

Calculation of the Ki67 index in pancreatic neuroendocrine tumors: a comparative analysis of four counting methodologies

Michelle D Reid^{1,12}, Pelin Bagci^{2,12}, Nobuyuki Ohike³, Burcu Saka⁴, Ipek Erbarut Seven², Nevra Dursun⁵, Serdar Balci⁶, Hasan Gucer⁷, Kee-Taek Jang⁸, Takuma Tajiri⁹, Olca Basturk¹⁰, So Yeon Kong¹¹, Michael Goodman¹¹, Gizem Akkas¹ and Volkan Adsay¹

¹Department of Pathology, Emory University School of Medicine, Atlanta, GA, USA; ²Department of Pathology, Marmara University, Istanbul, Turkey; ³Department of Pathology, Showa University School of Medicine, Tokyo, Japan; ⁴Department of Pathology, Medipol University, Istanbul, Turkey; ⁵Department of Pathology, Istanbul Education and Training Hospital, Istanbul, Turkey; ⁶Department of Pathology, Yildirim Beyazit University, Ankara, Turkey; ⁷Department of Pathology, RTE University, Rize, Turkey; ⁸Department of Pathology, Samsung Medical Center, Seoul, Korea; ⁹Department of Pathology, Tokai University Hachioji Hospital, Tokyo, Japan; ¹⁰Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA and ¹¹Department of Pathology, Department of Epidemiology, Emory University, Atlanta, GA, USA

Ki67 index is now an essential part of classification of pancreatic neuroendocrine tumors. However, its adaptation into daily practice has been fraught with challenges related to counting methodology. In this study, three reviewers used four counting methodologies to calculate Ki67 index in 68 well-differentiated pancreatic neuroendocrine tumors: (1) ‘eye-ball’ estimation, which has been advocated as reliable and is widely used; (2) automated counting by image analyzer; (3) manual eye-counting (eye under a microscope without a grid); and (4) manual count of camera-captured/printed image. Pearson’s correlation (*R*) was used to measure pair-wise correlation among three reviewers using all four methodologies. Average level of agreement was calculated using mean of *R* values. The results showed that: (1) ‘eye-balling’ was least expensive and fastest (average time < 1 min) but had poor reliability and reproducibility. (2) Automated count was the most expensive and least practical with major impact on turnaround time (limited by machine and personnel accessibility), and, more importantly, had inaccuracies in overcounting unwanted material. (3) Manual eye count had no additional cost, averaged 6 min, but proved impractical and poorly reproducible. (4) Camera-captured/printed image was most reliable, had highest reproducibility, but took longer than ‘eye-balling’. In conclusion, based on its comparatively low cost/benefit ratio and reproducibility, camera-captured/printed image appears to be the most practical for calculating Ki67 index. Although automated counting is generally advertised as the gold standard for index calculation, in this study it was not as accurate or cost-effective as camera-captured/printed image and was highly operator-dependent. ‘Eye-balling’ produces highly inaccurate and unreliable results, and is not recommended for routine use.

Modern Pathology (2015) 28, 686–694; doi:10.1038/modpathol.2014.156; published online 21 November 2014

Pancreatic neuroendocrine tumors constitute a group of neoplasms with phenotypic and ultrastructural neuroendocrine differentiation. As studies

with longer follow-up and better defined cohorts are published, it has become increasingly clear that pancreatic neuroendocrine tumors are low-grade malignancies. Stage of tumor, naturally, predicts the outcome of pancreatic neuroendocrine tumors; however, it has been difficult to predict which tumors are prone to recurrence and metastasis by pathologic examination alone.

Among a variety of potential prognostic parameters analyzed in numerous studies, proliferation index has consistently stood out as having strong prognostic value, along with mitotic count. Ki67 has

Correspondence: Professor V Adsay, MD, Department of Pathology, Emory University School of Medicine, Emory University Hospital, 1364 Clifton Road NE, Room H-180B, Atlanta, 30322 GA, USA.

E-mail: nadsay@emory.edu

¹²These are co-contributing first authors of this manuscript.

Received 3 July 2014; revised 29 September 2014; accepted 30 September 2014; published online 21 November 2014

been shown to be a reliable marker of proliferation index. It has been in practice for more than two decades, and is widely used as a prognostic factor in a variety of tumors including breast, hematolymphoid, and glial malignancies. In fact, as it covers a wider spectrum of the mitotic process than the mitotic spindle observed histologically, it is believed that Ki67 is more effective than the tedious and ever irreproducible mitotic count.

Thus, Ki67, which was originally described as a classification/grading parameter by the World Health Organization in 2000,¹ was later adopted as the main grading parameter for neuroendocrine tumors by the European Neuroendocrine Tumor Society and World Health Organization 2004,² and is now endorsed by numerous organizations including the North American Neuroendocrine Tumor Society,³ College of American Pathologists,⁴ American Joint Committee on Cancer,⁵ and World Health Organization 2010,⁶ as well as an international consensus group.⁷ Accordingly, pancreatic neuroendocrine tumors are graded as well-differentiated (grade 1 (G1) and grade 2 (G2)) neuroendocrine tumors and poorly differentiated neuroendocrine carcinomas (grade 3 (G3)) based on the count of a minimum of 500 cells in Ki67-positive tumor hot spots and mitoses per 10 high-power fields.

Although the Ki67 index is now a requirement for the grading of pancreatic neuroendocrine tumors, the best method for calculating the index is still the subject of debate. It has been advocated by the leaders of the European Neuroendocrine Tumor Society and has been widely used in the assessment of pancreatic neuroendocrine tumors for more than a decade.⁸ Some members of the European Neuroendocrine Tumor Society consider the 'eye-balling' method of counting adequate, and other sources such as North American Neuroendocrine Tumor Society have also endorsed 'eye-balling' in their guidelines.^{9,10} However, recent studies have questioned the accuracy and reproducibility of this method, especially for tumors with equivocal 'gray zone' grades.^{11,12}

In this study, we investigated the reproducibility, applicability, and practicality of the four most widely used methods for Ki67 index calculation.

Materials and methods

After approval by the Institutional Review Board, 68 cases of well-differentiated pancreatic neuroendocrine tumors were retrieved. There were 33 males and 35 females, with a mean age of 51 years and mean tumor size of 3.9 cm. Ki67 immunohistochemical staining was then performed and the Ki67 index was calculated using four of the most widely used counting methodologies listed below.

Ki67 Immunostaining and Evaluation

Full-thickness, 5- μ m sections were cut from formalin-fixed, paraffin-embedded blocks from all 68 cases. These

were then stained with the Ki67 antibody (Mib-1 clone, 1:160; Dako Corporation, Carpinteria, CA, USA) using tonsillar tissue as a positive control. Negative controls were run simultaneously with the primary antibody replaced with a buffer. Antigen retrieval was conducted in citrate buffer at pH 6 under pressure for 3 min. Envision Dual Link Kit (Dako) was used for detection, with diaminobenzidine as the chromogen and hematoxylin as the counterstain. Staining was considered positive when brown nuclear labeling was observed.

A standard Olympus BX41 microscope was used to identify tumor 'hot spots' in each case. The percentage of tumor cell staining was counted by three reviewers (PB, BS, and K-TJ) and with four different methodologies as shown below. For each counting method, one hot spot was counted. For cases approaching categorical cutoffs, an additional hot spot was counted and the average score was used.

Counting Methodologies

'Eye-balling'. This was in essence an estimation of the percentage of Ki67-positive tumor cells on a slide section and it was carried out by scanning or 'eye-balling' the entire slide at intermediate power ($\times 10$ objective) without actually counting individual cells. This appears to be the most widely used method and is advocated by some as a reliable counting method, including by some of the original authors of European Neuroendocrine Tumor Society⁸ and the North American Neuroendocrine Tumor Society⁹ guideline papers.

Automated counting. Ki67 labeling index was obtained using an automated cellular image cytometer (ACIS[®] III) by Dako Corporation. The automated scope scanned the entire slide at $\times 4$ objective. Whole mount images were then scanned for Ki67 pattern of staining and three areas of highest nuclear labeling ('hot spots') were selected by trained technicians who also routinely use Automated Cellular Imaging System (ACIS[®]) for quantitative Ki67 immunohistochemical analysis for breast carcinomas (performed routinely in our laboratory), glial tumors, hematopoietic malignancies, and other endocrine neoplasms, in addition to other quantitative immunohistochemistry such as estrogen and progesterone receptors. Ki67 index quantification was carried out by a left-to-right, up-down visual slide scan. The areas to be counted were selected by the technician.

Eye-counting with microscope. This method consisted of 'real-time' counting of Ki67-positive cells/500 or more tumor cells within a microscopic field. This method is similar to the manual differential counting method used on peripheral smears. After an initial slide scan was performed to identify tumor 'hot spots', the microscope was then focused on the $\times 20$ objective and tumor cells (including

Ki67-positive cells) were counted in real time. The Ki67 index was then calculated based on this real-time 'eye-counting' method.

Manual counting of camera-captured/printed image. For this method, each tumor slide was manually scanned with a microscope at $\times 10$ objective, and the area of greatest Ki67 positivity (hot spot) was selected for photographing and printing. The static color image of the hot spot was captured via camera snapshot, printed on plain white paper, and Ki67-negative and -positive tumor cells were then visualized and immediately marked/crossed off once counted by each of the three reviewers, with circling of the Ki67-positive dark brown tumor nuclei as shown in Figure 1. Each reviewer independently selected tumor hot spots and calculated the Ki67 index based on this. Light brown or pale staining nuclei were ignored during counting.

Careful, extensive, and exhaustive analysis by an expert. In this study, the Ki67 labeling index in all the cases was carefully analyzed by an expert pancreatic pathologist (NO) with special interest in neuroendocrine tumors. This reviewer counted six different hot spots each composed of a minimum of 2000 cells and determined the 'final' gold standard Ki67 index to which all other counting methods were compared.

Comparative analysis of applicability of different counting methods

- (1) The ease or difficulty as well as cost of performing each method were recorded.
- (2) The time it took to perform each counting method and its potential impact on the turnaround time of the sign-out of the case was also calculated.

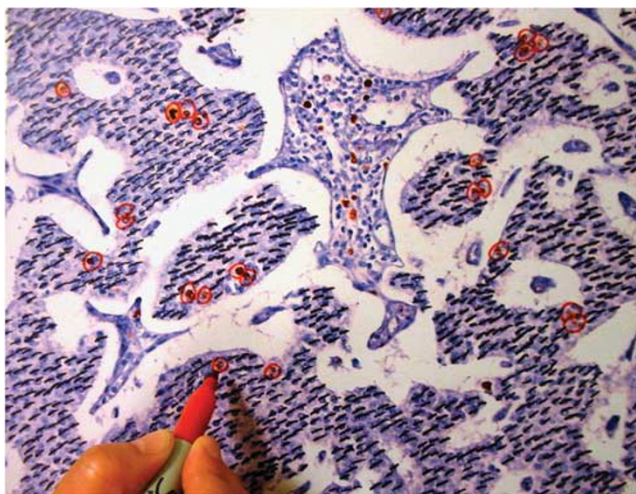


Figure 1 A printed color image of a well-differentiated pancreatic neuroendocrine tumor is used for manual cell count by crossing off the Ki67-negative tumor cells (black lines) and highlighting the Ki67-positive ones (circled in red) (Ki67 stain).

- (3) Comparative reproducibility analysis was performed for all methods and among observers. For this, Pearson's correlation (R) was used to measure pair-wise correlation among three observers. An average level of agreement for all three observers was calculated using the mean of the R values from each pair of observers. The results of comparative analysis of the different methodologies are summarized in Table 1.

Results

Using the gold standard Ki67 index calculated by the expert pathologist, there were 26 G1, 39 G2, and 3 G3 tumors. In terms of the distribution of Ki67 among the cases, and the relevance of accuracy and reproducibility of the counting, it was noted that more than 50% of the cases had an index in the numerical vicinity of the categorical cutoffs ($>50\%$ of the cases fell between 1 and 5% range (for G1 versus G2), and about 10% of the cases fell within the 15 and 25% range (for G2 versus G3) (Table 2). Along the same lines, cases with very low and very high indices were also very uncommon ($<10\%$ of the cases; see Table 3).

Cost

The cost of the automated system was estimated to be US\$50 000–150 000 (depending on the vendor), not accounting for the expense of technician salary and time, space allocation, and use of Health Insurance Portability and Accountability Act—compliant software. While in the United States there is a separate billing code for the quantitative analysis performed through automated systems, the difference is fairly small (billing code 88360, US\$379 for manual quantification, versus billing code 88361, US\$463 for automated systems).

The camera-captured/printed image method, which was found to be the most practical and reproducible (see below), used a camera with image capture ability, as well as a color printer, the cost of which was estimated to be US\$500–15 000. Considering that image capturing has become even more feasible with commercially available cell phone-microscope adapters, this cost could be significantly lower and limited to the purchase of an eyepiece adapter only (which on average costs $<US\$100$). Although black and white printers could also be used, these were not as reliable, and therefore color printers were used (average cost US\$100). The need for a color printer could potentially be negated by using free counting software available on the web.¹³ Eye-balling and eye-counting methodologies did not add any additional cost as they were performed by routine microscopic methods; however, they proved inaccurate or impractical (see below).

Table 1 Comparison of the Ki67 index counting methodologies

Methodology	Impact on turnaround time	Average time it takes (min)		Accuracy	Miscounting of non-target cells	Interobserver agreement (Pearson's correlation)	Additional cost
		Practicality					
Eye-balling	None	< 1	Highest	Very low	Unlikely	R = 62.1%	None
Eye-counting of cells (through microscope)	None	6	Low	High	Unlikely	R = 55.5%	None
Manual count of camera-captured/printed image	Minimal (depending on accessibility of a camera/printer setup)	8.1	Very high	Highest	Unlikely	R = 81.7%	Printer + camera US\$5400
Automated counting	Highest (depends on the technician availability)	5	Low (accessibility issues)	Moderate	Very likely	N/A	Image analyzer US\$150 000

Pearson's correlation (*R*) was used to measure pair-wise correlation among three observers (shown in bold). An average level of agreement for all three observers was calculated using the mean of the *R* values from each pair of observers. Abbreviation: N/A, not applicable.

Table 2 Cases with Ki67 indices approaching categorical cut-offs

	Counting Method				Automated
	Expert ^a	Eye-balling	Eye-counting	Camera-captured/printed image	
Ki67 index (%)	Average for three reviewers				
G1–G2 (1–5%)	35	42	32	35	46
G2–G3 (15–25%)	6	4	4	5	2
Total (%)	41 (60%)	46 (68%)	36 (53%)	40 (59%)	48 (71%)

Abbreviations: G1, grade 1; G2, grade 2; G3, grade 3.

^aThe expert counted six different hot spots each composed of a minimum of 2000 cells.

Table 3 Cases with very low (<1%) and very high (≥30%) Ki67 indices

	Counting method				Automated
	Expert ^a	Eye-balling	Eye-counting	Camera-captured/printed image	
Ki67 index (%)	Average for 3 Reviewers				
<1%	7	2	5	4	1
≥30%	0	0	1	1	4
Total (%)	7 (10%)	2 (3%)	6 (9%)	5 (7%)	5 (7%)

Abbreviations: G1, grade 1; G2, grade 2; G3, grade 3.

^aThe expert counted six different hot spots each composed of a minimum of 2000 cells.

Test Duration and Impact on Turnaround Time

'Eye-ball' estimation and real-time eye-counting through the microscope were both rapid methods and had no effect on sign-out turnaround time; however, their inaccuracy significantly limited their value (Table 1).

Manual counting of a camera-captured/printed image did not impact the turnaround time as it could be performed during the routine sign-out of a case, provided that the image-capture setup was readily available at the microscope. It took, on average, 8 min per case (range 2.16–23 min). The longer turnaround times were seen when reviewers did not have ready access to equipment.

Use of the automated system had the greatest impact on turnaround time as the process occurred outside of routine sign-out hours, was pathologist-independent, and relied almost entirely on laboratory technician availability and instrument accessibility. This created a significant delay that was even more protracted when the specimens were batched. However, once the machine became available, the calculation of the index was completed in ~5 min.

Practicality and Applicability

Although the eye-balling method theoretically appeared to be the most practical, its lack of accuracy

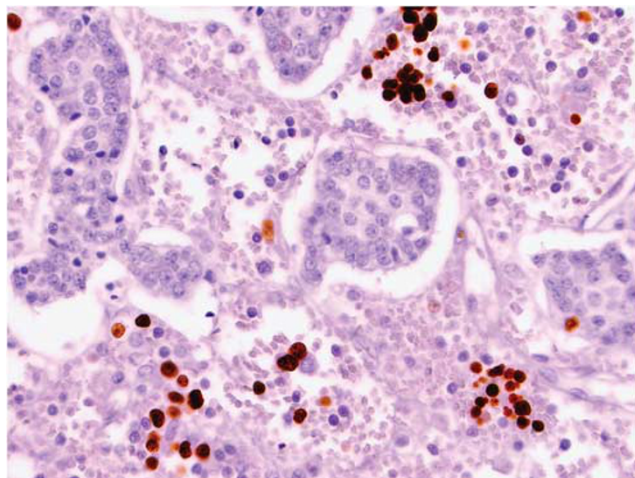


Figure 2 Unless software modification is carried out on individual cells, the Automated Cellular Imaging System (ACIS[®]) cannot distinguish tumoral from non-tumoral cells. In this Ki67-stained well-differentiated pancreatic neuroendocrine tumor, there are islands of tumor cells with rare Ki67-positive nuclei and a central cluster of Ki67-positive lymphocytes, a potential pitfall in automated analysis ($\times 200$ magnification).

limited its applicability (see below). Eye-counting, which is widely used in hematology, was very difficult to perform for this purpose, especially on thick tissue sections. This is because cellular density and overlapping is notoriously high in pancreatic neuroendocrine tumors, making it very difficult to distinguish the cells that had previously been counted from those that remained to be counted, even when a grid was used. Although there was some overlapping of tumor nuclei with the $5\ \mu\text{m}$ sections, we nevertheless performed the study with this tissue section thickness because this is typically what is used in routine practice and our goal was to determine (in as close a manner as possible) the applicability of the four methods in daily practice standards.

Automated imaging was especially challenging because it used a pathway that was independent of the general sign-out process. The lack of instrument accessibility and cost of a trained technician, instrument, and software made it an impractical test for routine use.

Although the camera-captured/printed image method was time-consuming at times, it was fairly practical and applicable to routine sign-out and took a relatively short period of time (Table 1). It also had the advantage of keeping a record of the area that was counted for future reference.

Accuracy and Reproducibility

Interobserver agreement was highest (R:81.7%) and interobserver variability was the lowest (R:43%) for the camera-captured/printed image method. As the

count was performed by the pathologist, inaccuracies due to tissue assessment and overcounting of non-tumoral cells was much less an issue for this method, unlike the automated method in which counting was performed by the technicians. Additionally, as this involved a physical count, it was much more reliable than 'guesstimate' methods of eye-balling and eye-counting.

While eye-balling was practical and usable for extreme ends of the labeling spectrum (very high or very low cell positivity), its overall accuracy was found to be very low rendering it inapplicable, especially for borderline cases.

The automated system had two problems in terms of accuracy. First, it relied on the personnel who were not familiar with the histopathologic subtleties of pancreatic neuroendocrine tumors, which often led to overcounting of lymphocytes, endothelial, and stromal cells, as well as hemosiderin-laden macrophages (Figure 2). Second, with the automated system, the technicians did not use a preset field area for each case, but instead arbitrarily selected the field area to be counted. This significantly altered the denominator (number of cells counted within a field) of the calculation and can thus alter the final index in some cases.

Overall Assessment

Considering the balance of reproducibility, practicality, and accuracy, the method that was found to be by far the most preferable (Table 1) was the manual count of the camera-captured/printed image.

Discussion

Prognostication of pancreatic neuroendocrine tumors has been a challenge. Along with mitotic count, the proliferation index calculated by Ki67 has been shown to be one of the most reliable prognostic factors for grading pancreatic neuroendocrine tumors and thus it is now a requirement in the College of American Pathologists, European Neuroendocrine Tumor Society, North American Neuroendocrine Tumor Society, and American Joint Commission on Cancer protocols used in diagnosing these tumors. However, there remain several challenges in determining the proliferation index in a given tumor, including the method of counting, tumor heterogeneity, and defining what constitutes a hot spot and what constitutes positive staining. Among these, the best method to do the counting of Ki67 is probably the most elemental.

The distribution of Ki67 index in this cohort highlights the importance of an accurate and reproducible counting method. More than half the cases analyzed in our study had Ki67 indices that were fairly close to the established categorical cutoffs (Table 2), and cases with extremely high or low indices (where a casual estimation method such

as 'eye-balling' may have been applicable) were unfortunately few and far between, and <10% of the overall cases (Table 3).

In this study, we focused on the assessment of four of the most widely used Ki67 index counting methodologies and assessed their practicality, applicability, and reproducibility, and the following conclusions were reached.

Eye-Balling

Although we found that this method was the fastest (averaging <1 min), least expensive (with no additional cost), and was by far the most practical, unfortunately with an 'R'-value of only 62%, its reliability and reproducibility were far too low to make it even remotely applicable to clinical management. Tumor cellularity and tissue section thickness, as well as pattern and quality of staining all, contributed to subtle deceptions when estimating tumor grade, and thus hampered the accurate measurement of percentage positivity/density of Ki67 staining. In 2011, the College of American Pathologists stated that 'estimation' of the Ki67 index was acceptable; however, this is especially impractical as precise three-tiered grading is required to best predict tumor behavior and dictate treatment algorithms.⁴ In a recent study, Tang *et al*¹² showed similar pitfalls in the 'eye-balling' method as evidenced by poor interobserver agreement when grading well-differentiated neuroendocrine tumors.¹² Additionally, in Tang's study all 18 reviewers (including experts in Ki67 analysis in neuroendocrine, brain, and hematopoietic tumors) looked at the same identical field/hot spot (via PowerPoint images), and despite this, their level of agreement was extremely low. 'Eye-balling' is now strongly discouraged unless a tumor shows an unequivocally low or high index.

Eye-Counting

While this method seems very practical in theory, for the purposes of Ki67 index calculation in well-differentiated neuroendocrine tumors, it proved to be the most cumbersome and difficult to employ. Although frequently used in hematopathology and other fields (with satisfactory and reproducible results), we found that, unlike hematologic smears in which the cells are often widely separated and thinly spread, pancreatic neuroendocrine tumors are fairly cellular with numerous overlapping cells, making 'real-time' counting a challenge. Even though grids have been recommended for index calculation in pancreatic neuroendocrine tumors,¹⁴ in our study we found that within a given field, even with grids and other orientation aids, it was almost impossible to separate accurately cells that had been previously counted from the ones that had not.

Manual Count of Camera-Captured/Printed Image

This method requires a digital camera and a printer, costs US\$4000–5500, and is readily available in most institutions, especially at tertiary care centers where most pancreatic neuroendocrine tumor resections are performed. The cost can also be kept to a minimum by using handheld devices with or without simple, low-cost adapters. If all components are preconfigured, connected, and readily available, then the calculation process can be completed within minutes, during routine sign-out. In addition, in our hands, the accuracy of this approach was incomparably higher than that of all other methods tested. Furthermore, it is far simpler than the dizzying eye-counting approach. Most importantly, the camera-captured/printed image method had the lowest interobserver variability and the closest correlation with the final 'gold standard' exhaustive count performed by the expert pathologist. It is important to note that the low interobserver variability was seen despite the fact that hot spots were not preselected. In fact, each reviewer chose their own hot spot for calculation, but despite this there was still strong correlation in grading. This makes the camera-captured/printed image method a highly reproducible and consistent scoring method that is applicable to routine use. In our current practice, we have started using this method on a routine basis.^{11,15} In 2010, Verbeke¹⁴ also recommended using printed photomicrographs for Ki67 index calculation but did not detail the methodology.

Recently, free software has become available online that allows for the uploading of camera-captured photomicrographs to a National Institutes of Health-based website that uses a public domain image analysis algorithm to analyze images and calculate the percentage of positive cells within a given area.¹³ It is more frequently used for estrogen and progesterone receptor quantification, and can also be used for Ki67 calculation (authors' personal observation). Other similar websites are also emerging online. We used one such website to calculate the Ki67 index in a camera-captured image and found that the software clearly identified and counted the Ki67-positive cells.

Automated Counting

Although this method has been advocated as the 'gold standard' for Ki67 calculation, with the idea that the count can be performed accurately by the instrument,^{12,16–18} our study and others^{19,20} highlighted several shortcomings in its presumed accuracy. First, we noted a relative inability of the instrument to distinguish Ki67-positive tumor cells from other Ki67-labeling cell types unless it was manually calibrated to disregard such non-tumor cells, with the application of special software. In particular, lymphocytes, endothelial cells, and stromal cells contributed significantly to overcounting and

erroneously high Ki67 indices in 13% of our cases. During the study, we also noted that molding of tumor nuclei and overly thick sections contributed to miscalculations by the instrument. Additionally, the presence of abundant background pigment (often in the form of hemosiderin-laden macrophages) led to serious instrument miscalculation and inflated Ki67 indices in some tumors. In one such example (not included in this study), there was an Automated Cellular Imaging System (ACIS®)-generated count of 45% *versus* a camera-captured/printed image-generated count of 3%. Conversely, the automated system also seemed to undercount occasionally in some cases as well, presumably because it was disregarding moderately staining nuclei, due to lack of proper calibration of the instrument. It is possible that these shortcomings could be avoided or minimized if the hot spots are preselected and highlighted by a pathologist before technician review, or if a pathologist performs the entire analysis along with technical personnel who can help calibrate the software. Furthermore, it is possible that in the future the automated systems will develop ways to avoid these shortcomings; however, currently, it does not seem practical and accurate enough. In fact, our study highlights the need for one or more of the following in order for an automated system to be used in current practice (a) preselection of hot spots for analysis, (b) technician training in histology to ensure accurate recognition of tumor cells and avoidance of non-tumoral tissue analysis, (c) instrument calibration or cell-specific analytic software to avoid analysis of non-tumoral tissue, and (d) slide scanning and analysis by a pathologist instead of a technician. In their 2012 study on Ki67 quantification, Tang *et al*¹² stated that manual and automated counting had comparable accuracy and used automated counting as their 'gold standard'. Although we would agree with this to some extent, we believe that the accuracy of automated counting is more technician-dependent than all other methodologies, and even when performed by experienced technicians, should be interpreted with caution and correlated with other methods and quality assurance mechanisms. Interestingly, in the Tang *et al*¹² study a pathologist reviewed the scanned images and then selected representative images for Ki67 calculation. It would be interesting to note if this was a routine practice at their institution or a one-time phenomenon carried out for the purpose of their study.

For laboratories interested in automated counting, the current price range for the most popular instruments is US\$100 000–200 000. Instrument and technician availability, as well as equipment cost (and maintenance), are therefore major considerations when using this method in routine practice. More importantly, although the generated count can be billed with a separate medical billing code, the final 'overhead' costs may still make it impractical for many laboratories. If these

challenges can be overcome, automated counting may prove to be fairly reliable in the future.

Other Problems with Counting Highlighted in this Study

There are various other issues in the determination of Ki67 labeling index that we experienced while conducting this study. First, intratumoral heterogeneity (a known problem with pancreatic neuroendocrine tumors), as well as the subjectivity of hot spot selection, can lead to marked variation in tumor grade.^{7,17–19,21} The issue of whether to interpret pale brown tumor nuclei as positive is another vexing point. For this, a comparison with background non-neoplastic tissue may be helpful as stromal cells should not stain positively if the assay is correctly performed; in a similar manner, we subscribe to the view that light brown nuclei should generally be disregarded. In cases that prove to be very close to the categorical cutoffs, it is advisable to count multiple hot spots and obtain their average to account for staining variability. This may not be necessary when hot spots seem to be uniform and the index is not close to the categorical cutoffs.

Despite these challenges, Ki67 index remains a strong prognosticator, and is in fact superior to, and far more reproducible than, mitotic count in our experience. As the Ki67 index captures all phases of cellular proliferation, it is invariably higher than the mitotic count, and is a bigger determinant of the grade and prognosis.^{22,23} It is as easy, if not easier to interpret, than numerous other quantitative immunohistochemical assays in current use (estrogen and progesterone receptors/her2neu for breast and epidermal growth factor receptor for gastric cancers). As some management protocols are designed based on tumor grade, accurate grading is critical for patient stratification for a variety of treatment protocols,³ including surgery, targeted drug therapy,²⁴ and a variety of clinical trials.²⁵ This is particularly applicable to the emerging group of grade-discordant, well-differentiated neuroendocrine tumors with G3 range Ki67 indices. These tumors are more clinically aggressive than typical well-differentiated G2 neuroendocrine tumors,²⁶ but not as aggressive as 'true' poorly differentiated G3 neuroendocrine carcinomas. Their identification by a correctly performed counting methodology is paramount to patients' receipt of more aggressive therapy.

Conclusions

Among the four methodologies for counting Ki67 in pancreatic neuroendocrine tumors, the method that was found to be the most accurate, reproducible, and practical is the one in which manual counting is performed on camera-captured printed images of tumor hot spots. Although automated counting is very accurate and easy to perform, this study

showed that the camera-captured/printed image counting method was more accurate, cheaper, counted *only* the targeted tumor cells, and had the highest interobserver agreement. It also had the lowest cost/benefit ratio, especially when compared with the 'eye-balling' method. Disadvantages of the camera-captured/printed image counting method included the length of time required for completion of counting and the requirement for equipment (camera and printer), which might not be available in every institution. However, it is still much cheaper and more accessible than the automated method.

Disclosure/conflict of interest

A preliminary analysis involving part of this cohort was reported at the 101st annual meeting (in 2012) of the United States and Canadian Academy of Pathology in Vancouver, Canada. Dr Nevra Dursun is the recipient of a scholarship from Türkiye Bilimsel ve Teknik Arastırma Kurumu (State Department of Scientific and Technical Investigation) of the Republic of Turkey. The other authors declare no conflict of interest.

References

- Kloppel G, Heitz PU, Capella C, *et al*. Endocrine tumors of the pancreas, In: Solcia E, Kloppel G, Sobin LH (eds) World Health Organization International Histological Classification of Tumours, 2nd edn. Springer: Berlin, Germany; 2000, pp 56–60.
- Heitz PU, Komminoth P, Perren A, *et al*. Pancreatic endocrine tumours: introduction, In: DeLellis A, Lloyd RV (eds) World Health Organization Classification of Tumors. Pathology and Genetics. Tumors of Endocrine Organs. IARC Press: Lyon, France; 2004, pp 177–182.
- Kulke MH, Anthony LB, Bushnell DL, *et al*. NANETS treatment guidelines: well-differentiated neuroendocrine tumors of the stomach and pancreas. *Pancreas* 2010;39:735–752.
- Tang LH, Berlin J, Branton P, *et al*. Protocol for the examination of specimens from patients with carcinoma of the endocrine pancreas, In: Cancer Protocols. College of American Pathologists: Northfield, IL, USA, 2012.
- Edge SB, Byrd DR, Compton CC, *et al*. Exocrine and endocrine pancreas, In: AJCC Cancer Staging Manual, 7th edn. Springer: New York, NY, USA; 2010, pp 241–249.
- Rindi G, Arnold R, Bosman FT, *et al*. Nomenclature and classification of neuroendocrine neoplasms of the digestive system, In: Bosman FT, Carneiro F (eds) WHO Classification of Tumours of the Digestive System, 4th edn. IARC Press: Lyon, France; 2010, pp 13–14.
- Klimstra DS, Modlin IR, Adsay NV, *et al*. Pathology reporting of neuroendocrine tumors: application of the Delphic consensus process to the development of a minimum pathology data set. *Am J Surg Pathol* 2010; 34:300–313.
- Rindi G, Kloppel G, Alhman H, *et al*. TNM staging of foregut (neuro)endocrine tumors: a consensus proposal including a grading system. *Virchows Arch* 2006;449: 395–401.
- Anthony LB, Strosberg JR, Klimstra DS, *et al*. The NANETS consensus guidelines for the diagnosis and management of gastrointestinal neuroendocrine tumors (nets): well-differentiated nets of the distal colon and rectum. *Pancreas* 2010;39:767–774.
- Boudreaux JP, Klimstra DS, Hassan MM, *et al*. The NANETS consensus guideline for the diagnosis and management of neuroendocrine tumors: well-differentiated neuroendocrine tumors of the jejunum, ileum, appendix, and cecum. *Pancreas* 2010;39:753–766.
- Adsay V. Ki67 labeling index in neuroendocrine tumors of the gastrointestinal and pancreatobiliary tract: to count or not to count is not the question, but rather how to count. *Am J Surg Pathol* 2012;36: 1743–1746.
- Tang LH, Gonen M, Hedvat C, *et al*. Objective quantification of the Ki67 proliferative index in neuroendocrine tumors of the gastroenteropancreatic system: a comparison of digital image analysis with manual methods. *Am J Surg Pathol* 2012;36: 1761–1770.
- NIH Image. Available at: <http://rsb.info.nih.gov/nih-image/> accessed 10 September 2014.
- Verbeke CS. Endocrine tumours of the pancreas. *Histopathology* 2010;56:669–682.
- Reid MD, Balci S, Saka B, *et al*. Neuroendocrine tumors of the pancreas: current concepts and controversies. *Endocr Pathol* 2014;25:65–79.
- Dhall D, Frishberg DP, Galliano G, *et al*. Interobserver variability in assessing ki-67 proliferative index in gastrointestinal well-differentiated neuroendocrine neoplasms. *Mod Pathol* 2009;22:116A.
- Goodell PP, Krasinskas AM, Davison JM, *et al*. Comparison of methods for proliferative index analysis for grading pancreatic well-differentiated neuroendocrine tumors. *Am J Clin Pathol* 2012;137:576–582.
- Yang Z, Tang LH, Klimstra DS. Effect of tumor heterogeneity on the assessment of Ki67 labeling index in well-differentiated neuroendocrine tumors metastatic to the liver: implications for prognostic stratification. *Am J Surg Pathol* 2011;35:853–860.
- Fung AD, Cohen C, Kavuri S, *et al*. Phosphohistone h3 and ki-67 labeling indices in cytologic specimens from well-differentiated neuroendocrine tumors of the gastrointestinal tract and pancreas: a comparative analysis using automated image cytometry. *Acta Cytol* 2013; 57:501–508.
- Ong CW, Kim LG, Kong HH, *et al*. Computer-assisted pathological immunohistochemistry scoring is more time-effective than conventional scoring, but provides no analytical advantage. *Histopathology* 2010;56: 523–529.
- Couvelard A, Deschamps L, Ravaud P, *et al*. Heterogeneity of tumor prognostic markers: a reproducibility study applied to liver metastases of pancreatic endocrine tumors. *Mod Pathol* 2009;22:273–281.
- McCall CM, Shi C, Cornish TC, *et al*. Grading of well-differentiated pancreatic neuroendocrine tumors is improved by the inclusion of both Ki67 proliferative index and mitotic rate. *Am J Surg Pathol* 2013;37: 1671–1677.
- Klimstra DS. Pathology reporting of neuroendocrine tumors: essential elements for accurate diagnosis,

- classification, and staging. *Semin Oncol* 2013;40:23–36.
- 24 Yao JC, Phan AT, Chang DZ, *et al*. Efficacy of RAD001 (everolimus) and octreotide LAR in advanced low- to intermediate-grade neuroendocrine tumors: results of a phase II study. *J Clin Oncol* 2008;26:4311–4318.
- 25 NANETS. Advanced islet cell/pancreatic neuroendocrine tumors. Available at: <http://www.nanets.net/research/current-clinical-trials/advanced-islet-cellpancreatic-neuroendocrine-tumors>, accessed 10 September 2014.
- 26 Basturk O, Yang Z, Tang LH, *et al*. Increased (>20%) Ki67 proliferation index in morphologically well differentiated pancreatic neuroendocrine tumors (PanNETs) correlates with decreased overall survival. *Mod Pathol* 2013;26:423A.