Comparison of Methods for Proliferative Index Analysis for Grading Pancreatic Well-Differentiated Neuroendocrine Tumors

Pamela P. Goodell, MD, Alyssa M. Krasinskas, MD, Jon M. Davison, MD, and Douglas J. Hartman, MD

Key Words: Neuroendocrine tumor; Pancreatic neuroendocrine tumor; Grading neuroendocrine tumors; Proliferative index; Ki-67; Digital image analysis

Abstract
Assessment of proliferative activity is required for grading well-differentiated pancreatic neuroendocrine tumors. However, a standardized method for obtaining the Ki-67 proliferative index is lacking. This study compared proliferative activity obtained by 3 methods: single-field hot spot (Ki-67 HS) and 10 consecutive field average (Ki-67 CFA) using the Ventana image analysis system (Ventana Medical Systems, Tucson, AZ) and mitotic index (MI). These methods resulted in discrepant grades in 30 (67%) of our 45 cases. With the current Ki-67 cutoff of more than 2% for intermediate-grade tumors, MI, CFA, and HS resulted in specificities of 91%, 94%, and 31%, respectively, for detecting metastasis, with positive predictive values (PPVs) of 25%, 67%, and 31%, respectively. At a higher Ki-67 cutoff of 7.5%, HS analysis resulted in a specificity of 94% and PPV of 71% for predicting metastasis. While single-field HS analysis may be practical and reliable at a higher cutoff, this study emphasizes the variability that can exist when different methods of assessment are used.

Well-differentiated pancreatic neuroendocrine tumors (PanNETs) are uncommon tumors, representing 1% to 2% of all pancreatic neoplasms with a prevalence of about 0.2 to 2 per million persons per year. The behavior of these tumors is difficult to predict and has been the subject of many studies. In addition to tumor size, histologic features such as necrosis, lymphovascular invasion, mitotic count, Ki-67 index, and immunohistochemical expression of cytokeratin 19 have been used as prognostic indicators. However, stage and grade seem to be the most predictive of prognosis in most studies. Recently, the World Health Organization (WHO) 2010 classification of neuroendocrine neoplasms of the digestive system along with the AJCC Cancer Staging Manual, 7th edition, have adopted the use of mitotic count and Ki-67 index to determine tumor grade. The WHO 2010 classification of neuroendocrine neoplasms of the digestive system incorporates recent recommendations made by the European Neuroendocrine Tumor Society (ENETS) to use 2 classification systems. Instead of using a single hybrid grade- and stage-based system, the new WHO classification recommends a purely grade-based classification system and a separate site-specific (TNM) staging system. All neuroendocrine tumors, including PanNETs (with the exception of neuroendocrine microadenomas <0.5 cm), are regarded as malignant tumors and the clinical behavior is largely determined by site-specific tumor biologic characteristics and stage at diagnosis. The proliferative index has emerged as a significant prognostic indicator and directly determines the tumor grade for PanNETs, especially in patients with same-stage...
tumors.\textsuperscript{1,2,11,12} Recommendations for scoring these parameters includes counting 50 high-power fields (HPF; 1 HPF = 2 mm\textsuperscript{2}) for the mitotic index (MI) and counting 500 to 2,000 cells in the area with the highest fraction of positive tumor cells (“hot spot” [HS]) for the Ki-67 index.\textsuperscript{2,7,10} One previous study suggested use of a grid or printed picture of a selected field to facilitate accurate counting.\textsuperscript{13}

Various digital image analysis platforms have been shown in previous studies to be an effective method for scoring immunohistochemical results, including the Ki-67 index.\textsuperscript{14-17} Although digital imaging offers the opportunity to streamline the potentially labor-intensive process of scoring the proliferative index, its correlation with conventional analysis and tumor behavior requires validation. We set out in this pilot study to do the following: (1) compare MI with Ki-67 index results obtained by 2 digital imaging methods, a single HS field count, and a random 10 consecutive field count; (2) assess the implications on tumor grading; and (3) determine which method may better predict metastatic disease.

Materials and Methods

This study was designed as a pilot study to assess potential variability in assessing proliferative indexes when different methods are used. A retrospective review of PanNETs at the University of Pittsburgh Medical Center, Pittsburgh, PA, was performed. High-grade neuroendocrine carcinomas were not included in this study. This project is the result of a quality improvement initiative approved by the institutional Total Quality Council. A total of 45 PanNETs were evaluated from the majority of cases (n = 41); only 4 cases had an MI of 2 or greater (range, 2-4).

Our analysis yielded 3 main groups (from raw data in Table 2): Group 1 tumors (n = 13; cases 1-13) had a Ki-67 index of 2% or less by HS and CFA; group 2 (n = 26; cases 14-39) had a Ki-67 of 3% to 20% by HS but a Ki-67 index of 2% or less by H and CFA; and group 3 (n = 6; cases 40-45) had a Ki-67 index of 3% to 20% by HS and CFA. The clinicopathologic data for these 3 groups are summarized in Table 3.

HS Counting vs CFA

By HS counting, an average of 243 nuclei were counted, with 21 cases having fewer than 200 nuclei and 24 cases having more than 200 nuclei. For CFA, an average of 191...
nuclei per field (range, 84-377) and an average total of 1,909 nuclei (range, 836-3,774) were counted. In group 1, the HS analysis yielded fewer than 200 counted nuclei in 4 cases and more than 200 counted nuclei in 9 cases (average nuclei, 249; range, 134-407); the CFA yielded an average nuclei count of 1,861 (range, 1,302-2,450). In group 2, HS yielded fewer than 200 counted nuclei in 11 cases and more than 200 counted nuclei in 15 cases (average nuclei, 259; range, 86-463); the CFA yielded an average nuclei count of 2,004 (range, 836-3,774). In group 3, HS yielded fewer than 200 counted nuclei in all 6 cases (average nuclei, 164; range, 141-181); the CFA yielded an average nuclei count of 1,602 (range, 906-1,957).

When comparing HS counting with CFA, the results were concordant in 19 (42%) of 45 cases. All 13 cases in group 1 were classified as low grade by HS and CFA. However, only 6 (19%) of the 32 remaining cases in groups 2 and 3 were concordant (cases 40-45); the remaining 81% of cases (26/32) in groups 2 and 3 were intermediate grade by HS and low grade by CFA.

**HS Counting vs MI**

When comparing HS counting with MI, the results were concordant in 17 (38%) of 45 cases. All 13 cases in group 1 were classified as low grade by HS and MI. However, only 4 (13%) of the 32 remaining cases in groups 2 and 3 were concordant (cases 38, 39, 44, and 45); the remaining 88% of cases (28/32) in groups 2 and 3 were intermediate grade by HS and low grade by MI.
Overall, using the current cutoffs of an MI of less than 2 and/or 2% or less for the Ki-67 index for low-grade tumors and an MI of 2 to 20 and/or 3% to 20% for the Ki-67 index for intermediate-grade tumors, 30 (67%) of 45 cases had inconsistent grading based on the 3 methods: 2 (4%) of 45 tumors were classified as intermediate grade by MI and HS but low grade by CFA; 4 (9%) of 45 tumors were classified as intermediate grade by HS and CFA but low grade by MI; and 24 (53%) of 45 tumors were classified as intermediate grade by HS but low grade by CFA and MI.

Correlation With Metastatic Disease

Because metastatic disease is a well-documented predictor of prognosis in PanNETs,4,8,11 we used this as an indicator of more aggressive behavior in our cases. Of our cases, 13 had nodal and/or distant metastasis. The sensitivities and specificities of the 3 methods for detecting nodal and/or distant metastasis are summarized in Table 4. With a cutoff value of Ki-67 of more than 2% for intermediate-grade tumors, 10 (77%) of 13 were intermediate grade by HS, 4 (40%) of the 10 were intermediate grade by HS alone,
4 (40%) of 10 were intermediate grade by HS and CFA, and 2 (20%) of 10 were intermediate grade by HS and MI; 3 (23%) of 13 tumors with nodal and/or distant metastasis were low grade by all 3 methods. Of the 32 PanNETs that were intermediate grade by Ki-67 HS analysis, only 10 (31%) had nodal and/or distant metastasis (positive predictive value [PPV], 31%). However, of the 6 PanNETs that were intermediate grade by Ki-67 CFA analysis, 4 cases had nodal and/or distant metastasis (PPV, 67%).

With ROC analysis, the current recommended cutoff value for Ki-67 of more than 2% for intermediate-grade tumors using CFA analysis achieves the highest possible sensitivity of 31% and specificity of 94% for detecting lymph node and/or distant metastasis, similar to specificity results when using MI (sensitivity, 8%; specificity, 91%). In comparison, ROC analysis for HS analysis results in a sensitivity of 77% and specificity of 31% (PPV, 31%) for detecting metastasis. Based on the ROC analysis, using a higher cutoff for Ki-67 of more than 7.5% for intermediate-grade tumors by HS analysis achieves a sensitivity of 38% and specificity of 94% for detecting metastatic disease, similar to Ki-67 by CFA (Table 4). In HS analysis with a cutoff value for Ki-67 of more than 7.5%, rather than having a PPV of only 31% in intermediate-grade PanNETs, the Ki-67 HS index predicts nodal and/or distant metastasis in 5 of 7 intermediate-grade cases (PPV, 71%).

### Table 3

<table>
<thead>
<tr>
<th>Patient and PanNET Characteristics by Group*</th>
<th>Group 1 (n = 13)</th>
<th>Group 2 (n = 26)</th>
<th>Group 3 (n = 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (range) age (y)</td>
<td>62 (31-78)</td>
<td>58 (28-79)</td>
<td>61 (46-77)</td>
</tr>
<tr>
<td>Sex, F/M</td>
<td>7/6</td>
<td>13/13</td>
<td>3/3</td>
</tr>
<tr>
<td>Mean (range) size (cm)</td>
<td>3.4 (0.8-8.5)</td>
<td>2.8 (1.3-6.5)</td>
<td>3.4 (1.5-4.8)</td>
</tr>
<tr>
<td>LVI</td>
<td>5</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>PNI</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Necrosis</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>LN metastasis</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Distant metastasis</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>LN and/or distant metastasis</td>
<td>3</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>

Nodal and/or Distant Metastasis

CFA, consecutive field average analysis; HS, hot spot analysis with >2% cutoff for intermediate-grade tumors; MI, mitotic index analysis; NPV, negative predictive value; PanNET, well-differentiated pancreatic neuroendocrine tumor; PPV, positive predictive value.

### Table 4

<table>
<thead>
<tr>
<th>Sensitivity, Specificity, PPV, and NPV of the Three Methods for Predicting Nodal and/or Distant Metastasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodal and/or Distant Metastasis</td>
</tr>
<tr>
<td>PanNET Graded as Intermediate by</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Sensitivity (%)</td>
</tr>
<tr>
<td>Specificity (%)</td>
</tr>
<tr>
<td>PPV (%)</td>
</tr>
<tr>
<td>NPV (%)</td>
</tr>
</tbody>
</table>

Discussion

For PanNETs, current consensus statements set forth by several organizations (WHO, College of American Pathologists, and ENETS) recommend evaluating the MI and Ki-67 proliferation index to predict behavior of these rare lesions. However, the practical act of generating these values has been varied, with some methods seeming impractical for routine clinical purpose (such as counting 2,000 cells as proposed by ENETS). The College of American Pathologists Protocol for the Examination of Specimens From Patients With Carcinoma of the Endocrine Pancreas states that it is acceptable to “estimate” the Ki-67 index. However, separating low- and intermediate-grade PanNETs requires accurate scoring of low proliferative indexes to discriminate tumors with greater than 2% nuclear positivity from those with less. Because this distinction has potential implications for follow-up interval and receipt of additional therapy, it is important to accurately and reproducibly determine the Ki-67 index.

A recent consensus report by Klimstra et al, which involved a diverse panel of 12 pathologists and 8 clinicians, noted disagreement among participants on the optimal method of Ki-67 index determination of PanNETs. Only 53% of the participants thought that the Ki-67 index was reproducible among pathologists to discriminate clinically relevant subsets of PanNETs, 75% thought that manual counting should not be done (because it was too time-consuming), and yet, 63% also thought that image analysis was not ready for routine use. The group recommended HS counting and choosing a variety of areas within the tumor; a minimum number of cells to count was not offered.

While use of digital imaging to calculate Ki-67 proliferative indexes in neuroendocrine tumors is still not widely practiced, computer-assisted counting has been shown to be a viable alternative to pathologist counting in a manner that is practical and time-effective. The aim of this study was to compare different methods of determining proliferative index in PanNETs. The overall findings of our study reveal that the 3 different methods for determination of proliferative index (Ki-67 HS, Ki-67 CFA, and MI) can yield different results, with inconsistent grading in two thirds (67%) of the
PanNETs in this study. One source of inconsistency may be the total number of cells counted. The 2010 WHO recommends counting 500 to 2,000 cells in HS areas.\(^2\) In the present study, our HS analysis yielded an average of 243 nuclei per HS field. When compared with these HS counts, the CFA counts (which counted an average of 1,909 cells) demonstrated an overall reduction in proliferative index. Similarly, when compared with MI, HS analysis was discordant in 62% of cases, while CFA was discordant in only 13% of cases. These are reasonable observations, since the proliferative index is often variable within a tumor, with the highest index most commonly located at the leading edge of the tumor.\(^7,25\)

Based on this information, it seems that single-field Ki-67 HS analysis may overestimate the proliferative index. To determine which method would best predict behavior in our group of cases, we correlated the results with the presence or absence of lymph node/distant metastases.

As expected, when the established cutoff of 2% is used, the Ki-67 HS analysis was the most sensitive method for detecting metastatic disease, but it had the lowest specificity, at 31%. The CFA had the highest PPV for intermediate-grade PanNETs of 67%, while Ki-67 HS and MI had PPVs of 31% and 25%, respectively. If clinical decisions, such as to treat or not to treat with chemotherapy or to alter follow-up surveillance intervals, rely on the grade of the PanNET (low vs intermediate grade), the specificity and the PPV of the test used to decide this difference should be optimized.

While HS analysis of a single field using a computer-assisted system such as VIAS seems to be an efficient method for determining the proliferative index of PanNETs, a different cutoff may be required to separate indolent tumors from those with a higher likelihood of aggressive behavior (lymph node and/or distant metastasis). Our data suggest that when using HS analysis, a cutoff value for Ki-67 of 7.5% for intermediate-grade PanNETs increases the specificity to 94% and the PPV to 71% for predicting lymph node and/or distant metastasis, similar to that seen when using the current cutoff with CFA analysis.

This study was designed as a preliminary assessment of the variability that can arise when different methods for determining the proliferative index are used. We provided evidence that enough variability exists to warrant additional studies that use larger populations of patients with PanNETs with long-term outcome data. Even though our study population is small, the findings shed light on the difficulty of determining which method to use to detect proliferation rates in PanNETs and of determining the significance of the result. Moreover, the study raises serious concerns about our ability to accurately discriminate between low- and intermediate-grade well-differentiated PanNETs.

Until better prognostic indicators are developed for PanNETs, Ki-67 proliferative indexes will be used. Although digital imaging can be used to calculate Ki-67 proliferative indexes and will likely provide a more practical and time-effective assessment in the near future, our results demonstrate that the counting method can significantly impact the score and that different cutoff values will most likely be required for each method. This study also suggests, however, that if a higher Ki-67 cutoff for intermediate-grade tumors is used when applying the HS method, HS analysis of a single 250-cell field provides a similar or an increased PPV for predicting aggressive tumor behavior (metastasis) when compared with other methods, without the need to count 2,000 cells or manually count mitotic figures. We recommend that individual practices adopt a uniform method to measure the Ki-67 proliferative index. If single-field HS analysis is used, the results need to be interpreted with caution. Results of less than 2% can be regarded as low grade with any method. If a result greater than 2% is obtained with HS counting, additional fields should be scored to ensure that the result is not biased by a single microscopic cluster of proliferative activity. Based on our experience, we believe that the WHO recommendations that stipulate counting 500 to 2,000 cells in HS areas should be followed for any case that is borderline, because the 2% cutoff seems to be relevant to this method of counting. Pathologists should consider specifying their scoring method in the pathology report because the method clearly impacts the score for a given tumor.

References


