Published in final edited form as:

AJR Am J Roentgenol. 2013 April ; 200(4): W345–W352. doi:10.2214/AJR.12.9521.

Increasing Rate of Detection of Wrong-Patient Radiographs: Use of Photographs Obtained at Time of Radiography

Srini Tridandapani^{1,2}, Senthil Ramamurthy¹, Samuel J. Galgano³, and James M. Provenzale^{1,4,5,6}

¹Department of Radiology and Imaging Sciences, Winship Cancer Institute, Emory University, 1701 Uppergate Dr NE, Ste 5018, Atlanta, GA 30322.

²School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA.

³Department of Internal Medicine, Baptist Health System, Birmingham, AL.

⁴Department of Oncology, Emory University School of Medicine, Winship Cancer Institute, Atlanta, GA.

⁵Department of Biomedical Engineering, Emory University School of Medicine, Atlanta, GA.

⁶Department of Radiology, Duke University Medical Center, Durham, NC.

Abstract

OBJECTIVE—The objective of our study was to evaluate whether facial and chest photographs obtained simultaneously with radiographs increase radiologists' detection rate of labeling errors.

MATERIALS AND METHODS—We obtained simultaneous portable radiographs and photographs of 34 patients. We generated 88 pairs of chest radiographs (one recent radiograph, one prior radiograph) and compiled a set of 20 pairs for reader review. Two, three, or four mismatched pairs (i.e., pairs containing radiographs of different patients) were introduced into each list. Ten radiologist readers blinded to the presence of mismatches interpreted the 20 radiograph pairs. Readers then reviewed a second set of 20 pairs containing mismatches but photographs of the patients obtained at the time of imaging were attached to the radiographs. Readers were not instructed regarding the purpose of the photographs. The mismatch detection rate and time for interpretation was recorded for both sessions. The two-tailed Fisher exact test was used to evaluate differences in mismatch detection rates between sessions, with a *p* value of less than 0.05 being considered significant.

RESULTS—The error detection rates without (3/24 = 12.5%) and with (16/25 = 64%) photographs significantly differed (*p* = 0.0003). The average interpretation times without and with photographs were 35.73 and 26.51 minutes, respectively (two-tailed Student *t* test, *p* = 0.1165).

CONCLUSION—The use of photographs increased the detection of errors without a concomitant increase in film interpretation time, which may translate into improvements in patient safety without an increase in interpretation time.

Keywords

DICOM; digital photography; medical errors; PACS; patient identification

Patient safety issues have gained prominence in the national dialogue particularly since publication of the Institute of Medicine's 2001 report on health care quality [1]. The first specific requirement in the Joint Commission's 2012 National Patient Safety Goals (NPSG) is that at least two patient identifiers be used when providing care, treatment, and services [2] with the rationale being the following:

[W]rong-patient errors occur in virtually all stages of diagnosis and treatment....
Acceptable identifiers may be the individual's name, an assigned identification number, telephone number, or other person-specific number.

The National Quality Forum [3] has also recognized that wrong-patient errors affect medical imaging and, with support from the Agency for Healthcare Research and Quality, has specifically endorsed the implementation of a "standardized protocol to prevent mislabeling of radiographs" in its "30 Safe Practices for Better Health Care Fact Sheet."

It is quite difficult to obtain data about labeling or identification errors. A report by Kuzmak and Dayhoff [4] showed an error rate of 0.73% in 93,000 imaging examinations in one study; in another study, Gale and Gale [5] reported an error rate of 0.26% in 48,800 studies. These error rates can lead to substantial problems in patient management. Even more troubling is that such studies may, in reality, underreport the true error rate. Many errors may not even be noticed and erroneous radiology reports may be issued. Many other errors are likely detected by technologists at the time of completing the study and are unreported because errors can be corrected at that stage before a clinical decision based on the imaging study has been made. Even if we assume a low error rate of 0.01%, at a large institution such as ours where more than 1 million imaging studies are performed annually, this rate could result in nearly 100 potentially preventable wrong-patient errors per year. In 2009, 652 "wrong events" in radiology that led to patient harm were reported to the Pennsylvania Patient Safety Authority and 196 (30.1%) of those events were wrongpatient errors [6]. Investigators at one institution reported that 10 of 62 "near-miss events" (16%) in radiology between 2007 and 2009 were because of incorrect patient identity information [7]; during a similar period, investigators at another large radiology department reported that 313 (19%) departmental errors were related to patient identification [8].

Further aggravating identification errors is that obtaining and assigning the acceptable identifiers noted in the Joint Commission's NPSG requirements can be problematic, particularly when patients are unconscious, uncooperative, or noncommunicative and cannot verify their identity. A means to more confidently identify patients based on facial appearance, a long-standing method for identification of individuals, could potentially decrease identification errors.

To minimize or prevent mislabeling of medical imaging studies, we introduced a scheme to obtain digital photographs of patients simultaneously with all medical imaging studies including, but not limited to, radiography, CT, ultrasound, MRI, scintillation imaging, and PET (Tridandapani S, et al., presented at the 2010 Institute of Electrical and Electronics Engineers [IEEE] Conference on Biomedical Engineering and Sciences [IECBES]; Ramamurthy S, et al., presented at the 2011 American Medical Association-IEEE Medical Technology Conference). The technical details of the system for obtaining simultaneous images with portable radiography have been presented elsewhere (Ramamurthy S, et al., presented at the 2012 annual meeting of the Society of Imaging Informatics in Medicine [SUM]). The system was designed to ensure that photographs are matched with radiographs and thus no new errors are introduced by using this system. Furthermore, no additional time is invested by the technologist in obtaining these photographs because the photographs are obtained automatically by the system. In fact, the technologist is unaware that photographs are also being obtained simultaneously.

These digital photographs are small additions to the imaging study similar to the scout or localizer images that are obtained with CT studies. We do not intend these digital photographs to entirely replace numeric identifiers or other identifiers based on new technologies such as radiofrequency identification or bar code technologies; rather, we envision photographs as a means to supplement and strengthen those identifiers. However, in some cases, such as unconscious trauma patients who cannot provide identification information, these photographs may indeed be the only available identifier. Aakre and Johnson [9] discussed direct patient verification and bar code techniques to reduce wrong-patient errors; again, our scheme can work in conjunction with these other schemes and can provide additive benefits in improving patient safety.

The purpose of this study was to evaluate in a simulated setting whether facial and chest photographs obtained simultaneously with portable chest and abdominal radiographs in an ICU environment increase radiologists' rate of detection of mislabeling.

Materials and Methods

This study was approved by Emory University's Institutional Review Board and written informed consent was obtained from patients recruited to the study or from one of their family members authorized to provide consent. The study was compliant with HIPAA.

Study Population

Data were gathered between August 5, 2011, and October 12, 2011, in two cardiothoracic surgery ICUs in Emory University Hospital. Most of the data were gathered between 2:00 and 6:00 am (i.e., when most portable radiographs are obtained in ICUs). Some of the data were obtained in step-down care units or other regular hospital floors if the patient was transferred out of the ICU during hospitalization. This data collection method is a convenience sampling method in which radiograph-photograph combinations were obtained only on the days that the researcher assigned to data acquisition was available. Thus, although the study extended for 68 days, data were obtained on only 44 days of the study period.

We initially recruited 34 patients to the study. However, we obtained only one radiograph-photograph combination for seven of these patients. Thus, radiograph pairs—with pairs defined as new and old radiographs—could not be generated for those patients. A radiograph from one of these patients was, however, used to create two erroneous pairs, as described later. One of the patients had more than the 10 radiograph-photograph combinations that were included in the study (Fig. 1).

The final study cohort consisted of 28 patients (13 men, 15 women; mean age $\pm$ SD, 61 ± 15.16 years; age range, 22–89 years). These patients were admitted for a variety of diagnoses, with the four most common being aortic stenosis ($n = 10$), congestive heart failure ($n = 7$), mitral valve regurgitation ($n = 3$), and coronary artery disease ($n = 3$). The four most common surgeries that the patients underwent during the current hospitalization or previously were aortic valve replacement ($n = 9$), left ventricular assist device placement ($n = 6$), mitral valve replacement ($n = 3$), and coronary artery bypass grafting ($n = 3$).

Data Acquisition and Storage

All portable radiographs were single-view radiographs of the chest or abdomen obtained in the standard fashion with the technologist confirming patient identity verbally or by checking wristband information. Immediately before obtaining the radiograph, a single photograph of the patient's face and chest was obtained by placing a camera adjacent to the properly positioned radiographic equipment. All photographs were obtained by a single

individual with a 5-megapixel camera on a cell phone (iPhone 4, Apple) with the use of a flash. Photographs were initially stored in JPEG format. Before photograph acquisition, a photograph of the paper requisition form was also obtained to ensure that no errors were made in matching the radiographic study with the patient photograph. On occasion, studies were added to the work list and the technologist was paged with the request when he or she was already on the floor; thus, some radiographs were obtained without the aid of the requisition form. If an acquisition number for a study was unavailable at the time of the examination, an acquisition number was acquired from the PACS using other patient identifiers including name, date of birth, medical record number, and date and time the study was performed. This process was followed for this simulated study only; the hardware that we have discussed elsewhere can obtain photographs and radiographs simultaneously in an automated fashion (Ramamurthy S, et al., 2012 SUM meeting). All photographs and patient information were stored on a research computer with password encryption to ensure protection of patient information.

DICOM Integration of Photographs

The JPEG photographs were converted into DICOM format. The photograph and corresponding radiograph were “stitched” together using custom-developed software. A composite image was generated with the color photograph located to the left of the gray-scale radiograph. The photograph was approximately one fourth of the size of the radiograph, as shown in Figure 2.

Removal of Patient Identifiers and Pairing of Images

Study sets without patient identifiers were generated by combining two sequential radiographs of the same patient—that is, a “current” radiograph and the most recent previous radiograph. These images were presented as a pair of images to each reader for interpretation. When more than two radiographs existed for the same patient, every two consecutive images were paired. We ensured that no radiograph appeared in two different pairs so that some degree of independence between pairs was maintained. Some patients underwent only one combined portable radiograph-photograph examination during hospitalization, whereas others underwent several (Fig. 1). Data from patients who underwent only one portable radiograph-photograph examination could be used only for the erroneous sets. One such radiograph-photograph combination was used, resulting in the 28 patients that contributed to the final dataset.

A total of 176 radiographs were obtained, from which we created 88 unique pairs of matched radiographs. In addition, we created 10 mismatched sets by combining current and comparison studies from different patients. From this set of 88 pairs, we randomly selected the radiographs to show to readers for interpretation. Each reader was shown 20 pairs of radiographs that included up to 20% mismatched pairs.

Each of the 10 mismatched pairs was created as follows: One subject from the 28 subjects in the cohort was selected randomly using a uniform distribution. One of that subject's radiographs was selected randomly using a uniform distribution. A second subject was also then selected randomly from the remaining 27 subjects using a uniform distribution, and one of that subject's radiographs was selected randomly. The temporally later radiograph from these two selected radiographs was used as the “current” study and the temporally earlier radiograph was used as the “comparison” study. These subjects and radiographs were replaced in the pool before selecting the next mismatched pair. The same 10 mismatched pairs were used with and without photographs in the two phases of the study. The mismatched pairs were not generated using age or sex matching.

Reader Group

Ten recently trained radiologists served as readers. Nine of these radiologists had been certified by the American Board of Radiology within 2 years before the study. One radiologist had been trained in the United Kingdom and was pursuing his second year of fellowship training in the United States. All readers were either pursuing fellowship training in subspecialties that did not include cardiothoracic radiology or were first-year faculty members in a division other than cardiothoracic radiology. We chose this population of readers because it more closely represents the skill level of general radiologists than the subspecialist radiologists at our institution.

Image Presentation

The study was conducted in two phases, the first consisting of observations without photographs and the second with photographs. In neither phase were the readers told that the intent of the study was to detect mismatches.

First film assessment session—In phase 1, 20 randomly selected pairs of radiographs were presented to each reader on a DICOM viewer (ClearCanvas Workstation 2.0 SP1), running on a dual-monitor workstation. Within these 20 pairs, between two and four mismatched pairs were included in random order. Aggregating across all readers, a total of 10 mismatched pairs were shown. A total of 58 pairs were shown without photographs and 70 pairs were shown with photographs. However, because mismatched figure pairs were randomly distributed, not all mismatched pairs were shown without and with photographs to the same reader. We recorded the detection rate of mismatched pairs without photographs and the same rate for those with photographs. In addition, for cases in which readers individually assessed the same erroneous pair of radiographs one time without photographs and another time with photographs, we calculated the improvement in reader performance for each pair. The total time for interpretation of this set of 20 radiographs was recorded.

As we noted earlier, readers were not informed about the presence of mismatched pairings or that the intent of the study was to detect mismatches between radiographs. Instead, readers were allowed to assume that paired radiographs were of the same or other similar patients. The DICOM viewer provided the reader with basic capabilities, such as windowing and inversion. Readers evaluated the images using the form shown in Appendix 1.

The fourth item in the form given to readers (Appendix 1)—that is, a space for “Other Comments”—was used as a means for readers to indicate that they had detected mismatched pairs. The time for interpretation of the entire set was spontaneously (i.e., without prompting) recorded for each reader.

Second film assessment session—In phase 2, the same readers were asked to assess an additional 20 pairs of radiographs using the same criteria as in the first reading session. However, for this phase, color photographs obtained simultaneously with the radiographs were also shown. Again, readers were not told that the intent of the study was to detect mismatched pairs or that such pairs existed. Readers were told that additional information (i.e., photographs) would be shown and that this information may or may not help with their assessment. However, they were not told that the photographs were specifically intended to enhance detection rate of mismatched radiographs. As in phase 1, the total time for interpretation of the set of 20 pairs was recorded.

Reader Questionnaire

After the two phases of image assessment, readers were asked to complete the questionnaire shown in Appendix 2.

Statistical Analysis

The two-tailed Fisher exact test was used to compare error detection rates without and with photographs; a p value of less than or equal to 0.05 was used to indicate a significant difference. The Student t test was performed to compare the average times taken by readers in phase 1 and phase 2. A p value of less than or equal to 0.05 was considered to indicate a significant difference. Statistical testing was performed using statistics software (QuickCalcs, GraphPad Software, La Jolla, CA).

Results

Qualitative Results: Samples of Mismatched Pairings

Figures 2 and 3 show two examples of the mismatched pairs used in the study. Figure 2A shows an 81-year-old man who underwent aortic valve replacement and was being imaged after coronary artery bypass grafting and aortic valve replacement; the characteristic median sternotomy wires are seen. Figure 2B, which is the comparison ("previous") radiograph from 3 days earlier, shows an 89-year-old white man with aortic stenosis admitted for aortic valve replacement surgery; the radiograph also shows a calcified aortic knob and calcified mediastinal lymph nodes not seen in the patient in Figure 2A. In addition, given a difference of only 3 days between the two radiographs, it is unlikely that the postoperative changes would show median sternotomy wires only and no support lines and tubes. The photographs, despite being edited to protect patient identity for this article, clearly show differences in facial hair and baldness between the two patients.

In Figure 3A, the radiograph (the "current" examination) shows a 65-year-old white woman with a history of bronchiolitis obliterans who underwent imaging after bilateral lung transplant. The radiograph shows a normal heart size, clamshell sternotomy wires associated with bilateral lung transplant, and no median sternotomy wires or aortic atherosclerotic calcification. Bibasilar atelectasis and a small left pleural effusion are shown. A new feeding tube and a new right peripherally inserted central catheter are also seen. Figure 3B, which serves as a comparison radiograph from 43 days earlier, shows a 75-year-old white man with aortic stenosis and a history of aortic valve replacement. The radiograph shows cardiomegaly, median sternotomy wires, aortic knob calcification, and left lung base atelectasis. The photographs clearly show sex differences despite being edited to protect patient identity for this article. The readers were shown the photographs without such editing.

Quantitative Results

Table 1 provides the assessment results for the readers. In phase 1 (i.e., without photographs), only one reader correctly reported the presence of mismatched radiographs. In phase 2 (i.e., with photographs), four of the same readers correctly reported all mismatched radiographs and an additional four readers reported some mismatched pairs. One reader later mentioned that he actively ignored the photographs because he thought the intent of the photographs was to distract readers from providing a proper radiologic assessment. Despite this belief, he did note the last mismatch as he was finishing the study. Two readers reported that they used the photographs to confirm the presence or absence of IV lines and tubes but did not compare the two photographs to see if they were of the same patient.

Overall there was a significant difference in the mismatch detection rates without and with photographs ($p = 0.0003$). In the absence of photographs, only three of 24 (12.5%) mismatched pairs were reported; with all three photographs, reader performance improved and 16 of 25 (64%) mismatched pairs were reported.

The mean time for interpretation without photographs was 35.73 minutes (SD, 9.4) and with photographs was 26.51 minutes (SD, 11.6), with a mean difference of 9.23 minutes (95% CI, -2.66 to 21.12); this difference was not statistically significant (unpaired Student *t* test, two-tailed $p = 0.1165$).

Improvement in Reader Performance

In all, 10 mismatched pairs were used in the 10 work lists with two, three, or four of these errors randomly introduced in each work list. In 11 instances, readers were shown the same pair without and with photographs. In nine of these 11 instances (81.82%), readers did not identify the mismatch without photographs but did identify the mismatch with photographs in another session. In the remaining two instances, the readers failed to identify the mismatch both without and with photographs.

The sensitivity for the detection of errors was 12.5% without photographs and 64% with photographs when these performance values are calculated without regard to the fact that some error pairs were shown to some readers twice—once with and once without photographs. When we excluded the repeated error pairs, sensitivity was 20% without photographs and 56.25% with photographs. Specificity in all cases was 100% because no reader made a false-positive identification.

Thus, when considering only those radiograph pairs that were shown to individual readers both without and with photographs, reader performance improved in nine of 11 instances (81.82%).

Poststudy Questionnaire

Two of the 10 readers (20%) indicated that the photographs were a distraction. Seven readers (70%) reported that they spent more time per case when photographs were present. In fact, objective times, as we noted earlier, did not reveal any significant difference between the two phases when the data were considered for each group. There was a trend toward decreased interpretation time after the introduction of photographs, although this difference was not statistically significant.

Eight readers (80%) thought that photographs helped with the interpretation. Of these eight, five thought the photographs assisted with interpretation of lines and tubes, four explicitly mentioned that the photographs helped with identifying mislabeled patients, and five noted that the photographs helped with evaluating patient status.

Nine readers spontaneously (i.e., without prompting) stated that if they noted mismatched photographs they went back to check the radiographs for mismatches. Some readers who saw a mismatch in the photographs assumed that the errors were software-related in matching the photographs with the radiographs (i.e., they thought the radiographs in the pair belonged to the same patient and the errors were only in the photographs).

Discussion

Our results show that the presence of patient photographs obtained simultaneously with ICU portable radiographs significantly increases the error detection rate when radiologists are presented with mismatched radiographs.

Introducing photographs into the workflow could potentially result in an increase in interpretation time. Some of the readers thought that they were spending more time because of the photographs. We tested this hypothesis by measuring interpretation times, and found that the time actually decreased between phase 1 and phase 2 (although not in a statistically

significant manner). Possibly the decrease in time may have been because of a learning effect—that is, readers became accustomed to the task during phase 1 and were thus quicker during phase 2. Nonetheless, some readers thought that the supplementary photographic information regarding lines and tubes sped up their evaluation which is an argument against a learning effect in all cases.

For this study, one author obtained the photographs at the time the radiographs were obtained. We reiterate that the system that we have designed does not lead to any increase in the technologists' effort and acquisition time (Ramamurthy S, et al., 2012 SIIM meeting). The system can automatically obtain the photographs and radiographs simultaneously and transmits that information to PACS without the technologist being aware of it. The system also ensures that the radiograph and photograph are tightly coupled by an innovative use of time-stamping and radiofrequency identification tags.

Relationship to Prior Studies

To our knowledge, only one published study [10] has addressed the problem of establishing patient identity from improperly labeled portable chest radiographs. In that study, Bhalla et al. [10] discussed various radiographic features such as the following for establishing patient identity:

...characteristic location and configuration of surgical material, fractures, and dense parenchymal/pleural scars with or without calcifications....

However, they also recognized that most patients lack characteristic surgical and pathologic features and suggested that some anatomic features such as the transverse processes of the first thoracic vertebra and the adjoining tubercles of the first ribs and the spinous processes may help radiologists to correctly identify patients. However, given the large volume of imaging studies that clinical radiologists encounter on a daily basis, use of such a method during radiograph interpretation would be onerous. The goal of our study was to show that the addition of patient photographs may provide a simple means to overcome this problem.

Our results indicate that even the presence of metallic hardware and characteristic IV line and tube positions did not enhance the rate of detection of mismatched pairs when only radiographs were provided. For instance, consider the case shown in Figure 3 in which a mismatched pair initially showed a patient who had undergone aortic valve replacement with the characteristic median sternotomy wires and subsequently showed a patient who had undergone bilateral lung transplant with the characteristic clamshell sternotomy wires. The rate of detection of this specific mismatched pair before and after the introduction of photographs was 0% (0/2) and 100% (2/2), respectively.

After radiograph assessment sessions, readers were shown the mismatched cases they did not detect. All readers said they retrospectively attributed the change between the radiographs to the incorrect assumption that surgical intervention had occurred in the interval between radiographs. We hypothesize that radiologists are trained to rationalize differences to account for differences between radiographs. In addition, we hypothesize that radiologists' training to identify mislabeled radiographs is relatively weak because these errors are relatively uncommon.

The exact mechanism by which the use of photographs increased the rate of detection of mismatched pairs is not entirely clear. One likely mechanism is that readers first noticed the mismatch between the photographs and then looked more closely at the radiographs to note whether they showed discordant findings. However, another possibility is suggested by one of the few other studies that have incorporated photographs into medical imaging examinations. In that study, the investigators reported that a poststudy questionnaire

answered by readers showed that radiologists' empathy for patients increased with the introduction of photographs [11]. Interestingly, that study also found that the number of incidental findings reported increased when photographs were used. This finding suggests that the presence of photographs may promote a greater degree of involvement by film readers, resulting in a higher detection rate not only of incidental findings but also of discrepant imaging studies.

Whether radiologists welcome the addition of photographs to medical images is a matter of debate. In one study, Weiss and Safdar [12] conducted a single-institution survey of 21 radiologists regarding supplementary use of digital photographs; 24% of respondents believed that a digital picture would help inform their radiologic decisions. However, 67% of those surveyed thought that facial pictures should not be included. These findings are at variance with those in our study, which was not simply a survey but a working scenario in which radiologists actually used photographs in a setting very similar to what could be seen in a future clinical experience if the use of photographs in radiology became prevalent. In our study, eight of 10 readers (80%) thought that the photographs were not a distraction and eight of 10 readers (80%) thought that the photographs helped with their interpretation. In contrast to the readers in the study by Weiss and Safdar [12], the readers in our study looked at photographs of actual patients along with the radiographs in a simulated study before they responded to a questionnaire. Experiencing first-hand the advantage of having access to photographs included in the study may have influenced their preference for such inclusion. In contrast to Turner and Hadas-Halpern [11] who performed a study in a real-life workflow setting, our study was a simulated one; however, their study was not designed to evaluate the error-detection capabilities of radiologists.

Limitations

As in all studies, a number of limitations of this simulated study can be noted.

This study was a simulated study in which we introduced a mismatch error rate of between 10% and 20% in each reader's work list—that is, we used an enriched sampling technique. This error rate is significantly higher than the error rates reported by Kuzmak and Dayhoff [4] and Gale and Gale [5], who reported 0.73% and 0.26%, respectively. We had anticipated that the difference in error detection rates without and with photographs would be small; thus, we used error rates between 10% and 20% to power the study adequately. In retrospect, we may have been able to show the same results even with a smaller injected error rate. Radiologist recruitment also was a significant consideration in choosing the error rate. Use of an error rate of 1% would have required readers assessing 100 radiographs in each phase. Our ability to recruit 10 readers would have been severely hampered and the study would not have been feasible. Although a 10–20% error rate does not reflect the real-world error rate, it may not be a true limitation in evaluating radiologists' perceptual capabilities. If radiologists' mismatch error detection rate is low when they are presented with errors up to 20% of the time, their error detection rate when the error rate is less than 1% is likely going to be even lower because their level of suspicion will be lower.

We used color photographs because we believe color photographs show distinguishing facial features better than gray-scale photographs. The error detection rate for mismatches might have been even lower if gray-scale photographs are used. Many radiology departments have only black-and-white monitors; if such departments were to implement this photographic technique in their workflow, they may not note the same benefits that were seen in this simulated study. A future study to compare reader performance when gray-scale photographs are used with reader performance when color photographs are used is warranted.

We recruited only fellows in training and junior faculty from divisions other than cardiothoracic radiology because we wanted to assess the perceptual capabilities of general radiologists. We note that these recently trained radiologists may not truly represent the average general radiologist nationally who may have several years of experience; however, we believe that they more closely reflect general radiologists than the subspecialized radiologists in our institution who either are experts in thoracic radiology solely or had not interpreted any chest radiographs in several years. Subspecialized cardiothoracic radiologists may perform better than general radiologists in detecting mismatch errors with portable chest radiographs and evaluating this population is a legitimate topic for future study.

It can be argued that if the readers had known about the possible presence of identification errors in phase 1, the detection rate would have been higher even without the photographs. However, we believe that withholding that information more accurately reflects actual film interpretation conditions in which wrong-patient errors are low-frequency events and interpreters' level of suspicion is not high. If the readers had been informed specifically about the potential for labeling errors, their approach to interpretation would have been biased and would not reflect reality. For the same reason, we avoided a crossover study design in which some of the readers could have been asked to interpret the cases with photographs in phase 1 and then asked to interpret the cases without photographs in phase 2. The difficulty with a crossover study design is that readers typically become biased once they deduce the objective of the study. On the other hand, if photographs were to be implemented in a real-world workflow as a tool to aid with patient identification at the time of radiograph interpretation, then it is reasonable to presume that the interpreting radiologists would be informed about the presence of and rationale for such a tool.

Some of the readers were shown the same error pairs in phase 1 and phase 2 to allow us to study reader improvement after the introduction of photographs. However, there is potential confounding in that reader sensitivity may have increased because they spent more time on a specific error pair as a result of seeing it twice rather than an effect of the photographs. Thus, we also computed sensitivity values after excluding duplicates in the work list.

Another limitation of this study is generalizability. Error detection rates may have been low because all the radiographs were of patients who were in the cardiothoracic ICU and had similar diseases and conditions. In smaller institutions where radiologists interpret pediatric, adult, inpatient, and outpatient radiographs all from one work list, misidentified patients may be easily detected—for example, if a neonate's chest radiograph is included in an adult patient's image folder. In large institutions where there is separation of the reading work lists and all ICU studies are found in one work list without other types of studies in the work list, error detection may be lower. On the other hand, because errors tend to occur in clusters (e.g., patients from a single floor or ICU may have their studies erroneously exchanged), it is less likely that pediatric radiographs would be erroneously placed in an adult patient's imaging folder.

Future Work

The scheme we have introduced may need modifications before implementation in the clinical environment, particularly for studies without prior examinations, for comparison. In these cases, patient photographs from the electronic medical record may need to be retrieved and presented along with the photographs obtained at the point of care of medical imaging. Technical challenges also need to be solved to ensure that the photograph and the medical images are tightly coupled and the photographs do not introduce other errors into the system; we have presented a scheme elsewhere to address this issue for portable radiography machines (Tridandapani S, et al., 2010 IEEE IECBES; Ramamurthy S, et al., 2011 AMA-IEEE Medical Technology Conference; Ramamurthy S, et al., 2012 SIIM meeting). These

issues need to be investigated in a clinical environment setting for other modalities in addition to resolving any legal and ethical issues that arise with storing and viewing patient photographs. We have addressed some privacy concerns previously (Tridandapani S, et al., 2010 IEEE IECBES).

Conclusion

The use of photographs obtained at the point of care—that is, simultaneously with the acquisition of medical imaging studies—can significantly increase detection rate of mislabeled radiographic studies without increasing interpretation time. This addition could have a significant impact on patient care and safety in health care delivery.

Acknowledgments

We thank the following radiologists for their participation in this study: Michael Collins, Abhijit Datir, Hemali Desai, Paul Harkey, Travis Henry, John Holbrook, William Auffermann, Faisal Khosa, Brent Little, Kiran Maddu, Sadhna Nandwana, Nimesh Patel, Tarak Patel, Sanjit Peter, Edward Richer, Joanna Rossi, Aarti Sekhar, Abdul-Rahman Tarabishy, Stefan Tigges, and Keith Tomich. We also thank anonymous reviewers from earlier versions who contributed significantly to the presentation.

S. Tridandapani was supported in part by grants (UL1 RR025008, KL2 RF025009) from the Clinical and Translational Science Award Program of the Public Health Service, National Institutes of Health, Naciona Center for Research Resources, and in part by an award (K23EB013221) from the National Institute of Biomedical imaging and Bioengineering.

References

1. Committee on Quality of Health Care in America. Institute of Medicine, crossing the quality chasm: a new health system for the 21st century. Washington, DC: National Academy Press; 2001.
2. Joint Commission on Accreditation of Healthcare Organizations. [Accessed June 19, 2012] Hospital national patient safety goals. 2012. www.jointcommission.org/assets/1/6/2012_NPSG_HAP.pdf
3. Agency for Healthcare Research and Quality Website. [Accessed June 23, 2012] National Quality Forum. 30 Safe practices for better health care fact sheet. www.ahrq.gov/qual/30safe.htm
4. Kuzmak P, Dayhoff R. The use of digital imaging and communications in medicine (DICOM) in the integration of imaging into the electronic patient record at the Department of Veterans Affairs. J Digit Imaging. 2000; 13:133–137. [PubMed: 10847382]
5. Gale ME, Gale DR. DICOM modality worklist: an essential component in a PACS environment. J Digit Imaging. 2000; 13:101–108. [PubMed: 15359747]
6. Pennsylvania Patient Safety Authority, ECRI Institute, Institute for Safe Medication Practices. Pennsylvania patient safety advisory: applying the universal protocol to improve patient safety in radiology. Pa Patient Saf Advis. 2011; 8:63–69.
7. Thornton RH, Miransky J, Killen AR, Solomon SB, Brody LA. Analysis and prioritization of near-miss adverse events in a radiology department. AJR. 2011; 196:1120–1124. [PubMed: 21512079]
8. Schultz SR, Watson RE Jr, Prescott SL, et al. Patient safety event reporting in a large radiology department. AJR. 2011; 197:684–688. [PubMed: 21862812]
9. Aakre K, Johnson C. Plain-radiographic image labeling: a process to improve clinical outcomes. J Am Coll Radiol. 2006; 3:949–953. [PubMed: 17412207]
10. Bhalla M, Noble E, McLean F, et al. Radiographic features useful for establishing patient identity from improperly labeled portable chest radiographs. J Thorac Imaging. 1994; 9:35–40. [PubMed: 8114163]
11. Turner, Y.; Hadas-Halpern, I. Radiological Society of North America scientific assembly and annual meeting. Oak Brook, IL: Radiological Society of North America; 2008. The effects of including a patient's photograph to the radiographic examination; p. 576
12. Weiss F, Safdar N. Do digital pictures add to the interpretation of diagnostic imaging: a survey along the spectrum of experience. (abstract) AJR. 2010; 194(suppl):A4–A7.

APPENDIX 1: Form Readers Used to Evaluate the Image Pairs

1) Image Quality:	OK		Not OK
2) Lines and Tubes:	OK		Not OK
3) Patient Status:	Improved	No Change	Worsened
4) Other Comments:			

APPENDIX 2: Questionnaire Readers Completed After Two Phases of Image Assessment

1) Were the photographs a distraction?	Yes	No
2) Did you feel you spent more time because of the photographs?	Yes	No
3) Did the photographs help with the interpretation?	Yes	No
	If yes,how?	
4) If you noted mismatched photographs, did you go back and check the radiographs?	Yes	No

Fig. 1.

Histogram shows number of patients from whom different numbers of composite image pairs were obtained for study: 27 patients contributed 88 pairs.

Fig. 2.

First example of mismatched pair of two radiographs of different patients along with corresponding facial and chest photographs. Readers were not told that radiographs were of different patients.

A, This radiograph-photograph pair was designated as “current” study when presented to readers. Images are of 81-year-old white man.

B, This radiograph-photograph pair served as older, comparison (“previous”) radiograph of patient shown in **A**. This image set is of 89-year-old white man. Differences in facial features are more noticeable than differences in radiographic appearances.

Fig. 3. Second example of mismatched pair of two radiographs in our study. Each radiograph along with corresponding photograph is of different patient.

A, This radiograph-photograph pair was designated as “current” study when presented to readers. Radiograph and photograph show 65-year-old white woman.

B, This radiograph-photograph pair served as older, comparison (“previous”) radiograph for patient shown in **A**. Radiograph obtained 43 days before radiograph in **A** shows 75-year-old white man. Again, photographic differences are more obvious than radiographic differences.

TABLE 1

Results From Reader Study Involving 10 Readers

Reader No.	Without Photographs			With Photographs		
	No. of Mismatches Introduced	No. of Mismatches Reported	Assessment Time(min:s)	No. of Mismatches Introduced	No. of Mismatches Reported	Assessment Time (min:s)
1	2	0	20:30	3	2	20:02
2	3	0	30:20	3	1	14:30
3	2	0	—	2	2	—
4	3	3	37:14	2	2	—
5	2	0	27:30	2	0	13:43
6	2	0	45:28	2	1	35:57
7	2	0	39:31	2	2	36:35
8	2	0	40:21	2	0	38:15
9	2	0	45:00	3	3	—
10	4	0	—	4	3	—
Total	24	3 ^a		25	16 ^b	
Average assessment time (min)			35.73			26.51

Note—Dash (—) indicates missing data.

^aPercentage of errors detected = $3/24 = 12.5\%$.

^bPercentage of errors detected = $16/25 = 64\%$.