HHS Public Access

Author manuscript

Acad Radiol. Author manuscript; available in PMC 2017 October 01.

Published in final edited form as:

Acad Radiol. 2016 October ; 23(10): 1221–1229. doi:10.1016/j.acra.2016.05.009.

Validation of a Projection-domain Insertion of Liver Lesions into CT Images

Baiyu Chen, PhD, **Chi Ma, PhD**, **Shuai Leng, PhD**, **Jeff L. Fidler, MD**, **Shannon P. Sheedy, MD**, **Cynthia H. McCollough, PhD**, **Joel G. Fletcher, MD**, and **Lifeng Yu, PhD**

Department of Radiology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905

Abstract

Rationale and Objectives—The aim of this study was to validate a projection-domain lesion-insertion method with observer studies.

Materials and Methods—A total of 51 proven liver lesions were segmented from computed tomography images, forward projected, and inserted into patient projection data. The images containing inserted and real lesions were then reconstructed and examined in consensus by two radiologists. First, 102 lesions (51 original, 51 inserted) were viewed in a randomized, blinded fashion and scored from 1 (absolutely inserted) to 10 (absolutely real). Statistical tests were performed to compare the scores for inserted and real lesions. Subsequently, a two-alternative-forced-choice test was conducted, with lesions viewed in pairs (real vs. inserted) in a blinded fashion. The radiologists selected the inserted lesion and provided a confidence level of 1 (no confidence) to 5 (completely certain). The number of lesion pairs that were incorrectly classified was calculated.

Results—The scores for inserted and proven lesions had the same median (8) and similar interquartile ranges (inserted, 5.5–8; real, 6.5–8). The means scores were not significantly different between real and inserted lesions (P value = 0.17). The receiver operating characteristic curve was nearly diagonal, with an area under the curve of 0.58 ± 0.06. For the two-alternative-forced-choice study, the inserted lesions were incorrectly identified in 49% (25 out of 51) of pairs; radiologists were incorrect in 38% (3 out of 8) of pairs even when they felt very confident in identifying the inserted lesion (confidence level ≥4).

Conclusions—Radiologists could not distinguish between inserted and real lesions, thereby validating the lesion-insertion technique, which may be useful for conducting virtual clinical trials to optimize image quality and radiation dose.

Keywords

Computed tomography (CT); image quality assessment; lesion insertion; projection domain; virtual clinical trials

Address correspondence to: L.Y. Yu.Lifeng@mayo.edu.

INTRODUCTION

To optimize computed tomography (CT) image quality and radiation dose for liver lesion detection tasks, patient images containing proven liver lesions are required. Proof of lesion presence and etiology may be obtained from biopsy, surgical extirpation, or regression or progression of hepatic disease on cross-sectional imaging in patients with known hepatic malignancy. Although images containing proven liver lesions can be collected via clinical trials, the process is time-consuming and expensive. An alternative to conventional clinical trials is a virtual clinical trial, which acquires images by inserting lesions into patient images at specified locations. With virtual clinical trials, the image data collection process becomes substantially more time-efficient and less costly. Moreover, creating images containing inserted lesions would permit control of lesion characteristics and locations.

The pathway toward a virtual clinical trial in low-dose liver CT would necessitate several milestones, including (1) the ability to insert lesions into designated locations to ensure that lesions obey anatomic boundaries, (2) that inserted and actual proven liver lesions appear indistinguishable to experienced radiologists, and (3) that lesion detection and characterization for the inserted and real lesions are similar over a range of acquisition and reconstruction conditions. A lesion-insertion method has been recently developed (1), which inserts lesions via the projection domain (ie, before the image reconstruction) and is compatible with a state-of-the-art commercial CT scanner in both axial and helical modes, with various tube potential settings and focal spot movement patterns. The projection-domain insertion method is more sophisticated than conventional image domain (ie, after the image reconstruction) lesion-insertion methods (2–7) because the resulting inserted lesions reflect the impact of reconstruction method and parameters on lesion appearance, which is critical to the evaluation of lower dose images with iterative reconstruction.

The aforementioned projection-domain lesion-insertion method has been previously validated in terms of CT number accuracy and high contrast spatial resolution (1). The lesion-insertion method must, however, be shown to perform satisfactorily in a clinical task by determining if inserted and real liver lesions can be distinguished by experienced radiologists. This study aims to validate our projection-domain lesion-insertion method with observer studies.

MATERIALS AND METHODS

Projection-domain Lesion-insertion Program

The recently developed projection-domain lesion-insertion program (1) is summarized in Figure 1. First, patient CT raw data acquired on a commercial scanner (Somatom Definition Flash, Siemens Healthcare, Forchheim, Germany) were decoded with the help of the vendor to acquire two types of information: the patient projections (sinogram), which recorded the attenuation information about the patient; and the CT acquisition parameters, which described the X-ray spectrum, gantry geometry, gantry rotation, and table movement. The patient raw data were also reconstructed on the commercial scanner to acquire patient CT images, which were used to visually identify a lesion-insertion location. Next, a voxelized liver lesion (previously segmented from patient CT images where voxel intensity represents

CT numbers in Hounsfield unit) was forward projected to simulate lesion projections using a program written in MATLAB (version R2013b, The Mathworks Inc., Natick, MA). The forward projection program mathematically simulated the CT acquisition process with inputs of CT acquisition parameters and the desired lesion-insertion location, such that the lesion projections were similar to what would be acquired by physically scanning the lesion at the given location using the given acquisition parameters (8). The lesion projections were subsequently inserted into the patient projections. The modified patient projection data were then reconstructed on a commercial CT scanner to generate the CT image datasets with simulated lesions. The computation time of the lesion-insertion program varied as a function of lesion size and CT acquisition parameters, but was always below 10 minutes on a computer with an Intel Xeon X5690 3.46 GHz 6 Core processor and 96 GB RAM. This is much shorter than the time needed to retrospectively collect cases in conventional clinical trials.

The image quality aspects of the insertion technique, such as CT number accuracy and spatial resolution, have been validated in a previous study (1). In this study, the validation focused on the realism of lesions inserted into patient images, as perceived by radiologists.

Patient and Lesion Database

To prepare for lesion insertions, a database of 30 patients who underwent contrast-enhanced abdominal CT was created. The patients were retrospectively selected, and all had proven benign or malignant liver lesions. A total of 51 liver lesions were selected from the 30 patients (one to three lesions per patient) to cover a wide range of lesion sizes, contrasts, shapes, and pathologies. Details of the patient and lesion characteristics are provided in Appendix A. All patient images were acquired on commercial CT scanners (Somatom Definition AS and Somatom Definition Flash, Siemens Healthcare) at clinical dose levels with tube current modulation and a z-flying focal spot technique (9), and reconstructed using filtered back projection (10). The selection of tube voltage, spiral pitch, reconstruction slice thickness, and slice interval varied from patient to patient, as provided in Appendix B. Each lesion was segmented from the images using Seg3D (version 2.1.5, University of Utah, Salt Lake City, UT), where the lesion boundaries were manually drawn image-by-image under the instruction of a radiologist (with 24 years of experience in abdominal CT). If the lesion had a fading boundary, the boundary was drawn generously to include the entire lesion.

Subsequently, the lesion projections, created by forward projecting the segmented lesions, were inserted back into the patient projections at a new location to yield 51 inserted lesions. The new locations, which are specified by the radiologist who oversaw the lesion segmentation, were chosen to emulate the location and appearance of liver lesions and to avoid vascular structures. Nonspherical lesions with flat borders secondary to the liver capsule were placed at similar peripheral locations in the liver so that the orientation of the flat edge of the lesion was similar to the adjacent capsule. The realism of the inserted lesions was assessed by two observer studies: a randomized likelihood study and a two-alternative-forced-choice (2AFC) study.

Observer Study 1: Likelihood Scores

A total of 102 lesions (51 proven and 51 inserted) were randomized and viewed on a dual-monitor workstation (Advantage Workstation, GE Healthcare, Waukesha, WI) with controlled lighting. Two radiologists (13 and 17 years of experience in abdominal CT) blinded to all clinical data rated the realism of each lesion in consensus fashion, using a likelihood score to reflect the possibility of the lesion being an inserted one (from 1 [absolutely inserted] to 10 [absolutely real]). A detailed explanation of scores is provided in Appendix C. Because this study was a discrimination task rather than a detection task, the lesion location was provided to the observers. The observers reviewed and discussed each case before rating the lesion, with no set time limit. The lesions were reviewed in the context of the entire liver. The observers were allowed to scroll through the respective lesions, but not through the entire series to avoid potential distractions from other lesions. The default display window/level setting was an abdominal window (window center = 40, window width = 400), but the observers were allowed to adjust the window/level settings.

Observer Study 2: Two-alternative-forced-choice Study

The second observer study was performed 2 hours after the first one. The reading room, workstation, and observers remained the same. A 2AFC test was conducted, which is an established controlled measure of observers' ability to distinguish between two features (11,12). The two alternatives consisted of a real lesion and an inserted lesion, and the inserted lesion was created by inserting the real lesion back into the same patient at a different location. The image volume that contained both real and inserted lesions was loaded onto two monitors, side-by-side. On one monitor (randomly selected and blinded to the radiologists), the central slice through the real lesion was displayed, and on the other monitor, the central slice through the inserted lesion was displayed. The radiologists were informed of the lesion locations and were allowed to scroll through both lesions and adjust the window/level settings. The observers then chose, in consensus, the inserted lesion and provide a reason for the decision. For each choice, the observers were also asked to provide a confidence level from 1 (no confidence) to 5 (completely certain). A detailed explanation of confidence levels is provided in Appendix C.

Statistical Analysis

For the likelihood scoring study, the scores of real and inserted lesions were compared in three ways. First, descriptive statistics were calculated, including the number of lesions with likelihood scores of ≤3 (ie, probably inserted lesions) or ≥7 (ie, probably proven, real lesions) and the median and interquartile range of the scores. Second, a Wilcoxon signed-rank test was performed with MATLAB (version R2013b, The Mathworks Inc., Natick, MA) at a significance level of 0.05 to examine the null hypothesis that the mean likelihood scores for real and inserted lesions were not different. Third, a receiver operating characteristic (ROC) curve was computed for the scores. The ROC curve characterized how well the readers can differentiate the inserted lesions from real lesions. For example, if the ROC curve is a diagonal line (area under the ROC curve [AUC] is 0.5), it means that the reader cannot differentiate between real and inserted lesions at all. The ROC curve was calculated using an established ROC calculator (13), which was a direct translation of the ROCFIT

program developed by Metz and Kronman (14). The ROC calculated estimated the maximum likelihood of a binomial ROC curve from categorical rating data and reported a fitted ROC curve along with its 95% confidence interval and the AUC. Even though multiple lesions might have originated from the same patient, because they were randomized within and across patients, we treated all lesions as independent cases for the ROC analysis.

For the 2AFC study, the percentage of incorrectly identified lesion pairs and the distribution of the confidence levels were calculated. Histograms of the confidence levels were calculated for correctly and incorrectly identified lesion pairs. The percentage of incorrectly identified lesions was also calculated for lesion pairs that received a confidence level of ≥4, ie, pairs for which the radiologists were confident that they had correctly identified the inserted lesion.

This is a retrospective study approved by our institutional review board. It complied with the Health Insurance Portability and Accountability Act and did not require informed consent.

RESULTS

Observer Study 1: Likelihood Scores

A total of 35 inserted and 38 real lesions were scored ≥7 (ie, probably real), whereas 6 inserted and 1 real lesions were scored ≤3 (ie, probably inserted). The median likelihood score was 8 for both real and inserted lesions. The interquartile ranges were similar for real and inserted lesions (real, 6.5–8; inserted, 5.5–8). The Wilcoxon signed-rank test of the likelihood scores showed that the means of the scores were not significantly different between real lesions and inserted lesions (P value = 0.17). Figure 2 shows the fitted ROC curve along with its 95% confidence interval and the points making up the empirical ROC curve. The AUC of the curve was 0.58 with a standard deviation of 0.06.

Figure 3 shows examples of inserted lesions that received a score of ≥7 (ie, probably real), including hypo- and hyper-attenuating lesions, isolated and vessel-attached lesions, invasive and well-circumscribed lesions, homogeneous and heterogeneous lesions, small and large lesions, and high- and low-contrast lesions. Figure 4 shows examples of inserted lesions that received a score of ≤3 (ie, probably inserted). The radiologists commented that these lesions were given low likelihood scores predominantly due to unrealistic relationships with the surrounding structures (vessels and liver boundaries) or uncommon shapes (eg, nonsphericity).

Observer Study 2: Two-alternative-forced-choice Study

The inserted lesions were incorrectly classified in 25 out of 51 lesion pairs (49%). Figure 5 shows the distribution of the confidence levels among correctly and incorrectly classified lesion pairs. Overall, 32 out of 51 pairs (63%) were randomly guessed (confidence level of ≤2). Among lesion pairs that were confidently classified (confidence level of ≥4), three out of eight pairs (38%) were incorrect. Examples of lesion pairs incorrectly classified at a high confidence level are shown in Figure 6.

DISCUSSION

Because the optimization of diagnostic performance and radiation dose is expensive and time-consuming—often due to the expense associated with obtaining a reference standard, virtual clinical trials that use patient CT images with inserted lesions present a desirable alternative. This study used observer studies to validate a projection-domain liver lesion-insertion method, and showed that the method could prospectively simulate clinically realistic lesions for virtual clinical trials.

Our observer study confirmed the realistic appearance of the inserted lesions by showing that experienced radiologists could not differentiate inserted lesions from real ones. The likelihood scores of real and inserted lesions had similar distributions (same median and similar interquartile range), nonsignificantly different means ($P = 0.14$), and almost equal probability of being higher than the other (near-diagonal ROC curve with an AUC of 0.58). Although six inserted lesions (and only one real lesion) were scored toward the unrealistic end of the scoring range, this scoring was largely based on the relationship of the lesion to the surrounding structures rather than the appearance of the lesion itself. This can be largely avoided in future insertions by meticulous selection of lesion-insertion site. In the 2AFC study, half of the choices were incorrectly made regardless of the associated confidence levels (25 out of 51 not considering the confidence level, or 3 out of 8 for a confidence level of ≥4). Because this percent correct was close to that of random guessing, the real and inserted lesions were considered indistinguishable. In addition, the majority of the choices (63%) were made at low confidence levels, which again confirms the realism of the inserted lesions.

Multiple lesion-insertion methods have been developed and validated in previous studies. Li et al (3) developed an image-domain method to insert lung nodules into pediatric CT images, and validated the technique by instructing the observers to rate the appearance of any nodule they could detect (either real or inserted). Compared to Li et al's study, our observer study had lesion locations pointed out to the observers, such that the evaluation focused on the discrimination performance rather than the detection performance. Our observer study was also more challenging because our lesions were larger and more internally complex compared to Li et al's nodules (Li et al's lesions, 2.5–6 mm; our lesions, 5–20 mm), therefore exposing more details to observers for examination. Despite the increased difficulty, our inserted lesions were more indistinguishable from real lesions in terms of an AUC closer to 0.5 from the ROC analysis. Solomon and Samei (4) also developed an image-domain insertion method, which inserted liver lesions, lung lesions, and kidney stones into adult CT images. Compared to Solomon and Samei's study, our validation is stricter because our observers viewed the lesions in the context of the entire liver instead of small regions of interest. Our observers were also allowed to scroll through the lesions to relate the lesions to adjacent structures, instead of viewing a single image. Furthermore, our projection-domain method forced the lesion to undertake the appearance dictated by the reconstruction parameters because the lesions were inserted into CT projection data before the image reconstruction. Our method can therefore be used for the evaluation of reconstruction parameters, which is not possible using Li et al's or Solomon and Samei's image-domain techniques.

Many interesting opportunities may arise with this projection-domain lesion-insertion technique. For example, for clinical trials that evaluate the impact of iterative reconstruction on detection performance, lesions of desired characteristics (the contrast, size, orientation, and location of the lesions can all be selected) can be inserted into existing patient CT raw data and reconstructed with iterative reconstruction to create new cases. Additionally, the lesion-insertion process finished within 10 minutes, therefore saving time and cost as compared to conventional case collection and pathology verification.

This study has several limitations. First, the lesions were inserted into the patients from whom they were segmented, instead of new patients. However, this design was necessary to ensure that the real and inserted lesions were compared in controlled conditions, where noise and background brightness were very similar. Second, only liver lesions were inserted and validated in this study, even though the projection-domain insertion program has the potential to insert any voxelized lesions into patient CT raw data. Ongoing work will test the insertion of lung nodules (into a complex background) (15) and kidney stones (highly attenuating object). Finally, we employed consensus reading in each phase of our study because we wanted both radiologists to be aware of potential shape, textural, or anatomic issues that they or another radiologist might detect to “recognize” the inserted pseudolesions. The use of consensus reading is meant to verify these initial observations and provide impetus for future independent multireader assessment.

To summarize, two observer studies (a likelihood score and a 2AFC study) were conducted to validate a projection-domain liver lesion-insertion method. The results demonstrated the realism of the inserted lesions by showing that radiologist observers could not effectively distinguish between real and inserted liver lesions. Future work is required to ensure that archived, segmented lesions can be inserted into different patients under a variety of conditions.

Acknowledgments

This publication was supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under award numbers U01EB017185 and R01EB017095.

References

1. Chen B, Leng S, Yu L, et al. Lesion insertion in the projection domain: methods and initial results. Med Phys. 2015; 42:7034–7042. [PubMed: 26632058]
2. Madsen MT, Berbaum KS, Ellingson AN, et al. A new software tool for removing, storing, and adding abnormalities to medical images for perception research studies. Acad Radiol. 2006; 13:305–312. [PubMed: 16488842]
3. Li X, Samei E, Delong DM, et al. Three-dimensional simulation of lung nodules for paediatric multidetector array CT. Br J Radiol. 2009; 82:401–411. [PubMed: 19153182]
4. Solomon J, Samei E. A generic framework to simulate realistic lung, liver and renal pathologies in CT imaging. Phys Med Biol. 2014; 59:6637–6657. [PubMed: 25325156]
5. Hoe CL, Samei E, Frush DP, et al. Simulation of liver lesions for pediatric CT. Radiology. 2006; 238:699–705. [PubMed: 16371579]
6. de Sisternes L, Brankov JG, Zysk AM, et al. A computational model to generate simulated three-dimensional breast masses. Med Phys. 2015; 42:1098–1118. [PubMed: 25652522]

7. Karantzavelos K, Shin HO, Jordens S, et al. Development and evaluation of a software tool for the generation of virtual liver lesions in multidetector-row CT datasets. Acad Radiol. 2013; 20:614–620. [PubMed: 23477827]
8. Siddon RL. Fast calculation of the exact radiological path for a three-dimensional CT array. Med Phys. 1985; 12:252–255. [PubMed: 4000088]
9. Flohr TG, Stierstorfer K, Ulzheimer S, et al. Image reconstruction and image quality evaluation for a 64-slice CT scanner with z-flying focal spot. Med Phys. 2005; 32:2536–2547. [PubMed: 16193784]
10. Stierstorfer K, Rauscher A, Boese J, et al. Weighted FBP—a simple approximate 3D FBP algorithm for multislice spiral CT with good dose usage for arbitrary pitch. Phys Med Biol. 2004; 49:2209. [PubMed: 15248573]
11. Lapid E, Ulrich R, Rammsayer T. On estimating the difference limen in duration discrimination tasks: a comparison of the 2AFC and the reminder task. Percept Psychophys. 2008; 70:291–305. [PubMed: 18372750]
12. Burgess AE. Visual perception studies and observer models in medical imaging. Semin Nucl Med. 2011; 41:419–436. [PubMed: 21978445]
13. Eng, J. ROC analysis: web-based calculator for ROC curves. Baltimore: Johns Hopkins University; Available at: http://www.jrocfit.orgupdated March 19, 2014 [Accessed May 11, 2015]
14. Metz C, Kronman H. Statistical significance tests for binormal ROC curves. J Math Psychol. 1980; 22:218–243.
15. [blinded].

APPENDIX A. PATIENT AND LESION CHARACTERISTICS

Table A1

Summary of Patient and Lesion Characteristics. For Lesions that Are Inhomogeneous, Contrast Values Are Provided Separately for the Hypo-attenuating Part and for the Hyper-attenuating Part

Patient Characteristics		Lesion Characteristics		
Patient Number	**Water-equivalent Diameter (cm)**	**Diameter (cm)**	**Contrast (HU)**	**Type**
1	24.6	1.5	−51, 35	Metastasis from leiomyosarcoma
2	31.6	0.6	−95	Cyst
3	29.9	1.0	−30,110	Hemangioma
4	26.8	0.6	−33	Metastasis from carcinoid
5	26.7	1.9	−91,130	Hemangioma
		2.3	−49,111	Hemangioma
6	35.9	2.2	−51	Metastasis from colon
		1.9	−47	Metastasis from colon
7	32.0	1.6	15	Metastasis from rectum
		1.7	14	Metastasis from rectum
		2.2	31	Metastasis from rectum
8	26.0	0.7	−93	Cyst
9	30.2	1.9	−34	Metastasis from melanoma
		0.6	−24	Metastasis from melanoma
10	23.7	1.2	−10, 59	Hemangioma
11	28.0	3.1	−60	Metastasis from colon
12	29.5	0.7	−91	Metastasis from colon
		1.3	−96	Metastasis from colon

Patient Characteristics		Lesion Characteristics		
Patient Number	**Water-equivalent Diameter (cm)**	**Diameter (cm)**	**Contrast (HU)**	**Type**
		0.5	−71	Indeterminate
13	29.0	1.1	−132	Cyst
14	28.8	1.5	−24	Metastasis from pancreas
		1.3	−41	Metastasis from pancreas
		0.7	59	Hemangioma
15	28.7	1.8	−129	Cyst
16	32.2	1.7	−46	Metastasis from melanoma
17	26.2	0.7	−37	Metastasis from pancreas
		0.7	−42	Metastasis from pancreas
		1.8	−43	Metastasis from pancreas
18	23.8	1.4	−61	Metastasis from rectum
		0.5	−67	Metastasis from rectum
19	27.3	1.9	−65	Hemangioma
		0.4	−81	Cyst
20	22.6	1.3	−8	Metastasis from carcinoid
		1.0	−19	Metastasis from carcinoid
21	30.4	1.4	−29	Metastasis from leiomyosarcoma
22	33.5	1.6	−42	Metastasis from colon
23	23.7	1.2	−99	Metastasis from thyroid
		1.0	−68	Metastasis from thyroid
24	32.5	1.8	−18	Metastasis from rectum
		1.9	−10	Metastasis from rectum
25	32.2	0.9	−43	Metastasis from pancreatic neuroendocrine
		1.1	−50	Metastasis from pancreatic neuroendocrine
26	30.2	1.0	−76	Metastasis from colon
		1.1	−78	Metastasis from colon
27	27.2	0.9	−21	Metastasis from ovary
28	35.8	0.7	−65	Cyst
		0.5	−69	Cyst
29	29.3	1.4	−56	Metastasis from leiomyosarcoma
		2.0	−69	Metastasis from leiomyosarcoma
30	29.9	0.6	−73	Cyst
	24.6	1.6	−33	Metastasis from adenoid cystic carcinoma

APPENDIX B. ACQUISITION AND RECONSTRUCTION PARAMETERS

Table B1

Summary of Acquisition and Reconstruction Parameters. The Reconstruction Kernel is B40f (Filtered Backprojection) for all Reconstructions and Therefore Not Listed in the Table

	Acquisition Parameters			Reconstruction Parameters	
Patient Number	**Tube Voltage (kV)**	**$CTDI_{vol}$ (mGy)**	**Pitch**	**Slice Thickness (mm)**	**Slice Interval (mm)**
1	100	7.6	0.8	5	5
2	100	18.4	0.35	2	2
3	100	13.0	0.8	5	5
4	100	11.8	0.8	3	2.5
5	120	11.6	1	3	2
6	120	25.5	0.8	5	5
7	120	16.5	0.8	5	5
8	100	10.0	0.8	5	5
9	100	11.4	0.8	5	5
10	100	7.1	0.6	5	5
11	100	9.9	0.8	5	5
12	120	23.4	0.75	3	2.5
13	100	12.0	0.8	5	5
14	100	16.8	0.6	2	1
15	100	20.1	0.6	2	1
16	120	26.4	0.6	3	2.5
17	100	17.7	0.6	2	1
18	100	5.5	0.8	5	5
19	100	8.4	0.6	5	5
20	100	6.3	0.6	5	5
21	120	16.8	0.7	5	5
22	120	22.5	0.8	3	2.5
23	100	8.0	0.8	5	5
24	120	16.6	0.8	5	5
25	120	29.9	0.35	2	1
26	120	17.5	0.8	5	5
27	120	14.6	0.6	5	5
28	120	24.9	0.6	5	5
29	120	16.8	0.8	5	5
30	120	18.6	0.8	5	5

APPENDIX C. LESION REALISM ASSESSMENT PERFORMED BY RADIOLOGIST READERS

Table C1

The Definition of the Likelihood Scores for Lesion Realism

Score	Definition
1	100% confidence that the lesion is not real
2	Very doubtful that the lesion is real
3–4	Unlikely that the lesion is real
5–6	Very unsure whether the lesion is real or not
7–8	Likely that the lesion is real
9	Very likely that the lesion is real
10	100% confidence that the lesion is real

Table C2

The Definition of the Confidence Levels for the Choices Made in the 2AFC Test

Confidence Level	Definition
1	No confidence in determination
2	Possibly correct
3	Probably correct
4	Likely correct
5	100% confidence in determination

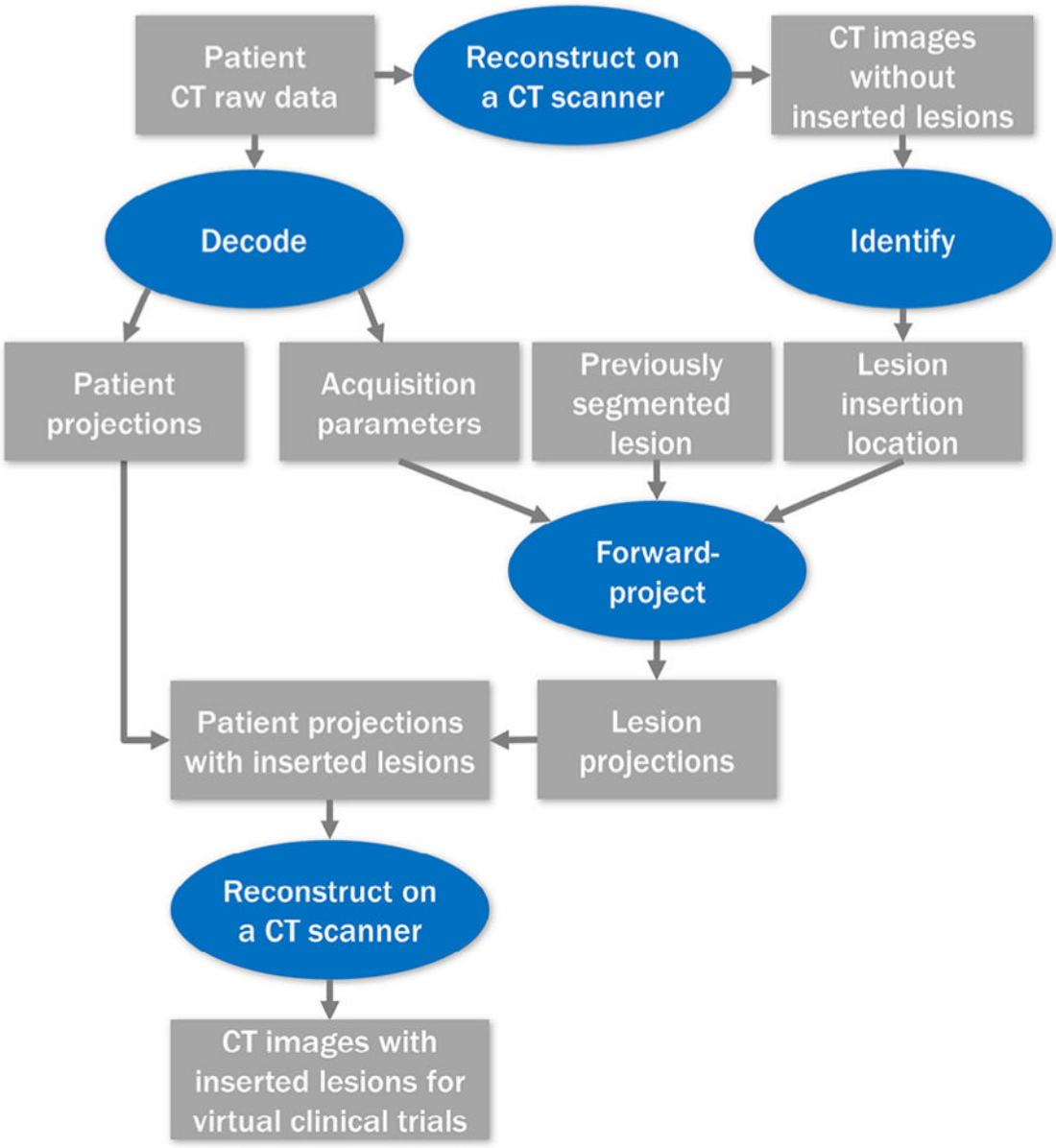

Figure 1.
Illustration of the projection-domain lesion-insertion process.

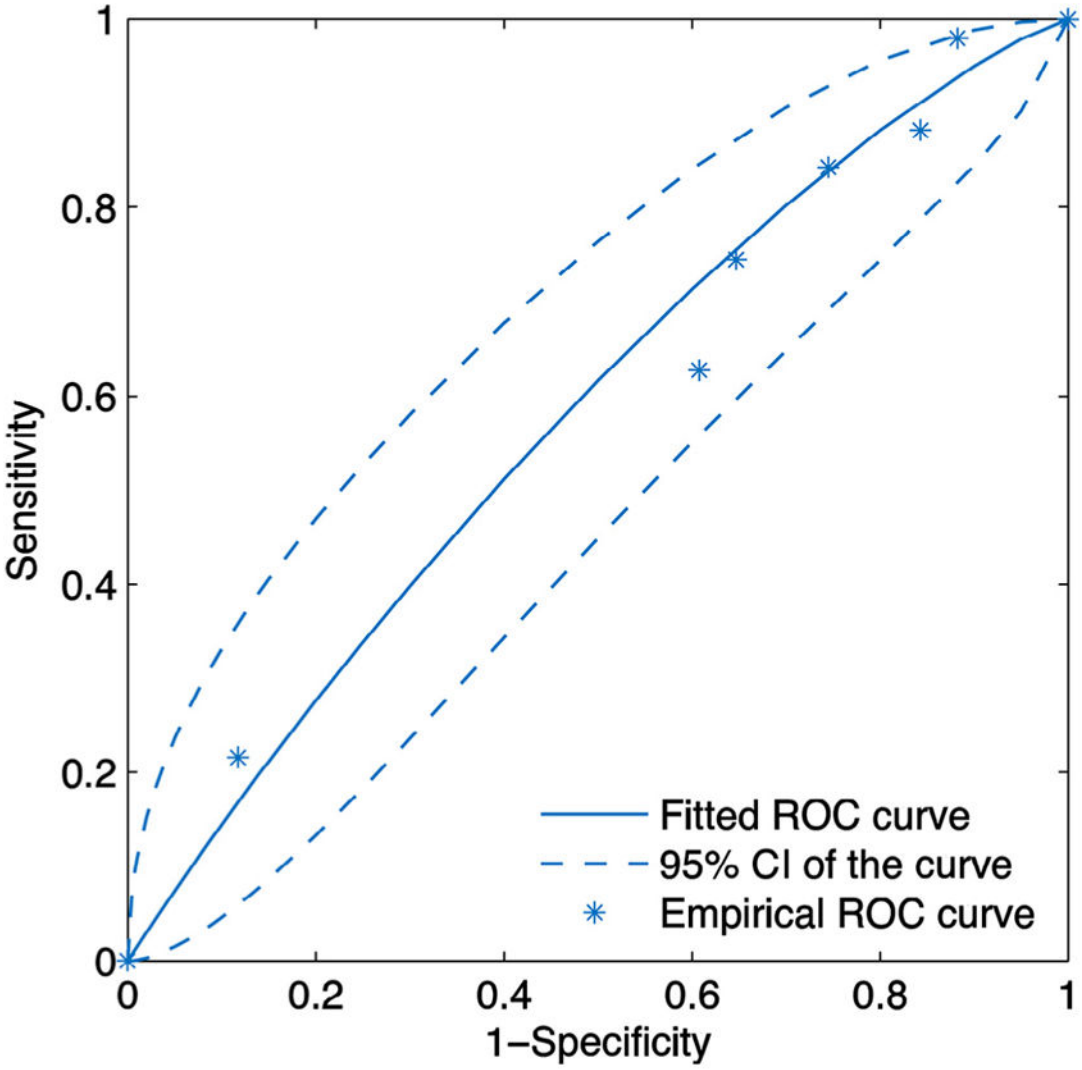

Figure 2.
Fitted receiver operating characteristic (ROC) curve of likelihood scores along with the 95% confidence interval of the curve and the points making up the empirical ROC curve.

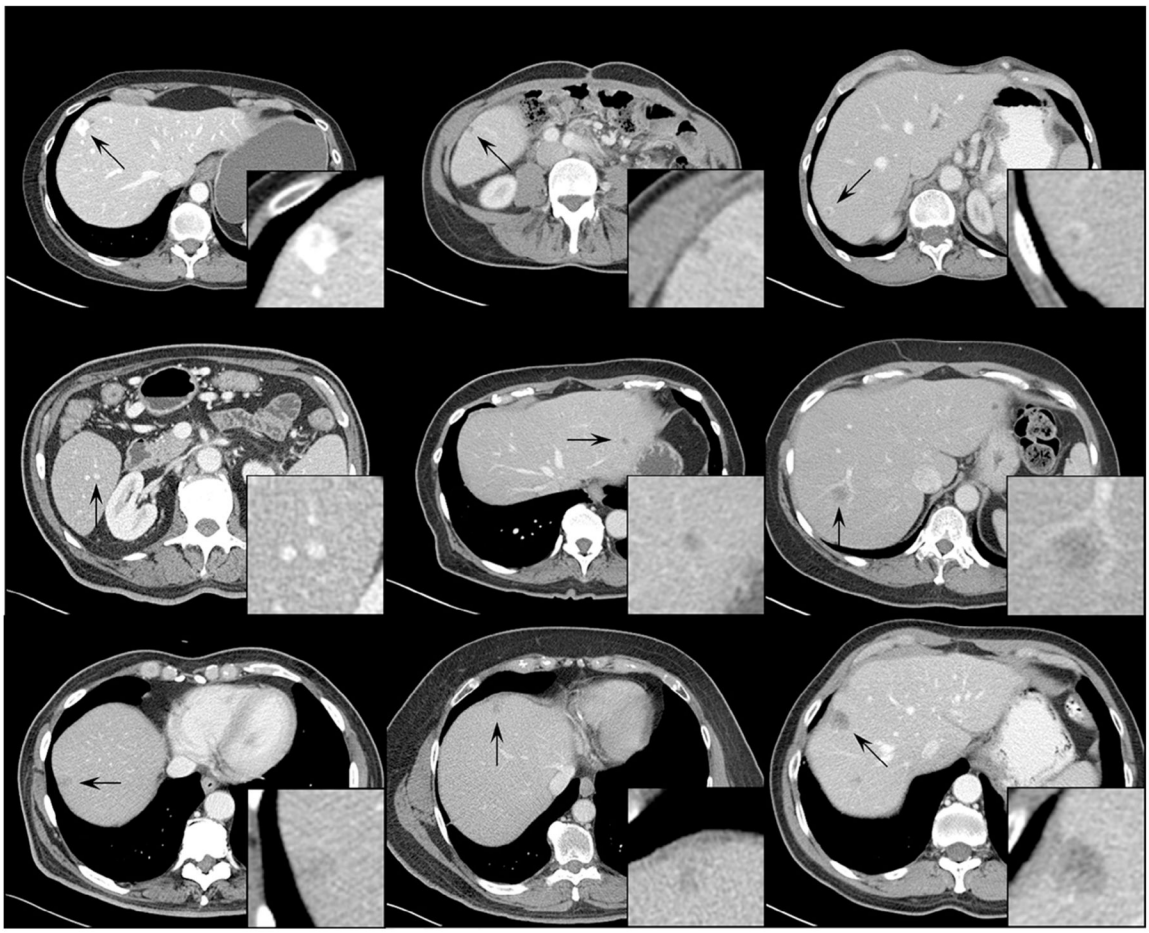

Figure 3.
Examples of inserted lesions with a likelihood score of ≥7 (ie, likely real). The display window width setting is 400 Hounsfield unit (HU) and the level setting is 40 HU.

Figure 4.
Examples of inserted lesions with a likelihood score of ≤3 (ie, likely inserted). The display window width setting is 400 Hounsfield unit (HU) and the level setting is 40 HU.

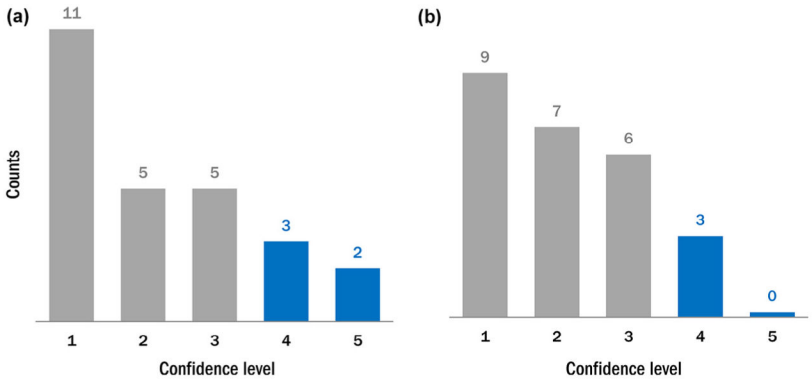

Figure 5.

The histograms of the confidence levels for (**a**) correctly identified inserted lesions and (**b**) incorrectly identified inserted lesions. Low score indicates lack of confidence in classification as inserted vs. real, whereas a high score indicates a high level of certainty in making this classification.

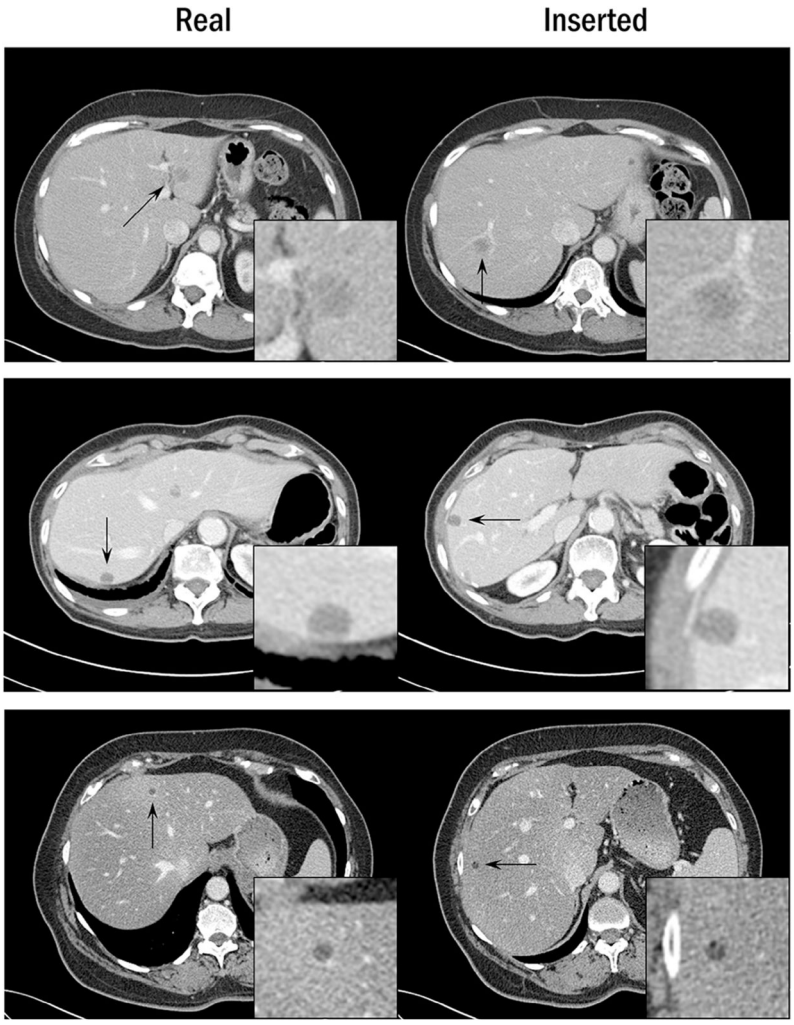

Figure 6.

Examples of lesion pairs when the inserted lesion was incorrectly identified at a high confidence level. The display window width setting is 400 Hounsfield unit (HU), and the level setting is 40 HU.