Published in final edited form as:

Arch Pathol Lab Med. 2016 October ; 140(10): 1116–1120. doi:10.5858/arpa.2015-0316-OA.

Morphologic Accuracy in Differentiating Primary Lung Adenocarcinoma From Squamous Cell Carcinoma in Cytology Specimens

Maureen F. Zakowski, MD, Natasha Rekhtman, MD, Manon Auger, MD, Christine N. Booth, MD, Barbara Crothers, DO, Mohiddean Ghofrani, MD, Walid Khalbuss, MD, Rodolfo Laucirica, MD, Ann T. Moriarty, MD, Z. Laura Tabatabai, MD, and Güliz A. Barkan, MD

Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York (Drs Zakowski and Rekhtman); the Department of Pathology, Royal Victoria Hospital, Montreal, Quebec, Canada (Dr Auger); the Department of Pathology, Cleveland Clinic Main Campus, Cleveland, Ohio (Dr Booth); the Department of Pathology, Walter Reed Army Medical Center, Washington, DC (Dr Crothers); the Department of Pathology, Peacehealth Laboratory, Vancouver, Washington (Dr Ghofrani); the Department of Pathology, University of Pittsburgh Medical Center Shadyside, Pittsburgh, Pennsylvania (Dr Khalbuss); the Department of Pathology, Baylor College of Medicine, Houston, Texas (Dr Laucirica); the Department of Pathology, AmeriPath, Indianapolis, Indiana (Dr Moriarty); the Department of Pathology, University of California, San Francisco (Dr Tabatabai); and the Department of Pathology, Loyola University Medical Center, Maywood, Illinois (Dr Barkan). Dr Zakowski is no longer affiliated with Memorial Sloan Kettering Cancer Center. Dr Khalbuss is now with GE Clariant Diagnostic Services, Riyadh, Saudi Arabia. Dr Moriarty is no longer affiliated with AmeriPath

Abstract

Context—The National Cancer Care Network and the combined College of American Pathologists/International Association for the Study of Lung Cancer/Association for Molecular Pathology guidelines indicate that all lung adenocarcinomas (ADCs) should be tested for epidermal growth factor receptor (*EGFR*) mutations and anaplastic lymphoma kinase (*ALK*) rearrangements. As the majority of patients present at a later stage, the subclassification and molecular analysis must be done on cytologic material.

Objective—To evaluate the accuracy and interobserver variability among cytopathologists in subtyping non–small cell lung carcinoma using cytologic preparations.

Design—Nine cytopathologists from different institutions submitted cases of non–small cell lung carcinoma with surgical follow-up. Cases were independently, blindly reviewed by each cytopathologist. A diagnosis of ADC or squamous cell carcinoma was rendered based on the Diff-Quik, Papanicolaou, and hematoxylin-eosin stains. The specimen types included fine-needle aspiration from lung, lymph node, and bone; touch preparations from lung core biopsies; bronchial

Reprints: Güliz A. Barkan, MD, Department of Pathology, Loyola University Medical Center, 2160 South First Ave, Maywood, IL 60153 (gbarkan@lumc.edu).

Supplemental digital content is available for this article at www.archivesofpathology.org in the October 2016 table of contents.

The authors have no relevant financial interest in the products or companies described in this article.

washings; and bronchial brushes. A major disagreement was defined as a case being misclassified 3 or more times.

Results—Ninety-three cases (69 ADC, 24 squamous cell carcinoma) were examined. Of 818 chances (93 cases × 9 cytopathologists) to correctly identify all the cases, 753 correct diagnoses were made (92% overall accuracy). Twenty-five of 69 cases of ADC (36%) and 7 of 24 cases of squamous cell carcinoma (29%) had disagreement ($P = .16$). Touch preparations were more frequently misdiagnosed compared with other specimens. Diagnostic accuracy of each cytopathologist varied from 78.4% to 98.7% (mean, 91.7%).

Conclusion—Lung ADC can accurately be distinguished from squamous cell carcinoma by morphology in cytologic specimens with excellent interobserver concordance across multiple institutions and levels of cytology experience.

About 85% to 90% of primary lung cancers are non-small cell lung cancer (NSCLC). There are 2 main subtypes of NSCLC, adenocarcinoma (ADC) and squamous cell carcinoma (SqCC), which together account for most of the lung cancers currently diagnosed. Although there are many differences in the 2 subtypes, these tumors had been grouped together until very recently because the approach to treatment and prognosis had been very similar. It is now known that patients with ADC, but not SqCC, may have mutations in the epidermal growth factor receptor (EGFR) and show significant tumor shrinkage on *EGFR* tyrosine kinase inhibitors. The National Cancer Care Network guidelines¹ and the combined College of American Pathologists/International Association for the Study of Lung Cancer/Association for Molecular Pathology guidelines² indicate that all lung ADCs, metastatic or recurrent, should be tested for the presence of the activating mutations of *EGFR*, as well as for *ALK* rearrangements. In order to study samples for these targetable abnormalities, NSCLC must be accurately subclassified.

A current US Food and Drug Administration–approved standard treatment for NSCLC is carboplatin/paclitaxel/bevacizumab; however, based upon survival benefit, patients with SqCC should not receive bevacizumab because of a 30% mortality rate due to fatal hemoptysis.³ Furthermore, selection of therapies such as vascular endothelial growth factor receptor and *EGFR* inhibitors depends on the correct differentiation of SqCC from ADC of the lung.

Accurate morphologic separation of ADC and SqCC is necessary but can be difficult, particularly at the poorly differentiated end of the spectrum, where diagnostic features can overlap.⁴ This is partly due to the fact that up to 70% of lung cancers are diagnosed on limited material such as fine-needle aspiration (FNA) or transbronchial and transthoracic biopsy specimens.⁵ In addition, NSCLC is often heterogeneous and can exist in an adenosquamous subtype; one subtype may predominate in limited biopsy samples and provide misleading information in subtyping or treatment decisions.

In contrast to decades-old practice, pathologists now need to subclassify NSCLC into these 2 distinct subtypes.^{6–8} Accurate subtyping depends heavily on conventional histology and cytology, sometimes supplemented with immunohistochemical (IHC) stains. The majority of ADC and SqCC cases can be diagnosed by cytomorphology alone,⁹ but there remain

difficult cases that defy this approach. High interobserver agreement in distinguishing small cell lung cancer from NSCLC in cytologic material is well established,¹⁰ but less information is available on interobserver variability in subtyping NSCLC.

The aim of this study was to evaluate the accuracy of subtyping NSCLC based on morphology.

MATERIALS AND METHODS

Nine cytopathologists from 9 different institutions, both academic and nonacademic, were asked to contribute cytology specimen slide sets from patients in whom a diagnosis of NSCLC was rendered, where the morphologic subtype was known from matched surgical biopsy material or repeat cytology material. The study slides included Diff-Quik, Papanicolaou, and hematoxylineosin stains of FNAs from lung, lymph node, and bone. Some specimens were obtained by endobronchial ultrasound. Some of the cases consisted mainly of core needle biopsy touch preparations (TPs) and FNA needle rinses prepared by the ThinPrep method.

Cases were collected and independently reviewed by each cytopathologist without knowledge of the histologic or repeat cytologic diagnoses. No prestudy test set was given and no instruction regarding diagnostic criteria or cytologic characteristics of ADC or SqCC took place before the review. The slides were not prescreened for this exercise, but some of the slides still had the original cytotechnologists' screening ink dots. The cytology volume and practice characteristics of each institution were not included in this study.

The cytopathologists were asked to distinguish ADC from SqCC based on their own practice patterns. Basically, the pathologists sat around a table, each with his or her own microscope, and passed the cases around by study set. All cases were reviewed at one sitting. The cytopathologists were asked to indicate whether they thought the specimens were ADC or SqCC and to state if the diagnosis was made by morphology alone or if IHC was needed to determine the subtype. Determination was made of how many cases were unanimously and correctly called ADC or SqCC. A major disagreement was defined as a case being misclassified 3 or more times, and a disagreement was considered minor when a case was misclassified fewer than 3 times.

There was no attempt made to apply the International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification to the diagnoses of these cases as it was considered irrelevant to the study.

Two-tailed *P* values were calculated using an online Fisher exact test Web site, <http://www.langsrud.com/fisher.htm>. Interobserver variability was calculated using Fleiss κ (R software version 3.2.3 and irr package version 0.84; <https://www.r-project.org>).

RESULTS

The 9 sets comprised a total of 93 cases, and the number of cases in a set ranged from 9 to 15; the average number of slides per set was 10. The slides included FNA of lung (59), FNA

of lymph nodes (15), FNA of bone (1), bronchial brushing (2), bronchial washing (4), and TP from core biopsy (1). There were 69 ADC cases and 24 SqCC cases. Immunohistochemical stains, mostly p63 and TTF-1, were available in 36 of the 93 cases (38%). Not all cytopathologists saw all 93 cases because of time constraints. All 9 cytopathologists reviewed at least 7 different sets of cases (77%). Six of 9 cytopathologists (66%) saw all 9 sets, 2 (22%) saw 8 sets, and 1 (11%) saw 7 sets of slides. Therefore, the total number of times all 93 slides were seen was 6. Tallying all the cases that were seen by any pathologist resulted in 818 opportunities to make a correct diagnosis.

The Supplemental Table (see supplemental material file at www.archivesofpathology.org in the October 2016 table of contents) shows the diagnosis given by the cytopathologists, the years of experience, as well as specimen type and presence/absence of available IHC stains. The experience of the cytopathologists ranged from 6 to 22 years. To determine overall group accuracy, simple calculations were used. Of 818 chances to correctly identify all the cases, 753 correct diagnoses were made, resulting in 91.7% overall group accuracy, or accuracy of all the cytologists calculated together.

Twenty-five of 69 cases of ADC (36%) had some level of disagreement. Seven of 24 cases of SqCC (29%) had some level of disagreement. There was no significant disagreement in diagnosing ADC compared with SqCC, with $P = .16$. Of the cases with some level of disagreement, 6 of 69 cases of ADC (9%) had major disagreements and 3 of 24 cases of SqCC (12%) had major disagreements; $P = .25$; see Figures 1 through 4 showing minor and major disagreements.

Of 472 chances to correctly diagnose ADC, 50 cases (11%) were called SqCC; in other words, 420 of 472 cases of ADC were accurately diagnosed. The group accuracy in diagnosing ADC was thus 89%. Of 346 chances to correctly identify SqCC, 15 cases (4%) were called ADC; in other words, 331 of 346 cases of SqCC were accurately diagnosed, for a group accuracy of 95% in diagnosing SqCC; $P = .004$.

Immunohistochemical stains were included in 36 cases; 25 of these cases were ADC and 11 were SqCC. We examined the role of these IHC stains in making accurate diagnoses by asking the pathologists to indicate if the IHC stains were used in arriving at a diagnosis. The cytopathologists indicated that IHC was helpful or necessary in reaching a diagnosis in 81 of 376 cases, or 21% of the time. In some cases where IHC was available, the pathologists indicated that it was not needed or useful in making their diagnosis. In those cases where IHC was not available (442) there were 53 misclassifications, for an error rate of 12% (53 of 442). There were 5 errors in the 376 cases where IHC was available but not used in diagnosis, for an error rate of 13% (5 of 376). The lowest error rate (0%) was in those cases where IHC was available and was used to classify the tumors.

Looking at diagnostic accuracy by grouping cases according to the presence or absence of IHC stains provided the following results: the total number of correct diagnoses where IHC was included with the conventional stains was 371 (371 of 376; 98%) and the total number of correct diagnoses where IHC was not available was 385 (385 of 442; 87%), $P < .001$. The interobserver variability among the 9 cytopathologists for morphology-only cases was

slightly higher (Fleiss $\kappa = .617$), compared with that of cases with IHC (Fleiss $\kappa = .464$). The slightly low κ value may be attributed to the cases where IHC was available but not used for diagnosis.

This study thus also reaffirmed the suitability of IHC stains in cytologic material.

Diagnostic accuracy of each reviewing cytopathologist varied from 78.4% to 98.7% (mean, 91.7%). This did not seem to be related to years of experience, although the number of pathologists was too small for a statistically meaningful data analysis to be performed on this.

DISCUSSION

Because of differences in mutation profiles and therapeutic responses, there is heightened emphasis on the accurate pathologic subtyping of NSCLC into ADC and SqCC. The accuracy of such subtyping in cytologic material is not well established,¹¹ particularly in reference to newer IHC stains such as p40, TTF-1, and napsin A. A few studies are available in the literature,^{4,10–15} but most are reports from single institutions with relatively small numbers of cases. One large series of 1182 transbronchial and transthoracic aspirates demonstrated that the positive predictive value of FNA in subtyping NSCLC was 92% for ADC and 82% for SqCC.¹⁵ The Table outlines previous publications in this area.

The distinction between ADC and SqCC using cytologic material varies, but the overall accuracy is good. The greatest accuracy was achieved after the introduction of targeted therapy and with the use of IHC stains,^{4,10,13–15} suggesting that the criteria necessary to separate ADC from SqCC may not have been carefully applied prior to the recognition of the increased need for subtyping.

We have analyzed the largest interinstitutional series of 93 cases of NSCLC diagnosed in a ringlike fashion by 9 cytopathologists from 9 different laboratories. Our results show an overall accuracy rate of 92%. This is a remarkably high rate of accuracy, especially when considering that the cases originated from different laboratories with different handling, processing, and staining methods. As in other reports, our accuracy rate increased with the use of IHC stains. Some recently reported series⁵ show cytologic accuracy ranging from 64% to 96% before the application of IHC markers such as TTF-1 and p63; in our series the specificity rate was 100% with the use of IHC. In one series, IHC was used less frequently in cytologic than in histologic specimens.¹ This may be due to the routine use of up to 3 stains in cytology: Papanicolaou, Diff-Quik, and hematoxylin-eosin, each highlighting different diagnostic features. For example, the Papanicolaou stain, which has Orange G, serves to highlight the keratin of squamous cells. Similarly, the Diff-Quik stain highlights extracellular material, such as mucin, which may be helpful in diagnosing ADC. Histology generally uses hematoxylin-eosin stains only. Lack of sufficient cytologic material, such as cell blocks for IHC, may also play a role.

Immunohistochemical stains were included in 36 cases: 25 of 69 total ADC cases and 11 of 24 total SqCC cases. In other words, 36% of ADC cases had IHC and 45% of SqCC cases had IHC. This was not statistically significant at $P = .47$. This suggests equal difficulty in

diagnosing ADC and SqCC. The group accuracy results of 95% for ADC and 89% for SqCC are consistent with the findings of Sigel et al,¹¹ who reported that the agreement between cytology specimens and surgical biopsies was higher for ADC than for SqCC. Sigel et al¹¹ also note that lack of IHC was considered a contributing factor to misdiagnoses, and that IHC was used more often in biopsy material than in cytologic material, perhaps supporting our impression that the use of 3 stains in cytology is superior to the use of 1 stain in histology. Manucha et al¹² also found that accuracy in diagnosing ADC was greater than in SqCC in unspecified cytology specimens, as did Cho et al¹⁶ when reporting on cytology fluid specimens.

In our study IHC was used twice as often in ADC cases as in SqCC cases. We considered that this might be due to the need to distinguish epithelial lesions metastatic to the lung from primary NSCLC. A positive TTF-1 stain is used to confirm both ADC and lung origin and is useful in ruling out metastasis from nonpulmonary sites. Another consideration was that when keratinizing, SqCC can be a straightforward diagnosis, although information regarding the degree of keratinization was not included in our series.

The type of cytology specimen did seem to influence the diagnostic accuracy, although we did not attempt to statistically confirm this observation. A larger study would be needed to confirm this impression. As can be seen in the Supplemental Table, specimens submitted from institution 2, which consisted mainly of core needle biopsy TPs with needle rinses prepared by the ThinPrep method, were more frequently misdiagnosed than cases from other laboratories. This might be due to the lack of available hematoxylin-eosin and Diff-Quik stains in the TPs, which are Papanicolaou-stained preparations, or possibly due to less experience using TPs by the cytopathologists from the other laboratories. Again, a larger series is needed to confirm this impression.

CONCLUSIONS

Lung ADC can accurately be distinguished from SqCC in cytologic material with excellent interobserver concordance across multiple institutions and levels of cytology experience. Immunohistochemical stains improve the accuracy of diagnosis, which, in this study, was greater for SqCC than for ADC.

Certain types of cytology specimens were associated with greater error rates, specifically cytology TPs of core biopsies. In fact, the use of TPs from core biopsies of lung specimens can result in the deletion of cellularity and DNA content of the cores.¹⁷ Because molecular analysis is vital to treatment in lung cancer, this study discourages the use of these preparations for subclassification of ADC and SqCC of the lung.

Excellent accuracy is obtained when trained cytopathologists interpret the material, but levels of experience did not seem to affect outcome.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Demetri G, Elias A, Gershenson D, et al. NCCN small-cell lung cancer practice guidelines: the National Comprehensive Cancer Network. *Oncology* (Williston Park). 1996; 10(11 suppl):179–194. [PubMed: 8953602]
2. Lindeman NI, Cagle PT, Beasley MB, et al. Molecular testing guideline for selection of lung cancer patients for EGFR and ALK tyrosine kinase inhibitors: guideline from the College of American Pathologists, International Association for the Study of Lung Cancer, and Association for Molecular Pathology. *Arch Pathol Lab Med*. 2013; 137(6):828–860. [PubMed: 23551194]
3. Whithaus K, Fukuoka J, Prihoda TJ, Jagirdar J. Evaluation of napsin A, cytokeratin 5/6, p63, and thyroid transcription factor 1 in adenocarcinoma versus squamous cell carcinoma of the lung. *Arch Pathol Lab Med*. 2012; 136(2):155–162. [PubMed: 22288962]
4. da Cunha Santos G, Lai SW, Saieg MA, et al. Cyto-histologic agreement in pathologic subtyping of non small cell lung carcinoma: review of 602 fine needle aspirates with follow-up surgical specimens over a nine year period and analysis of factors underlying failure to subtype. *Lung Cancer*. 2012; 77(3):501–506. [PubMed: 22656670]
5. Rekhtman N, Brandt SM, Sigel CS, et al. Suitability of thoracic cytology for new therapeutic paradigms in non-small cell lung carcinoma: high accuracy of tumor subtyping and feasibility of EGFR and KRAS molecular testing. *J Thorac Oncol*. 2011; 6(3):451–458. [PubMed: 21266922]
6. Gridelli C. Histology-based treatment: a new scenario in the management of advanced nonsmall cell lung cancer. *Curr Opin Oncol*. 2009; 21(2):97–98. [PubMed: 19532008]
7. Scagliotti G, Hanna N, Fossella F, et al. The differential efficacy of pemetrexed according to NSCLC histology: a review of two Phase III studies. *Oncologist*. 2009; 14(3):253–263. [PubMed: 19221167]
8. Tiseo M, Bartolotti M, Gelsomino F, Ardizzoni A. First-line treatment in advanced non-small-cell lung cancer: the emerging role of the histologic subtype. *Expert Rev Anticancer Ther*. 2009; 9(4): 425–435. [PubMed: 19374597]
9. Kimbrell HZ, Gustafson KS, Huang M, Ehya H. Subclassification of non-small cell lung cancer by cytologic sampling: a logical approach with selective use of immunocytochemistry. *Acta Cytol*. 2012; 56(4):419–424. [PubMed: 22846197]
10. Omland SH, Henrik H, Olsen EK, Birthe T, Guldhammer SB. Subtyping of nonsmall cell lung cancer on cytology specimens: reproducibility of cytopathologic diagnoses on sparse material. *Diagn Cytopathol*. 2014; 42(2):105–110. [PubMed: 23687077]
11. Sigel CS, Moreira AL, Travis WD, et al. Subtyping of non-small cell lung carcinoma: a comparison of small biopsy and cytology specimens. *J Thorac Oncol*. 2011; 6(11):1849–1856. [PubMed: 21841504]
12. Manucha V, Wang C, Huang Y. Non-small-cell lung carcinoma subtyping on cytology without the use of immunohistochemistry—can we meet the challenge? *Acta Cytol*. 2012; 56(4):413–418. [PubMed: 22846184]
13. Righi L, Graziano P, Fornari A, et al. Immunohistochemical subtyping of nonsmall cell lung cancer not otherwise specified in fine-needle aspiration cytology: a retrospective study of 103 cases with surgical correlation. *Cancer*. 2011; 117(15):3416–3423. [PubMed: 21246522]
14. Ocque R, Tochigi N, Ohori NP, Dacic S. Usefulness of immunohistochemical and histochemical studies in the classification of lung adenocarcinoma and squamous cell carcinoma in cytologic specimens. *Am J Clin Pathol*. 2011; 136(1):81–87. [PubMed: 21685035]
15. Nizzoli R, Tiseo M, Gelsomino F, et al. Accuracy of fine needle aspiration cytology in the pathological typing of non-small cell lung cancer. *J Thorac Oncol*. 2011; 6(3):489–493. [PubMed: 21258246]
16. Cho A, Hur J, Hong YJ, et al. NSCLC subtype prediction using cytologic fluid specimens from needle aspiration biopsies. *Am J Clin Pathol*. 2013; 139(3):309–316. [PubMed: 23429366]
17. Rekhtman N, Kazi S, Yao J, et al. Depletion of core needle biopsy cellularity and DNA content as a result of vigorous touch preparations. *Arch Pathol Lab Med*. 2015; 139(7):907–912. [PubMed: 25521802]

Figure 1. Fine-needle aspiration of lung. Adenocarcinoma called squamous cell carcinoma 3 times and favor squamous cell carcinoma 2 times. No immunostains available. Major disagreement (Papanicolaou, original magnification $\times 40$).

Figure 2. Fine-needle aspiration of lung. Adenocarcinoma called squamous cell carcinoma 1 time and favor squamous cell carcinoma 1 time. No immunostains available. Minor disagreement (Diff-Quik, original magnification $\times 60$).

Figure 3. Fine-needle aspiration of lung. Cell block section. Adenocarcinoma called squamous cell carcinoma 2 times. No immunostains available. Minor disagreement (hematoxylin-eosin, original magnification $\times 60$).

Figure 4.
Fine-needle aspiration of lung. Adenocarcinoma called squamous cell carcinoma 2 times.
No immunostains available. Minor disagreement (Diff-Quik, original magnification $\times 60$).

Table

Previous Studies Outlining Accuracy of Cytologic Subclassification of Non–Small Cell Carcinoma of the Lung

Type of Specimen	Cytology Accuracy Compared With Histology	No. of Cases	Source, y
FNA, BB, BW, BL	96%–100% with IHC	128	Rekhtman et al, ¹⁷ 2015
FNA	85%	158	Nizzoli et al, ¹⁵ 2011
FNA cell blocks	64%–87% using IHC	103	Righi et al, ¹³ 2011
FNA, BB, BW, BL	93%; IHC needed less frequently in cytology	101	Sigel et al, ¹¹ 2011

Abbreviations: BB, bronchial brushing; BL, bronchial lavage; BW, bronchial washing; FNA, fine-needle aspiration; IHC, immunohistochemistry.