Interrater Agreement and Diagnostic Accuracy of a Novel Computer-Aided Detection Process for the Detection and Prevention of Retained Surgical Instruments

Theodore C. Marentis[1]
Matthew S. Davenport[2,3]
Jonathan R. Dillman[4]
Ramon Sanchez[2]
Aine M. Kelly[2]
Paul Cronin[2]
Mariana R. DeFreitas[5]
Lubomir Hadjiiski[2]
Heang-Ping Chan[2]

Keywords: computer-aided detection, gossypiboma, never-event, patient safety, retained surgical instrument

doi.org/10.2214/AJR.17.18576

Received June 6, 2017; accepted after revision September 12, 2017.

T. C. Marentis is the founder of Kalyspo, LLC, a start-up company that has licensed from the University of Michigan the intellectual technology pertaining to the μTag and the computer-aided detection software. All data were independently analyzed and controlled by all the other coauthors. M. S. Davenport is a paid consultant to the U.S. Food and Drug Administration and the National Cancer Institute and receives book royalties from Wolters Kluwer.

[1]Mercy Radiology Group, Dignity Health, Sacramento, CA.

[2]Department of Radiology, University of Michigan Hospital and Health Systems, 1500 E Medical Center Dr, Ann Arbor, MI 48109. Address correspondence to M. S. Davenport (matdaven@med.umich.edu).

[3]Michigan Radiology Quality Collaborative, Ann Arbor, MI.

[4]Department of Radiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH.

[5]University of Michigan Medical School, Ann Arbor, MI.

AJR 2018; 210:709–714

0361–803X/18/2104–709

OBJECTIVE. The purpose of this study is to evaluate the diagnostic accuracy of a process incorporating computer-aided detection (CAD) for the detection and prevention of retained surgical instruments using a novel nondeformable radiopaque μTag.

MATERIALS AND METHODS. A high-specificity CAD system was developed iteratively from a training set ($n = 540$ radiographs) and a validation set ($n = 560$ radiographs). A novel test set composed of 700 thoracoabdominal radiographs (410 with a randomly placed μTag and 290 without a μTag) was obtained from 10 cadavers embedded with confounding iatrogenic objects. Data were analyzed first by the blinded CAD system; radiographs coded as negative ($n = 373$) were then independently reviewed by five blinded radiologists. The reference standard was the presence of a μTag. Sensitivity and specificity were calculated. Interrater agreement was assessed with Cohen kappa values. Mean (± SD) image analysis times were calculated.

RESULTS. The high-specificity CAD system had one false-positive (sensitivity, 79.5% [326/410]; specificity, 99.7% [289/290]). A combination of the CAD system and one failsafe radiologist had superior sensitivity (98.5% [404/410] to 100% [410/410]) and specificity (99.0% [287/290] to 99.7% [289/290]), with 327 (47%) radiographs not requiring immediate radiologist review. Interrater agreement was almost perfect for all radiologist pairwise comparisons ($\kappa = 0.921–0.992$). Cumulative mean image analysis time was less than one minute (CAD, 29 ± 2 seconds; radiologists, 26 ± 16 seconds).

CONCLUSION. The combination of a high-specificity CAD system with a failsafe radiologist had excellent diagnostic accuracy in the rapid detection of a nondeformable radiopaque μTag.

Errors in medicine are common, costly, often preventable, and sometimes life threatening. According to the National Quality Forum [1], retained surgical instruments are among the most common "never events"—that is, adverse events that are preventable and should never occur in a hospital. The retained surgical instrument rate has been estimated to be 1.78 per 10,000 operations, with approximately two-thirds of retained surgical instruments being surgical sponges and towels [2–4]. The most common complications are adhesions, abscess, fistula, and the need for reoperation [2, 5]; in some patients, these complications can lead to death [6]. The reported total cost of a retained surgical instrument is $300,000–450,000 [7–9]. Retained surgical instruments commonly prolong hospitalization and can lead to multiple readmissions [5]. The Centers for Medicare & Medicaid Services determined in 2008 that the added costs of retained surgical instruments would no longer be reimbursed and estimated that hospitalizations related to retained surgical instruments cost an average of $63,631 [10].

The current approach to the prevention of retained surgical instruments is centered on the surgical count. In this process, objects are counted by operating room staff as the objects enter and leave the operative field. When there is a discrepancy between the number of objects reported to have entered the field and the number of objects reported to have left the field, radiographs are acquired to determine whether there is a retained surgical instrument in the patient. This is facilitated by the presence of radiopaque markers in many surgical materials. However, this method is flawed. It has been shown that most confirmed cases of retained surgical instruments are associated with a correct surgical count and that the surgical count is not a good discriminator of the pres-

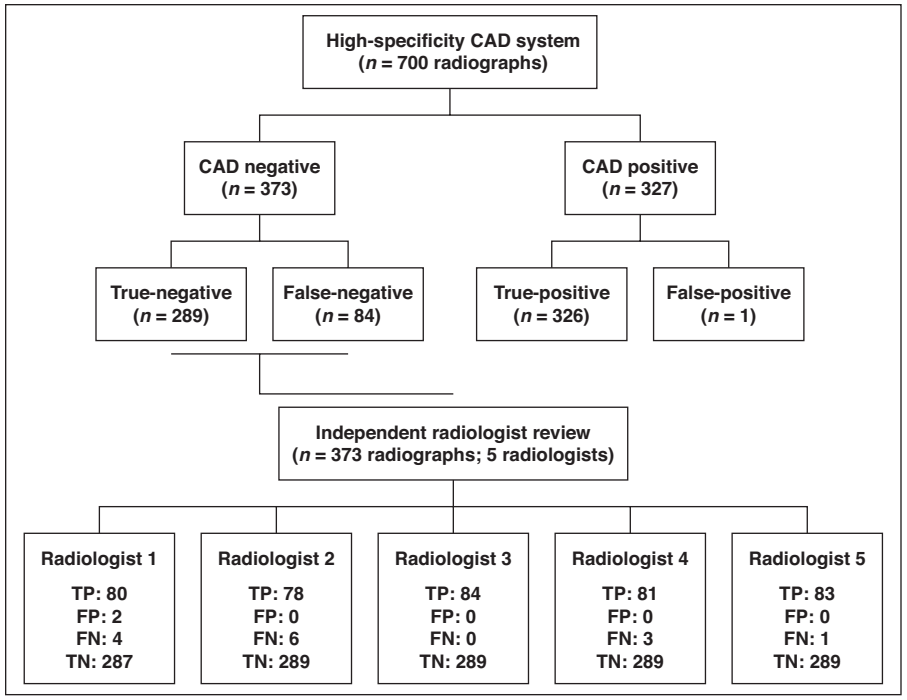

Fig. 1—Study population flowchart for test set. All numbers in boxes refer to numbers of radiographs. CAD = computer-aided detection, TP = true-positive, FP = false-positive, FN = false-negative, TN = true-negative.

ence of a retained surgical instrument [2, 5]. At the Mayo Clinic, where postoperative radiographs are performed after body-cavity surgeries even if the surgical count is normal, 62% (21/34) of retained surgical instruments were discovered after surgical closure despite a correct surgical count [2]. Even if the surgical count is discrepant, radiographs have a low real-world sensitivity for retained surgical instruments (≈ 50–90%) [2, 9, 11]. This is because, even though retained surgical instruments often have features visible by radiography, such as barium strips or metal filaments, the embedded materials are not standardized across vendors, and their shape changes within the body, creating a perceptual challenge for the radiologist. In addition, postoperative radiographs performed for retained surgical instrument detection are infrequently performed at most centers and the frequency of retained surgical instruments is very low, creating an environment of insufficient exposure for interpreting radiologists.

Given the shortcomings of both the surgical count and the postoperative radiographs, we developed a high-specificity computer-aided detection (CAD) algorithm to be used in conjunction with radiologists for the detection of an embedded radiopaque μTag (Kalyspo). We hoped that this process could improve on the challenges that exist with the standard algorithm. The purpose of this study was to evaluate the diagnostic accuracy of a process incorporating CAD for the detection and prevention of retained surgical instruments using a novel nondeformable radiopaque μTag.

Materials and Methods

Approval was obtained from the University of Michigan institutional review board and the institutional Anatomic Donations Program for this HIPAA-compliant prospective randomized interrater agreement and diagnostic accuracy study. Because of the use of cadaver subjects, the requirement for informed consent was waived by the institutional review board.

Study Population

A high-specificity CAD system was developed iteratively from a training set (n = 540 radiographs) and a validation set (n = 560 radiographs) using established methods [12, 13]. A novel test set composed of 700 thoracoabdominal radiographs obtained using standard diagnostic technique (410 with a randomly placed μTag and 290 without a μTag) was obtained from 10 cadavers embedded with confounding iatrogenic objects. Data were analyzed first by the blinded CAD system. The CAD system had not been exposed previously to the test set. Radiographs coded as negative by the CAD system (n = 373; 289 true-negatives plus 84 false-negatives) were then independently reviewed by five blinded radiologists. The reference standard was the presence or absence of a μTag.

This study design was chosen to mimic a hypothetical future state in which a high-specificity CAD system could be used initially in the operating room without immediate radiologist input (i.e., if the high-specificity CAD system identified a μTag, then the surgeon would look for a retained surgical instrument without needing to speak with a radiologist; this would be possible because a high-specificity CAD system would have rare false-positives). High-specificity CAD was used instead of high-sensitivity CAD to mitigate the number of patients who might undergo on-the-table exploration for a false-positive CAD result. All negative results in the hypothetical future state would be screened by a radiologist to identify any false-negatives that were not detected by the high-specificity system. Although the CAD system is capable of detecting multiple μTags, for the purposes of this study, each cadaver had only one or zero μTags embedded. The study population flowchart is shown in Figure 1.

Confounding Iatrogenic Objects

Preliminary data were obtained to determine the frequency of iatrogenic objects that are typically visible on an intraoperative radiograph of retained surgical instruments and that might confound identification of a radiopaque μTag. Intraoperative radiographs (n = 104) obtained at the study institution for evaluation of retained surgical instruments in 104 subjects unrelated to the diagnostic accuracy study were identified by query of the institutional electronic medical record system. The number and type of iatrogenic objects visible on the radiographs was tabulated, and frequency data were recorded (Table 1). Frequency data from this cohort were used to populate the cadavers with a similar prevalence of iatrogenic objects.

Cadavers

Ten adult cadavers with a mean age at death of 85.1 years and a mean body mass index of 21.4 kg/m^2 were used as test subjects. All of the subjects had extraluminal air visible on the radiographs. The positioning of each cadaver was changed by randomly translating (up to 15 cm), rotating (up to 10°), and tilting (up to 10°) the cadaver with respect to the imaging detector and table, to increase the variability of the anatomic background of the radiographs. The cadavers were maintained in a generally supine position for all radiographs. To blind the investigator placing the μTag, the μTag was encapsulated in an opaque sphere that obscured its

TABLE 1: Frequency of Iatrogenic Objects Visible on 104 Radiographs of 104 Subjects Performed for Evaluation of a Retained Surgical Instrument

Device or Instrument	Percentage (No.) of Objects
Percutaneous drain	49 (51)
Nasogastric tube	39 (41)
Surgical clip	35 (36)
Surgical staple	30 (31)
Urinary catheter	27 (28)
Plastic endobiliary stent	15 (16)
Intravascular catheter	11 (11)
Surgical needle	9 (9)
Surgical clamp	9 (9)
Surgical sponge	8 (8)
Endotracheal tube	7 (7)
Nasoenteric feeding tube	7 (7)
ECG lead	3 (3)
Gastrojejunostomy catheter	3 (3)
Intraluminal barium	3 (3)
Surgical retractor	2 (2)
Orthopedic hardware	2 (2)
Metal endobiliary stent	2 (2)
Sternotomy wires	2 (2)
Inferior vena cava filter	1 (1)

Note—These data were used to populate the cadavers with common confounding surgical material.

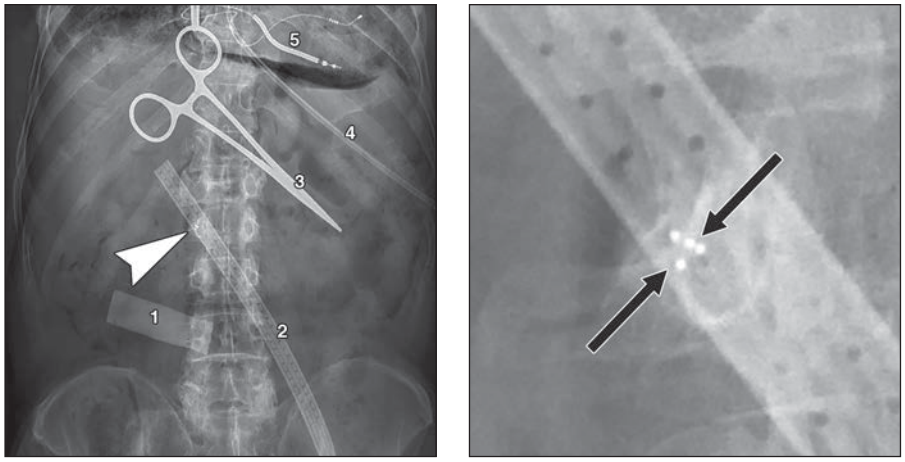

Fig. 2—Example of μTag (Kalyspo) projected within cadaver abdomen.
A and **B,** Abdominal radiograph (**A**) and magnified view (**B**) show that μTag (*arrowhead*, **A**; *arrows*, **B**) is approximately 3 mm^3, contains four radiopaque microbeads, and projects over right L2 vertebral body. There is overlapping surgical drain (diagonal white stripe with parallel circular lucencies; 2, **A**). Small size of μTag enables it to be embedded within variety of surgical objects, including radiopaque tape in surgical sponge (1, **A**), surgical drain (2, **A**), scissors (3, **A**), overlying catheter (4, **A**), and intracardiac leads (5, **A**).

orientation; this enabled random orientation of the μTag during placement. A 4 × 5 (20 × 25 cm) cellophane grid with 20 coordinates was superimposed over the cadaver thorax and abdomen. The μTag was placed on the grid using prespecified randomly allocated coordinates. A maximum of one μTag was placed per radiograph. In this manner, we ensured random orientation and random location of the μTag over the cadaver anatomy (Fig. 2). Each μTag was placed over the cadaver abdomen on the side farthest from the x-ray detector to create the greatest degree of geometric blurring and scatter degradation. This was done to simulate a worst-case scenario in terms of distance of the μTag from the detector, depth of intervening tissue, and number of overlapping objects. A total of 700 thoracoabdominal radiographs were obtained (410 with a randomly placed μTag and 290 without a μTag).

Iatrogenic objects with a frequency of 5% or more on the preliminary retained surgical instruments radiographs (Table 1) were embedded into the cadavers with the same frequency. Objects with an expected anatomic position (e.g., urinary catheter or endobiliary stent) were placed into anatomic position. The set of objects, the position of the objects, and the orientation of the objects were otherwise randomly configured for every image. The placement of these objects was done to mimic the clinical environment of an intraoperative radiograph.

To simulate intraoperative portable radiography technique, the cadavers were placed in a generally supine position and an 8-megapixel flat-panel detector was positioned behind their back. Radiographs were acquired using standard technique with a portable x-ray machine used for intraoperative imaging at the study institution. Radiographs were transmitted and stored at the research PACS and subsequently were downloaded to the CAD laboratory servers.

μTag

The 3D μTag used in this study has been described in detail elsewhere [14]. The μTag is approximately 3 mm^3 and contains four high-contrast microspheres arranged at the vertices of a tetrahedron. The design of the μTag has been optimized through experimentation on phantom and cadaver image datasets unrelated to this study. The 3D geometry of the μTag facilitates distinction from naturally occurring high-contrast anatomic features, such as bone edges of the spine, femurs, and pelvis. Unlike radiopaque markers commonly embedded in surgical sponges, the μTag is nondeformable and does not change shape in the human body. The result is that, regardless of its orientation in space, the μTag produces a 2D radiographic projection that can be recognized by a CAD system or radiologist.

CAD System

The CAD system used in this study has been described in detail elsewhere [12]. Briefly, it uses the following steps to detect the μTag: image preprocessing, μTag candidate labeling, segmentation, feature analysis, classification, false-positive reduction, and an artificial neural network for μTag detection. This is an automated process and does not require specialized user training. The CAD system used for this study was developed iteratively from a training set (n = 540 radiographs) and a validation set (n = 560 radiographs) and was designed to achieve a high specificity (> 99%) for intraoperative utilization. The high-specificity mode in the validation set had a sensitivity of 80.1% with an average of 0.004 false-positives per radiograph (i.e., specificity, 99.6%) [12]. The CAD system was not trained on the test set and was exposed to the test set only a single time.

Radiologist Review

Five board-certified radiologists fellowship-trained in cardiothoracic or abdominal radiology were familiarized with the display workstation, an in-house graphical user interface, and the appearance of the μTag on 10 training radiographs that were not part of the test set. The radiologists were then shown the 373 test images that the CAD system had scored as negative (289 true-negatives and 84 false-negatives). The review was conducted by each radiologist independently in two to five sessions. Each radiologist was blinded to the μTag

TABLE 2: Diagnostic Accuracy of the High-Specificity Computer-Aided Detection (CAD) System and Independent Radiologists in Evaluation of the Test Set

Variable	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
CAD system (*n* = 700 radiographs)	79.5 (326/410)	99.7 (289/290)	99.7 (326/327)	77.5 (289/373)
Radiologists (*n* = 373 radiographs)				
Radiologist 1	95.2 (80/84)	99.3 (287/289)	97.6 (80/82)	98.6 (287/291)
Radiologist 2	92.9 (78/84)	100 (289/289)	100 (78/78)	98.0 (289/295)
Radiologist 3	100 (84/84)	100 (289/289)	100 (84/84)	100 (289/289)
Radiologist 4	96.4 (81/84)	100 (289/289)	100 (81/81)	99.0 (289/292)
Radiologist 5	98.8 (83/84)	100 (289/289)	100 (83/83)	99.7 (289/290)
CAD plus radiologists (*n* = 700 radiographs)				
CAD plus radiologist 1	99.0 (406/410) [98.1–100]	99.0 (287/290) [97.8–100]	99.3 (406/409)	98.6 (287/291)
CAD plus radiologist 2	98.5 (404/410) [97.4–99.7]	99.7 (289/290) [99.0–100]	99.8 (404/405)	98.0 (289/295)
CAD plus radiologist 3	100 (410/410) [100–100]	99.7 (289/290) [99.0–100]	99.8 (410/411)	100 (289/289)
CAD plus radiologist 4	99.3 (407/410) [98.4–100]	99.7 (289/290) [99.0–100]	99.8 (407/408)	99.0 (289/292)
CAD plus radiologist 5	99.8 (409/410) [99.3–100]	99.7 (289/290) [99.0–100]	99.8 (409/410)	99.7 (289/290)

Note—Data are percentage (number/total) [95% CI]. The test set included 700 radiographs, 410 with μTag (Kalyspo) and 290 without μTag. The high-specificity CAD system reviewed all radiographs. The radiologists only reviewed radiographs that were determined by the high-specificity CAD system to be negative (*n* = 373 radiographs; 84 with μTag and 289 without μTag).

frequency and the potential number of μTag*s* per radiograph. The radiographs were shown in random order. Radiologists marked the suspected location of a μTag by placing a digital mark on the radiograph using the in-house graphical interface.

Sample Size Estimation

Sample size estimation was performed according to methods described by Eng [15]. To show 99% sensitivity with a 95% CI of ± 1% for the combination of the high-specificity CAD system plus one failsafe radiologist, 380 radiographs with a μTag would be required; 410 radiographs with a μTag were used in our study. To show 99.5% specificity with a 95% CI of ± 1% for the combination of the high-specificity CAD system plus one failsafe radiologist, 191 radiographs without a μTag would be required; 290 radiographs without a μTag were used in our study.

Reference Standard

The reference standard was the presence or absence of a μTag. This information was recorded by the individual who placed the μTag on the cadavers. This individual was not involved in the image review. The location of each μTag was marked during study design with a 3 × 3 mm digital square based on the known location of the μTag. The digital square marking the location of the μTag was not displayed during image review. Radiologist and CAD system review was performed blinded to the reference standard. A true-positive was recorded if the CAD system or radiologist marked the radiograph with a digital marker within 3 mm of the centroid of the reference square marking the location of a μTag. A false-negative was recorded if the CAD system or radiologist did not place a mark, or if a mark was not placed within 3 mm of the centroid of the reference square.

Statistical Analysis

Sensitivity, specificity, positive predictive value, and negative predictive value were calculated for the CAD system alone, each radiologist alone, and the combination of the CAD system with each radiologist. Interrater agreement was assessed with McNemar test and Cohen kappa scores. Kappa scores were interpreted as follows: 0.01–0.20 (slight agreement), 0.21–0.40 (fair agreement), 0.41–0.60 (moderate agreement), 0.61–0.80 (substantial agreement), and 0.81–0.99 (almost perfect agreement) [16]. Image analysis times for each radiologist and the CAD system were measured and are expressed as mean (± SD).

Results

The high-specificity CAD system had a sensitivity of 79.5% (326/410) and a specificity of 99.7% (289/290) for the previously unseen test set (Table 2). This indicates that 326 radiographs were correctly identified by the CAD system as containing a retained surgical instrument with only one false-positive.

Radiologists reviewing the radiographs determined to be negative (*n* = 373; 289 true-negatives and 84 false-negatives) by the high-specificity CAD system had a sensitivity of 92.9–100% and a specificity of 99.3–100% (Table 2). The combination of the high-specificity CAD system plus one failsafe radiologist had a sensitivity of 98.5% (404/410) to 100% (410/410) and specificity of 99.0% (287/290) to 99.7% (289/290), with 327 (47%) radiographs not requiring immediate radiologist review (Table 2). Pairwise interradiologist agreement was almost perfect (κ = 0.921–0.992 for radiologists alone; κ = 0.968–0.994 for radiologists plus the high-specificity CAD system) (Table 3). Pairwise comparisons of diagnostic accuracy for each radiologist-CAD couplet were not significantly different (McNemar p = 0.13–0.99), indicating statistically indistinguishable diagnostic accuracy.

Mean image analysis time was less than one minute (CAD, 29 ± 2 seconds per radiograph; radiologists, 26 ± 16 seconds per radiograph). Mean image review time by radiologist ranged from 20 ± 15 to 35 ± 12 seconds per radiograph.

Discussion

Retained surgical instruments are considered sentinel events by The Joint Commission [2] and are always wrong according to the Leapfrog group [17]. Unfortunately, in spite of wide recognition, clinical efforts, and technologic advancements, retained surgical instruments continue to occur. Advanced technologies, such as barcodes [18] and radiofrequency identification chips [19, 20], have incomplete penetrance into operating rooms in the United States, possibly because of in-

TABLE 3: Interrater Agreement Expressed With Cohen Kappa Values

Variable	Interrater Agreement				
	Radiologist 1	Radiologist 2	Radiologist 3	Radiologist 4	Radiologist 5
Radiologists (n = 373 radiographs)					
Radiologist 1	—	0.921	0.954	0.945	0.961
Radiologist 2	—	—	0.953	0.976	0.960
Radiologist 3	—	—	—	0.977	0.992
Radiologist 4	—	—	—	—	0.984
Radiologist 5	—	—	—	—	—
CAD plus radiologists (n = 700 radiographs)	CAD Plus Radiologist 1	CAD Plus Radiologist 2	CAD Plus Radiologist 3	CAD Plus Radiologist 4	CAD Plus Radiologist 5
CAD plus radiologist 1	—	0.968	0.980	0.977	0.983
CAD plus radiologist 2	—	—	0.980	0.988	0.983
CAD plus radiologist 3	—	—	—	0.988	0.994
CAD plus radiologist 4	—	—	—	—	0.991
CAD plus radiologist 5	—	—	—	—	—

Note—The high-specificity CAD system reviewed all radiographs (n = 700 radiographs; 410 with μTag [Kalyspo], 290 without μTag). Radiologists only reviewed radiographs that were determined by the high-specificity CAD system to be negative (n = 373 radiographs; 84 with μTag, 289 without μTag). Dashes (—) indicate that no comparison was made.

complete effectiveness, cost, and difficulties integrating into existing operating room environments [18, 20, 21]. The approach proposed in this study—using a high-specificity CAD system coupled with failsafe radiologist review to detect a novel nondeformable μTag—achieves outstanding diagnostic accuracy in an efficient manner with the potential to integrate seamlessly into existing retained surgical instruments algorithms. In a hypothetical future state, if the CAD system identifies a radiopaque μTag while operating in high-specificity mode, a surgeon could immediately remove that object without waiting for a radiologist to identify it. The μTag is sufficiently small to be embedded in a wide variety of iatrogenic objects and can be detected reliably despite confounding instruments in the surgical field. These results suggest that the proposed combined CAD-radiologist algorithm may help dramatically reduce the risk and occurrence of retained surgical instruments.

The expected cost to add one μTag to one iatrogenic object (≈ $10) is less than the expected cost of a radiofrequency identification chip ($75) [9] and similar to the expected cost of a barcode ($9) [9]. It has been estimated that bar-coding prevents 97.5% of cases of retained surgical instruments at a cost of $95,000/retained surgical instrument and that radiofrequency identification chips may prevent a similar number of retained surgical instruments at a greater cost of $720,000/retained surgical instrument [9]. Formal cost-effectiveness and integration studies would be needed to evaluate these questions with respect to the μTag. However, because the proposed approach makes use of an existing hospital system for retained surgical instrument detection (i.e., postoperative radiography), does not require a proprietary barcode or radiofrequency chip reader, and incorporates CAD software that can be loaded on any standard computer (hypothetical estimated onetime cost of roughly $20,000, with hypothetical estimated annual maintenance cost of $1000), the cost may be less and adoption may be easier than for these other advanced technologies.

The high-specificity CAD system we developed had superior specificity (99.7% [289/290]) with a single false-positive and moderate sensitivity (79.5% [326/410]) on a novel test set with randomly positioned common confounding iatrogenic objects. The superior specificity of this approach allows integration of the CAD system into the operating room workflow in a hypothetical future state in which positive cases identified by the CAD system would invoke immediate search for a retained surgical instrument without radiologist involvement. All negative cases would prompt radiologist review. This approach has the potential to eliminate surgeon delay in positive cases waiting for a radiologist to respond and review the radiographs [18] and would minimize the effect of radiologist inexperience. When attending radiologists independently reviewed the radiographs designated by the high-specificity CAD system to be negative, they had almost perfect agreement (κ = 0.968–0.994 with the CAD system) and superior sensitivity (98.5–100% with the CAD system) with few false-positives (n = 1–3). These few false-positives and false-negatives may be minimized or eliminated with assistance provided by a high-sensitivity CAD system [12] during radiologist review, but we did not test that approach in this study. Both the CAD review and the radiologist review were completed individually and in combination, on average, in less than 1 minute.

There were several limitations of our study. Although we conducted a prospective randomized diagnostic accuracy study, the subjects were a small (n = 9) number of cadavers. It is possible that the radiologists and the CAD system overperformed because they were able to recall the shape and general habitus of the small number of cadavers. To mitigate that risk, we randomly incorporated confounding iatrogenic objects according to their frequency in radiographs of retained surgical instruments, randomly altered the position and occurrence of the μTag, and randomly altered the position of the cadaver on the table. Although it would not be predicted to affect the CAD performance, radiologists may have overperformed on their task because they were attending to the problem of retained surgical instruments over a period of two to five condensed read-

ing sessions. Future investigations may consider incorporating study-related radiographs of retained surgical instruments into routine clinical workflow to mitigate that possibility. We expect that the μTag would be inexpensive to implement and would not require major changes to existing operating room systems for detection of retained surgical instruments, but we did not explicitly test those assumptions in this diagnostic accuracy study. Finally, it is possible that multiple retained surgical instruments can occur simultaneously. Therefore, testing the CAD system and radiologists in the detection of a range of potentially embedded μTags is important.

In conclusion, the combination of a high-specificity CAD system with a failsafe radiologist had excellent diagnostic accuracy in the rapid detection of a nondeformable radiopaque μTag. The superior sensitivity, specificity, and reproducibility of the algorithm have the potential to dramatically reduce the frequency of retained surgical instruments and to do so in a low-cost efficient fashion. Future research is warranted in the clinical setting to investigate the effectiveness and cost of this strategy and the performance of the CAD system in patients with multiple retained surgical instruments. Such research would ideally combine the advantages of a high-specificity CAD system (few false-positives) with those of a high-sensitivity CAD system (few false-negatives) and further clarify the optimal role of the supervising radiologist.

References

1. National Quality Forum. *Serious reportable events in healthcare: a consensus report*. Washington, DC: National Quality Forum, 2002
2. Cima RR, Kollengode A, Garnatz J, Storsveen A, Weisbrod C, Deschamps C. Incidence and characteristics of potential and actual retained foreign object events in surgical patients. *J Am Coll Surg* 2008; 207:80–87
3. Pennsylvania Patient Safety Authority. Use of x-rays for incorrect needle counts. Pennsylvania Patient Safety Reporting System website. patientsafety.pa.gov/ADVISORIES/documents/200406_05.pdf. Published 2004. Accessed December 7, 2017
4. Barrow CJ. Use of x-ray in the presence of an incorrect needle count. *AORN J* 2001; 74:80–81
5. Gawande AA, Studdert DM, Orav EJ, Brennan TA, Zinner MJ. Risk factors for retained instruments and sponges after surgery. *N Engl J Med* 2003; 348:229–235
6. Gibbs VC, Coakley FD, Reines HD. Preventable errors in the operating room: retained foreign bodies after surgery. Part 1. *Curr Probl Surg* 2007; 44:281–337
7. [No authors listed.] Study offers evidence on items left behind during surgical cases. *OR Manager* 2003; 19:10–11
8. Kennedy J. Prevention of retained foreign objects. Western States Healthcare Materials Management Association website. www.wshmma.org/events/2010/Prevention_of_Retained_Foreign_Objects.pdf. Published 2008. Accessed April 4, 2013
9. Regenbogen SE, Greenberg CC, Resch SC, et al. Prevention of retained surgical sponges: a decision-analytic model predicting relative cost-effectiveness. *Surgery* 2009; 145:527–535
10. Centers for Medicare & Medicaid Services (CMS). CMS proposes additions to list of hospital-acquired conditions for fiscal year 2009. CMS website. www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2008-Fact-sheets-items/2008-04-14.html. Published 2008. Accessed March 30, 2013
11. Stawicki SP, Moffatt-Bruce SD, Ahmed HM, et al. Retained surgical items: a problem yet to be solved. *J Am Coll Surg* 2013; 216:15–22
12. Hadjiiski L, Marentis TC, Chaudhury AR, Rondon L, Chronis N, Chan HP. Computer aided detection of surgical retained foreign object for prevention. *Med Phys* 2015; 42:1213–1222
13. Marentis TC, Hadjiiyski L, Chaudhury AR, Rondon L, Chronis N, Chan HP. Surgical retained foreign object (RFO) prevention by computer aided detection (CAD). In: Aylward S, Hadjiiski L, eds. *Medical imaging 2014: computer-aided diagnosis—proceedings of SPIE*, vol. 9035. Bellingham, WA: SPIE, 2014:903529
14. Tripathi A, Marentis T, Chronis N. Microfabricated instrument tag for the radiographic detection of retained foreign bodies during surgery. In: Pelc NJ, Nishikawa RN, Whiting BR, eds. *Medical imaging 2012: physics of medical imaging—proceedings of SPIE*, vol. 8313. Bellingham, WA: SPIE, 2012:83134H
15. Eng J. Sample size estimation: how many individuals should be studied? *Radiology* 2003; 227:309–313
16. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005; 37:360–363
17. Copeland AW. Retained surgical sponge (gossypiboma) and other retained surgical foreign bodies: prevention and management. UpToDate website. www.uptodate.com/contents/retained-surgical-sponge-gossypiboma-and-other-retained-surgical-items-prevention-and-management. Accessed June 6, 2012
18. Greenberg CC, Diaz-Flores R, Lipsitz SR, et al. Bar coding surgical sponges to improve safety: a randomized controlled trial. *Ann Surg* 2008; 247:612–616
19. [No authors listed.] Radio-frequency surgical sponge detection: a new way to lower the odds of leaving sponges (and similar items) in patients. *Health Devices* 2008; 37:193–202
20. Macario A, Morris D, Morris S. Initial clinical evaluation of a handheld device for detecting retained surgical gauze sponges using radiofrequency identification technology. *Arch Surg* 2006; 141:659–662
21. Rogers A, Jones E, Oleynikov D. Radio-frequency identification (RFID) applied to surgical sponges. *Surg Endosc* 2007; 21:1235–1237