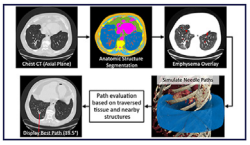

Artificial Intelligence–Aided Selection of Needle Pathways: Proof-of-Concept in Percutaneous Lung Biopsies

Meridith A. Kisting, MA, J. Louis Hinshaw, MD, Giuseppe V. Toia, MD, Timothy J. Ziemlewicz, MD, Adrienne L. Kisting, BS, Fred T. Lee Jr., MD, and Martin G. Wagner, PhD

ABSTRACT

Purpose: To evaluate the concordance between lung biopsy puncture pathways determined by artificial intelligence (AI) and those determined by expert physicians.

Materials and Methods: An AI algorithm was created to choose optimal lung biopsy pathways based on segmented thoracic anatomy and emphysema in volumetric lung computed tomography (CT) scans combined with rules derived from the medical literature. The algorithm was validated using pathways generated from CT scans of randomly selected patients (n = 48) who had received percutaneous lung biopsies and had noncontrast CT scans of 1.25-mm thickness available in picture archiving and communication system (PACS) (n = 28, mean age, 68.4 years ± 9.2; 12 women, 16 men). The algorithm generated 5 potential pathways per scan, including the computer-selected best pathway and 4 random pathways (n = 140). Four experienced physicians rated each pathway on a 1–5 scale, where scores of 1–3 were considered safe and 4–5 were considered unsafe. Concordance between computer and physician ratings was assessed using Cohen's κ.

Results: The algorithm ratings were statistically equivalent to the physician ratings (safe vs unsafe: $\overline{\kappa} = 0.73$; ordinal scale: $\overline{\kappa} = 0.62$). The computer and physician ratings were identical in 57.9% (81/140) of cases and differed by a median of 0 points. All least-cost "best" pathways generated by the algorithm were considered safe by both computer and physicians (28/28) and were judged by physicians to be ideal or near ideal.

Conclusions: AI-generated lung biopsy puncture paths were concordant with expert physician reviewers and considered safe. A prospective comparison between computer- and physician-selected puncture paths appears indicated in addition to expansion to other anatomic locations and procedures.

ABBREVIATIONS

AI = artificial intelligence, CI = confidence interval, CT = computed tomography, DICO = digital imaging and communications in medicine, IRB = institutional review board, PACS = picture archiving and communication system

Computed tomography (CT)-guided lung biopsies are generally safe and effective, but adverse events such as pneumothorax, hemothorax, air embolism, and pulmonary hemorrhage can be catastrophic, particularly in patients with pre-existing compromised lung function (1–4). Many adverse events from lung biopsies arise because of inadvertent traversal of vulnerable anatomic structures, making the choice of needle pathway critical. Choosing an optimal pathway is currently a complex and manual process owing to the large number of variables that physicians must consider. For example, structures that need to be simultaneously avoided include large airways, vessels, bones, fissures, heart, and focal emphysematous lungs (5–8). Other considerations include choosing a short skin-to-target distance, avoiding the underside of the ribs, and maximizing biopsy needle travel across the long axis of the tumor (8,9). Biopsy paths are ideally performed in the scanning plane to ensure that the needle and anatomical structures can be simultaneously visualized. Most CT systems have the ability to tilt the gantry (generally by ±20°), which gives physicians a large number of off-axial planes to choose an optimal pathway. However, it is not feasible for a physician to evaluate all anatomic structures in every potential path at every possible gantry tilt during the planning process. Thus,

Figure E1 can be found by accessing the online version of this article on *www.jvir.org* and selecting the Supplemental Material tab.

J Vasc Interv Radiol 2024; 35:770–779
https://doi.org/10.1016/j.jvir.2023.11.016

RESEARCH HIGHLIGHTS

- An artificial intelligence algorithm was created to simulate and evaluate pathways for percutaneous lung biopsy.
- A reader study was performed to validate the safety and quality of computer-generated pathways, with physicians and computer rating 140 pathways on a 1–5 rating scale specifically created and deployed for this study.
- Correlation between physician and computer ratings was $r = 0.86$, with high sensitivity and specificity for "safe" vs "unsafe" pathways.

STUDY DETAILS

Study type: Artificial intelligence study

anatomically favorable out-of-plane pathways may go unrecognized.

Recent advances in computer-aided segmentation allow for rapid and accurate CT localization of relevant anatomic structures such as airways, bones, blood vessels, focal emphysema, and mediastinum (8,9). In addition, many cognitive judgments that physicians make during the biopsy planning phase, such as choosing the shortest puncture path, are opportunely suitable to computer-aided decision-making. An artificial intelligence (AI) platform was recently developed that analyzes preprocedural volume CT data sets, segments vulnerable anatomic structures, considers a large number of puncture pathways not constrained by plane (50,000), and generates ranked potential puncture pathways for the operating physician to consider during the procedure planning process.

The purpose of this proof-of-concept study was to determine the concordance rate between ratings generated by this AI system and those provided by a group of experienced physicians for choosing safe lung biopsy pathways. A high concordance between computer and physician grading would favor prospective testing in patients undergoing percutaneous lung biopsy as well as extension of the algorithm into other anatomical locations.

MATERIALS AND METHODS

Biopsy Path Algorithm Design

This study was approved by the institutional review board (IRB) of the University of Wisconsin-Madison (IRB protocol 2016-0418). A needle trajectory planning algorithm was designed to minimize risk and increase effectiveness of percutaneous lung biopsy (**Fig 1**). A machine learning–based algorithm was developed using a convolutional neural network (3D U-net) (10) with an encoder depth of 3 and filter size of 3 × 3 × 3 voxels, which was pretrained using CT images of virtual patients based on the 4D CT phantom XCAT (11). This was followed by training with manually annotated chest CT images from patients (n = 50) retrospectively collected at the institution with IRB approval. All CT images were reconstructed with a 1.25-mm slice thickness and 0.98 × 0.98-mm pixel spacing. Training was performed on randomly extracted image blocks (128 × 128 × 32 voxels) for 80 epochs using a cross-entropy loss function and a minibatch size of 5. The algorithm was applied to the patient data set to segment organs, tissue types, and anatomic structures of interest, including the heart, lungs, pulmonary vessels and airways, muscle, bone, fat, and fissures.

To account for areas of focal emphysema, a post-processing step was developed to analyze CT attenuation values (HU) from the immediate preprocedural scan (**Fig 2a**). The CT volume was analyzed using a median filter with a 5 × 5 × 5-voxel kernel size. Emphysematous tissue was segmented using a global threshold for all lung voxels, where voxels with a CT value of less than −950 HU were considered emphysematous (**Fig 2b**) (12). A connected component analysis was performed to determine the size of individual blebs, and all blebs with an equivalent diameter of ≥4 mm were considered critical.

The lung nodule target was manually identified, and then, the algorithm simulated n = 50,000 equidistant paths (angular sampling distance = 0.87°) within ±20° of the axial plane. For each path, an associated risk value was calculated based on a set of known risk factors from the medical literature (**Table 1**). Some risk factors were not readily available from the medical literature owing to infrequency (ie, needle crossing aorta or spine) or lack of quantification (ie, more extreme needle angles traversing the pleura). Therefore, weighting decisions were made in consensus between 2 radiologists (J.L.H., F.T.L.) with 16 and 31 years of lung biopsy experience. The most important factors considered were the risk of puncture of mediastinal organs, large vessels or airways, crossing of lung fissures or emphysematous blebs, length of the trajectory through tissue (longer trajectories received higher risk scores), and the angle to the pleural surface (angles further from perpendicular received higher risk scores) (2,5,6,13–17). The individual risk factors were weighted and added to calculate the overall risk value for a given puncture pathway (**Table 1**). The individual weights (within the range of 0–1) for each risk factor were manually selected a priori by a radiologist with 31 years of lung biopsy experience. The algorithm was implemented in Matlab R2021b (Mathworks, Natick, Massachusetts).

Experimental Design

A reader study was conducted to compare computer-generated path ratings with expert opinions to validate the algorithm. IRB approval was obtained to create a deidentified database for this single-center prospective, comparative, transversal, and observational study. The initial cohort included 48 random patients who underwent percutaneous CT fluoroscopy–guided biopsy (single-rotation) for a lung

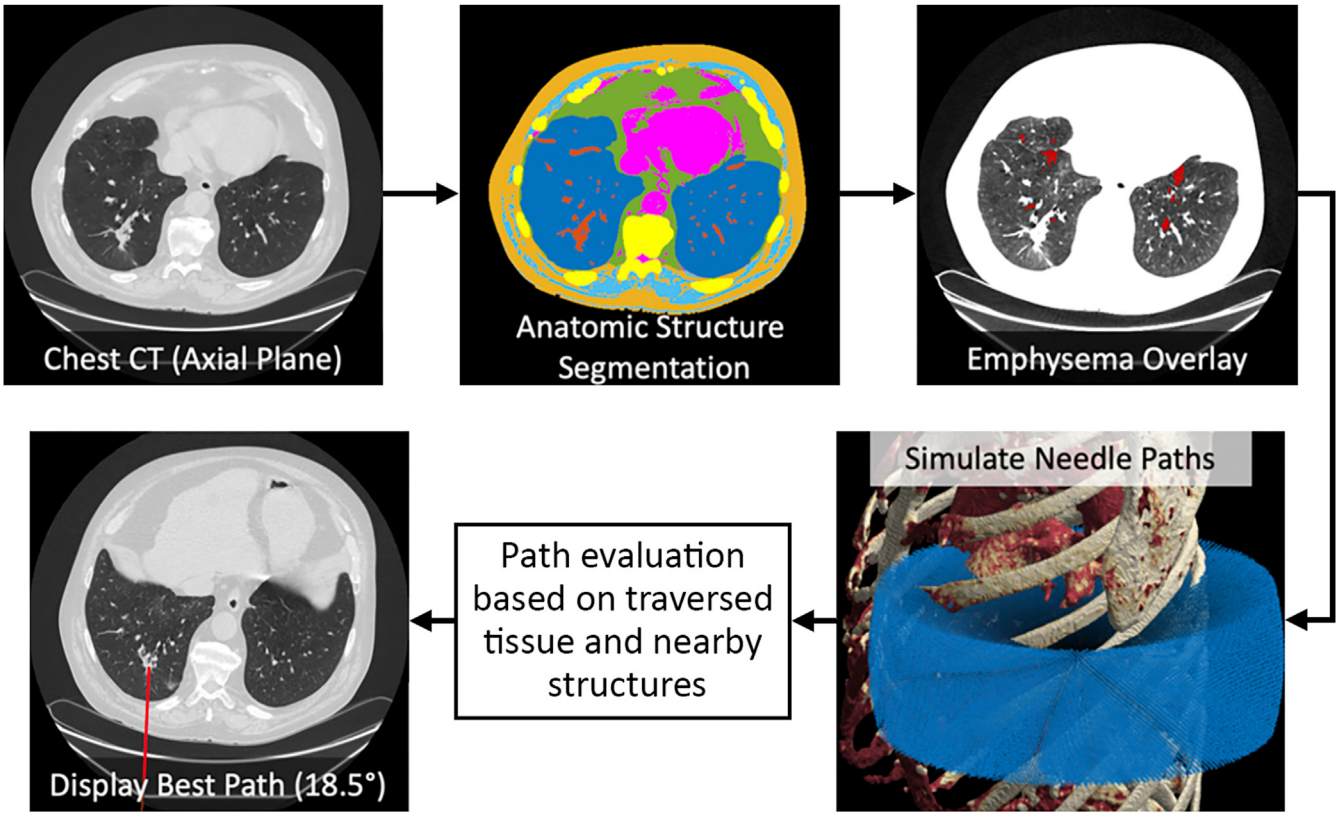

Figure 1. Biopsy path algorithm. A machine learning algorithm was created to segment anatomy in chest computed tomography (CT) scans including organs, bone, vessels, airways, and fissures. Areas of emphysema were identified by detecting focal areas of low attenuation. The computer then simulated over 50,000 pathways within 20° of the biopsy target. Pathway safety was assessed using a set of rules based on physician logic and the medical literature (**Table 1**). These rules were used to calculate the cost for each of the 50,000 pathways, with the least-cost pathway displayed as the best path. The best pathway shown in this example has a gantry tilt of 18.5°.

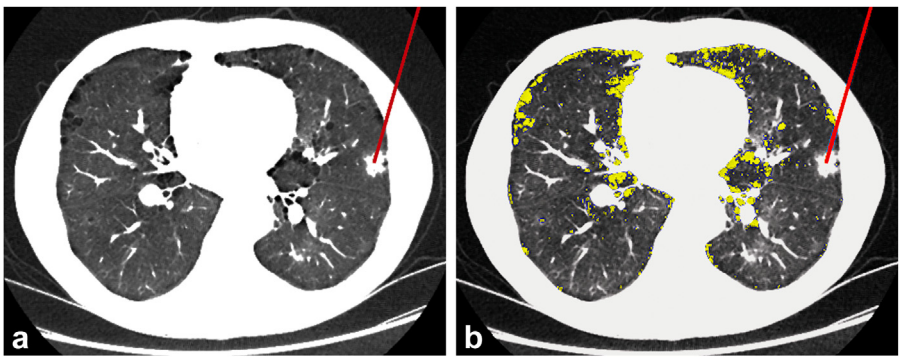

Figure 2. Emphysema identification. **(a)** The algorithm identified areas of low attenuation (less than −950 HU). **(b)** Emphysematous blebs were categorized as areas of low attenuation greater than 4 mm in diameter (yellow spots). Trajectories crossing these areas were assigned a higher cost owing to the risk of pneumothorax. The proposed pathways (red line) avoided blebs while also traversing >1 cm of "normal" lung parenchyma (a rule incorporated into the algorithm).

nodule between February 2012 and June 2020. Inclusion criteria included a preprocedural noncontrast CT examination available in picture archiving and communication system (PACS) with a minimum of 1.25-mm slices. There were no additional exclusion criteria. For each case, the path with the lowest risk score and 4 additional random paths (n = 140 paths in total) were selected and rated by both AI algorithm and each physician reviewer. Reviewers were blinded to ratings from both computer algorithm and other reviewers.

Table 1. Risk Factors and Weighting

Category	Structure or organ	Metric	Weight	Minimum ranking
Path is crossing critical structures or organs (yes = 1, no = 0)	Heart/aorta	Yes/no	Infinity	5
	Large vessels or airways	Radius ≥ 1.5 mm	1	3
	Other organs	Yes/no	3	4
	Fissure	Yes/no	0.6	4
	Emphysematous bleb	Radius ≥ 2 mm	0.8	4
	Ribs or bone	Yes/no	3	4
	Spine	Yes/no	Infinity	5
Path is close to critical organs	Heart	(10 mm to minimum distance)/10 mm	0.8	NA
	Large vessels or airways	(10 mm to minimum distance)/10 mm	0.8	NA
Trajectory angle	Angle (±90°) to axial plane	Abs (angle/90)	0.05	NA
	Angle (0°–90°) to pleural surface	(90 – angle)/90	0.25	3 if metric > 0.7
Length of needle path	Path length (mm) through lung tissue (at least 1 cm)	Maximum (0 [10 mm – length]/10 mm)	0.5	NA
	Total path length (mm) through tissue	Maximum (0 [length– 60 mm]/40 mm)	1	3 if metric > 100
	Path length through minor vessels or airways	0.5 × length	1	NA

Patient Characteristics

A total of 28 patients met the inclusion criteria and were included in the test cohort (12 female and 16 male patients). All patients underwent percutaneous CT fluoroscopy–guided lung biopsy of a single lung nodule, and all procedures were technically successful (Table 2). The mean age at the time of procedure was 68.4 years ± 9.2, and 64% (18/28) of the patients had a history of emphysema. The diagnostic yield was 89.3% (25/28), with malignant diagnoses accounting for 57.1% (16/28) of biopsies, benign or infectious/inflammatory diagnoses accounting for 32.1% (9/28), and nondiagnostic biopsies accounting for 10.7% (3/28). The mean diameter of the biopsied nodules was 2.0 cm ± 1.6, with a mean skin-to-target distance of 6.5 cm ± 1.9. The actual needle path distance was 8.1 cm ± 2.3 for the procedures.

Table 2. Patient and Nodule Characteristics

Category	Value
Patients (N = 28)	
Sex, n (%)	
Female	12 (43)
Male	16 (57)
Age (y)	
Mean	68.4
Range	44–87
History of emphysema, n (%)	18 (64)
Nodules (N = 28)	
Location, n (%)	
Right upper lobe	7 (25)
Right middle lobe	2 (7)
Right lower lobe	7 (25)
Left upper lobe	9 (32)
Left lower lobe	3 (11)
Diameter (cm), mean ± SD	2.0 ± 1.6
Skin-to-target depth (cm), mean ± SD	6.5 ± 1.9
Technical success	28 (100)

Computer Algorithm Biopsy Path Ratings

Computer algorithm rankings were determined using a scale that combined the calculated cost and certain safety criteria (Table 3). Pathways were rated 1 if the cost was <1.7 and no dangerous conditions were met (Fig 3a). Pathways receiving a rating of 2 (Fig 3b) either had a calculated cost of >1.7 or the angle between the needle and the pleura-to-lesion axis was >10° (18). Pathways receiving a rating of 3 crossed large vessels or airways, had a target depth of >10 cm, or had needle-to-pleura lesion angle of >27° (Fig 3c). Angles were chosen based on previous work showing decreased forced expiratory volume at needle-to-pleura angles of <80° and the mean value between groups with and without chest tubes reported by Saji et al (18,19). Pathways receiving a rating of 4 crossed ribs, emphysematous blebs, fissures, or other organs (Fig 3d). Pathways receiving a rating of 5 crossed through the heart, aorta, or spine (Fig 3e).

Physician Biopsy Path Ratings

A digital imaging and communications in medicine (DICOM) viewer for the visualization of needle paths in CT images was built in Matlab (R2021b; Mathworks). The viewer provided cut plane views and maximum intensity projections in axial, coronal, and sagittal planes and in cut planes along the proposed needle paths. Four radiologists (G.V.T., J.L.H., T.J.Z., and F.T.L. with 2, 15, 16, and 31 years of experience in percutaneous lung biopsies, respectively) were shown 5 computer-generated puncture pathways (highest rated and 4 random pathways) for each patient for a total of 140 pathways. Each pathway was rated independently on an ordinal scale of 1–5 based on criteria noted in Table 3. Reviewers were shown cases in an identical order and were given the option to return to previous cases and make changes before submitting ratings. No additional clinical information was provided to physicians. Any physician scores of 4 or 5 were re-reviewed, and physicians were asked for a reason for the poor rating using a combination of pick list and free text.

Table 3. Rating Scales

Score	Algorithm rating rationale	Physician rating rationale
Safe		
1	None of the conditions for 2–5 apply	Ideal or near-ideal pathway
2	Cost > 1.7 or pleural angle < 36°	Good, but not quite ideal pathway
3	Path crosses through large vessels or airways, length through tissue is >10 cm or pleural angle of <27°	Potential pathway if necessary, but better pathways are available for safety or technical reasons
Unsafe		
4	Path crosses ribs, emphysematous blebs, fissures, or other organs	Dangerous or technically challenging pathway
5	Path crosses through the heart, aorta, or spine	Extremely dangerous (ie, potentially fatal) or technically extremely challenging pathway

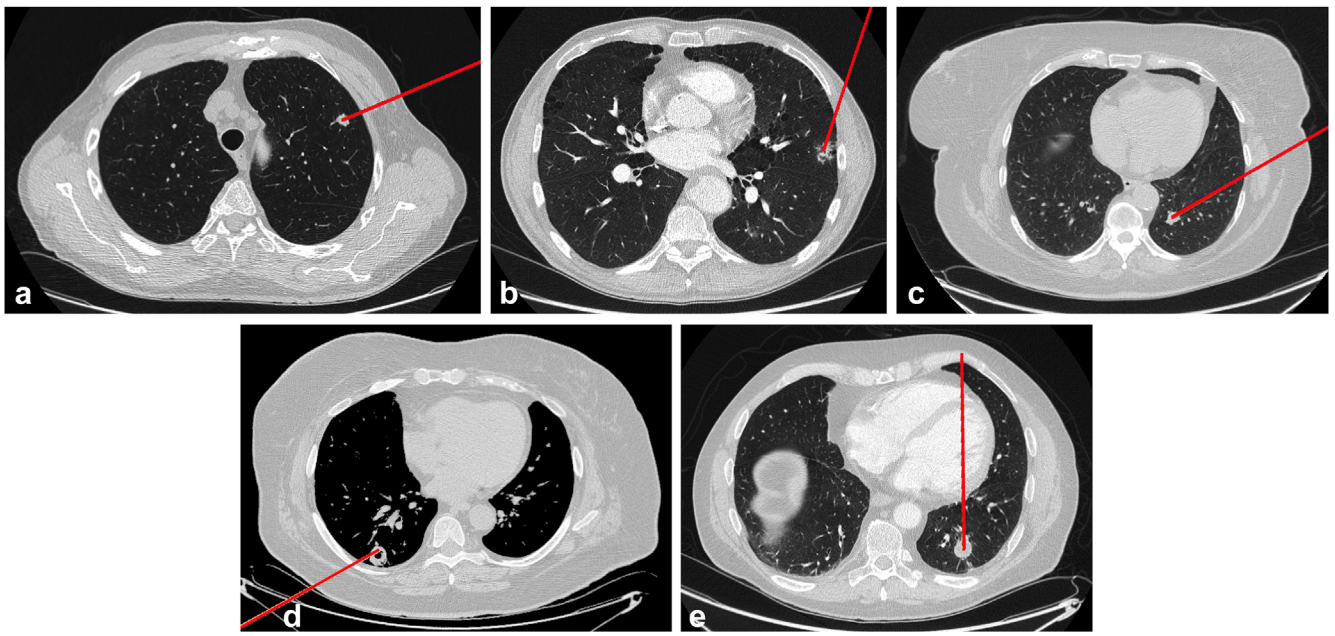

Figure 3. Examples of computer-generated pathways. For this study, both high- and low-cost pathways were generated for evaluation (red line). **(a)** Pathway rated "1" by computer by not meeting any of the criteria listed in **Table 1**. **(b)** Pathway rated "2" by computer owing to an angle of <10° to pleural surface. **(c)** Pathway rated "3" by computer owing to an unnecessarily long pathway. **(d)** Pathway rated "4" by computer owing to a crossing rib. **(e)** Pathway rated "5" by computer for crossing heart.

Statistical Analysis

The required sample size (n = 139 pathways) was estimated to achieve a 95% confidence interval (CI) width of ≤0.2 for κ, assuming a target agreement of 0.8 for binary (safe vs unsafe) ratings and a marginal probability of 0.5 for safe pathways (20). Interobserver concordance was measured using Fleiss' κ for ratings on a 1–5 scale as well as binary ratings (safe vs unsafe), where pathways with a rating of ≤3 were considered safe (21). In addition, Cohen's κ was calculated to measure agreement between all combinations of 2 reviewers as well as between the proposed algorithm and each reviewer (22). Linear weighting was used for ordinal scale ratings (23,24). A noninferiority test according to Walker et al (25) was performed to compare agreements (Cohen's κ) between physicians with agreements between the proposed algorithm and physicians using an inferiority margin of 0.1. For all analyses, a significance level of $\alpha = .05$ was used. A confusion matrix was calculated to get a detailed understanding of the relationship between computer and physician ratings. Sensitivity and specificity were calculated for each individual class and for the binary classification problem to detect safe (Classes 1–3) and unsafe (Classes 4 and 5) trajectories. A bootstrapping approach (10,000 random samples) was used to calculate 95% CIs for sensitivity, specificity, and ratios (26).

RESULTS

Physician versus Computer Ratings

Interobserver agreement measured as Fleiss κ between the 4 physician ratings was 0.36 (CI: 0.35–0.36; $P < .001$) for the 1–5 scale and 0.75 (CI: 0.73–0.76; $P < .001$) for safe vs unsafe decisions. Cohen's κ between all pairs of reviewers ranged from 0.54 to 0.69 for 1–5 ratings and from

Table 4. Physician versus Computer Ratings

Rating	No. of pathways receiving each rating		No. of identical ratings: computer vs physician (n = 81)	No. of pathways with computer rating concordant with physicians (n = 136)	No. of pathways with computer rating discordant with physicians (n = 14, 10%)
	Computer (n = 140)	Physician (n = 140)			
1	46 (32.8)	34 (24.3)	27 (19.3)	84 (users: safe, computer: safe)	9 (users: safe, computer: unsafe)
2	7 (5.0)	36 (25.7)	3 (2.1)		
3	36 (25.7)	23 (16.4)	17 (12.1)		
4	35 (25.0)	33 (23.6)	23 (16.4)	42 (users: unsafe, computer: unsafe)	5 (users: unsafe, computer: safe)
5	16 (11.4)	14 (11.4)	11 (7.9)		

Note–Values are given as n (%).

0.66 to 0.86 for binary ratings. As a comparison, the agreements between algorithm and physician ranged from 0.53 to 0.69 for ordinal scale ratings and between 0.69 and 0.78 for binary ratings. In both cases, CIs for the agreement between algorithm and physician (safe vs unsafe: $\overline{\kappa} = 0.73$, CI: 0.67–0.79; ordinal scale: $\overline{\kappa} = 0.62$, CI: 0.55–0.70) were within the inferiority margin of the agreements between individual physicians (safe vs unsafe: $\overline{\kappa} = 0.75$, CI: 0.70–0.79; ordinal scale: $\overline{\kappa} = 0.63$, CI: 0.57–0.69). Computer and physician scores were identical in 57.9% (81/140, CI: 49.3%–65.7%) of cases (**Table 4**). The median difference between physician and computer ratings was 0 points. The confusion matrix between algorithm and physician ratings is shown in **Figure 4a**. The sensitivities of Classes 1–5 ranged from 42.9% (class 2, CI: 0.0%–85.7%) to 68.8% (Class 5, CI: 41.7%–88.9%) and the corresponding specificities from 75.2% (Class 2, CI: 67.2%–82.0%) to 97.6% (Class 5, CI: 93.3%–99.2%). For the binary classification problem (safe vs unsafe) the sensitivity and specificity were 94.4% (CI: 87.6%–97.8%) and 82.4% (CI: 69.4%–90.9%), respectively. The corresponding confusion matrix is shown in **Figure 4b** and the receiver operating characteristic curve (ROC) is shown in **Figure 4e**.

Quality of Computer-Selected "Best" Pathway

All 28 computer-generated best pathways were considered safe by both computer and physicians (**Fig 4d**). There were no best pathways rated as 4 or 5 by the computer or physicians (**Fig 4c**). The median computer and physician ratings of the best pathways were both 1, with a median difference between computer and physician of 0 points. The mean interobserver difference between physicians in the same cases was 0.61 ± 0.68. The highest computer-rated paths had a mean skin-to-nodule distance of 7.6 cm ± 3.8 compared with 8.1 cm ± 2.3 for the actual needle path during the procedure.

Many of the 28 best pathways used trajectories outside of the axial plane (**Fig 5c**), with 39.3% (11/28) of pathways within 5° of the axial plane, 28.6% (8/28) between 5° and 10°, and 32.1% (9/28) between 10° and 20°. The mean gantry tilt of the best computer-rated path was 8.5° ± 6.2. **Figure 5a, b** show an example were access to the nodule was challenging in the axial plane, while a 15° gantry tilt allowed easy access.

Safety of the Computer Algorithm

The computer rated 63.6% (89/140; CI: 55.0%–71.4%) of pathways as "safe" (rating 1–3), whereas 66.4% (93/140; CI: 58.6%–74.3%) of pathways were considered safe by physicians (**Fig 3**). No pathway scored 5 by physicians was considered safe by the computer, and no pathway scored 4 by physicians was given a 1 or 2 by the computer. However, 5 pathways met the criteria to receive a computer rating of 3 (safe, but not ideal or near ideal) while receiving a median physician rating of 4 (dangerous or technically challenging) (**Fig E1**, available online on the article's Supplemental Material page at *www.jvir.org*). After additional physician review, all 5 of these pathways (**Fig E1a-e**) were determined to be excessively long, with 2 pathways crossing vessels (**Fig E1b, e**) and 1 pathway crossing a rib and an area of emphysema (**Fig E1d**). A computer score of 3.2 appeared to be the cutoff between safe and unsafe, ie, all cases considered unsafe by physicians had computer scores higher than 3.2.

Of the 140 pathways, 47 (33.6%; CI: 25.7%–41.4%) received an "unsafe" median physician rating. The results of the physician re-review of unsafe pathways are listed in **Table 5**. The median computer rating for these pathways was 3, and the median physician rating was 2.5. None of these 47 pathways were presented as the best pathway by the computer. Most of these 47 pathways were given a computer rating of 4 (26/47, 55.2%; CI: 40.4%–68.1%), followed by a computer rating of 5 (16/47, 34.0%; CI: 21.3%–48.9%). The remaining 5 pathways received a computer rating of 3, as detailed earlier (**Table 5** and **Fig 4**).

DISCUSSION

Choosing a needle pathway for percutaneous interventions is currently a manual process, during which physicians assess critical anatomy and apply "rules" from previous experience and the medical literature in an attempt to find the safest and most effective route to the target. The complex 3D anatomy of the chest and large number of rules fit naturally with the strengths of AI, particularly when combined with recent advances in computer-based segmentation of chest anatomy. The design of the current AI algorithm takes into account both anatomy from a volumetric 3D data set and rules garnered from both literature and experienced

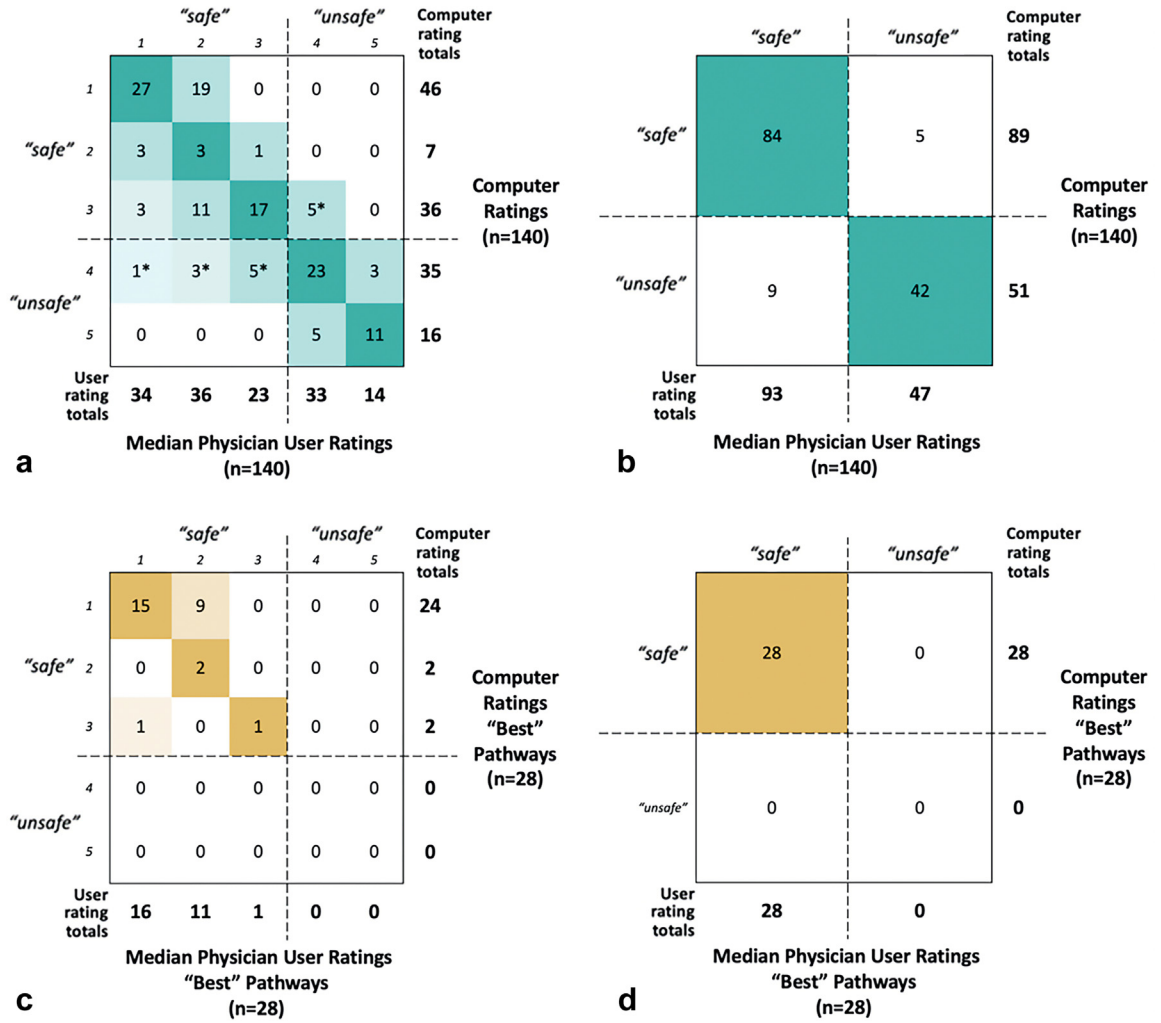

Figure 4. Confusion matrix of algorithm and physician ratings. The value in each cell corresponded to the total number of pathways receiving this rating from both computer and physician user (ie, 27 pathways were rated 1 by the computer and physicians). The total number of pathways receiving a score 1–5 are indicated in bold on the bottom and right edge. The shading of cells corresponds to the degree of disagreement between the computer and physician ratings. **(a)** Overall, 60% of pathways (84/140) received the same rating from both computer and physician users (darkest shading). Totals marked with an asterisk represent pathways that were classified as "safe" by 1 system of evaluation and "unsafe" by the other. **(b)** All 140 pathways categorized as safe and unsafe. **(c, d)** Confusion matrices of the 28 lowest cost "best" pathways. **(e)** Receiver operating characteristic curve showing the false-positive rate and the true-positive rate for safe vs unsafe decisions made by the algorithm.

physicians to determine the "cost" for 50,000 possible biopsy pathways. The results of this study demonstrated that the highest rated computer-generated pathways are safe and generally deemed ideal or near ideal by a cohort of experienced physicians. Known unsafe pathways were also detected by the algorithm with high concordance with physician reviewers. In unusual cases (5/47) of discordance between physicians and computer concerning pathway safety, the computer-scored pathway was not the best pathway and, thus, unlikely to even be presented to physician operators for consideration.

Computer-aided navigation has been used for years to assist endoluminal imaging and interventions such as angiography, bronchoscopy, and virtual colonoscopy. Key differences between the current and previous computer-aided navigation systems include choosing pathways for percutaneous procedures vs mapping pathways along existing luminal structures such as airways, blood vessels, or the gastrointestinal tract. A further difference is the consideration of both anatomy and physician logic and the creation of a cost score for each pathway, wherein the least-cost pathway is considered most desirable. An example of the use of physician logic includes favoring pathways that contact the lung surface at more perpendicular angles. One of the benefits of this approach is that there is no limit to the rules that can be added as new discoveries take place, allowing the system to improve over time. A machine learning opportunity is to compare pathways chosen by expert physicians with those chosen by computers. Algorithmic improvements will accelerate if combined with a

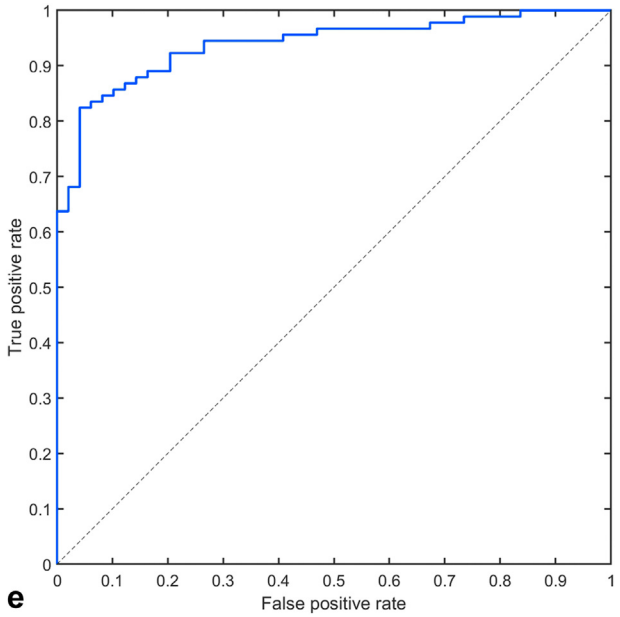

Figure 4. (Continued).

Table 5. Results of Physician Review of 47 Pathways Receiving an Unsafe Median Physician Rating

	Pathways meeting criteria (n)	Pathways meeting criteria (%)
Traverses a rib or bone	26	55.3
Traverses heart	4	8.5
Traverses large or important vessel	17	36.2
Traverses large airway	1	2.1
Traverses emphysematous lung	2	4.3
Traverses spinal column	5	10.6
Traverses pleura (fissure)	15	31.9
Long biopsy pathway	39	83.0
Very oblique angle of approach to the pleura	1	2.1
Other/unknown	1	2.1
Traverses liver	2	4.3
Traverses esophagus	3	6.4

feedback loop where successes (ie, needle reaching the target nodule) and failures (autodetection of adverse events such as pulmonary hemorrhage and pneumothorax) are automatically fed back into the system.

A previous work (27) has investigated the potential benefits of applying computer assistance to CT-guided lung biopsy decision-making. Automated AI-assisted lung nodule segmentation has been used for choosing biopsy targets. The algorithm had high agreement with physicians, but, although a novel concept, it is unlikely that most operators will need assistance in target identification. Another work (28) has demonstrated 3D motion tracking models to follow lung nodules subject to respiratory variability with some success. Although this may have utility in biopsy pathway assessment, the study did not directly implement AI or machine learning into the design, making practical integration an arduous task. Despite over a decade's worth of AI research in radiology, there remain a paucity of data combining computer-assisted applications for any type of extraluminal image-guided procedure.

The thorax was the first location chosen for application of the study algorithm owing to the large number of lung biopsies, known anatomical and procedural characteristics that can affect effectiveness and safety, the importance of physician logic amenable to programming, potentially catastrophic consequences of errant needle pathways, and high intrinsic tissue contrast that aids computer segmentation. Overall, the ability for the computer to rapidly interrogate the position of ribs and intrathoracic structures to choose an ideal pathway and gantry angle from 50,000 possibilities is something not possible for a physician. Other abdominopelvic, extremity, spine, or intracranial procedures may also benefit from computer-aided pathways, and these will be explored in the future.

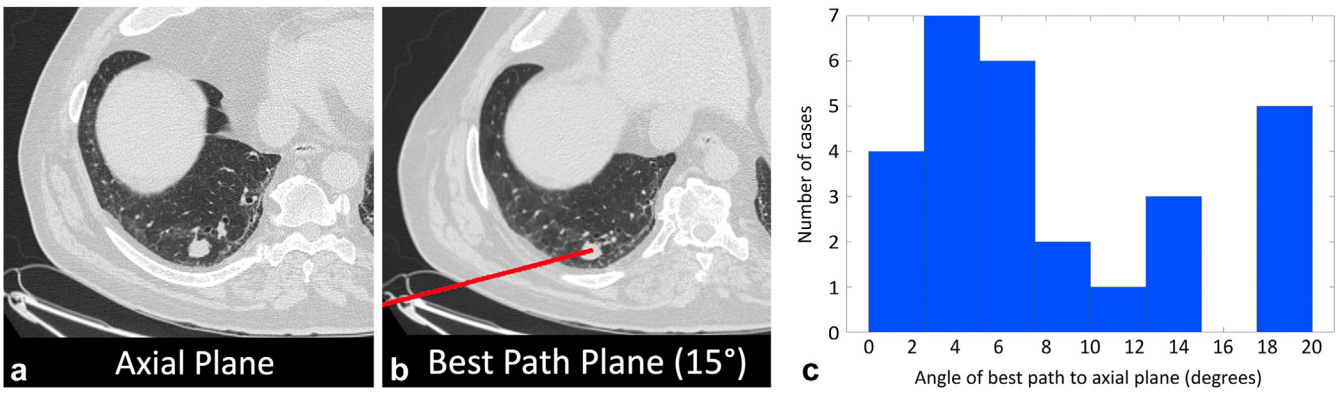

Figure 5. Gantry tilt of best pathways. Many of the 28 best pathways used trajectories outside of the axial plane. **(a, b)** The same nodule was easily accessed with 15° of tilt away from axial. Only 11 (39.2%) of the highest rated pathways were within 5° of axial compared with 9 (32.1%) between 10° and 20°. The ability of the computer algorithm to rapidly assess multiple pathways in any plane was a particular strength of this algorithm.

A future enhancement of the computer-based algorithm is integration with an anatomic guidance system so that skin punctures can be performed at the location and angle chosen by the algorithm. Currently, the system can choose an ideal pathway, but the physician needs to then find the puncture site and angle of approach. Fortunately, several such systems (electromagnetic, optical, robotic, and laser) are already in clinical use, use onboard computers to process image data, and potentially could be rapidly adapted for use with the study algorithm (29–33). Similarly, this software program could be incorporated directly into CT systems for procedural enhancement.

There were several limitations inherent in this proof-of-concept study. The CT data sets used for determining puncture pathways were supine diagnostic scans obtained days or weeks before the biopsy. This was because thin slices (1.0–1.25 mm) are necessary to construct an adequate 3D data set for anatomic segmentation, and the standard biopsy setup scan at the study institution was performed with 2.5-mm slices. Because of this limitation, a prospective study comparing the computer-chosen vs actual (physician-chosen) biopsy pathways with the patient in the treatment position could not be performed. However, data are currently being gathered for a future such study. Another limitation was the early software version. The authors are already planning several anatomic and logic improvements that were not incorporated in this version of the software. For example, a preference for pathways that traverse the long axis of the tumor, puncturing of fat vs muscle tissues, and segmenting (and avoiding) the internal mammary and axillary vessels are improvements already in development. The ability to iteratively improve the targeting software is an important feature, but, despite the early version of the program, the results are already closely comparable to the judgment of expert physicians. Finally, in this study, the scoring system used to evaluate risk is novel and lacks validation outside of high level of concordance seen with physician ratings.

In summary, this proof-of-concept study demonstrates that a computer algorithm was able to select safe biopsy pathways for percutaneous lung biopsies, regardless of puncture plane, by incorporating segmentation of thoracic anatomy along with logic derived from physician experience and the literature. These pathways were rated by the algorithm in concordance with expert physicians. In the future, this approach may also be used to assist percutaneous procedures in other anatomic locations. The algorithm is designed for rapid improvements through incorporation of additional rules and machine learning and has the potential to be paired with existing electromagnetic, optical, or robotic guidance devices that can direct the physician to the chosen puncture site and angle.

AUTHOR INFORMATION

From the Departments of Radiology (M.A.K., J.L.H., G.V.T., T.J.Z., A.L.K., F.T.L., M.G.W.), Urology (J.L.H., F.T.L.), Medical Physics (G.V.T., M.G.W.), and Biomedical Engineering (F.T.L.), University of Wisconsin-Madison, Madison, Wisconsin. Received August 3, 2023; final revision received November 9, 2023; accepted November 17, 2023. Address correspondence to M.G.W., Department of Radiology and Medical Physics, University of Wisconsin-Madison, Madison, Wisconsin; E-mail: mwagner9@wisc.edu

J.L.H. is a consultant for Neuwave Medical and shareholder in Elucent, Accure, Histosonics, and Cellectar. G.V.T. is a consultant for GE Medical and is on the medical advisory board of Canon Medical. T.J.Z. reports research support from and is a shareholder in Histosonics, and receives research support from and is a consultant for Neuwave Medical. F.T.L. is one of the board of directors of Histosonics; is a consultant for Elucent Medical and Neuwave Medical; is a stockholder in Histosonics; reports research support from Histosonics and patents and royalties from Medtronic; and is a member of the medical advisory board of Canon Medical. M.G.W. is a consultant for Histosonics. None of the other authors have identified a conflict of interest.

From the 2023 SIR Annual Scientific Meeting (Abstract No. 568, "Computer-Aided Biopsy of Lung Nodule: Proof-of-Concept of an Artificial Intelligence (AI) Algorithm to Choose Puncture Pathways").

REFERENCES

1. Sharma A, Shepard JAO. Lung cancer biopsies. Radiol Clin North Am 2018; 56:377–390.
2. Huang MD, Weng HH, Hsu SL, et al. Accuracy and complications of CT-guided pulmonary core biopsy in small nodules: a single-center experience. Cancer Imaging 2019; 19:51.
3. Wu CC, Maher MM, Shepard JAO. Complications of CT-guided percutaneous needle biopsy of the chest: prevention and management. AJR Am J Roentgenol 2011; 196:W678–W682.
4. Tomiyama N, Yasuhara Y, Nakajima Y, et al. CT-guided needle biopsy of lung lesions: a survey of severe complication based on 9783 biopsies in Japan. Eur J Radiol 2006; 59:60–64.
5. Nour-Eldin NEA, Alsubhi M, Emam A, et al. Pneumothorax complicating coaxial and non-coaxial CT-guided lung biopsy: comparative analysis of determining risk factors and management of pneumothorax in a retrospective review of 650 patients. Cardiovasc Interv Radiol 2016; 39: 261–270.
6. Ozturk K, Soylu E, Gokalp G, Topal U. Risk factors of pneumothorax and chest tube placement after computed tomography-guided core needle biopsy of lung lesions: a single-centre experience with 822 biopsies. Pol J Radiol 2018; 83:e407–e414.
7. Bourgouin PP, Rodriguez KJ, Fintelmann FJ. Image-guided percutaneous lung needle biopsy: how we do it. Tech Vasc Interv Radiol 2021; 24:100770.
8. He L, Meng Y, Zhong J, Tang L, Chui C, Zhang J. Preoperative path planning algorithm for lung puncture biopsy based on path constraint and multidimensional space distance optimization. Biomed Signal Process Control 2023; 80:104304.
9. Liu Q, Zhou G, Zhong J, et al. Path planning for percutaneous lung biopsy based on the loose-Pareto and adaptive heptagonal optimization method. Med Biol Eng Comput 2023; 61:1449–1472.
10. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol. 9351. Cham: Springer International Publishing; 2015. p. 234–241.
11. Segars WP, Sturgeon G, Mendonca S, Grimes J, Tsui BMW. 4D XCAT phantom for multimodality imaging research. Med Phys 2010; 37: 4902–4915.
12. Wang Z, Gu S, Leader JK, et al. Optimal threshold in CT quantification of emphysema. Eur Radiol 2013; 23:975–984.
13. Li Y, Du Y, Yang HF, Yu JH, Xu XX. CT-guided percutaneous core needle biopsy for small (≤20 mm) pulmonary lesions. Clin Radiol 2013; 68: e43–e48.
14. Hiraki T, Mimura H, Gobara H, et al. Incidence of and risk factors for pneumothorax and chest tube placement after CT fluoroscopy–guided percutaneous lung biopsy: retrospective analysis of the procedures conducted over a 9-year period. AJR Am J Roentgenol 2010; 194: 809–814.
15. Zlevor AM, Mauch SC, Knott EA, et al. Percutaneous lung biopsy with pleural and parenchymal blood patching: results and complications from 1,112 core biopsies. J Vasc Interv Radiol 2021; 32:1319–1327.
16. Mills M, Choi J, El-Haddad G, et al. Retrospective analysis of technical success rate and procedure-related complications of 867 percutaneous CT-guided needle biopsies of lung lesions. Clin Radiol 2017; 72: 1038–1046.
17. Kuban JD, Tam AL, Huang SY, et al. The effect of needle gauge on the risk of pneumothorax and chest tube placement after percutaneous computed tomographic (CT)-guided lung biopsy. Cardiovasc Interv Radiol 2015; 38:1595–1602.

18. Saji H, Nakamura H, Tsuchida T, et al. The incidence and the risk of pneumothorax and chest tube placement after percutaneous CT-guided lung biopsy: the angle of the needle trajectory is a novel predictor. Chest 2002; 121:1521–1526.
19. Ko JP, Shepard JAO, Drucker EA, et al. Factors influencing pneumothorax rate at lung biopsy: are dwell time and angle of pleural puncture contributing factors? Radiology 2001; 218:491–496.
20. Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: review and new results. Stat Methods Med Res 2004; 13:251–271.
21. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 1973; 33:613–619.
22. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960; 20:37–46.
23. Kwiecien R, Kopp-Schneider A, Blettner M. Concordance analysis: part 16 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2011; 108:515–521.
24. Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. Psychol Bull 1968; 70:213–220.
25. Walker J. Non-inferiority statistics and equivalence studies. BJA Educ 2019; 19:267–271.
26. Davison AC, Hinkley DV. Bootstrap methods and their application, Vol. 1. Cambridge (United Kingdom): Cambridge University Press; 1997.
27. Chen D, Wang M, Li W, Li X, Hou H, Shan H. Computed tomography image: guided needle biopsy in the diagnosis of lung malignant tumors under artificial intelligence algorithm. Sci Program 2022; 2022:1–10.
28. Xu S, Fichtinger G, Taylor RH, Cleary KR. 3D motion tracking of pulmonary lesions using CT fluoroscopy images for robotically assisted lung biopsy. In: Galloway RL Jr, editor. Medical Imaging 2004: visualization, image-guided procedures, and display; 2004. San Diego, CA: SPIE; 2004. p. 394.
29. Challacombe B, Patriciu A, Glass J, et al. A randomized controlled trial of human versus robotic and telerobotic access to the kidney as the first step in percutaneous nephrolithotomy. Comput Aided Surg 2005; 10:165–171.
30. Zlevor AM, Kisting MA, Couillard AB, et al. Percutaneous CT-guided abdominal and pelvic biopsies: comparison of an electromagnetic navigation system and CT fluoroscopy. J Vasc Interv Radiol 2023; 34:910–918.
31. Grasso RF, Cazzato RL, Luppi G, et al. Percutaneous lung biopsies: performance of an optical CT-based navigation system with a low-dose protocol. Eur Radiol 2013; 23:3071–3076.
32. Ritter M, Rassweiler MC, Michel MS. The Uro Dyna-CT enables three-dimensional planned laser-guided complex punctures. Eur Urol 2015; 68:880–884.
33. Zhang Z, Shao G, Zheng J, et al. Electromagnetic navigation to assist with computed tomography-guided thermal ablation of liver tumors. Minim Invasive Ther Allied Technol 2020; 29:275–282.

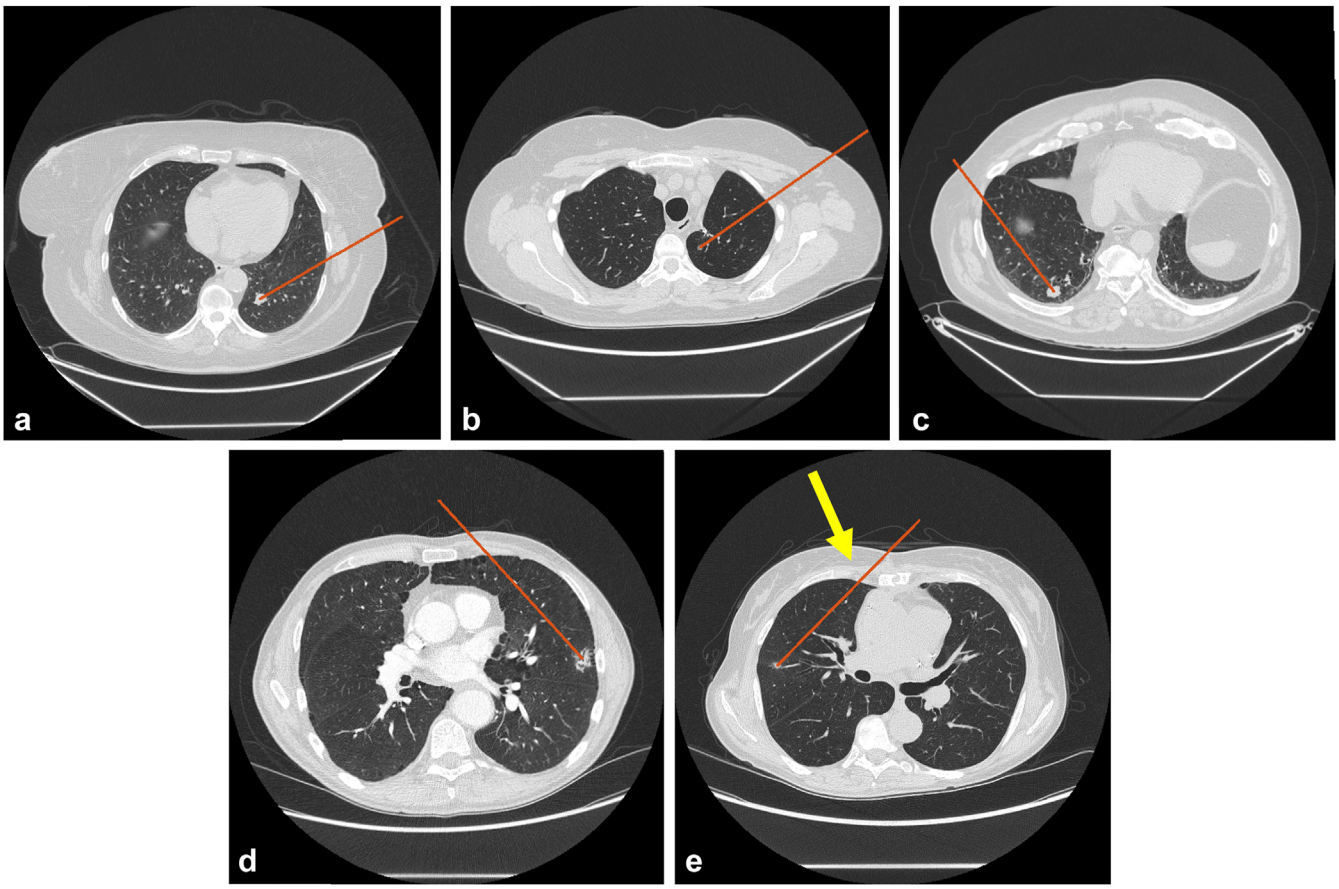

Figure E1. Cases considered safe by computer that were scored as unsafe by physicians. These 5 pathways met the criteria to receive a "safe" rating of 3 from the computer while receiving an "unsafe" physician rating (all received a score of 4). On physician review, all pathways **(a–e)** were determined to be excessively long. Additional reasons for unsafe ratings included the following: **(b)** traversal of vessels, **(d)** traversal of both a rib and an area of emphysema, and **(e)** traversal of vessels. The internal mammary artery (yellow arrow) was not included in the initial software version (Table 1), but this will be corrected in later versions.