



Diagnostic Accuracy of Percutaneous Transthoracic Needle Lung Biopsies: A Multicenter Study

Kyung Hee Lee, MD, PhD¹, Kun Young Lim, MD², Young Joo Suh, MD, PhD³, Jin Hur, MD, PhD^{3, 4}, Dae Hee Han, MD, PhD⁵, Mi-Jin Kang, MD⁶, Ji Yung Choo, MD, PhD⁷, Cherry Kim, MD, PhD⁷, Jung Im Kim, MD, PhD⁸, Soon Ho Yoon, MD⁹, Woojoo Lee, PhD¹⁰, Chang Min Park, MD, PhD^{9, 11}

¹Department of Radiology, Seoul National University Bundang Hospital, Seongnam, Korea; ²Department of Radiology, National Cancer Center, Goyang, Korea; ³Department of Radiology, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea; ⁴Research Institute of Radiological Science, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea; ⁵Department of Radiology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea; ⁶Department of Radiology, Inje University Sanggye Paik Hospital, Seoul, Korea; ⁷Department of Radiology, Korea University Ansan Hospital, Korea University College of Medicine, Ansan, Korea; ⁸Department of Radiology, Kyung Hee University Hospital at Gangdong, College of Medicine, Kyung Hee University, Seoul, Korea; ⁹Department of Radiology, Seoul National University College of Medicine, Seoul, Korea; ¹⁰Department of Statistics, Inha University, Incheon, Korea; ¹¹Institute of Radiation Medicine, Seoul National University Medical Research Center, Seoul, Korea

Objective: To measure the diagnostic accuracy of percutaneous transthoracic needle lung biopsies (PTNBs) on the basis of the intention-to-diagnose principle and identify risk factors for diagnostic failure of PTNBs in a multi-institutional setting.

Materials and Methods: A total of 9384 initial PTNBs performed in 9239 patients (mean patient age, 65 years [range, 20–99 years]) from January 2010 to December 2014 were included. The accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of PTNBs for diagnosis of malignancy were measured. The proportion of diagnostic failures was measured, and their risk factors were identified.

Results: The overall accuracy, sensitivity, specificity, PPV, and NPV were 91.1% (95% confidence interval [CI], 90.6–91.7%), 92.5% (95% CI, 91.9–93.1%), 86.5% (95% CI, 85.0–87.9%), 99.2% (95% CI, 99.0–99.4%), and 84.3% (95% CI, 82.7–85.8%), respectively. The proportion of diagnostic failures was 8.9% (831 of 9384; 95% CI, 8.3–9.4%). The independent risk factors for diagnostic failures were lesions ≤ 1 cm in size (adjusted odds ratio [AOR], 1.86; 95% CI, 1.23–2.81), lesion size 1.1–2 cm (1.75; 1.45–2.11), subsolid lesions (1.81; 1.32–2.49), use of fine needle aspiration only (2.43; 1.80–3.28), final diagnosis of benign lesions (2.18; 1.84–2.58), and final diagnosis of lymphomas (10.66; 6.21–18.30). Use of cone-beam CT (AOR, 0.31; 95% CI, 0.13–0.75) and conventional CT-guidance (0.55; 0.32–0.94) reduced diagnostic failures.

Conclusion: The accuracy of PTNB for diagnosis of malignancy was fairly high in our large-scale multi-institutional cohort. The identified risk factors for diagnostic failure may help reduce diagnostic failure and interpret the biopsy results.

Keywords: Image-guided biopsy; Sensitivity and specificity; Lung neoplasms; Multicenter study

INTRODUCTION

Percutaneous transthoracic needle lung biopsies (PTNBs) are useful procedures for patients with pulmonary lesions considered to be malignant. The accuracy of PTNBs for

diagnosis of malignancy has been reported to be high as 90–99% (1–8). However, diagnostic failures often occur, and they hinder proper and timely management of patients. Diagnostic failures of PTNBs can be divided into false-positive and false-negative results and non-evaluable

Received March 15, 2019; accepted after revision April 22, 2019.

This study was supported by a grant by Korea Health Industry Development Institute (HC15C3390).

Corresponding author: Chang Min Park, MD, PhD, Department of Radiology, Seoul National University College of Medicine, Institute of Radiation Medicine, Seoul National University Medical Research Center, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea.

• Tel: (822) 2072-7364 • Fax: (822) 763-0511 • E-mail: cmpark.morphius@gmail.com

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

results due to insufficient specimens (3, 4).

When biopsy results are non-evaluable due to insufficient specimens, the uncertainty over diagnoses remains unsolved, and additional medical procedures are required to confirm diagnoses. Therefore, non-evaluable results due to insufficient specimens should be considered as diagnostic failures irrespective of whether the final diagnosis is benign or malignant. According to the intention-to-diagnose principle, non-evaluable results due to insufficient specimens should be counted as false-negative findings when calculating sensitivity and false-positive findings when calculating specificity (9). Thus, an understanding of the actual, realistic diagnostic accuracy of PTNBs is essential for deciding whether to perform a biopsy or not. However, there is a possibility that the sensitivity and specificity of PTNBs are overestimated because the majority of the studies did not follow the intention-to-diagnose principle (9). Indeed, previous studies excluded non-evaluable results (1, 4, 8), regarded non-evaluable results as negative results (7), or did not mention how they handled non-evaluable results (5, 6) when they calculated the diagnostic accuracies.

To date, multiple studies have measured the diagnostic accuracy of PTNBs and investigated the risk factors for diagnostic failures (1-8, 10). Nonetheless, the generalizability of these studies is limited, because the majority were single-center studies (1-4, 6, 8, 10), had small numbers of patients (1, 2, 4, 8, 10), and limited patient inclusion to specific groups using particular guiding modalities (1-8, 10), biopsy needles (1-3, 6, 7, 10), and lesion sizes (2, 4). Moreover, the results of previous studies were inconsistent in relation to the effects of lesion size (3, 6, 8), lesion location (3, 4, 6), lesion depth (2, 6), lesion type (6, 10), biopsy needle (1, 4), number of tissue samplings (3, 6), and final diagnosis (3, 4, 8) on diagnostic failures. Therefore, this study aimed to measure the diagnostic accuracy based on the intention-to-diagnose principle and identify risk factors for diagnostic failure of PTNBs in a multi-institutional setting.

MATERIALS AND METHODS

Study Design and Study Population

This retrospective multi-institutional study was conducted by the Korean Registry of Percutaneous Transthoracic Needle Lung Biopsy group. Eight teaching hospitals in metropolitan Seoul (six tertiary and two secondary hospitals, with a

median bed number of 979, ranging from 555 to 2320) participated in this study. The study was approved by the Institutional Review Board of each institution, and the requirement for informed consent was waived.

We enrolled all consecutive patients at least 20 years of age who had undergone PTNBs from January 2010 to December 2014. During this period, 10568 consecutive PTNBs were performed in 9823 patients. Among the 10568 PTNBs, 1184 biopsies were excluded because they did not fulfill the predefined criteria for determining the final diagnosis ($n = 590$) or were repeat biopsies or re-biopsies of the same target lesion ($n = 594$). Finally, the remaining 9384 PTNBs on 9239 patients were included. In two recent studies, we used the same patient cohort to report the malignancy risk of non-diagnostic results from PTNBs and complications of PTNBs by using CT-guidance modalities (11, 12). Unlike these previous studies, the current study primarily focuses on measuring diagnostic accuracy and identifying risk factors for diagnostic failures of PTNBs.

Procedures

The biopsy procedures were performed in accordance with the routine clinical practice at each institution. The attending physicians requested the biopsy, and the final decision to perform PTNB was made at the discretion of dedicated thoracic radiologists in each institution. The selection of guiding modality and biopsy needle was made based on each institution's resources, preference of the performing thoracic radiologists, and target lesion characteristics. Most procedures were performed by attending radiologists. In cases wherein the procedures were performed by radiology residents or fellows, attending radiologists closely supervised the procedure. On-site cytopathology technologists were not available at any hospital.

Data Collection

Site investigators and trained research assistants recorded the variables related to patients, target lesions, and biopsy procedures. The patient variables were age, sex, familial history of lung cancer, smoking history (never smoker, former smoker, or current smoker) (13), and the presence of pulmonary emphysema on CT images. The target lesion variables included size (long-axis diameter on axial CT images), type (solid or subsolid), location (upper and middle lobes or lower lobe), and depth (distance from the pleura to the target). The procedure variables included

biopsy needle (fine needle aspiration [FNA] or core needle biopsy [CNB], or combined), guiding modality (fluoroscopy, conventional CT, CT fluoroscopy, cone-beam CT [CBCT], or ultrasonography), and the number of tissue samplings.

The original pathological reports of the PTNBs were also recorded. PTNB results were categorized as positive, negative, or non-evaluable due to insufficient specimens based on the original pathological reports, while the reference standard results were blinded. Malignancy, atypical cells suggestive of malignancy, atypical cells suspicious for malignancy, and atypical cells of indeterminate malignancy were considered positive, on the basis of our previous work (11). Specific benign, nonspecific benign pathological results, and atypical cells favoring benignity were considered negative. Non-evaluable results due to insufficient specimens was considered when the original pathologic reports indicated that the specimen was inadequate or insufficient for pathologic diagnosis.

Reference Standards

For each lesion, the final diagnosis was categorized as malignant or benign by site investigators using predefined criteria. The reference standard for each target lesion was identified as follows (3, 6). First, the diagnosis was based on a surgical pathologic report if the lesion was surgically resected. Second, it was based on pathologic analysis of a non-surgical biopsy of the lesion if it revealed a malignant or specific benign result. Because atypical cell results were not considered as specific diagnoses, a final diagnosis was not determined based on the atypical cell results. Third, the lesion was identified as benign when its diameter decreased by 20% or more or its size was stable for at least two years without treatment. Fourth, the lesion was identified as malignant when the clinical behavior showed obvious malignant processes. Lesions that did not fit the criteria were classified as having incomplete reference standards, and they were excluded from this study.

Statistical Analyses

The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for the diagnosis of malignancy were calculated with 95% confidence intervals (CIs) on a per-lesion basis. According to the intention-to-diagnose principle, non-evaluable results due to insufficient specimens were considered as false negatives when calculating sensitivity and false positives when calculating specificity (9).

All biopsies were categorized as diagnostic successes or diagnostic failures. Diagnostic successes consisted of the true-positive and true-negative results. Diagnostic failures consisted of the false-positive, false-negative, and non-evaluable results due to insufficient specimens. To determine the risk factors for diagnostic failures, univariable and multivariable analyses were performed by using a generalized linear mixed model. The institutions were entered into the model as a random effect to account for the clustering effect. In the univariable analyses, variables related to the patients, target lesions, and biopsy procedures, a final diagnosis of benignity/malignancy, and final diagnosis of lymphomas were tested based on clinical contexts. Significant variables with *p* values less than 0.05 in the univariable analyses were used as input variables for the multivariable analysis to identify independent risk factors. Because missing data consisted of only 3.3% (308 of 9384) of biopsy cases, multivariable analysis was performed based on a complete-case analysis. *P* values less than 0.05 were considered statistically significant. All statistical analyses were performed using STATA 15.0 (StataCorp, College Station, TX, USA).

RESULTS

Patient Demographics and Biopsy Procedures

The patient demographics and biopsy procedure characteristics grouped by institution are presented in Table 1. Mean patient age was 65 years (age range, 20–99 years) and the proportion of women was 38.0% (3510 of 9239). The main guiding modalities were fluoroscopy (used in two institutions), conventional CT (three institutions), CBCT (two institutions), and CT fluoroscopy (one institution).

Diagnostic Accuracy

Of the 9384 biopsies, 6774 (72.2%) showed positive results, 2170 (23.1%) showed negative results, and 440 (4.7%) were non-evaluable due to insufficient specimens. As for the final diagnosis, 7268 lesions were confirmed as malignant and 2116 lesions were confirmed as benign. Therefore, the overall accuracy, sensitivity, specificity, PPV, and NPV for diagnosis of malignancy were 91.1% (8553 of 9384; 95% CI, 90.6–91.7%), 92.5% (6723 of 7268; 95% CI, 91.9–93.1%), 86.5% (1830 of 2116; 95% CI, 85.0–87.9%), 99.2% (6723 of 6774; 95% CI, 99.0–99.4%), and 84.3% (1830 of 2170; 95% CI, 82.7–85.8%), respectively.

The diagnostic categories and accuracy stratified by

Table 1. Patient Demographics and Biopsy Procedure Characteristics

Variables	Total	Institution							
		A	B	C	D	E	F	G	H
Patient, No.	9239	186	1831	306	1425	429	3133	828	1101
Age (years)									
Mean (range)	65 (20–99)	65 (27–91)	65 (21–89)	64 (21–90)	65 (20–91)	71 (20–99)	64 (20–94)	65 (21–93)	65 (20–91)
Women	3510 (38.0%)	60 (32%)	674 (36.8%)	111 (36.3%)	552 (38.7%)	144 (33.6%)	1254 (40.0%)	319 (38.5%)	396 (36.0%)
Smoking history									
Never smoker	4082 (44.2%)	115 (61.8%)	690 (37.7%)	145 (47.4%)	642 (45.1%)	163 (38.0%)	1460 (46.6%)	392 (47.3%)	475 (43.1%)
Former smoker	3293 (35.6%)	19 (10.2%)	726 (39.7%)	66 (21.6%)	487 (34.2%)	130 (30.3%)	1047 (33.4%)	305 (36.8%)	513 (46.6%)
Current smoker	1763 (19.1%)	51 (27.4%)	393 (21.5%)	95 (31.1%)	278 (19.5%)	126 (29.4%)	603 (19.3%)	105 (12.7%)	112 (10.2%)
Familial history of lung cancer									
No	8635 (93.5%)	179 (96.2%)	1662 (90.8%)	301 (98.4%)	1351 (94.8%)	408 (95.1%)	2937 (93.7%)	798 (96.4%)	999 (90.7%)
Yes	450 (4.9%)	3 (1.6%)	118 (6.4%)	4 (1.3%)	59 (4.1%)	6 (1.4%)	172 (5.5%)	18 (2.2%)	70 (6.4%)
Lesions, No.	9384	188	1853	307	1442	432	3196	854	1112
Lesion size									
Median (IQR)	3.0 (2.0–4.5)	3.8 (2.4–5.4)	3.2 (2.3–4.7)	3.6 (2.5–5.6)	3.2 (2.2–4.7)	3.8 (2.6–5.5)	2.7 (1.8–4.0)	3.3 (2.3–4.9)	2.6 (1.8–4.0)
Biopsy needle									
FNA only	4194 (44.7%)	1 (1%)	1698 (91.6%)	0 (0%)	1366 (94.7%)	157 (36.3%)	63 (2%)	17 (2%)	892 (80.2%)
CNB or combined	5190 (55.3%)	187 (99.5%)	155 (8.4%)	307 (100.0%)	76 (5%)	275 (63.7%)	3133 (98.0%)	837 (98.0%)	220 (19.8%)
Main guiding modality		CBCT (96.8%)	Fluoroscopy (100.0%)	CT (86.3%)	CT (97.0%)	Fluoroscopy (93.1%)	CBCT (99.9%)	CT (99.8%)	CTF (75.2%)

Unless otherwise specified, data represent number of biopsies (and percentages). CBCT = cone-beam CT, CNB = core needle biopsy, CTF = CT fluoroscopy, FNA = fine needle aspiration, IQR = interquartile range

lesion size, lesion type, lesion depth, biopsy needle, and guiding modality are presented in Tables 2 and 3. Target lesions ≤ 2 cm in size had a sensitivity, specificity, PPV, and NPV of 88.4% (1508 of 1705; 95% CI, 86.8–89.9%), 85.0% (599 of 705; 95% CI, 82.1–87.5%), 98.8% (1508 of 1527; 95% CI, 98.1–99.2%), and 82.8% (599 of 723; 95% CI, 79.9–85.5%), respectively. The sensitivity, specificity, PPV, and NPV of subsolid pulmonary lesions were 86.1% (304 of 353; 95% CI, 82.1–89.6%), 85.7% (42 of 49; 95% CI, 72.8–94.1%), 99.0% (304 of 307; 95% CI, 97.2–99.8%), and 60.0% (42 of 70; 95% CI, 47.6–71.5%), respectively.

In comparison with FNAs, the use of CNBs resulted in a lower proportion of non-evaluable results due to insufficient specimens (1.3% with CNBs and 8.9% with FNAs). The sensitivity, specificity, PPV, and NPV of CNB were 94.9% (3833 of 4037; 95% CI, 94.2–95.6%), 96.1% (1108 of

1153; 95% CI, 94.8–97.1%), 99.7% (3833 of 3846; 95% CI, 99.4–99.8%), and 86.7% (1108 of 1278; 95% CI, 84.7–88.5%), respectively. The sensitivity, specificity, PPV, and NPV of FNA were 89.4% (2890 of 3231; 95% CI, 88.3–90.5%), 75.0% (722 of 963; 95% CI, 72.1–77.7%), 98.7% (2890 of 2928; 95% CI, 98.2–99.1%), and 80.9% (722 of 892; 95% CI, 78.2–83.5%), respectively.

Diagnostic Failures and Risk Factors

The proportions of PTNBs with diagnostic successes and diagnostic failures were 91.1% (8553 of 9384; 95% CI, 90.6–91.7%) and 8.9% (831 of 9384; 95% CI, 8.3–9.4%), respectively. The diagnostic success group (n = 8553) consisted of 6723 true-positive and 1830 true-negative results (Fig. 1). The diagnostic failure group (n = 831) consisted of 51 false-positive, 340 false-negative results,

Table 2. PTNB Diagnostic Categories for Diagnosis of Malignancy Stratified by Lesion Characteristics, Biopsy Needle, and Guiding Modality

Factor	TP	TN	FP	FN	Non-Evaluable*
Lesion size (cm)					
≤ 1.0	162 (57.9%)	85 (30.4%)	2 (0.7%)	16 (5.7%)	15 (5.4%)
1.1–2.0	1346 (63.2%)	514 (24.1%)	17 (0.8%)	108 (5.1%)	145 (6.8%)
2.1–3.0	1716 (70.7%)	508 (20.9%)	14 (0.6%)	71 (2.9%)	118 (4.9%)
> 3.0	3498 (77.0%)	722 (15.9%)	18 (0.4%)	145 (3.2%)	162 (3.6%)
Lesion type					
Solid	6418 (71.5%)	1788 (19.9%)	48 (0.5%)	312 (3.5%)	415 (4.6%)
Subsolid	304 (75.6%)	42 (10.5%)	3 (0.8%)	28 (7.0%)	25 (6.2%)
Lesion depth (cm)					
0	1710 (70.7%)	506 (20.9%)	14 (0.6%)	88 (3.6%)	102 (4.2%)
0.1–1.0	926 (70.1%)	264 (20.0%)	10 (0.8%)	57 (4.3%)	65 (4.9%)
1.1–2.0	1340 (71.9%)	350 (18.8%)	12 (0.6%)	57 (3.1%)	104 (5.6%)
2.1–3.0	1147 (73.4%)	288 (18.4%)	9 (0.6%)	55 (3.5%)	64 (4.1%)
> 3.0	1600 (72.2%)	422 (19.0%)	6 (0.3%)	83 (3.8%)	105 (4.7%)
Biopsy needle					
FNA only	2890 (68.9%)	722 (17.2%)	38 (0.9%)	170 (4.1%)	374 (8.9%)
CNB or combined	3833 (73.9%)	1108 (21.4%)	13 (0.3%)	170 (3.3%)	66 (1.3%)
Guiding modality					
Fluoroscopy	1906 (75.3%)	305 (12.1%)	24 (1.0%)	90 (3.6%)	206 (8.1%)
CT					
Conventional CT	1855 (72.9%)	519 (20.4%)	9 (0.4%)	76 (3.0%)	86 (3.4%)
CTF	512 (61.2%)	143 (17.1%)	6 (0.7%)	74 (8.8%)	102 (12.2%)
CBCT	2379 (70.5%)	848 (25.1%)	11 (0.3%)	96 (2.8%)	42 (1.2%)
Ultrasonography	71 (74.7%)	15 (15.8%)	1 (1.1%)	4 (4.2%)	4 (4.2%)
Total	6723 (71.6%)	1830 (19.5%)	51 (0.5%)	340 (3.6%)	440 (4.7%)

Unless otherwise specified, data represent number of biopsies (and percentages). *Non-evaluable due to insufficient specimens. FN = false negative, FP = false positive, PTNB = percutaneous transthoracic needle lung biopsy, TN = true negative, TP = true positive

and 440 non-evaluable results due to insufficient specimens. There was one patient who had false-positive PTNB among the 6222 PTNBs that showed malignancy. Although non-small cell carcinoma was suggested on PTNB, this patient was finally diagnosed as having a benign disease (pulmonary actinomycosis).

The diagnostic failure and success groups showed differences related to patient age, the presence of emphysema, lesion size, lesion type, biopsy needle, guiding modality, tissue sample numbers, final diagnosis of benignity/malignancy, and final diagnosis of lymphomas (all *p* values < 0.05) (Table 4). Lesion depth and lesion location did not significantly increase the risk of diagnostic failure. In the multivariable analysis, independent risk factors for diagnostic failures were lesion size ≤ 1.0 cm (adjusted odds ratio [AOR], 1.86; 95% CI, 1.23–2.81; *p* = 0.003), lesion size of 1.1–2.0 cm (AOR, 1.75; 95% CI, 1.45–2.11; *p* < 0.001), subsolid lesions (AOR, 1.81; 95% CI, 1.32–2.49; *p* < 0.001), FNA only (AOR, 2.43; 95% CI, 1.80–3.28; *p* < 0.001), final

diagnosis of benign lesions (AOR, 2.18; 95% CI, 1.84–2.58; *p* < 0.001), and final diagnosis of lymphomas (AOR, 10.66; 95% CI, 6.21–18.30; *p* < 0.001). CBCT (AOR, 0.31; 95% CI, 0.13–0.75; *p* = 0.009) and conventional CT-guidance (AOR, 0.55; 95% CI, 0.32–0.94; *p* = 0.03) reduced diagnostic failures when compared with the use of fluoroscopy.

A final diagnosis of lymphoma was associated with the highest proportion of diagnostic failures (36.2%, 25 of 69). Among the 29 lymphoma cases in which FNAs were performed, nearly half (*n* = 15) showed nonspecific benign results, two cases showed non-evaluable results due to insufficient specimens, and six cases showed atypical cell results. Only six were confirmed as lymphoma with FNA (Fig. 2). However, diagnoses were made in 75% (30 of 40) of lymphoma cases with CNB.

DISCUSSION

In this study, we measured the diagnostic accuracy in

Table 3. PTNB Accuracy for Diagnosis of Malignancy Stratified by Lesion Size, Lesion Type, Biopsy Needle, and Guiding Modality

Factor	Accuracy	Sensitivity	Specificity	PPV	NPV
Lesion size (cm)					
≤ 1.0	88.2 (83.8–91.7) [247/280]	87.1 (81.4–91.6) [162/186]	90.4 (82.6–95.5) [85/94]	98.8 (95.7–99.9) [162/164]	84.2 (75.6–90.7) [85/101]
1.1–2.0	87.3 (85.8–88.7) [1860/2130]	88.6 (86.9–90.2) [1346/1519]	84.1 (81.0–86.9) [514/611]	98.8 (98.0–99.3) [1346/1363]	82.6 (79.4–85.5) [514/622]
2.1–3.0	91.6 (90.5–92.7) [2224/2427]	93.4 (92.2–94.5) [1716/1837]	86.1 (83.0–88.8) [508/590]	99.2 (98.6–99.6) [1716/1730]	87.7 (84.8–90.3) [508/579]
> 3.0	92.8 (92.1–93.6) [4220/4545]	93.9 (93.1–94.7) [3498/3725]	88.0 (85.6–90.2) [722/820]	99.5 (99.2–99.7) [3498/3516]	83.3 (80.6–85.7) [722/867]
Lesion type					
Solid	91.4 (91.0–91.9) [8206/8981]	92.8 (92.2–93.4) [6418/6914]	86.5 (85.0–87.9) [1788/2067]	99.3 (99.0–99.5) [6418/6466]	85.1 (83.5–86.6) [1788/2100]
Subsolid	86.1 (82.3–89.3) [346/402]	86.1 (82.1–89.6) [304/353]	85.7 (72.8–94.1) [42/49]	99.0 (97.2–99.8) [304/307]	60.0 (47.6–71.5) [42/70]
Lesion depth (cm)					
0	91.6 (90.4–92.6) [2216/2420]	93.0 (91.7–94.1) [1710/1839]	87.1 (84.1–89.7) [506/581]	99.2 (98.6–99.6) [1710/1724]	85.2 (82.1–87.9) [506/594]
0.1–1.0	90.0 (88.3–91.6) [1190/1322]	91.6 (89.7–93.2) [926/1011]	84.9 (80.4–88.7) [264/311]	98.9 (98.0–99.5) [926/936]	82.2 (77.6–86.3) [264/321]
1.1–2.0	90.7 (89.3–92.0) [1690/1863]	93.1 (91.7–94.4) [1340/1439]	82.5 (78.6–86.0) [350/424]	99.1 (98.5–99.5) [1340/1352]	86.0 (82.2–89.2) [350/407]
2.1–3.0	91.8 (90.3–93.1) [1435/1563]	92.8 (91.2–94.2) [1147/1236]	88.1 (84.1–91.4) [288/327]	99.2 (98.5–99.6) [1147/1156]	84.0 (79.6–87.7) [288/343]
> 3.0	91.2 (90.0–92.4) [2022/2216]	91.8 (90.4–93.0) [1600/1743]	89.2 (86.1–91.9) [422/473]	99.6 (99.2–99.9) [1600/1606]	83.6 (80.0–86.7) [422/505]
Biopsy needle					
FNA only	86.1 (85.0–87.2) [3612/4194]	89.4 (88.3–90.5) [2890/3231]	75.0 (72.1–77.7) [722/963]	98.7 (98.2–99.1) [2890/2928]	80.9 (78.2–83.5) [722/892]
CNB or combined	95.2 (94.6–95.8) [4941/5190]	94.9 (94.2–95.6) [3833/4037]	96.1 (94.8–97.1) [1108/1153]	99.7 (99.4–99.8) [3833/3846]	86.7 (84.7–88.5) [1108/1278]
Guiding modality					
Fluoroscopy	87.4 (86.0–88.6) [2211/2531]	90.5 (89.2–91.7) [1906/2106]	71.8 (67.2–76.0) [305/425]	98.8 (98.2–99.2) [1906/1930]	77.2 (72.8–81.3) [305/395]
CT					
Conventional CT	93.3 (92.2–94.2) [2374/2545]	94.8 (93.7–95.7) [1855/1957]	88.3 (85.4–90.8) [519/588]	99.5 (99.1–99.8) [1855/1864]	87.2 (84.3–89.8) [519/595]
CTF	78.3 (75.3–81.0) [655/837]	80.1 (76.8–83.2) [512/639]	72.2 (65.4–78.3) [143/198]	98.8 (97.5–99.6) [512/518]	65.9 (59.2–72.2) [143/217]
CBCT	95.6 (94.8–96.3) [3227/3376]	95.6 (94.7–96.4) [2379/2488]	95.5 (93.9–96.8) [848/888]	99.5 (99.2–99.8) [2379/2390]	89.8 (87.7–91.7) [848/944]
Ultrasonography	90.5 (82.8–95.6) [86/95]	91.0 (82.4–96.3) [71/78]	88.2 (63.6–98.5) [15/17]	98.6 (92.5–100.0) [71/72]	78.9 (54.4–93.9) [15/19]
Total	91.1 (90.6–91.7) [8553/9384]	92.5 (91.9–93.1) [6723/7268]	86.5 (85.0–87.9) [1830/2116]	99.2 (99.0–99.4) [6723/6774]	84.3 (82.7–85.8) [1830/2170]

Unless otherwise specified, data represent percentages, % (95% CI) [numerator/denominator]. CI = confidence interval, NPV = negative predictive value, PPV = positive predictive value

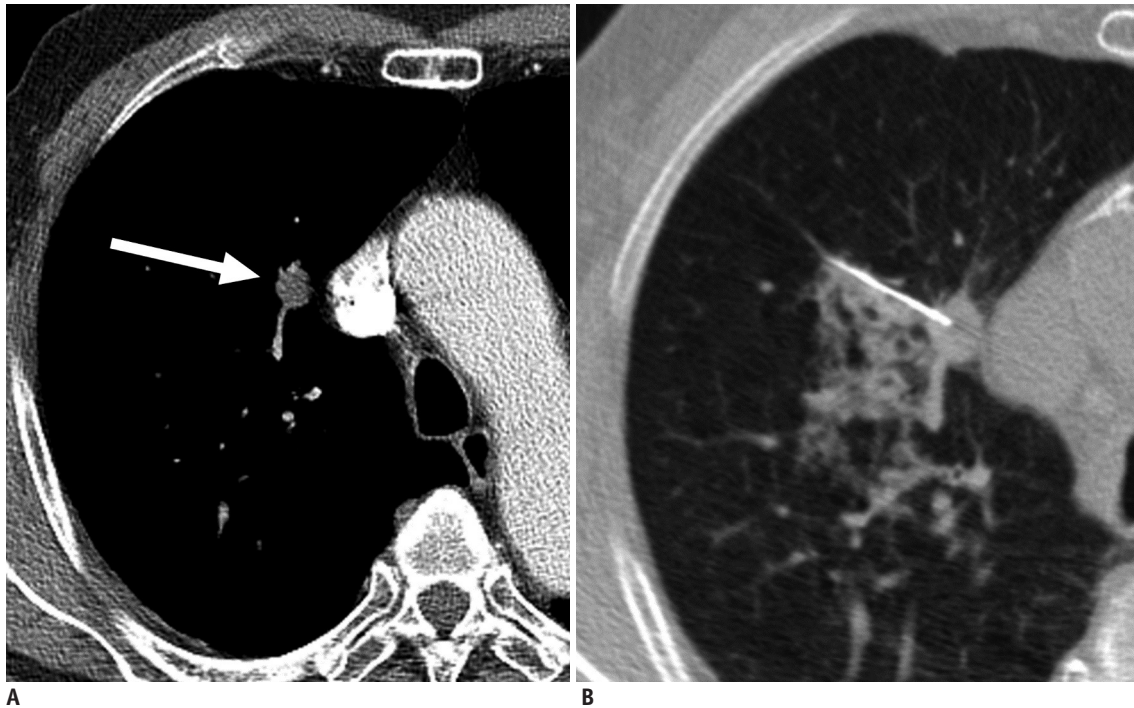


Fig. 1. Axial CT images in 78-year-old man with history of esophageal cancer.

A. Contrast-enhanced chest CT image shows 1.3-cm solid nodule in right upper lobe (arrow). **B.** CT-guided PTNB was performed by using 22-gauge aspiration needle, and pathologic examination showed adenocarcinoma of pulmonary origin (true-positive result). Patient underwent right upper lobectomy, and lesion was confirmed to be adenocarcinoma in pathologic examination. PTNB = percutaneous transthoracic needle lung biopsy

terms of the intention-to-diagnose principle and identified the risk factors of diagnostic failures in our dataset, which has the largest number of patients yet recorded in the literature. Overall, the accuracy, sensitivity, and specificity of PTNB for the diagnosis of malignancy were 91.1%, 92.5%, and 86.5%, respectively. Diagnostic failures occurred in 8.9% of the PTNB cases. The risk factors for diagnostic failures were target lesion size ≤ 2 cm, subsolid lesions, FNA only, a final diagnosis of benign lesions, and final diagnosis of lymphomas. Conventional CT and CBCT significantly lowered the proportion of diagnostic failures, when compared with fluoroscopy.

The specificity of PTNBs was 86.5%, which is much lower than the previously reported values of 98–100% (3–8). We think that the specificity in the previous studies was overestimated because the specificities in these studies were calculated by excluding the non-evaluable results (4, 5, 8) or regarding non-evaluable categories as negative results (7, 14). In this study, however, the specificity was calculated according to the intention-to-diagnose principle by including non-evaluable results due to insufficient specimens as false positives (9).

In a previous multicenter study (7), the NPV was only 51% when nonspecific benign pathologic results and

non-evaluable results due to insufficient specimens were regarded as negative results. Because the clinical meanings of nonspecific benign pathologic results and non-evaluable results due to insufficient specimens might be substantially different in actual clinical practice, we did not consider non-evaluable results due to insufficient specimens as negative results. Indeed, the proportion of final diagnosis of malignancy was much higher among the non-evaluable results due to insufficient specimens (46.6%) than that among the nonspecific benign results (21.3%). When we measured the NPV in the nonspecific benign category, the NPV was 78.7%.

Target lesion size ≤ 2 cm and the presence of subsolid pulmonary lesions were risk factors for diagnostic failures. These findings are consistent with studies that reported smaller lesions were associated with higher diagnostic failures (3, 6, 8), although the diagnostic accuracy of PTNBs in target lesions ≤ 2 cm in size remained at a relatively high level. Target lesions over 3 or 5 cm in size were not associated with high diagnostic failures in our study, in contrast to the previous studies (3, 15). Careful selection of the target area might have prevented diagnostic failures for large lesions with necrosis. With subsolid lesions, the NPV was only 60.0%, and the proportion of diagnostic failures

Table 4. Diagnostic Failures and Their Risk Factors

Variable	Percentage, % (95% CI) [Numerator/Denominator]	Diagnostic Failure		
		<i>P</i> *	Multivariable Analysis, AOR (95% CI)	<i>P</i> [†]
Age (years)		< 0.001		
≤ 65	9.9% (9.0–10.8%) [440/4451]		1 (Reference)	
> 65	7.9% (7.2–8.7%) [391/4933]		0.89 (0.76–1.04)	0.13
Sex		0.07	-	-
Male	8.6% (7.8–9.3%) [498/5822]			
Female	9.4% (8.4–10.4%) [333/3562]			
Emphysema		0.04		
No	8.7% (8.0–9.3%) [600/6936]		1 (Reference)	
Yes	9.4% (8.3–10.7%) [231/2447]		1.02 (0.86–1.22)	0.79
Lesion size (cm)		< 0.001		
≤ 1.0	11.8% (8.3–16.2%) [33/280]		1.86 (1.23–2.81)	0.003
1.1–2.0	12.7% (11.3–14.2%) [270/2130]		1.75 (1.45–2.11)	< 0.001
2.1–3.0	8.4% (7.3–9.5%) [203/2427]		1.03 (0.85–1.25)	0.77
> 3.0	7.2% (6.4–7.9%) [325/4545]		1 (Reference)	
Lesion type		0.002		
Solid	8.6% (8.1–9.2%) [775/8981]		1 (Reference)	
Subsolid	13.9% (10.7–17.7%) [56/402]		1.81 (1.32–2.49)	< 0.001
Biopsy needle		< 0.001		
FNA only	13.9% (12.8–15.0%) [582/4194]		2.43 (1.80–3.28)	< 0.001
CNB or combined	4.8% (4.2–5.4%) [249/5190]		1 (Reference)	
Guiding modality				
Fluoroscopy	12.6% (11.4–14.0%) [320/2531]		1 (Reference)	
CT		< 0.001		
Conventional CT	6.7% (5.8–7.8%) [171/2545]		0.55 (0.32–0.94)	0.03
CTF	21.7% (19.0–24.7%) [182/837]		0.88 (0.59–1.31)	0.52
CBCT	4.4% (3.7–5.2%) [149/3376]		0.31 (0.13–0.75)	0.009
Ultrasonography	9.5% (4.4–17.2%) [9/95]		0.88 (0.36–2.13)	0.78
No. of tissue samplings		0.01		
≤ 2	11.5% (10.7–12.3%) [675/5874]		1.07 (0.79–1.45)	0.65
≥ 3	4.4% (3.7–5.1%) [140/3203]		1 (Reference)	
Final diagnosis		< 0.001		
Benign	13.5% (12.1–15.0%) [286/2116]		2.18 (1.84–2.58)	< 0.001
Malignant	7.5% (6.9–8.1%) [545/7268]		1 (Reference)	
Lymphoma		< 0.001		
No	8.7% (8.1–9.2%) [806/9315]		1 (Reference)	
Yes	36.2% (25.0–48.7%) [25/69]		10.66 (6.21–18.30)	< 0.001

**p* values from univariable analyses, [†]*p* values from multivariable analysis. AOR = adjusted odds ratio

increased to 13.9%. Because part-solid nodules have a high probability of malignancy (16), our results support direct surgical resections in part-solid nodules based on CT assessments (17).

Although there was no control over the biopsy needle selection, the strength of our study was that the accuracies of FNA and CNB were compared in a multi-institutional setting. Since only small tissue samples can be obtained

with FNA, each institution's FNA accuracy may vary depending on the target lesions' characteristics, aspiration techniques, the presence of on-site pathologists, and the experience levels of the interpreting pathologists (18). In this study, CNB showed a much lower proportion of non-evaluable results due to insufficient specimens than did FNA, and FNA was found to be an independent risk factor for diagnostic failures. The specificity of PTNBs

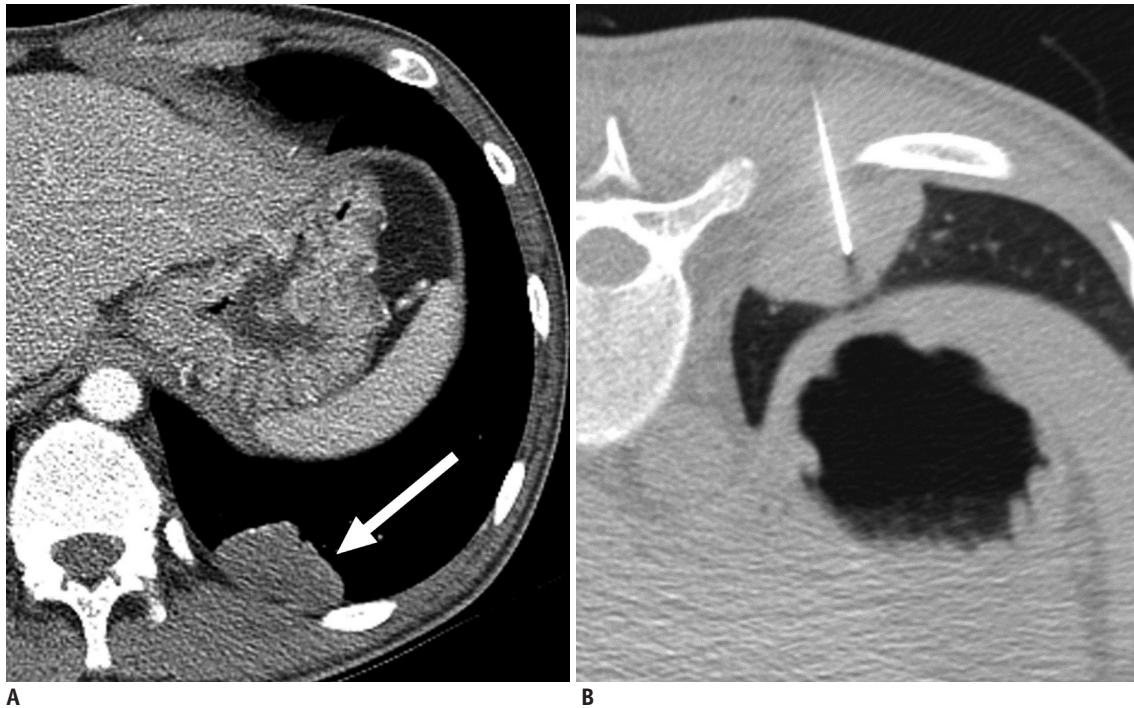


Fig. 2. Axial CT images in 21-year-old man diagnosed with acquired immunodeficiency syndrome. He was referred for incidentally detected pulmonary mass during treatment of *Pneumocystis jiroveci* pneumonia.
A. Contrast-enhanced chest CT image shows 4.0-cm mildly enhanced mass in left lower lobe (arrow). **B.** CT-guided PTNB was performed by using 22-gauge aspiration needle, and pathologic examination showed chronic inflammation (false-negative result). Patient underwent left lower lobectomy, and lesion was confirmed to be malignant lymphoma in pathologic examination.

was significantly higher with CNB (96.1%) than with FNA (75.0%). CNB can be a better option than FNA to reduce the risk of diagnostic failure of PTNB, especially for lung lesions with a high risk of diagnostic failure. Although we did not include the results of the molecular analysis in this study, CNB has been preferred over FNA to obtain adequate tissue for molecular analysis (19, 20). However, it should be noted that CNB has been associated with higher complication rates (21). Radiologists should choose the type of biopsy needles carefully in a given case, considering the possible impact of diagnostic failure and occurrence of complications.

The guiding modality is usually determined on the basis of accessibility, target lesion characteristics, and operator preferences. Each modality has its advantages and disadvantages; it is difficult to identify a single modality of choice for PTNBs. In this study, biopsies performed under CBCT were associated with the lowest percentage of diagnostic failures. However, there was no significant difference between CBCT and conventional CT. Because of the flexibility of CBCT with regard to orienting detectors around patients, this technique has shown high diagnostic accuracy even with small lesions (5). Nonetheless,

radiologists are exposed to fluoroscopic radiation during CBCT-guided biopsy. Conventional CT is accessible in most institutions and does not expose radiologists to radiation. However, conventional CT may be related to more patient radiation exposure than CBCT (22). The guidance under CT fluoroscopy did not show higher diagnostic performance when compared to fluoroscopy. We speculate that the high proportion of diagnostic failures under CT fluoroscopy may be institution-specific and related to lesion characteristics and aspiration techniques, but not an issue with CT fluoroscopy (11).

This study had the following limitations. First, it was retrospective, and the biopsy practice pattern differed across institutions depending on preferences, resources, and operators' experience. There was no control over the biopsy decision processes. Second, pre-test probabilities of malignancy and detailed biopsy indications were not evaluated in this retrospective study. However, the majority of the PTNBs were performed for the diagnosis of lung lesions considered to be malignant. Third, diagnostic failures or accuracies could not be evaluated in patients with incomplete reference standards for final diagnosis, although the proportion of such patients was small (5.9%).

Fourth, technical failure cases, operator-related factors, and procedure times were not evaluated. Fifth, no institution had on-site pathologists.

In conclusion, the accuracy of PTNB for diagnosis of malignancy was fairly high in our large-scale multi-institutional cohort. The identified risk factors for diagnostic failure may help reduce diagnostic failure and interpret the biopsy results.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

ORCID iDs

Chang Min Park
<https://orcid.org/0000-0003-1884-3738>
 Kyung Hee Lee
<https://orcid.org/0000-0003-2248-2525>
 Kun Young Lim
<https://orcid.org/0000-0002-8220-1476>
 Young Joo Suh
<https://orcid.org/0000-0002-2078-5832>
 Jin Hur
<https://orcid.org/0000-0002-8651-6571>
 Dae Hee Han
<https://orcid.org/0000-0002-5239-9207>
 Mi-Jin Kang
<https://orcid.org/0000-0001-8510-3268>
 Ji Yung Choo
<https://orcid.org/0000-0003-1244-0102>
 Cherry Kim
<https://orcid.org/0000-0002-3361-5496>
 Jung Im Kim
<https://orcid.org/0000-0001-6242-486X>
 Soon Ho Yoon
<https://orcid.org/0000-0002-3700-0165>
 Woojoo Lee
<https://orcid.org/0000-0001-7447-7045>

REFERENCES

- Staroselsky AN, Schwarz Y, Man A, Marmor S, Greif J. Additional information from percutaneous cutting needle biopsy following fine-needle aspiration in the diagnosis of chest lesions. *Chest* 1998;113:1522-1525
- Ohno Y, Hatabu H, Takenaka D, Higashino T, Watanabe H, Ohbayashi C, et al. CT-guided transthoracic needle aspiration biopsy of small (< or = 20 mm) solitary pulmonary nodules. *AJR Am J Roentgenol* 2003;180:1665-1669
- Hiraki T, Mimura H, Gobara H, Iguchi T, Fujiwara H, Sakurai J, et al. CT fluoroscopy-guided biopsy of 1,000 pulmonary lesions performed with 20-gauge coaxial cutting needles: diagnostic yield and risk factors for diagnostic failure. *Chest* 2009;136:1612-1617
- Choi SH, Chae EJ, Kim JE, Kim EY, Oh SY, Hwang HJ, et al. Percutaneous CT-guided aspiration and core biopsy of pulmonary nodules smaller than 1 cm: analysis of outcomes of 305 procedures from a tertiary referral center. *AJR Am J Roentgenol* 2013;201:964-970
- Choo JY, Park CM, Lee NK, Lee SM, Lee HJ, Goo JM. Percutaneous transthoracic needle biopsy of small (\leq 1 cm) lung nodules under C-arm cone-beam CT virtual navigation guidance. *Eur Radiol* 2013;23:712-719
- Lee SM, Park CM, Lee KH, Bahn YE, Kim JI, Goo JM. C-arm cone-beam CT-guided percutaneous transthoracic needle biopsy of lung nodules: clinical experience in 1108 patients. *Radiology* 2014;271:291-300
- Fontaine-Delaruelle C, Souquet PJ, Gamondes D, Pradat E, De Leusse A, Ferretti GR, et al. Negative predictive value of transthoracic core-needle biopsy: a multicenter study. *Chest* 2015;148:472-480
- Takeshita J, Masago K, Kato R, Hata A, Kaji R, Fujita S, et al. CT-guided fine-needle aspiration and core needle biopsies of pulmonary lesions: a single-center experience with 750 biopsies in Japan. *AJR Am J Roentgenol* 2015;204:29-34
- Schuetz GM, Schlattmann P, Dewey M. Use of 3x2 tables with an intention to diagnose approach to assess clinical performance of diagnostic tests: meta-analytical evaluation of coronary CT angiography studies. *BMJ* 2012;345:e6717
- Kim JI, Park CM, Kim H, Lee JH, Goo JM. Non-specific benign pathological results on transthoracic core-needle biopsy: how to differentiate false-negatives? *Eur Radiol* 2017;27:3888-3895
- Lee KH, Lim KY, Suh YJ, Hur J, Han DH, Kang MJ, et al. Nondiagnostic percutaneous transthoracic needle biopsy of lung lesions: a multicenter study of malignancy risk. *Radiology* 2019;290:814-823
- Yoon SH, Park CM, Lee KH, Lim KY, Suh YJ, Im DJ, et al. Analysis of complications of percutaneous transthoracic needle biopsy using CT-guidance modalities in a multicenter cohort of 10568 biopsies. *Korean J Radiol* 2019;20:323-331
- Lin HH, Ezzati M, Chang HY, Murray M. Association between tobacco smoking and active tuberculosis in Taiwan: prospective cohort study. *Am J Respir Crit Care Med* 2009;180:475-480
- Schreiber G, McCrory DC. Performance characteristics of different modalities for diagnosis of suspected lung cancer: summary of published evidence. *Chest* 2003;123(1 Suppl):115S-128S
- Yeow KM, Tsay PK, Cheung YC, Lui KW, Pan KT, Chou AS. Factors affecting diagnostic accuracy of CT-guided coaxial

- cutting needle lung biopsy: retrospective analysis of 631 procedures. *J Vasc Interv Radiol* 2003;14:581-588
16. Henschke CI, Yankelevitz DF, Mirtcheva R, McGuinness G, McCauley D, Miettinen OS; ELCAP Group. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. *AJR Am J Roentgenol* 2002;178:1053-1057
 17. Lee SM, Park CM, Song YS, Kim H, Kim YT, Park YS, et al. CT assessment-based direct surgical resection of part-solid nodules with solid component larger than 5 mm without preoperative biopsy: experience at a single tertiary hospital. *Eur Radiol* 2017;27:5119-5126
 18. Yao X, Gomes MM, Tsao MS, Allen CJ, Geddie W, Sekhon H. Fine-needle aspiration biopsy versus core-needle biopsy in diagnosing lung cancer: a systematic review. *Curr Oncol* 2012;19:e16-e27
 19. Kim H, Chae KJ, Yoon SH, Kim M, Keam B, Kim TM, et al. Repeat biopsy of patients with acquired resistance to EGFR TKIs: implications of biopsy-related factors on T790M mutation detection. *Eur Radiol* 2018;28:861-868
 20. Yoon HJ, Lee HY, Lee KS, Choi YL, Ahn MJ, Park K, et al. Repeat biopsy for mutational analysis of non-small cell lung cancers resistant to previous chemotherapy: adequacy and complications. *Radiology* 2012;265:939-948
 21. Heerink WJ, de Bock GH, de Jonge GJ, Groen HJ, Vliegenthart R, Oudkerk M. Complication rates of CT-guided transthoracic lung biopsy: meta-analysis. *Eur Radiol* 2017;27:138-148
 22. Abi-Jaoudeh N, Fisher T, Jacobus J, Skopec M, Radaelli A, Van Der Bom IM, et al. Prospective randomized trial for image-guided biopsy using cone-beam CT navigation compared with conventional CT. *J Vasc Interv Radiol* 2016;27:1342-1349