Radiographic and Clinical Characterization of False Negative Results from CT-Guided Needle Biopsies of Lung Nodules

Brian D. Gelbman, MD,* Mathew D. Cham, MD,† Won Kim, MD,* Daniel M. Libby, MD,* James P. Smith, MD,* Jeffrey L. Port, MD,† Nasser K. Altorki, MD,‡ Claudia I. Henschke, MD,† and David F. Yankelevitz, MD‡

Methods: We performed a consecutive series review of patients who had an initial benign result from an FNA between January 2002 and December 2004. Medical charts were reviewed to identify patients with false negative and true negative results and determine which variables were associated with a missed diagnosis. Previous work has reported that certain technical variables can affect the accuracy of a needle biopsy, such as the presence of a cytopathologist during the procedure, the size and type of needle (FNA versus core), and the number of passes.15,16 Our study sought to identify additional technical variables and clinical, radiologic, and pathologic findings that might assure the patient and physician that a benign biopsy result is truly indicative of a benign process. In addition, we analyzed the outcomes of patients with false negative results to determine if whether there was any clinical impact from a delay in diagnosis.

Subjects

With approval from the Institutional Review Board at Weill Cornell Medical center (Protocol # 0903010284), we reviewed the electronic medical records for all patients who underwent an FNA lung biopsy performed at New York Presbyterian...
Hospital, Weill Cornell Medical Center during a 3-year period (January 2002–December 2004).

To be included in the study, patients had to have an initial “benign” cytology result (meaning that no malignant cells were seen on cytologic review) and adequate clinical follow-up. For adequate clinical follow-up, the biopsied lesion must have been followed for at least 2 years by CT demonstrating resolution or no growth, or had a definitive surgical biopsy to determine the true nature of the lesion. Patients with initial cytology results that were classified as malignant (including atypical bronchoalveolar proliferation\textsuperscript{17} or bronchoalveolar carcinoma\textsuperscript{18,19} or those who underwent an FNA of a non-parenchymal lesion (e.g., mediastinum or rib) were excluded from the analysis.

Initial benign results were further classified into specific benign, nonspecific benign, and nondiagnostic. Specific benign is defined as a benign lesion (e.g., hamartoma and granuloma) or inflammatory cells with a positive bacterial, fungal, or mycobacterial culture that could explain the radiologic findings. Nonspecific benign is defined as the presence of normal respiratory elements (e.g., benign bronchial cells and macrophages) or blood. True negative cases were defined as those demonstrating CT stability for at least 2 years, resolution of the lesion of interest on follow-up imaging, or undergoing a surgical biopsy demonstrating a benign process.\textsuperscript{2,20} False negative biopsy cases were defined as those in whom the diagnosis of a malignant nodule was established by pathology from a subsequent surgical biopsy or repeat FNA.

For patients who were found to have a false negative biopsy, further clinical information was collected to determine the impact of the delay in diagnosis, the stage of the cancer at diagnosis, and the overall outcome of the patient.

### Needle-Biopsy Technique

For all biopsies, 1% lidocaine was administered into the subcutaneous tissues and pleural surfaces for local anesthesia. A General Electric Lightspeed CT scanner (GE Healthcare, Waukesha, WI) was used, with CT parameters set at 120kVp for tube voltage, 20 to 40 mAs for tube current, and 1.25 to 2.5 mm for slice thickness, depending on patient body habitus and nodule size. Under CT guidance, a 22-gauge Westcott needle (BD Medical, Franklin Lakes, NJ) was advanced into the pulmonary lesion. Samples were obtained and provided to an onsite cytologist and when a sufficient amount of specimen was available, to the Microbiology Laboratory when there was microscopic evidence of inflammation and no evidence of cancer. Additional biopsy samples were acquired using a new 22-gauge needle if the immediate cytologic evaluation revealed an insufficient sample. A postbiopsy CT image was obtained to document the development of a pneumothorax. All biopsies were performed by one of two radiologists.

### Data Collection

A retrospective chart review was performed to extract relevant clinical information using a standardized form for all benign biopsies. The clinical variables included prior medical conditions, underlying pulmonary disease, smoking status, number of pack years, history of prior malignancy, recent (within 3 months) positron emission tomography (PET) imaging, and reason for the initial CT which detected the lesion of interest. Reasons for the initial CT included pulmonary symptoms (e.g., cough and dyspnea), surveillance (patients with prior solid organ malignancies sent for yearly CT imaging), incidental (nodule noted on imaging performed for another reason, e.g., abdominal CT or routine chest radiograph), and screening (detect early stage lung cancer in asymptomatic high-risk subjects).

The radiologic variables analyzed for each patient included the size of the lesion of interest, the presence of other (multiple) lung nodules, and whether the nodules was solid, part-solid, or nonsolid in appearance.

CT images at the time of biopsy were retrospectively reviewed by a radiologist experienced in FNA, who was blinded to the outcome of the biopsy and who did not perform any of the biopsies. The reviewer collected information on the biopsy-related variables that may affect the results including the distance between the pleura and the nodule margin along the needle path, the number of passes with the needle, whether the needle tip appeared to be within the lesion (technical adequacy), and whether there were complications at any point during or immediately after the procedure (e.g., pneumothorax detectable by CT). In addition, as each adjustment of the needle was followed by an imaging series (typically five contiguous thin-section images), the total number of imaging series during the course of the procedure was recorded as an indirect measure of procedural difficulty, which can be because of nodule-related factors (e.g., size and location) or patient-related factors (e.g., motion and body habitus). Difficult-to-access nodules thus require more imaging attempts per needle pass. In addition, for each procedure the performing radiologist and cytologist were also recorded.

### Statistical Analysis

False negatives were compared with true negatives in terms of clinical, radiologic, biopsy-related, and pathologic variables. For categorical variables, a $\chi^2$ analysis was performed. For continuous variables (age, number of pack years, standardized uptake value (SUV) score on PET, and nodule size) a two-sided Student’s $t$ test was performed. As these variables are most likely not independent, a multivariate logistic analysis was performed to determine which variables jointly discriminated between false and true negatives. The statistical analysis was performed using SAS version 9.2.

### RESULTS

One hundred and seventy patients with an initial benign result and adequate clinical follow-up were included in the study. During the same time frame, 806 malignant diagnoses were rendered on the initial FNA biopsy, including 21 patients diagnosed with atypical bronchoalveolar proliferation (cytologically abnormal cells from pure bronchioloalveolar carcinoma). Among the 170 patients, 18 were “false negatives” as they were ultimately proved to be malignant, and 152 who...
Characterization of False Negative CT-FNA Biopsies

were “true negatives” (Figure 1). The ultimate diagnoses of these false negatives were lung cancer ($n = 12$), lymphoma ($n = 3$), and non-lung metastatic cancer ($n = 3$).

To identify the characteristics associated with the finding of a false negative, we compared the distribution of clinical, radiologic, and biopsy-related variables of the false negatives ($n = 18$) with that of the true negatives ($n = 152$) (Table 1). There were no significant differences in clinical variables such as age, sex, reason for initial CT scan, prior tobacco use or prior malignancy, between the two groups. The frequency of PET scans or average SUV score were not significantly different.

The radiologic variables revealed a significant difference in the size of the nodules, whereby false negatives were larger than the true negatives (mean, 27 mm versus 17 mm, respectively $p = 0.04$). False negatives were also more likely to be solitary (89% versus 69%) and less likely to be solid (78% versus 92%), although not statistically significant.

Comparison of the biopsy-related parameters showed that false negatives had fewer imaging adjustments per needle pass (4.5 versus 6.4, $p = 0.01$), a higher proportion of those in whom the needle tip could not be verified to be within the lesion (24% versus 7%, $p = 0.04$), and a higher pneumothorax rate (50% versus 22%, $p = 0.04$). The frequency of pneumothorax

TABLE 1. Characteristics Associated with False Negatives and True Negatives

<table>
<thead>
<tr>
<th></th>
<th>False Negatives ($n = 18$)</th>
<th>True Negatives ($n = 152$)</th>
<th>$p$ -value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age mean(range)</td>
<td>69 (range, 47–86)</td>
<td>65 (range, 21–89)</td>
<td>0.13</td>
</tr>
<tr>
<td>Sex M/F</td>
<td>8/10</td>
<td>74/78</td>
<td>0.73</td>
</tr>
<tr>
<td>Reason for initial CT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screening</td>
<td>2</td>
<td>22</td>
<td>0.74</td>
</tr>
<tr>
<td>Surveillance</td>
<td>3</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Symptoms</td>
<td>6</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Incidental</td>
<td>1</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>6</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>Smokers</td>
<td>11</td>
<td>100</td>
<td>0.69</td>
</tr>
<tr>
<td>Pack years mean(range)</td>
<td>44.9 (range, 15.0–75.0)</td>
<td>42.7 (range, 1.0–120)</td>
<td>0.81</td>
</tr>
<tr>
<td>Prior malignancy</td>
<td>8</td>
<td>58</td>
<td>0.60</td>
</tr>
<tr>
<td>Had PET scan</td>
<td>9</td>
<td>87</td>
<td>0.56</td>
</tr>
<tr>
<td>Average SUV mean</td>
<td>2.7 (range, 0–8.2)</td>
<td>3.8 (range, 0–14.5)</td>
<td>0.53</td>
</tr>
<tr>
<td>Solitary nodule</td>
<td>16</td>
<td>89</td>
<td>0.08</td>
</tr>
<tr>
<td>Solid nodule</td>
<td>14</td>
<td>78</td>
<td>0.07</td>
</tr>
<tr>
<td>Distance from pleura to nodule</td>
<td>21.2 (range, 0.0–77)</td>
<td>20.1 (range, 0.0–83)</td>
<td>0.83</td>
</tr>
<tr>
<td>Number of Passes</td>
<td>1.7 (range, 1.0–3.0)</td>
<td>1.7 (range, 1.0–4.0)</td>
<td>0.89</td>
</tr>
<tr>
<td># with needle tip in lesion</td>
<td>4.5 (range, 2.0–9.3)</td>
<td>6.4 (range, 1.0–16.0)</td>
<td>0.01</td>
</tr>
<tr>
<td>Positive Culture</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAI</td>
<td>1</td>
<td>6</td>
<td>0.04</td>
</tr>
<tr>
<td>Bacteria</td>
<td>0</td>
<td>8</td>
<td>0.90</td>
</tr>
<tr>
<td>Fungus</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

$^p$ values in bold and italics are significant.

CT, Computed tomography; PET, Positron emission tomography; MAI, Mycobacterium avium intracellulare.

Figure 1. Flow chart of CT-FNA biopsies from 2002–2004.
increased with increasing imaging attempts per needle pass ($p = 0.02$). Among the negative biopsy results for each radiologist, the proportion of false negatives varied between $8\%$ (11/131) and $18\%$ (7/39) ($p = 0.13$). Because biopsy-associated factors are interrelated, we performed a logistic regression analysis to determine which were significant. Logistic regression identified three variables which predicted a false negative result, pneumothorax ($p = 0.006$), solitary nodule ($p = 0.04$), and radiologist performing the procedure ($p = 0.04$). There was no significant difference between the reported cytology results between the patients with false negative results and true negatives, including the distribution of specific, nonspecific, and nondiagnostic results, which pathologist interpreted the specimen, and whether there were inflammatory, necrotic, or granulomatous features present.

Table 2 lists the reasons for the 18 missed diagnoses. In seven cases, the pathologist noted the presence of abnormal cells in the cytology specimen, but this was insufficient to render a malignant diagnosis. In three, later proven to be lymphomas, the cytology was believed to represent inflammation with a preponderance of lymphocytes being noted. In four, the needle tip was not documented to be inside the lesion. In seven cases, the pathologist noted the presence of abnormal cells in the cytology specimen, but this was insufficient to render a malignant diagnosis. In three, later proven to be lymphomas, the cytology was believed to represent inflammation with a preponderance of lymphocytes being noted. In four, the needle tip was not documented to be inside the lesion.
the lesion, acellular aspirate was seen in two cases, and another appeared to be a misread as an artifact during sample processing. In one case, the needle appeared to be in the postobstructive lung (Figure 2). This patient had the only positive biopsy culture among the false negatives, and that case grew Mycobacterium Avium Intracellulare, compared to 24 positive cultures among the true negatives (16%). Of the 18 false negatives, six had a repeat CT FNA which resulted in a diagnosis of malignancy: in one, the initial onsite cytological interpretation suggested fungal disease which prevented further passes; in another, the needle tip was not documented as being in the lesion. In the remaining four, the needle tip was documented to be in the lesion at the time of the biopsy and two passes were made. Subsequent needle biopsy in these four cases resulted in the diagnosis of metastatic renal carcinoma (bloody aspirate), lymphoma, adenocarcinoma in a nonsolid nodule, and adenocarcinoma in a solid nodule. In addition, for one patient, the cytology was reviewed 3 months later, which resulted in a correct diagnosis of adenocarcinoma.

Table 2 shows the initial and final clinical stages and the time between the FNA biopsy and surgery/repeat FNA. Median delay in diagnosis was 4.5 months, ranging from 0.2 to 70 months. Only one of the 18 patients had a higher clinical stage at diagnosis than the clinical stage of the initial biopsy (#1). This patient was clinical stage IB based on CT findings at the time of the FNA and at preoperative imaging and surgery 6 months later was found to have a malignant hilar lymph node, thus clinical and pathologic stage IIB at diagnosis.

DISCUSSION

Our first objective was to identify clinical, technical, pathologic, or radiologic characteristics that were associated with false negative results. Our time frame between 2002 and 2004 was selected to allow for adequate clinical follow-up after the biopsy.

The first variable that was significantly associated with false negative results was the size of the nodule. One might expect that larger lesions would be more accessible and therefore easier to biopsy and that smaller nodules would be more prone to missed diagnoses. However, our study demonstrated that false negatives tended to occur in larger lesions, a result noted previously. We speculate that the portion of the lesion with malignant cells was likely part of a larger consolidation, making it difficult to distinguish the primary lesion from surrounding atelectasis or inflammation. It may also be related to an increased amount of inflammation or necrosis in larger nodules.

Another variable that was significantly different in the false negative group was the number of imaging adjustments that were made for each needle pass during the biopsy. The true negative biopsies had on an average 6.4 imaging adjustments compared to 4.2 for the false negatives. One explanation for this could be more attentive and careful placement of the needle during the true negative biopsies, although documentation of the tip of the needle in the lesions should ultimately be the most important criteria. Another reason for the fewer imaging adjustments is that the larger lesion size of the false negative cases provided an easier target.

Our study revealed a significant association between the radiologist performing the procedure and the rate of false negatives, suggesting it is operator dependent. Although both radiologists were experienced with CT FNA, the operator with the lower incidence of false negatives had more years of experience with the procedure and performed more biopsies in general.

One limitation of our study is that all the biopsies were performed using BD Westcott single-shaft (non-coaxial) 22-gauge needles. The use of this single-shaft needle is based on the individual radiologist’s preference and experience. In a recent survey conducted among 139 thoracic and interventional radiologists, about half of the radiologists preferred this single-shaft needle over a coaxial needle. The commonly cited benefits of the single-shaft needle include a smaller caliber and increased needle flexibility. A commonly cited disadvantage of the single-shaft needle is the need for a new pleural puncture for each needle pass.

Previous reports have suggested that additional number of samples will improve the yield of this procedure. Our results showed that pneumothorax was the main limiting factor, as it limited the ability to put the needle tip in the lesion and also limited the number of passes. However, as six cases with an initial negative result underwent repeat FNA which yielded the diagnosis of malignancy, it does suggest that additional sampling at the time of the initial procedure might have yielded the correct result. The reason that additional samples were not obtained at the time of the procedure is not known as this was not specifically recorded. It could either be that the operator felt the amount of
tissue was sufficient or that the patient was unable or unwilling to continue to have additional samples obtained.

In the false negative group, there was one patient out of 18 (6%) with a positive culture, compared with the true negative group of which 24 of 152 patients (16%) had positive cultures. The positive culture in the false negative group occurred in a case where the postobstructed lung was suspected. Therefore, the predictive value that a positive culture from a benign biopsy is a true negative result is 97% (34 positive cultures that are true negative out of 35 positive cultures in total). Given the low likelihood of a false positive culture, one can conclude that the presence of a positive culture on an FNA (when there is no suspicion of postobstructive pneumonia) strongly supports the likelihood that this is truly a benign lesion.

One criticism of the FNA is that the procedure may be superfluous and delay resection. For the 18 false negatives cases, the median delay in diagnosis, from initial benign needle biopsy to definitive diagnosis of malignancy, was 4.5 months. A single case had a clinical stage IB and was found to be clinical stage IIB 6 months later. Previous studies have also demonstrated no harm to the patient when there is a delay in diagnosis provided that there is continued surveillance of an indeterminate lesion.\(^2,23,24\) The relatively large number of patients with true negative results suggests that a significant proportion of patients may have been spared unnecessary surgery. Therefore, the benefits of FNA, whereby patients avoid unnecessary surgery, may outweigh the harms (delay in diagnosis for a small percentage of patients, complication from the procedure).

In conclusion, this study demonstrates that factors associated with false negative results from an FNA include increased size of lung (which may have led to sampling error), fewer adjustments of the needle, the lack of positive cultures, and the occurrence of a pneumothorax. When a benign cytology result is obtained, the procedure should be reviewed to look for the presence of any of the above factors. False negative FNA in the diagnosis of pulmonary nodules is relatively uncommon and with careful monitoring rarely affects outcomes. We recommend that benign FNA biopsies have repeat imaging for at least 2 years to document stability or resolution of the lesion. If the nodule grows, then repeat biopsy or resection may be required to obtain a definitive diagnosis.

REFERENCES