Distribution Agreement

Signature:

____ _______________

Srini Tridandapani Date

Minimizing Patient Misidentification Errors in Radiology with Digital Photographs Obtained at the Point-of-Care of Medical Imaging

By

Srini Tridandapani
Master of Science
Clinical Research

____
Tracy Faber, PhD
Advisor

____
John Boring, PhD
Committee Member

____
John McGowan, MD
Committee Member

Accepted:

____
Lisa Tedesco, PhD
Dean of the James T. Laney School of Graduate Studies

____

Date

Minimizing Patient Misidentification Errors in Radiology with Digital Photographs Obtained at the Point-of-Care of Medical Imaging

By

Srini Tridandapani

BE, Anna University, 1988

MSEE, University of Washington, 1990

PhD, University of Washington, 1994

MD, University of Michigan, 2001

Thesis

The Master of Science in Clinical Research Program

Advisor: Tracy Faber, PhD

An Abstract of

A thesis submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Clinical Research

2012

ABSTRACT

Minimizing Patient Misidentification Errors in Radiology with Digital Photographs Obtained at the Point-of-Care of Medical Imaging

By Srini Tridandapani

OBJECTIVE: To evaluate whether facial photographs obtained simultaneously with radiographs increase radiologists' detection rate of mislabeling errors.

METHODS: After IRB approval, we obtained simultaneous portable radiographs and photographs from 41 patients. We generated 81 pairs of chest radiographs (one recent, one prior radiograph). We compiled lists of 20 pairs for reader review. Two to four mismatched pairs (i.e., containing radiographs from different patients) were introduced into each list.

Ten radiologists (Group 1), blinded to presence of mismatches, interpreted 20 radiograph pairs. Readers then reviewed a second list containing mismatches, but with photographs attached. Readers were not instructed regarding the purpose of the photographs. The mismatch detection rate was recorded for both sessions. Time for interpretation of the pairs was recorded for both sessions.

Another five radiologists (Group 2) reviewed radiograph pairs in two sessions but were informed that photographs were meant to correlate radiographs (but not that photographs were used to detect mismatches). The Fisher two-tailed exact test was used to evaluate differences in mismatch detection rates between sessions ($p < 0.05$).

RESULTS: For Group 1, detection rates without (3/24=12.5%) and with photographs (16/25=64%) significantly differed (P=0.0003). For Group 2, difference in detection rates without (0/20=0%) and with photographs (17/18=94.4%) was even greater (P=0.0001). For all readers, average interpretation time without photographs was 32.16 minutes and with photographs was 23.80 minutes (t-test two-tailed P=0.0401).

CONCLUSION: Use of photographs increased detection of errors and decreased film interpretation time, which may translate into improvements in patient safety and radiograph throughput.

Minimizing Patient Misidentification Errors in Radiology with Digital Photographs Obtained at the Point-of-Care of Medical Imaging

By

Srini Tridandapani

BE, Anna University, 1988

MSEE, University of Washington, 1990

PhD, University of Washington, 1994

MD, University of Michigan, 2001

Thesis

The Master of Science in Clinical Research Program

Advisor: Tracy Faber, PhD

A thesis submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Clinical Research

2012

ACKNOWLEDGEMENTS

I wish to place on record my heartfelt gratitude to my late mentor, Tracy Faber, PhD, for advising me through the MSCR and KL2 program. She served as a mentor through my KL2 program with the Atlanta Clinical and Translational Sciences Institute, and was instrumental in my obtaining a K23 grant. Unfortunately, she passed away a few weeks before I submitted my thesis. While her signature is not present on the thesis, she definitely left a mark on my career.

I am also grateful to John Boring, PhD and John McGowan, MD for serving on my thesis committee, and for giving me invaluable feedback, which significantly improved the final product. Several people contributed to the project discussed in this thesis. James Provenzale, MD, discussed the project with me over many hours and many of his thoughts and feedback are reflected in this work. Senthil Ramamurthy, MSEE, contributed extensively to this project by providing hardware and software support and assisting with the reader studies. Samuel Galgano, BS, spent many early morning hours in the hospital consenting and acquiring the patient data used for this project.

I thank the following radiologists for their participation in the reader portion of this study: Michael Collins, MD, Abhijit Datir, MD, Hemali Desai, MD, Paul Harkey, MD, Travis Henry, MD, John Holbrook, MD, Faisal Khosa, MD, Brent Little, MD, Kiran Maddu, MD, Sadhna Nandwana, MD, Nimesh Patel, MD, Tarak Patel, MD, Sanjit Peter, MD, Joanna Rossi, MD, Edward Richer, MD, Aarti Sekhar, MD, Abdul-Rahman Tarabishy, MD, Stefan Tigges, MD, and Keith Tomich, MD.

I would also like to thank everyone in the MSCR program, including Cheryl Sroka and all of the outstanding MSCR instructors.

Finally, I would like to thank my family for supporting me through this effort: My wife, Pamela Bhatti, PhD, whom I try to repay by supporting her through her MSCR work now. My older son, Shyam wanted to learn about biostatistics with me, even though he was only seven when I started this program. The start of my MSCR program coincided with the birth of my younger son, Rama, and while I'd like to think I spent enough time with him during the course of this program, I probably did not.

TABLE OF CONTENTS

INTRODUCTION

Patient safety issues have gained prominence in the national dialogue in the United States of America particularly since the publication of the 2001 Institute of Medicine's report on quality [1]. The Joint Commission on Accreditation of Healthcare Organizations (JCAHO) in its 2012 National Patient Safety Goals (NPSG) provides a specific requirement (NPSG.01.01.01) that at least two patient identifiers be used when providing care, treatment and services [2]. The rationale is that "wrong-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" [2]. Meanwhile, the National Quality Forum [3], with support from the Agency for Healthcare Research and Quality, has specifically endorsed, in its "30 Safe Practices for Better Health Care Fact Sheet," the use of standardized protocols to prevent mislabeling of radiographs.

It is quite difficult to obtain data on mislabeling or misidentification errors. In many cases, errors may be undetected. Many errors are likely detected by technologists at the time of completing the study and are unreported because a correction can be made at this stage before a clinical decision is made based on the imaging study. A report by Kuzmak and Dayhoff [4] showed an error rate of 0.73% in 93,000 imaging examinations in one study, and in another study Gale and Gale [5] reported an error rate of 0.26% in 48,800 studies. Such an error rate can lead to substantial problems in patient management. However, in reality, such studies may under-report the true error rate.

Obtaining and assigning the acceptable identifiers noted in the JCAHO NPSG requirements can be problematic, particularly when patients are unconscious, uncooperative or non-communicative and cannot verify their identity. A means to more confidently identify patients based on facial appearance, a longstanding method for identification of individuals, could potentially diminish identification errors.

To minimize or prevent mislabeling of medical imaging studies, we hereby introduce a scheme to obtain digital photographs of patients simultaneously with all medical imaging studies including, but not limited to radiography, Computed Tomography (CT), Ultrasound (US), Magnetic Resonance Imaging (MRI), Scintillation Cameras, and Positron-Emission-Tomography (PET) (Tridandapani S et al., presented at the 2010 IEEE IECBES meeting; Ramamurthy S et al., presented at the 2011 AMA-IEEE Medical Technology Conference). These digital photographs are small additions to the imaging study similar to the *scout* or *localizer* images that are performed with CT studies. *We do not intend these digital photographs to entirely replace numerical identifiers, but rather we envision them as a means to supplement and strengthen these identifiers*. However, in some cases, such as unconscious trauma patients who cannot provide patient identification information, these photographs may indeed be the only available identifiers.

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

METHODS

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

HYPOTHESIS

$H_{O,1}$: The *primary null hypothesis* for our study is that the addition of photographs to portable chest radiographs does not change the detection rate of mismatched (mislabeled) radiograph pairs.

$H_{A,1}$: The *primary alternative hypothesis* is that the addition of photographs to portable chest radiographs alters the rate of detection of mismatched radiographs pairs.

$H_{0,2}$: The *secondary null hypothesis* is that the addition of photographs does not alter the interpretation time of radiographs.

$H_{A,2}$: The *secondary alternative hypothesis* is that the addition of photographs affects the interpretation time.

STUDY POPULATION

The data was gathered between August 2011 and November 2011 in two cardiothoracic surgery intensive care units (ICUs) within Emory University Hospital. Most of the data was gathered during the hours between 2:00 AM and 6:00 AM when the majority of portable radiographs are obtained in ICUs. Some of the data was obtained in step-down care units or other regular hospital floors if the patient was transferred out of the ICU during their hospitalization.

We initially recruited 41 patients into the study. However, 11 of these patients either underwent only one chest radiograph during their hospitalization or had only chest radiographs for which concomitant photographs were not obtained. Thus radiograph pairs could not be generated from these patients. A radiograph from one of these patients was however used for creating an erroneous pair. Some of the patients had more than the 10 radiograph-photograph combinations that were included in the study. Of the 30 patients who had one or more radiograph pairs, we used data from only 27 (Fig. 1) patients because of technical difficulties with retrieving images from the PACS for the other three patients.

The final study cohort consisted of 28 patients (13 males, 15 females; mean age 61 years +/- 15.16; range 22-89 years). These patients were admitted for a variety of diagnoses with the four most common being aortic stenosis (n=11), congestive heart failure (n=9), mitral regurgitation (n=3), and coronary artery disease (n=3). The four most common surgeries that the patients had during their 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 performed in standard fashion with the technologist confirming patient identity verbally or by wristband information. Immediately prior to obtaining the radiograph, a single photograph of the patient’s face and chest was obtained by placing a camera immediately adjacent to the properly positioned radiographic equipment. All photographs were obtained by a single individual with a 5-megapixel camera on an Apple iPhone 4 (Apple Inc., Cupertino, CA), with the use of a flash. Photographs were initially stored in

Joint Picture Experts Group (JPEG) format. Prior to patient photograph acquisition, a photograph of the paper requisition form was also obtained to ensure that no errors were made in the matching of the radiographic study with the patient photograph. On occasion, studies were added to the worklist and the technologist was paged with the request when he/she was already on the floor; thus radiographs were obtained without the aid of the requisition form. For any study for which an acquisition number was unavailable at the time of the examination, acquisition numbers were acquired from the Picture Archiving and Communications Systems (PACS) using other patient identifiers, including name, date of birth, medical record number, and date and time the study was performed. 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 Digital Imaging and Communication in Medicine (DICOM) format. The photograph and the corresponding radiograph were “stitched” together using custom-developed software; one of the authors (S.R.) generated a composite image with the color photograph located to the left of the grayscale radiograph. The photograph was approximately one-fourth of the size of the radiograph (example shown in Fig. 2).

ANONYMIZATION AND PAIRING

Anonymized study sets were generated by combining two sequential radiographs of the same patient, i.e., a “current” radiograph and the most recent previous radiograph were presented as a pair of images to the 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 portable radiographic examination during their hospitalization, while others received several. Data from patients who only underwent one portable radiographic examination could only be used for the erroneous sets, as described below.

From the 166 radiographs obtained, we created 81 unique pairs of matched radiographs. In addition, we created 12 mismatched sets by combining current and comparison studies from different patients. From this set of 81 pairs, we selected a pool of 44 pairs from which radiographs were shown to readers, which included up to 20% mismatched pairs. Two academic cardiothoracic radiologists also independently categorized mismatched pairs according to degree of difficulty, with 1 designated as easily identifiable as a mismatch based solely on radiographic findings and 2 designated as very difficult to identify as a mismatch.

READER GROUPS

GROUP 1: Ten recently-trained radiologists served as readers. Nine of these radiologists were certified by the American Board of Radiology (ABR) within two years prior to the study. One radiologist had been trained in the United Kingdom (U.K.) and was pursuing his second year of fellowship training in the United States (U.S.). All readers were either pursing 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 we thought it would represent the overall population of general radiologists nationally.

SAMPLES OF MISMATCHED PAIRINGS

Figures 2 through 6 show some of the mismatched pairs used. Some differences between the patients in a mismatch were subtle, such as the presence or absence of calcified mediastinal lymph nodes (Fig. 2), or of a breast shadow (Fig. 3). Other pairs were markedly different (Figs. 4-6), for example, a mismatched pair in which the initial radiograph was that of a patient who had bilateral lung volume reduction surgery for chronic obstructive pulmonary disease and showed very different devices on the subsequent examination, i.e., median sternotomy wires and a left ventricular assist device (Fig. 4). Another example with a significant difference between the new and prior examination included a case of initial radiograph showing presence of metallic hardware associated with a left thoracotomy, which was absent on the subsequent radiograph obtained (in a different patient) one day later (Fig. 6).

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.

In Phase 1, twenty randomly selected pairs of radiographs were presented to each reader on a ClearCanvas DICOM viewer (ClearCanvas, Inc., Toronto, Ontario, Canada), running on a dual monitor workstation. Of these 20 pairs, between two and four mismatched pairs was included in random order. Aggregating across all readers, a total of 11 mismatched pairs were shown. Five of these were categorized as easy and six as difficult to detect. A total of 42 pairs were shown without photographs and 42 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 figures in which readers individually assessed the same pair of radiographs one time without photographs and another time with photographs, we calculated the improvement in reader performance for each figure. The total time for interpretation of this set of 20 radiographs was recorded.

Readers were not informed of 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 from 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 following form:

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

The fourth question was employed as a means for readers to indicate that they had detected mismatched pairs. The time for interpretation of the entire set was 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, the photographs which had been 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 to them, and that this information may or may not help them 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 following questionnaire:

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 If yes, how?	No
4) If you noted mismatched photographs, did you go back and check the radiographs?	Yes	No

GROUP 2: A second group of five recently-trained radiologists participated in a similar study, but here readers were explicitly informed before beginning phase 2 that photographs were added to aid with recognition of patients and correlation of intravenous lines and tubes. However, once again, readers were not specifically informed that mismatches existed and that photographs were intended to help identify mismatches.

In this group, four readers had received ABR certification within the last two years. One of the radiologists had been residency-trained in the U.K. and was pursuing fellowship training in the U.S. Another radiologist was recently ABR-certified after residency training in the U.K. and fellowship training in the U.S. and was a first-year faculty member. All interpreters were either pursuing fellowship training in subspecialties that did not include cardiothoracic radiology or were first-year faculty members in divisions other than cardiothoracic radiology.

STATISTICAL ANALYSIS

The two-tailed Fisher exact test was employed to compare error detection rates without and with photographs; a P value of 0.05 or less was used to indicate a significant difference. A t-test was performed comparing the times taken by each reader in phase 1 and phase 2. A P value of 0.05 or less was considered to indicate a significant difference. Statistical testing was performed using QuickCalcs (GraphPad Software, Inc., La Jolla, CA).

RESULTS

RESULTS FOR GROUP 1 READERS

Table 1 provides the assessment results for Group 1 readers. In Phase 1, (i.e., without the benefit of photographs), only one reader correctly reported the presence of mismatched radiographs. In Phase 2, i.e., when photographs were provided but readers were not made aware of the purpose of photographs, 5 of the same readers correctly reported the presence of all mismatched radiographs; an additional 3 readers reported the presence of some mismatched pairs. One reader later mentioned that he was actively ignoring the photographs because he thought the intent of their use was to distract readers from providing a proper radiologic assessment. Despite this belief, the reader 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 intravenous lines and tubes, but did not compare the two photographs to see if they were from the same patient.

Overall, there was a significant difference in the mismatch detection rates without and with photographs ($P = 0.0003$) in the first group of readers who were not informed of the purpose of the presence of photographs. In the absence of photographs, only 3 of 24 (12.5%) mismatched pairs were reported; with photographs detection rate improved to 16 of 25 (64%) mismatched pairs. Table 3, shows the sensitivity, specificity and accuracy for detection of mismatched pairs with and without photographs: we note an increase in accuracy from 78% to 91%.

The mean (standard deviation) time for interpretation without photographs was 35.73 (8.84) minutes and with photographs was 26.51 (11.65) minutes, with mean difference being 9.23 (95% CI: -2.66 to 21.12) minutes (unpaired t-test, two-tailed P = 0.1178), which was not statistically significant.

RESULTS FOR GROUP 2 READERS

For Group 2 readers, who were informed of the purpose of the photographs but not that mismatched pairs existed, in Phase 1, none of the 20 errors were detected. However, when photographs were provided, 94.4% (17/18) of mismatched pairs were reported (P = 0.0001). There was one false positive reading, i.e., one of 82 properly paired radiographs-photograph matches was erroneously reported as belonging to two different patients. As shown in Table 3, the sensitivity was even higher (88.89%) in the presence of photographs because the observers were incorporating the photographs in their interpretation. The false positive led to a small decrease in specificity to 98.78%.

Overall the accuracy increased from 80% to 97% (i.e., for the entire sample including matched and mismatched pairs) once photographs were introduced and readers were instructed to include them in their assessment. When both Group 1 and Group 2 were pooled, the accuracy still showed a sizeable increase from 78.22% to 94% once photographs were included.

Time was recorded for all 5 readers in Group 2. Time for interpretation of the 20 pairs did not increase with the introduction of photographs. In fact, a decrease in time for interpretation in phase 2 was seen (Table 2 and 3), although it was not statistically significant. The average (standard deviation) time for radiograph assessment for phase 1 and phase 2 for the 11 readers was 26.45 (±8.69) minutes and 20.55 (±3.40), respectively. The mean time difference between phase 1 and phase 2 was 5.90 (95% CI: -3.629 to 15.434) minutes (t-test, two-tailed P = 0.1911).

When the interpretation times for Groups 1 and 2 were pooled, the average (standard deviation) of interpretation times without photographs was 32.16 (±9.61) minutes and the average time with photographs was 23.80 (±9.06) minutes, for an average difference of 8.36 minutes (95% CI: 0.411 to 16.32) minutes (t-test two-tailed P=0.0401).

Improvement in Reader Performance

In all, 11 mismatched pairs were used. There were 21 instances where readers assessed the same mismatched radiographs without photographs and in another session with photographs. In 19 (90%) of these, readers did not identify the mismatch without photographs, but in another session they identified these as mismatches with photographs. In the remaining 2 instances, readers failed to identify the mismatch both without and with photographs.

For 9 mismatched pairs, all readers improved and in 2 pairs none of the readers improved. Of the 11 mismatched pairs, 5 that were categorized as “1” (easy to distinguish by radiography alone) by the consensus panel were shown to readers 12 times, and improvement was seen in 11 instances with failure to improve noted in one instance. Of the 11 mismatched pairs, 6 that were categorized as “2” (difficult to distinguish by radiography alone) by the consensus panel were shown to readers 9 times, and improvement was seen in 8 instances with failure to improve noted in one instance. Thus it appears that ease or difficulty of distinguishing the mismatch as judged by the consensus panel did not have much of an effect on the detection rate (Fischer exact test, two-tailed P=1.00).

Again when considering only those radiograph pairs that were shown to individual readers both without and with photographs, Group 1 readers improved 9 out of 11 times (81.82%), while Group 2 readers improved 10 out of 10 times (100%). Thus, instructing the readers leads to increase in improvement with introduction of photographs, although not in a statistically significant manner (Fischer's exact test, two-tailed P=0.4762).

Post-study Questionnaire

Post-study questionnaire results were as follows:

1) Three readers of the 15 readers (20%) indicated that the photographs were a distraction, while 12 (80%) did not find the photographs distracting.
2) Nine readers (60%) reported that they spent more time per case when photographs were present, while six readers (40%) indicated that they did not spend more time. In fact, objective times as noted above did not reveal any significant difference between the two phases when the data was considered for each group. In fact, when Groups 1 and 2 were pooled, a statistically significant decrease in interpretation time was seen after introduction of photographs.

3) Three readers (20%) indicated that photographs did not help with the interpretation, while 12 (80%) thought they did. Of these twelve, 7 thought photographs assisted with interpretation of lines and tubes, 6 explicitly mentioned that photographs helped with identifying mislabeled patients, and 7 noted that photographs helped with evaluating patient status.
4) Fourteen readers 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 felt the radiographs in the pair belonged to the same patient and the errors were only in the photographs.

DISCUSSION

Our primary result is that the presence of patient photographs obtained simultaneously with ICU portable radiographs significantly increases the detection rate when radiologists are presented with mismatched radiographs. The secondary result is that the addition of photographs may result in a decrease in the interpretation time.

Introducing photographs into the workflow could potentially result in an increase in interpretation time. Some of the readers felt that they were spending more time because of the photographs. However, we found that the time actually decreased between phase 1 and phase 2 (although not statistically significant). Quite possibly the decrease in time may have been due to learning effect, i.e., readers were getting accustomed to the task during phase 1 and were thus quicker during phase 2. On the other hand, some readers felt that the supplementary photographic information regarding lines and tubes sped up their evaluation.

RELATIONSHIP TO PRIOR STUDIES

To our knowledge, only one published study [5] has addressed the problem of establishing patient identity from improperly labeled portable chest radiographs. In that study, Bhalla et al. [5] discussed various radiographic features such as the "characteristic location and configuration of surgical material, fractures, and dense parenchymal/pleural scars with or without calcifications for establishing patient identity". However, they also recognized that the vast majority of patients lack such characteristic surgical and pathological features, and suggested that some anatomic features such as the transverse processes of the first thoracic vertebrae and the adjoining tubercles of the first ribs and the spinous processes may be 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 provide a simple means to overcome this problem.

Our results indicate that even the presence of metallic hardware and characteristic intravenous line and tube positions did not enhance rate of detection of mismatched pairs. For instance, the case in Fig. 4 where a mismatched pair initially showed a patient who had bilateral lung volume reduction surgery and subsequently showed a patient in heart failure did not result in increased detection in the absence of photographs. The rate of detection of the mismatch before and after introduction of photographs was (0/3) 0% and (3/3) 100%, respectively, in Group 2; this error was not shown to Group 1 readers.

After radiograph assessment sessions, readers were shown their undetected mismatched cases. All readers 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 in order to account for differences between radiographs. In addition, we hypothesize that radiologists' training to identify mislabeled radiographs is relatively weak because such 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. Clearly, one likely mechanism is that readers first noticed the mismatch between photographs and then looked more closely at radiographs to note whether they were discordant. 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 used a post-study questionnaire answered by readers and showed that radiologists' empathy increased with the introduction of photographs [6]. Interestingly, that study also found that the number of incidental findings reported increased when photographs were employed. This finding suggests that the presence of photographs may promote a greater degree of involvement by film readers, resulting in a higher detection rate of not solely 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 [7] conducted a single institution survey of 21 radiologists regarding supplementary use of digital photographs [7]; 24% of respondents were of the opinion 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 use of photography in radiology were to become prevalent. In our study 12 out of 15 (80%) readers felt that photographs were not a distraction and 12 out of 15 (80%) readers felt that photographs helped with their interpretation. In contrast with Weiss and Safdar's study [7], in our study readers looked at actual patients' photographs along with the radiographs in a simulated study before they responded to the questionnaire. Experiencing first-hand the advantage of including photographs may have influenced their preference for such inclusion. Our study was a simulated one, in contrast to Turner and Hadas-Halpern [7] who performed their study in a real-life workflow setting. However, their study was not designed to evaluate error-detection capabilities of the radiologists.

LIMITATIONS

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

The five readers in the second portion of the study were informed about the rationale for including patient photographs only before phase 2. A criticism of bias can be leveled at this study design, i.e., it can be argued that had the readers known about the possible presence of misidentification errors in Phase 1, the detection rate would be higher even without photographs. However, we believe that this more accurately reflects actual film interpretation conditions, and had the readers been specifically informed about the potential for mislabeling errors, their approach to interpretation would have been biased and not reflective of reality. For the same reason, we avoided a cross-over study design where 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 such a cross-over study design is that readers typically become biased once the objective of the study is deduced by the readers. However, 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 of the presence of, and rationale for, such a tool.

Another limitation of this study is generalizability. Error detection rates may have been low because all the radiographs were from patients who were in the cardiothoracic ICU and had similar pathologies. In smaller institutions, where radiologists interpret pediatric, adult, inpatient, and outpatient radiographs all from one worklist, misidentified patients may be easily detected, for example, if a neonate’s chest radiograph is included in an adult patient’s image folder. In larger institutions, where there is separation of the reading worklists, and all ICU studies are found in a separate worklist without other types of studies included in this worklist, error detection may be lower. On the other hand, since errors tend to occur in clusters, i.e., 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's imaging folder.

A second issue of generalizability is that detection rates in the absence of photographs may be very low because the actual error rate will be much smaller than the rates that we used. If the actual error rates are very low, we conjecture that the detection rates in the absence of photographs would be even lower since radiologists would have fewer opportunities to learn from these infrequent occurrences. Quite possibly the false positive rate may also increase in this setting if photographs were to be introduced.

CONCLUSION

The use of photographs obtained at the point-of-care, i.e., simultaneously with the acquisition of medical imaging studies, can significantly increase detection rate of mislabeled radiographic studies. This could have a significant impact on patient care and safety in healthcare delivery.

FUTURE WORK

The scheme we have introduced may need modifications before implementation in the real world, 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 some preliminary schemes elsewhere to address this issue (Tridandapani S et al., presented at the 2010 IEEE IECBES Meeting; Ramamurthy S et al., presented at the 2011 AMA-IEEE Medical Technology Conference; Ramamurthy S et al., accepted for presentation at the 2012 SIIM Annual Meeting). These issues need to be investigated in a real-world setting, in addition to resolving any legal and ethical issues that arise with storing and viewing patient photographs.

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 2010 National patient safety goals. Available at http://www.jointcommission.org/assets/1/6/NPSG_Chapter_Jan2012_AHC.pdf. Accessed on Jan 25th, 2012.

3. National Quality Forum. 30 Safe practices for better health care fact sheet. Available at http://www.ahrq.gov/qual/30safe.htm. Accessed on Jan. 25th, 2012.

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. *Journal of Digital Imaging* 2000; 13:133-137.

5. Gale M, Gale D. DIOM modality worklist: an essential component in a PACS environment. Journal of Digital Imaging 2000;13(3):101-108.

6. Bhalla M, Noble E, McLean F, *et al.*, "Radiographic features useful for establishing patient identity from improperly labeled portable chest radiographs," *J Thoracic Imaging* 1994; 9:35-40.

7. Y. Turner, I. Hadas-Halpern. "The effects of including a patient's photograph to the radiographic examination." In: Radiological Society of North America scientific assembly and annual meeting. Oak Brook, Ill: Radiological Society of North America, 2008; 576.

8. F. Weiss, N. Safdar. Do digital pictures add to the interpretation of diagnostic imaging: a survey along the spectrum of experience. *American Journal of Roentgenology* 2010; 194:A4-A7.

FIGURES

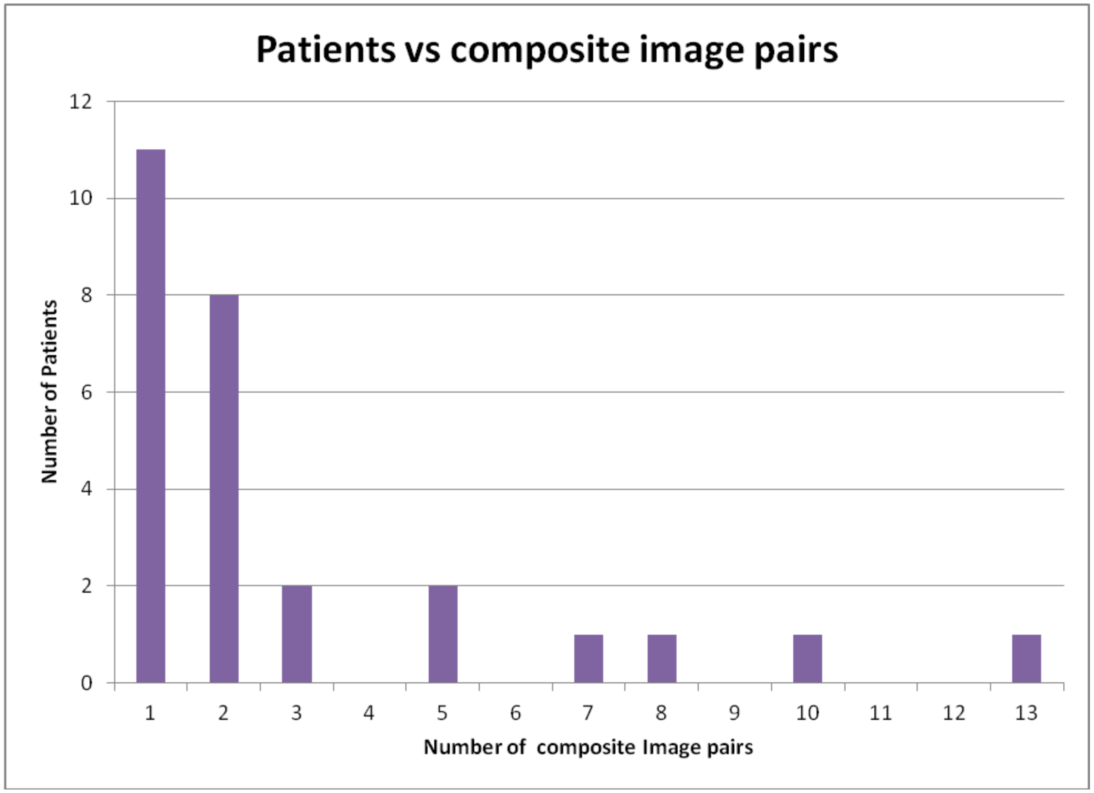

Figure 1. Histogram showing the number of patients from whom different numbers of composite image pairs were used for the study.

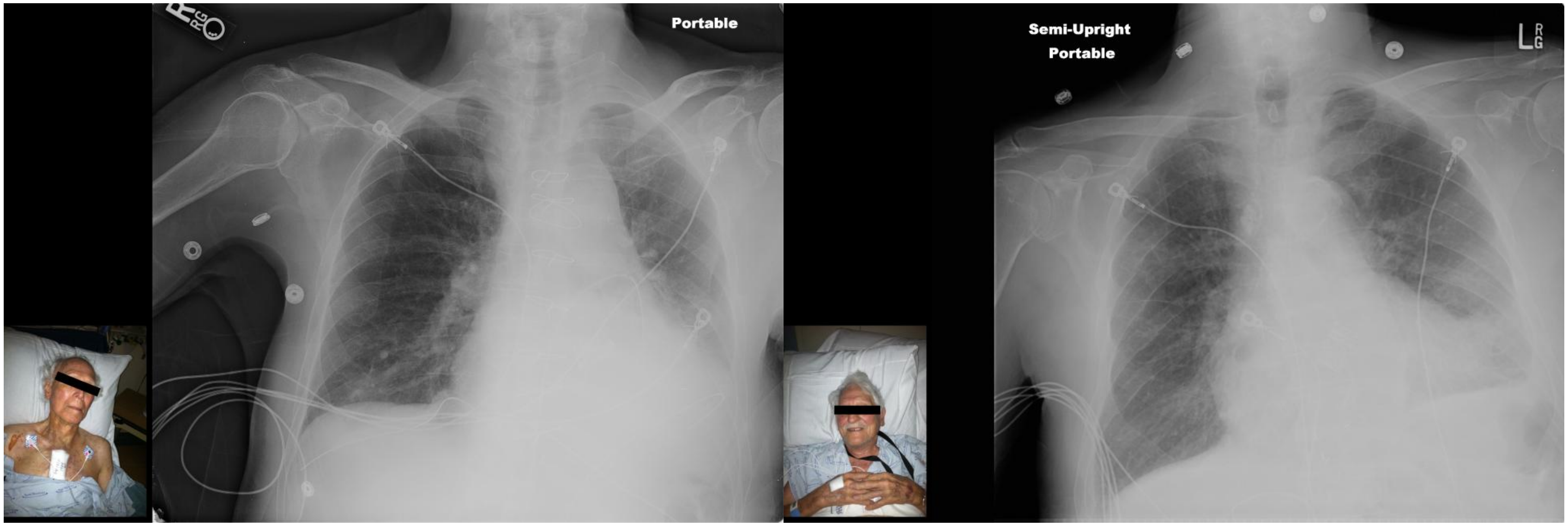

Figure 2. Sample mismatched study 1:

A) The initial radiograph (right) shows an 89 year-old white man with aortic stenosis admitted for aortic valve replacement; the aortic arch is calcified and calcified mediastinal lymph nodes are present.

B) The later radiograph (left) obtained three days after (A) shows an 81 year old man with aortic valve replacement status post coronary artery bypass grafting and aortic valve replacement; this patient does not have any noticeable aortic atherosclerotic calcification or calcified mediastinal lymph nodes, but does have median sternotomy wires. In addition, given a difference of only three days between the two radiographs, it is unlikely that the post-operative changes would show median sternotomy wires only and no support lines and tubes. These differences were overlooked by several readers when viewing radiographs alone.

The photographs (edited to protect patient identity) clearly show differences in facial hair and baldness pattern between the two patients, which prompted readers to more critically assess whether the pair was mismatched. Two readers viewed the radiographs without photographs and neither detected the mismatch. Five readers (including one who viewed radiographs without photographs) viewed the radiographs with photographs and three identified the mismatch. The single reader who viewed the radiographs both without photographs and with photographs failed to identify the mismatch in either case.

This pair was classified as “easy” to distinguish radiographically by the consensus panel.

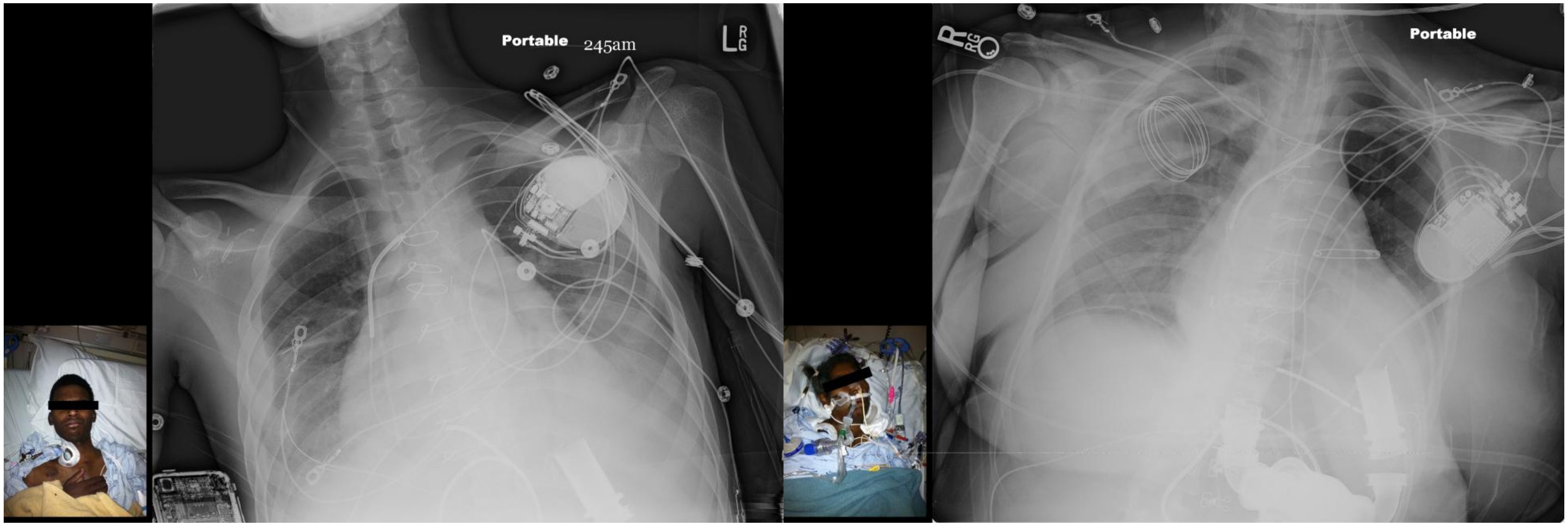

Figure 3. Sample mismatched study 2:

A) The initial radiograph from 9/15 on the right shows a 50 year-old African American woman with cardiomegaly and congestive heart failure who had a left ventricular assist device placed.

B) The later radiograph (left) obtained 12 days after (A) shows a 22 year old African American man with cardiomegaly, congestive failure and a left ventricular assist device. The differences between these two radiographs include different implantable defibrillator devices and a right internal jugular central venous catheter in the patient shown in B. In addition, there is a breast shadow on the radiograph on the right.

Three readers viewed the radiographs without photographs and only one reader identified the mismatch. Four readers viewed the radiographs with photographs (including one reader who viewed the radiographs both without and with photographs). Only one reader identified the mismatch with the assistance of photographs. This was an instance where the reader who viewed radiographs both without and with photographs failed to detect the mismatch in each instance.

This pair was classified as “difficult” to distinguish radiographically by the consensus panel.

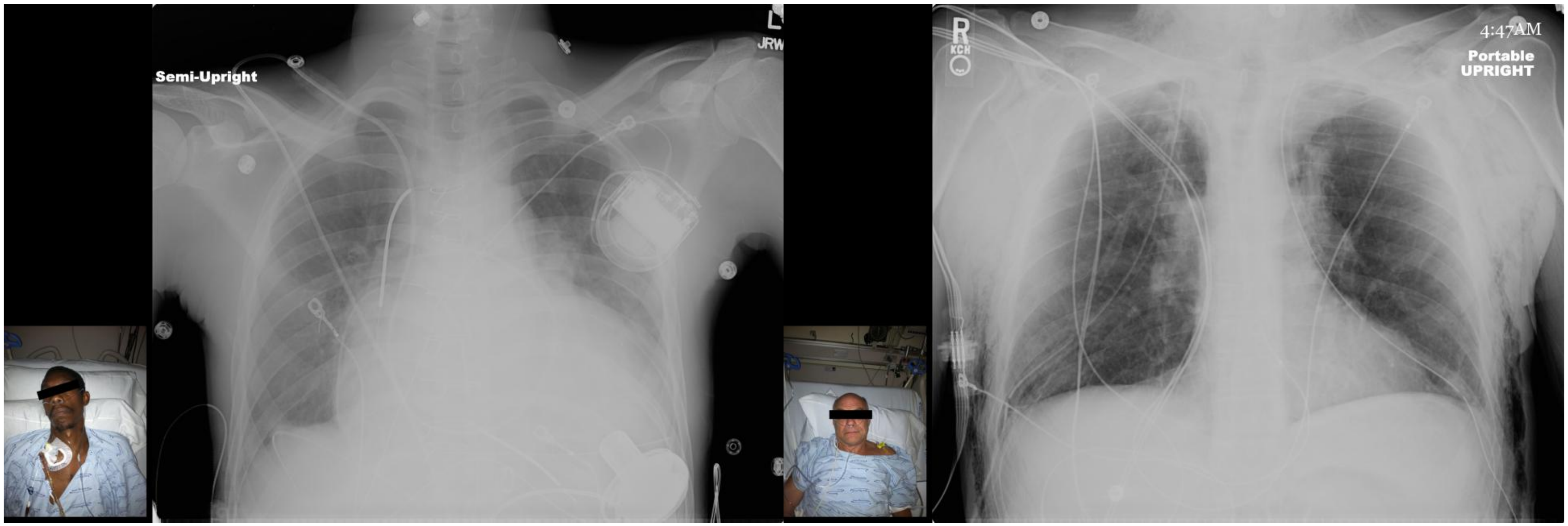

Figure 4. Sample mismatched study 3:

A) The initial radiograph (right) shows a 61 year-old man with a history of bullous emphysema who was status post bilateral lung volume reduction surgery; note the normal heart size and the presence of bilateral chest tubes.

B) The later radiograph (left) obtained 7 days after (A) shows a 33 year-old African American man with congestive heart failure who had a left ventricular assist device placed. The radiograph on the left demonstrates marked cardiomegaly, a left ventricular assist device and a left implantable defibrillator device; the chest tubes are no longer seen. The photographs (edited to protect patient identity) clearly show race and body habitus differences between the two patients and assisted with the detection of mislabeling.

Four readers viewed the radiographs without photographs and none identified the mismatch. Four readers viewed the radiographs with photographs (including three readers who viewed the radiographs both without and with photographs). All four readers identified the mismatch with the assistance of photographs. Thus, the performance of three readers improved with the use of photographs.

This pair was classified as “easy” to distinguish radiographically by the consensus panel.

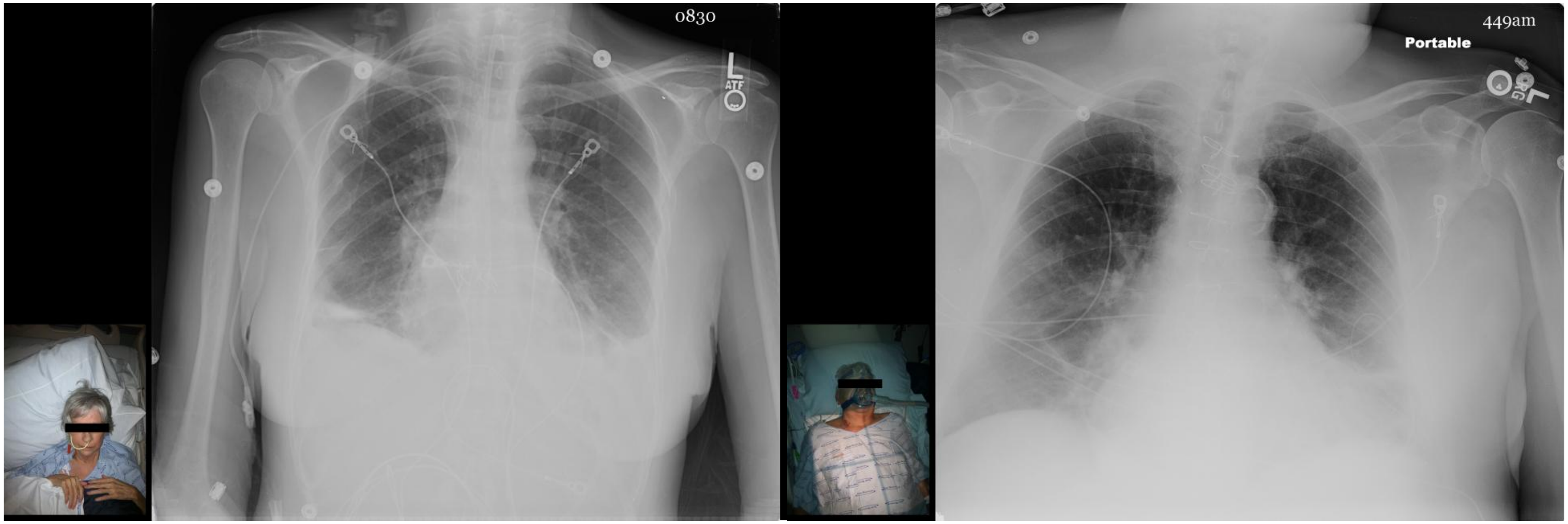

Figure 5. Sample mismatched study 4:

A) The comparison prior radiograph from 8/16 on the right 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.

B) The current radiograph on the left from 8/6 shows a 65 year-old white woman with a history of bronchiolitis obliterans status post bilateral lung transplant; the radiograph shows a normal heart size and clamshell sternotomy wires associated with bilateral lung transplant, and no median sternotomy wires or aortic atherosclerotic calcification as seen in the radiograph on the right. There is bibasilar atelectasis and a small left pleural effusion. A new feeding tube and a new right peripherally inserted central catheter are also seen.

Two readers viewed the radiographs without photographs and none identified the mismatch. Two readers viewed the radiographs with photographs (including one reader who viewed the radiographs both without and with photographs). Both readers identified the mismatch with the assistance of photographs. Thus, the performance of one reader improved with the use of photographs.

This pair was classified as “easy” to distinguish radiographically by the consensus panel.

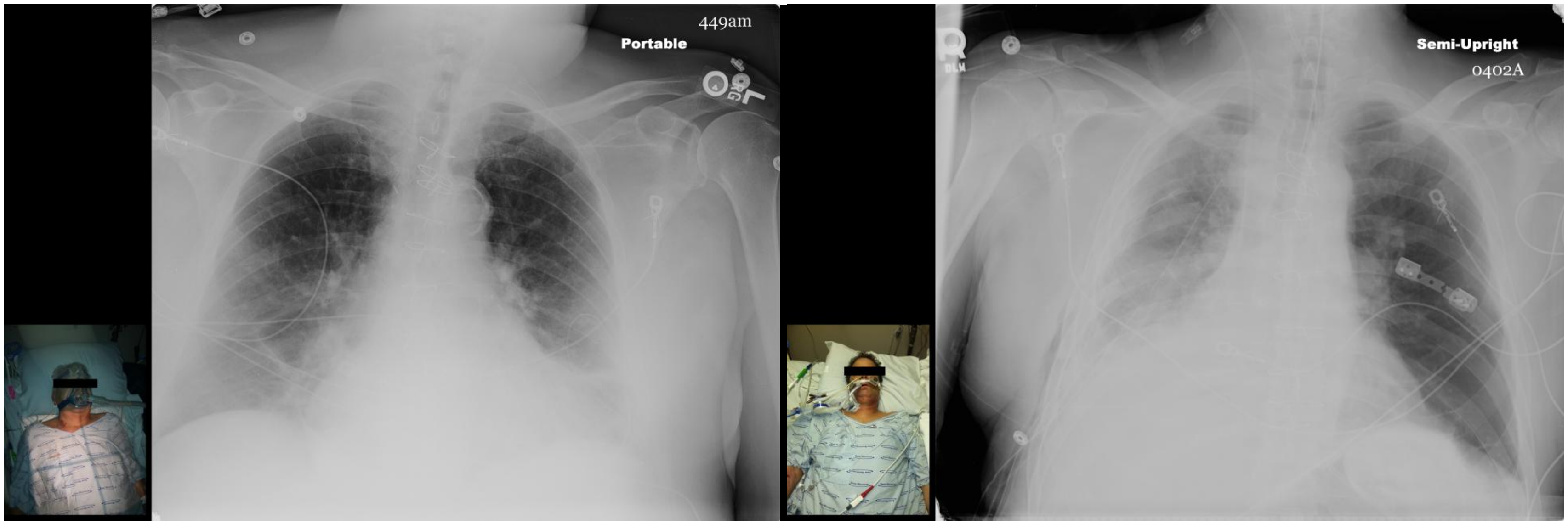

Figure 6. Sample mismatched study 5:

A) The initial radiograph (right) shows a 57 year-old man with idiopathic pulmonary fibrosis who had a prior left lung transplant and was recently status post right lung transplant. The radiograph shows median sternotomy wires, an endotracheal tube, a right internal jugular central venous catheter, and a right pleural effusion. In addition, there is evidence of a left thoracotomy with hardware bridging the thoracotomy defect.

B) The later radiograph obtained one day after (B) shows a 75 year-old 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; there is also a left chest tube and no right chest tube.

Two readers viewed the radiographs without photographs and none identified the mismatch. One of these two readers viewed the radiographs with photographs and correctly identified the mismatch. Thus, the performance of the reader improved with the use of photographs.

This pair was classified as “difficult” to distinguish radiographically by the consensus panel.

TABLES

Table 1. Results from Group 1 Study involving the first 10 readers. These readers did not know the reason for including photographs.

	Without Photographs			With Photographs		
Reader	Number of mismatches introduced	Number of mismatches reported	Assessment time (min:sec)	Number of mismatches introduced	Number of mismatches reported	Assessment time (min:sec)
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	Total = 3 % errors detected= 12.5%	Average = 35.73min	48	Total = 16 % errors detected= 64%	Average = 26.51min

*** - missing data

Table 2. Results from Group 2 Study involving 5 readers. These readers knew the reasons for including the photographs.

	Without Photographs			With Photographs		
Reader	Number of mismatches introduced	Number of mismatches reported	Assessment time (min:sec)	Number of mismatches introduced	Number of mismatches reported	Assessment time (min:sec)
11	4	0	18:47	4	4	17:18
12	4	0	28:34	3	3	24:21
13	4	0	16:09	4	4	16:46
14	4	0	34:46	4	3	23:04
15	4	0	34:00	3	2*	21:16
	Total = 20	Total = 0 % errors detected= 0%	Average = 26.45min	Total = 18	Total = 16 % errors detected= 94.4%	Average = 20.55min

*In addition, one false positive was noted by this reader

Table 3. Test Measures for Group 1, Group 2, and combined Groups 1 and 2.

	Test Measure	Without Photographs	With Photographs
Group 1	Sensitivity	12%	64%
	Specificity	100%	100%
	Accuracy	78%	91%
Group 2	Sensitivity	0%	88.89%
	Specificity	100%	98.78%
	Accuracy	80%	97%
Combined Group 1 and Group 2	Sensitivity	6.38%	74.42%
	Specificity	100%	99.26%
	Accuracy	78.33%	94%