Published in final edited form as:

J Am Soc Cytopathol. 2018 ; 7(3): 154–459. doi:10.1016/j.jasc.2018.02.001.

Fine-Needle Aspiration–Based Grading of Pancreatic Neuroendocrine Neoplasms Using Ki-67: Is Accurate WHO Grading Possible on Cytologic Material?

Vivian L. Weiss, MD PhD^a, Colleen Kiernan, MD^b, Jesse Wright, MD^b, Nipun B. Merchant, MD^b, Alice C. Coogan, MD^a, and Chanjuan Shi, MD PhD^a

^aDepartment of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center, Nashville, TN, 37232, United States

^bDepartment of Surgery, Surgical Oncology; Vanderbilt University Medical Center, Nashville, TN, 37232, United States

Abstract

Introduction—The World Health Organization (WHO) has provided grading guidelines for pancreatic neuroendocrine neoplasms (PanNENs) based on mitotic count and Ki-67 proliferation index. Due to the desire to provide earlier tumor grading for clinical management decisions, some groups have proposed grading PanNENs at the time of fine needle aspiration (FNA) using Ki-67 proliferation rates. While a Ki-67 can be performed on FNA cell blocks, there are potential sampling limitations with this technique that may affect the reliability of the Ki-67 result.

Materials and Methods—Forty-nine PanNENs with FNA cell blocks and corresponding resection material were evaluated by immunohistochemistry for expression of Ki-67. Ki-67 proliferation rate was calculated based on cell counts of >500 cells in the highest staining areas. Ki-67 scores from FNA cell blocks were correlated with Ki-67 scores from resection specimens.

Results—The FNA Ki-67 proliferation rates overall did not correlate well with the resection specimen. A linear regression analysis of the correlation between FNA %Ki-67 and resection %Ki-67 showed a slope of 3.2 and an $R^2=0.58$. The average difference in Ki-67 proliferation rate between FNA and resection was 5.9%. Thirty-nine percent (19/49 cases) of PanNENs showed discordant grading between the FNA cell block and resection specimen. Almost all (18/19) discordant cases demonstrated a lower FNA-based grade as compared to the resection grade.

Corresponding Author: Vivian Weiss, MD, PhD, Vanderbilt University School of Medicine, Department of Pathology, Microbiology and Immunology, Medical Center North, CC-2213, 1161 21st Ave. S. Nashville, TN 37232-2561, Phone: 615-875-3002, Fax: 615-343-7023, vivian.l.weiss@vanderbilt.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author contributions: Vivian Weiss and Chanjuan Shi reviewed all FNA slides and immunohistochemical stains, performed tumor grading and wrote the manuscript. Jesse Wright and Colleen Kiernan provided the clinical data. Nipun Merchant, and Alice Coogan reviewed the manuscript.

Conclusions—FNA cell block grading using Ki-67 frequently led to under-grading of the tumor. This finding is consistent with concerns that FNA may not provide accurate grading due to the limited sampling of the tumor.

Keywords

Pancreatic neuroendocrine neoplasm; Ki-67; fine needle aspiration; WHO grade

Introduction

Pancreatic neuroendocrine neoplasms (PanNENs) are rare tumors that represent approximately 2% of pancreatic primary tumors¹. For many years it has been felt that both the stage and grade of these tumors are important for predicting recurrence and metastasis. The mitotic count and Ki-67 proliferative index have been consistently found to be markers of patient prognosis². As a result, The European Neuroendocrine Tumour Society (ENETS) developed a grading system of PanNENs based on the mitotic count and Ki-67 proliferative index³. The World Health Organization (WHO) adapted these recommendations for the 2010 classification⁴. Under this current classification system, grade 1 neuroendocrine tumor is defined by <2 mitoses per 2mm² in 10 high-power fields and/or Ki-67 labeling index <3%; grade 2 is defined as 2-20 mitoses per 2mm² and/or 3-20% Ki-67 labeling index; and neuroendocrine carcinoma is defined by >20 mitoses per 2mm² and a Ki-67 proliferative index >20%. The Ki-67 proliferative index should be defined by the area with the highest percent nuclear staining and recommends at least 500-2000 nuclei to be counted.

Standard of care for pancreatic lesions now involves the use of endoscopic ultrasound-guided fine needle aspiration (EUS-FNA) to sample the lesion and provide an accurate diagnosis. However, there has been much debate as to whether these small FNA biopsies are sufficient for the accurate grading of PanNENs. Given the small amount of material aspirated in these biopsies, obtaining an accurate representation of heterogeneous and large neoplasms can be challenging. On the other hand, surgical resection and resection-based grading of the primary tumor may not always be feasible. A handful of studies have attempted to address the concordance of FNA-based grading and surgical resection-based grading.

To date, the majority of studies of PanNEN EUS-FNA have demonstrated concordance between the FNA grading using Ki-67% and the resection grading. One study of 22 PanNENs demonstrated 86% concordance between FNA grading and resection grading and suggested that EUS-FNA samples could provide accurate grading of the tumor⁵. Of the 22 PanNENs evaluated, 3 showed discordance of the FNA grading with 2/3 being underestimated of the grade. Unno et al evaluated 19 PanNENs and showed that 13 patients were concordant with only 6 patients that had discrepant FNA grading. They argue that the discrepant cases had significantly larger tumor size⁶.

We propose that a larger study is needed to evaluate an increased number of cases using a more systematic and reproducible Ki-67% evaluation method. We present 49 patients with both PanNEN FNA cell blocks and resection specimens to evaluate the concordance between FNA-based PanNEN grading and surgical resection-based grading. We propose that

using larger case numbers, as well as systematic and manageable counting regimens, will help to more accurately predict the concordance of PanNEN FNA-based grading in clinical practice.

Materials and Methods

Patient Selection

In this Institutional Review Board approved study, paired FNA cell blocks and resection material were retrieved from Vanderbilt University Medical Center pathology archives between 1998 and 2015. Forty-nine patients were identified in the archives with both adequate FNA cell block and resection material. FNA cell blocks were considered adequate if they contained at least 500 tumor cells. Clinical information including the patient age, gender, whether syndromes were diagnosed, and outcome were recorded. Tumor characteristics including WHO grade, size, functionality, presence of lymphovascular invasion, presence of perineural invasion, necrosis, degree of infiltration, and metastases were recorded.

Hematoxylin and eosin stained 4um sections from corresponding cell block and surgical resection specimens were reviewed by two pathologists (VW, CS) to confirm the cytopathologic and histopathologic diagnoses. Tumor differentiation was assessed using the 2010 WHO classification. On cytologic specimens, the diagnosis was confirmed using neuroendocrine markers including synaptophysin and chromogranin.

Immunohistochemistry

Four micron sections from both the FNA and corresponding resection material were evaluated by immunohistochemistry for expression of Ki-67. Ki-67 staining was performed according to manufacturer's protocols. FNA cell blocks were also stained using CD45 to confirm that the Ki-67 labeled cellularity represented tumor cellularity rather than tumor-infiltrating immune cells.

Unstained slides were first deparaffinized by routine methods. Antigen retrieval was performed as follows: slides were heated in citrate buffer (pH 6.0) at 100°C for 20 minutes or in EDTA (pH 9.0) at 98°C for 20 minutes, followed by a 10-minute cool down to room temperature. All slides were then quenched with 0.03% (v/v) H₂O₂ with NaN₃ for 5 minutes. Slides were blocked for 20 minutes with serum-free protein block (Dako, Carpinteria, CA), followed by application of primary antibodies. Slides were then incubated with Evison+HRP-labeled polymer (Dako) for 30 minutes, followed by 5 minute incubation with DAB (Dako).

Analysis

For the cases with abundant tumor cells in the cell block, Ki-67 immunostains were scanned under low power for areas with the highest nuclear labeling (hotspots). Only the hotspots were evaluated, and one to three 20× or 40× (original magnification: 200× or 400×) fields (containing >1000 tumor cells) were evaluated. The same selection method was also used for evaluating Ki67 index in resection specimens. For other cases with few tumor cells in the

cell blocks (approximately 500-1,000 cells), entire cell blocks were chosen to evaluate for Ki-67 index, with up to ten 20× or 40× fields being examined. Labeling for CD45 was also reviewed to ensure that the Ki-67 positive cells were tumor cells rather than lymphocytes present in the cell block. FNA cell blocks with less than 500 tumor cells following CD45 labeling, the number of tumor cells recommended by the WHO for accurate grading, were not included in this study.

Images were taken from the hotspots or entire cell blocks and printed in order to allow for practical and reproducible cell counting. The images were reviewed blindly by two pathologists (VW, CS) and each tumor cell was judged for positivity and marked/counted accordingly.⁷ Ki-67 scores from FNA cell blocks were correlated with Ki-67 scores from resection specimens. Statistical analysis was performed using student's t-test and Fisher Exact test.

Results

Patient characteristics

Forty-nine patients with PanNENs were evaluated in this study that had adequate FNA cell block material as well as a surgical resection. The average patient age was 57 years, with a range of 29-82 years. The cohort consisted of 22 males and 27 females. The average tumor size was 3.0 cm. Three of the tumors in this study were functional. Eight patients had multiple endocrine neoplasia type 1 (MEN1) syndrome and 3 patients had von Hippel-Lindau (VHL).

The majority of patients in the study had low stage, low grade tumors. Twenty-nine patients had stage I disease, 14 patients had stage II disease, and 6 patients had stage IV disease at the time of diagnosis. Twenty-three patients had grade 1 tumors, 19 patients had grade 2 tumors, and 7 patients had grade 3 tumors based on surgical resection Ki-67 proliferative index.

Thirty-five percent (17/49) of tumors were positive for lymphovascular invasion and only 16% (8/49) had perineural invasion. Thirty-seven percent (18/49) had infiltrative borders and 4% (2/49) showed necrosis on surgical resection.

Comparison of Ki-67 labeling of FNA and Resection Material

The FNA Ki-67 proliferation rates overall were not representative of the proliferation rate identified in the resection specimen (Table 1). A linear regression analysis of the correlation between FNA %Ki-67 and resection %Ki-67 showed a slope of 3.2 and an $R^2=0.58$ (Figure 1). The average difference in Ki-67 proliferation rate between FNA and resection was 5.9%. Thirty-nine percent (19/49 cases) of PanNENs showed discordant grading between the FNA cell block and resection specimen. Of the 19 discordant cases, 18 demonstrated a lower FNA-based grade as compared to the resection grade. Eleven cases were called grade 1 on FNA but grade 2 on resection, one case was grade 1 on FNA but grade 3 on resection, 6 cases were grade 2 on FNA but grade 3 on resection, and one case was grade 2 on FNA but grade 1 on resection (Table 2). Upon careful re-review of the discordant cases and comparison of the FNA cell blocks with the resection specimen, the difference in

proliferative index was most likely due to sampling (Figure 2). The average tumor size for discordant cases was 3.6 ± 0.5 cm, while the average tumor size for concordant cases was only slightly smaller at 2.7 ± 0.3 cm ($p=0.12$).

There was no difference in the Ki-67 index between cases with low versus high cell block cellularity; Ki67 index in the cases with 3 fields (high cellularity, $n=30$) evaluated and in those with 4 fields evaluated (low cellularity, $n=19$) was $6.6 \pm 2.2\%$ and $9.5 \pm 2.7\%$, respectively ($p=0.41$). In addition, a large Ki67 variation was not uncommon even within a single section from resection specimens; however, Ki-67 variability in cell blocks was not as prominent.

Correlation of Ki-67 labeling and patient outcome

The average patient follow-up in this study was 51 months, with the majority of patients alive with no evidence of disease (NED, 32/49, 65%). Sixteen percent (8/49) of patients were alive with disease at follow-up. Ten percent (5/49) died of their disease at an average of 28 months following surgical resection. Two patients died of other cancers and two patients died of surgical complications within a month of surgical resection.

Patients with stage II-IV tumors and grade 2-3 tumors had significantly worse survival and disease recurrence than those patients with Stage 1 and grade 1 tumors. Patients with resection grade 1 tumors had 89.5% disease-free survival compared with 57.7% disease-free survival with resection grade 2-3 tumors ($p<0.05$). Similarly, patients with FNA grade 1 tumors had 83.3% disease-free survival compared with 46.7% disease-free survival for FNA grade 2-3 tumors ($p<0.05$). While there is a trend toward slightly lower disease free survival for FNA-based grading when compared with resection-based grading, there are not enough patients in this study for further meaningful outcome-based statistical analysis.

Discussion

In a larger, more systematic and comprehensive study, we provide strong evidence against using FNA cell blocks to assess Ki-67 proliferation rates to determine grading. FNA cell block grading using Ki-67 labeling index frequently under-graded the tumors in our cohort. There are multiple factors potentially responsible for these discordant proliferation rates including tumor heterogeneity, tumor size, and lymphocytic infiltration of tumors.

While most studies thus far of FNA-based PanNEN grading have reported that this is a practical and useful technique to grade tumors prior to resection, a few studies have indicated that there might be flaws in this approach. A study of 33 PanNENs found good correlation for grade 3 tumors but discrepancies between histologic and cytologic grading in grade 2 tumors. As with other groups, FNAs appeared to underestimate the grade when evaluating grade 2 tumors⁸. Despite the discrepancies, that study still recommended EUS-FNA grading as an important part of the work-up for these lesions. A larger study by Rebours et al evaluated 48 PanNENs and demonstrated that while FNA grading of liver metastases correlated with resection grading, FNAs of primary pancreatic lesions did not fare as well. In fact, the pancreatic FNAs underestimated the grade in up to 30% of patients⁹. The authors suggested that this discordant grading was related to the size of the tumor.

Some groups have suggested that tumor heterogeneity and size may result in inaccurate grading of FNA biopsies. While Unno et al demonstrated that the discordant cases had significantly larger tumor sizes⁶, the size difference of tumors in this cohort was not statistically different. The heterogeneity of PanNENs has been well-described, with hotspots of Ki-67 labeling. In fact, both WHO and ENETS recommend that the area of highest Ki-67 labeling within the resection be used for prognosis. Hasegawa et al, in an attempt to resolve these discrepancies, recommended counting at least 2000 cells in an FNA cell block prior to estimating Ki-67 index. They studied 27 FNA cases and showed a concordance rate of 74% when counting all available cells, compared to a concordance rate of 90% when excluding all cases with less than 2000 cells in the cell block¹⁰. While this sampling cutoff improves the concordance rate, many cell blocks in clinical practice fail to contain 2000 cells. In addition, the ability of pathologists to accurately and efficiently count 2000 tumor cells in clinical practice is uncertain. Per WHO recommendations, we have counted 500 cells per cell block in order to perform accurate grading. Cases without sufficient tumor cells in the cell blocks were rare at our institution and were not included in this study. As aspirate smears are generally more cellular, the possibility also exists to stain smears for Ki-67 rather than cell blocks. While this may be useful for cases without sufficient cell block material, it would make CD45 staining difficult on the same slide. In addition, as the likely cause of the FNA vs resection discrepancy is the inherent challenge of limited FNA sampling, the use of FNA aspirate smears for Ki-67 labeling would not prevent this inaccuracy in grading.

Conclusions

In the current study we used a practical counting method which required a camera-captured and printed image in which each counted cell was judged for positivity and marked accordingly. This method is one of the most reproducible and cost-effective counting methodologies.⁷ At least 500 cells were counted in each cell block as per WHO recommendations on surgical specimens. Unlike using the “eye-ball” technique, this method provides a reproducible strategy for cell block assessment that can be performed in routine clinical practice. Despite this careful, but reproducible, approach, cytologic sampling frequently under-graded the PanNENs in our cohort. This finding is consistent with concerns that FNA may not provide accurate grading due to the limited sampling of the tumor. Additional larger studies are needed to determine if this under-grading will lead to improper treatment and worse patient outcomes. Until more detailed outcome-data is known, we do not recommend FNA-based grading of PanNENs; instead, we propose that the careful evaluation of the PanNEN resection specimens provides a more accurate and definitive WHO grade.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Sources: This study was supported by the National Institutes of Health [NIH/NIDDK 5P30 DK058404-13; NIH/NCI 5P50 CA095103-13]

References

- 1Fitzgerald TL, Hickner ZJ, Schmitz M, Kort EJ. Changing incidence of pancreatic neoplasms: a 16-year review of statewide tumor registry. *Pancreas*. 2008; 37:134–8. [PubMed: 18665072]
- 2La Rosa S, Klersy C, Uccella S, et al. Improved histologic and clinicopathologic criteria for prognostic evaluation of pancreatic endocrine tumors. *Hum Pathol*. 2009; 40:30–40. [PubMed: 18715612]
- 3Falconi M, Bartsch DK, Eriksson B, et al. ENETS Consensus Guidelines for the management of patients with digestive neuroendocrine neoplasms of the digestive system: well-differentiated pancreatic non-functioning tumors. *Neuroendocrinology*. 2012; 95:120–34. [PubMed: 22261872]
- 4Bosman FT, , Carneiro F, , Hruban RH, , Theise ND. *Pathology and Genetics of Tumors of the Digestive System Vol. 3*. Lyon, France: IARC Press; 2010World Health Organization Classification of Tumors.
- 5Farrell JM, Pang JC, Kim GE, Tabatabai ZL. Pancreatic neuroendocrine tumors: accurate grading with Ki-67 index on fine-needle aspiration specimens using the WHO 2010/ENETS criteria. *Cancer Cytopathol*. 2014; 122:770–8. [PubMed: 25044931]
- 6Unno J, Kanno A, Masamune A, et al. The usefulness of endoscopic ultrasound-guided fine-needle aspiration for the diagnosis of pancreatic neuroendocrine tumors based on the World Health Organization classification. *Scand J Gastroenterol*. 2014; 49:1367–74. [PubMed: 25180490]
- 7Reid MD, Bagci P, Ohike N, et al. Calculation of the Ki67 index in pancreatic neuroendocrine tumors: a comparative analysis of four counting methodologies. *Mod Pathol*. 2015; 28:686–94. [PubMed: 25412850]
- 8Weynand B, Borbath I, Bernard V, et al. Pancreatic neuroendocrine tumour grading on endoscopic ultrasound-guided fine needle aspiration: high reproducibility and inter-observer agreement of the Ki-67 labelling index. *Cytopathology*. 2014; 25:389–95. [PubMed: 24750272]
- 9Rebours V, Cordova J, Couvelard A, et al. Can pancreatic neuroendocrine tumour biopsy accurately determine pathological characteristics? *Dig Liver Dis*. 2015; 47:973–7. [PubMed: 26169284]
- 10Hasegawa T, Yamao K, Hijioka S, Bhatia V, Mizuno N, Hara K, Imaoka H, Niwa Y, Tajika M, Kondo S, Tanaka T, Shimizu Y, Kinoshita T, Kohsaki T, Nishimori I, Iwasaki S, Saibara T, Hosoda W, Yatabe Y. Evaluation of Ki-67 in EUS-FNA specimens for the assessment of malignancy risk in pancreatic neuroendocrine tumors. *Endoscopy*. 2014;32–8. [PubMed: 24218309]

Highlights

- A comprehensive and systematic evaluation of FNA-based WHO grading of pancreatic neuroendocrine neoplasms using Ki-67.
- Thirty-nine percent (19/49 cases) of PanNENs showed discordant grading between the FNA cell block and resection specimen.
- FNA-based grading of PanNENs is not recommended; instead, careful evaluation of the PanNEN resection specimens to obtain an accurate and definitive WHO grade is advised

Figure 1.
Linear regression analysis of the correlation between FNA %Ki-67 and resection %Ki-67.

Figure 2.

Ki67 labelling cytology and resection specimens from 2 PanNENs. Case 1-cytology specimen with a Ki67 index of <3% (A, WHO grade 1) and resection with a Ki67 index of <3% in a cold spot (B) and 15% in a hot spot (C, WHO grade 2). Case 2-cytology specimen with a Ki67 index of 5% (D, WHO grade 2) and resection with a Ki67 index of 5% in a cold spot (E) and >20% in a hot spot (F, WHO grade 3).

Table 1

FNA and Resection Grading. The Ki-67% and WHO grade for both the FNA cell block and resection specimen are presented for 49 tumors.

SPECIMEN NUMBER	FNA Ki67 (%)	RESECTION Ki67 (%)	FNA GRADE	RESECTION GRADE
1	0.0	4.4	1	2
2	0.0	2.0	1	1
3	0.9	5.6	1	2
4	1.5	15.8	2	2
5	0.0	4.0	1	2
6	0.0	22.9	1	3
7	1.8	1.0	1	1
8	0.0	0.3	1	1
9	0.0	0.9	1	1
10	1.5	9.4	1	2
11	1.6	1.0	1	1
12	5.0	22.1	2	3
13	1.2	2.3	1	2
14	1.5	1.2	1	1
15	1.6	1.0	1	1
16	1.7	3.4	1	2
17	3.0	30.0	2	3
18	17.4	64.7	2	3
19	2.3	2.9	1	2
20	3.2	21.0	2	3
21	1.9	2.0	1	1
22	3.2	5.0	2	2
23	1.7	5.0	1	2
24	0.6	2.0	1	1
25	4.3	21.0	2	3
26	1.0	15.6	2	2
27	0.4	1.9	1	1
28	1.6	2.0	1	1
29	0.3	1.0	1	1
30	0.8	7.0	1	2
31	3.2	7.0	2	2
32	3.0	5.0	2	2
33	1.6	3.0	1	2
34	0.8	1.0	1	1
35	0.4	2.5	1	2
36	1.5	2.0	1	1

SPECIMEN NUMBER	FNA Ki67 (%)	RESECTION Ki67 (%)	FNA GRADE	RESECTION GRADE
37	1.6	13.7	2	2
38	5.7	39.3	2	3
39	1.3	1.9	1	1
40	12.3	6.3	2	2
41	0.0	1.9	1	1
42	1.4	1.5	1	1
43	1.2	1.9	1	1
44	1.1	1.9	1	1
45	3.5	7.0	2	2
46	1.2	1.5	1	1
47	0.0	1.0	1	1
48	2.3	1.9	1	1
49	0.8	1.0	2	1

Table 2

Discordant sample grading. Thirty-nine percent of the cases were discordant. All discordant cases demonstrated a lower FNA-based grade compared to the resection grade.

FNA WHO GRADE	SURGICAL RESECTION GRADE			Total
	1	2	3	
1	22	11	1	34
2	1	8	6	15
3	0	0	0	0
Total	23	19	7	49