface that allowed M.H.S. to review and grade resident and fellow reports. Given that an average of 200 reports are interpreted by residents and fellows on-call each day, it was impossible to perform 100% review of trainee performance using Minerva. The process of having a single person grade all preliminary reports also introduced bias in the grading of discrepancies.

Ensuring 100% review of all resident and fellow performance on-call was a priority for this quality improve-

ment project, leading to creation of four standard review macros. These review macros served three purposes. First, they significantly reduced the workload required to ensure 100% review of on-call trainee performance by providing an initial grade of agreement, minor discrepancy, and major discrepancy at the time of final interpretation. Second, the faculty members issuing final reports have reviewed both the preliminary report and images associated with the study and are best able to judge whether a discrepancy had the potential to impact patient management or outcome. Finally, having the faculty responsible for initially grading preliminary reports as minor and major discrepancies significantly reduced the bias of having a single person review and grade all preliminary reports. Using the four standard macros when reviewing all on-call reports also provides a simple and straightforward mechanism to monitor faculty compliance.

In order to have accurate measures of major discrepancy rates for residents and fellows using Orion, we determined that faculty had to use one of the four standard review macros for at least 90% of modified reports based on the discrepancy results obtained using Minerva for the previous 20 months. As described in the **Results** section, it took approximately 6 months of monthly email reminders, summary reports, and a presentation at a departmental section chief meeting facilitated by the department chairman to obtain more than 90% compliance. Departmental leadership was critical for obtaining more than 90% compliance with use of the standard review macros. The ultimate goal is 100% compliance with standard review macros.

Automatic Detection of Discrepancies

The "AutoDiscrepancy" script used to automatically identify discrepancies in on-call reports was developed in parallel with the Orion application for several purposes. First and foremost, it was designed to detect discrepancies in reports where the reviewing faculty member had not used one of the standard review macros. As stated previously, it was determined that at least 90% of resident and fellow preliminary interpretations had to be reviewed to generate accurate major discrepancy rates. The "AutoDiscrepancy" script was a backup system designed to sort through a large number of reports and use a simple formula to determine the likelihood of a discrepancy. The script was successful in the sense that it identified 82–86% of major discrepancies for secondary review and grading by the residency program director. It failed in the sense that only 27% of reports identified by the script as "discrepancy" represented actual discrepancies. The second purpose of the AutoDiscrepancy script was to facilitate retrospective review of resident and fellow performance before instituting the standard review macros.

Resident Review of On-Call Discrepancies

Since residents and fellows interpreting studies on-call at our institution do not review these studies with faculty at the end of their call shift, it was a priority to document that every resident reviewed all of their minor and major discrepancies from call. A major limitation of the Orion application is that preliminary reports are not available to be reviewed by residents and fellows until 48 h after the preliminary interpretations have been issued. The 48 h delay is because faculty may not provide a final interpretation for up to 24 h (average is less than 10 h) and the duplicate RIS database is updated once every 24 h. Preliminary reports are saved in both the unique preliminary report table and in the main report table as the final report until a faculty member issues a new final report. When a delay of only 24 h was built into the queries, the preliminary report was entered into the MySQL database as both the preliminary and final reports and the “MacroType” and “AutoDiscrepancy” scripts assumed their agreement because no text difference was found between reports. This problem was solved by building in a 48-h delay, but resulted in residents and fellows not being able to obtain immediate feedback regarding cases from the prior evening on-call.

Orion as an Educational Tool

Residents are required to review all minor and major discrepancies on-call using Orion and compliance is monitored. The residency program director generates a list of all discrepancies for each resident, reviews this list with the resident during the semi-annual review, and places the list in the resident’s learning portfolio. Each resident is required to upload at least three missed cases to the department teaching file (MyPACS.net) every 6 months. In addition, every PGY-3 resident must give a missed-case conference to the entire residency class based on their misses after completing nightfloat call.

We initially used both minor and major discrepancy rates as performance metrics and in missed-case conferences. It became clear early in the process that minor discrepancies are not useful as performance metrics. One reason is because there is significant ambiguity inherent in trying to differentiate an agreement with addition, in which the faculty member agrees with the content of the report but has something to add, and a minor change, in which there is a minor discrepancy that does not have the potential to impact patient management. There was also considerable variability between faculty members in what was considered a minor discrepancy. Finally, none of the minor discrepancy cases were felt to be of sufficient educational value to be presented at a missed-case conference. Based on

these experiences, we record minor discrepancy rates but do not use them as performance metrics. However, we still require residents to review their minor discrepancies because they can provide subtle teaching points.

Missed Case Conferences

Resident and fellow discrepancies in on-call radiological studies are reviewed regularly by the residency program director and resident missed-case conference committee. In an early pilot project, we demonstrated that missed-case conferences could be used to significantly reduce the number of resident major discrepancies related to the types of cases reviewed [1]. In this study, trend analysis of missed cases was used to generate missed-case conferences reviewing acromioclavicular (AC) joint separation injuries, osteochondral fractures of the ankle and knee, and elbow joint effusions. These three topics were selected because all three injuries represented approximately 13% of the overall major discrepancies during the preceding 8-month time period (20 total), which was more than any other type of discrepancy. During the 8 months following the two missed-case conferences, there were only seven missed cases related to the injuries reviewed during the conferences. This represents a 64% decrease in the overall number of missed cases, with a 75% decrease in missed AC separation injuries, 50% decrease in missed elbow joint effusions, and 67% decrease in missed osteochondral fractures of the ankle and knee.

Since the initial pilot project, we have expanded the missed-case conference series to include nine dedicated conferences reviewing common and important misses on conventional radiography of the musculoskeletal system; conventional radiography of the chest, abdomen, and pelvis; CT and US of chest, abdomen and pelvis; obstetrical US; and CT of the neurological system. These conferences are given to the residents in June prior to taking independent call.

Discrepancy Rates as Performance Metrics

There are several retrospective and prospective studies evaluating discrepancy rates between radiology residents, fellows, and attending radiologists, with most studies reporting a major discrepancy rate between 0.5% and 2% [2–10]. Most of these studies focus on a particular modality such as cross-section imaging of the abdomen and pelvis or neurological CT. In comparison, our overall resident and fellow major discrepancy rates (0.85% and 1.47%, respectively) are similar to those previously reported. Using Orion, however, we are able to evaluate major discrepancy rates with significantly more granularity because all exam and interpreting physician information is integrated into the

database. With the use of filters, we can instantly evaluate the major discrepancy rates by modality, subspecialty, year of training, type of call shift, length of call shift, and various other parameters.

As an example, we were able to identify two residents with major discrepancy rates more than 1SD above the class average and one resident with a major discrepancy rate more than 2SD above the class average. All three residents have at least 1 year of residency training remaining and will receive additional training and one-on-one mentoring during core rotations including chest, neuroradiology, abdominal imaging, and ultrasound as part of a formal Clinical Competency Enhancement Program (CCEP). The residency program director will continue to monitor on-call performance and meet with these residents regularly to review major discrepancies and develop strategies for improvement.

Using Orion as Part of a Comprehensive Quality Assurance Process

There are a considerable number of benefits to using a software application such as Orion to monitor resident and fellow performance on-call. When implementing standard review macros for faculty to use during initial review, the work of identifying and tracking major discrepancies in preliminary interpretations is minimized, and 100% review of all trainee performance can be accomplished. The incredibly time-consuming process of manually recording these discrepant cases into a separate database is eliminated. Since Orion pulls all relevant study information directly from the RIS, major discrepancy rates can be generated by trainee type (resident or fellow), year of training, modality, subspecialty, type of call shift, and length of call shift. Orion is currently being used at our institution for the following activities:

- Individual major discrepancy rates are used to identify residents and fellows with major discrepancy rates significantly above that of their classmates
- Lists of major discrepancies by modality and subspecialty are generated for individual sections to help guide didactic and case-based curricula
- Major discrepancy rates and turnaround times are generated by hospital and patient type (ED or inpatient) for departmental quality improvement efforts
- Trend analysis of major discrepancies is used to develop missed-case conferences, which have been shown to significantly reduce the number of trainee major discrepancies [1]
- Major discrepancy rates are calculated by length of call shift and volume of studies to evaluate the impact of fatigue and workload on trainee performance on-call

- Trainee compliance with reviewing minor and major discrepancies in preliminary interpretations is monitored, and all major discrepancies are reviewed during the semi-annual review with the residency program director (MHS)

Conclusions

Orion is a simple software application developed using open-source tools to facilitate the identification and monitoring of resident and fellow discrepancies in on-call preliminary reports. This software application is a powerful IT tool that can be used by residency program directors, fellowship programs directors, residents, and fellows to improve educational and training efforts. When incorporated into residency and fellowship training programs, the data gathered can be used to establish benchmarks for individual resident and fellow performance by year of training, modality, and subspecialty, identify residents and fellows who demonstrate major discrepancy rates significantly greater than their classmates, identify subspecialty-specific major discrepancies to tailor the didactic and case-based curriculum, and identify trends in discrepancies that can be targeted with specific missed-case conferences. Major discrepancy rates for radiology residents and fellows can also be part of the Radiology departmental quality report to hospital administration.

Acknowledgments Regina O. Redfern is acknowledged for the development and implementation of the Minerva software application.

References

1. Itri JN, Redfern RO, Scanlon MH: Using a Web-based application to enhanced resident training and improve performance on-call. *Acad Rad* 17:917–920, 2010
2. Tieng N, Grimberg D, Li SF: Discrepancies in interpretation of ED body computed tomographic scans by radiology residents. *Am J Emerg Med* 25:45–48, 2007
3. Meyer RE, Nickerson JP, Burbank HN, Alsofrom GF, Linnell GJ, Filippi CG: Discrepancy rates of on-call radiology residents' interpretations of CT angiography studies of the neck and circle of Willis. *AJR Am J Roentgenol* 193:527–532, 2009
4. Wechsler RJ, Spettell CM, Kurtz AB, et al: Effects of training and experience in interpretation of emergency body CT scans. *Radiology* 199:717–720, 1996
5. Wysoki MG, Nassar CJ, Koenigsberg RA, Novelline RA, Faro SH, Faerber EN: Head trauma: CT scan interpretation by radiology residents versus staff radiologists. *Radiology* 208:125–128, 1998
6. Carney E, Kempf J, DeCarvalho V, Yudd A, Noshier J: Preliminary interpretations of after-hours CT and sonography by radiology residents versus final interpretations by body imaging radiologists at a level 1 trauma center. *AJR Am J Roentgenol* 181:367–373, 2003

7. Ruchman RB, Jaeger J, Wiggins 3rd, EF, et al: Preliminary radiology resident interpretations versus final attending radiologist interpretations and the impact on patient care in a community hospital. *AJR Am J Roentgenol* 189:523–526, 2007
8. Erly WK, Berger WG, Krupinski E, Seeger JF, Guisto JA: Radiology resident evaluation of head CT scan orders in the emergency department. *AJNR Am J Neuroradiol* 23:103–107, 2002
9. Cooper VF, Goodhartz LA, Nemcek Jr, AA, Ryu RK: Radiology resident interpretations of on-call imaging studies: the incidence of major discrepancies. *Acad Radiol* 15:1198–1204, 2008
10. Walls J, Hunter N, Brasher PM, Ho SG: The DePICTORS Study: discrepancies in preliminary interpretation of CT scans between on-call residents and staff. *Emerg Radiol* 16:303–308, 2009