

Synaptophysin-Ki67 double stain: a novel technique that improves interobserver agreement in the grading of well-differentiated gastrointestinal neuroendocrine tumors

Karen Matsukuma¹, Kristin A Olson¹, Dorina Gui¹, Regina Gandour-Edwards¹, Yueju Li² and Laurel Beckett²

¹Department of Pathology and Laboratory Medicine, University of California Davis, Sacramento, CA, USA and ²Department of Public Health Sciences, Division of Biostatistics, University of California Davis, Davis, CA, USA

A common problem in the assessment of Ki67 proliferative index in well-differentiated gastrointestinal neuroendocrine tumors is distinguishing tumor from non-tumor. This is because background stromal lymphocytes, entrapped non-neoplastic glands, and the delicate vascular network characteristic of neuroendocrine tumors frequently contain a subset of proliferating cells. Furthermore, in small biopsies, crush and cautery artifact can alter the morphologic appearance of tumor cells, making the Ki67 proliferative index more difficult to assess. To address these issues, we developed a synaptophysin-Ki67 double stain using a commercially available immunohistochemistry kit, allowing simultaneous visualization of tumor and proliferating nuclei. To test this method, three gastrointestinal pathologists individually graded 50 gastrointestinal neuroendocrine tumors first using synaptophysin-Ki67 double-stained slides and then, after a washout period, using Ki67-only stained slides (along with routine hematoxylin- and eosin-stained slides). Interobserver agreement on Ki67 proliferative index was moderate using the Ki67-only stained slides (intraclass correlation 0.51, 95% confidence interval: 0.35–0.66) and improved using the synaptophysin-Ki67 double stain (intraclass correlation 0.79, 95% confidence interval: 0.69–0.86). Similarly, interobserver agreement on tumor grade was fair with Ki67-only stained slides ($\kappa=0.39$, $P<0.001$) and improved with the double stain ($\kappa=0.58$, $P<0.001$). Analysis of individual pathologists' scores revealed that fewer total number of tumor cells counted correlated with higher grade designation and appeared to contribute to grade discordance. In tumors cited as particularly challenging to assess by the pathologists, three of four tumors were grade discordant with the Ki67-only stain, whereas all four tumors were grade concordant with the synaptophysin-Ki67 stain. The synaptophysin-Ki67 double stain is the first technique to address specifically the histomorphologic challenges of evaluating Ki67 proliferative index in well-differentiated gastrointestinal neuroendocrine tumors. Although further validation is needed, this study provides evidence that the synaptophysin-Ki67 double stain can improve interobserver agreement.

Modern Pathology (2017) 30, 620–629; doi:10.1038/modpathol.2016.225; published online 13 January 2017

The current edition of the World Health Organization Classification of Tumors of the Digestive System published in 2010 proposes a unified classification for neuroendocrine neoplasms of all gastrointestinal sites.¹ In this system, tumor grade is entirely dependent on proliferative rate, as determined by

mitotic count and/or Ki67 proliferative index. More recently, it has become clear that tumor grade based on Ki67 proliferative index is a better measure of tumor behavior than mitotic count.^{2–4} Thus, accurate assessment of Ki67 proliferative index is essential for determination of tumor grade. This information, in turn, is essential for accurate prognostication and optimal patient management.

A common problem in the assessment of Ki67 proliferative index in well-differentiated neuroendocrine tumors is that background stromal lymphocytes and entrapped non-neoplastic glands frequently contain proliferating cells. The delicate

Correspondence: Dr K Matsukuma, MD, PhD, Department of Pathology and Laboratory Medicine, University of California Davis, 4400 V Street, Sacramento, CA 95817 USA.
E-mail: kmatsukuma@ucdavis.edu

Received 20 April 2016; revised 13 November 2016; accepted 18 November 2016; published online 13 January 2017

vascular network characteristic of neuroendocrine tumors also contains a subset of proliferating cells. These non-neoplastic cells can be difficult to distinguish from tumor cells on the Ki67 stain and, if mistakenly counted, can artificially elevate the proliferative index. Furthermore, in small biopsies, crush and cautery artifact can alter the morphologic appearance of tumor cells, making the Ki67 proliferative index more challenging to assess.

Because the threshold between grade 1 and grade 2 gastrointestinal neuroendocrine tumors is quite low (by 2010 World Health Organization definition, grade 1 tumors have a Ki67 index of $\leq 2\%$, whereas grade 2 tumors have a Ki67 index of 3 to 20%),¹ mistaking even one or two non-tumor Ki67-positive cells as neoplastic or missing one or two Ki67-positive tumor cells could mean the difference between a grade 1 and grade 2 designation. Making matters more complex, the guidelines do not specify how to interpret Ki67 proliferative indices falling between 2 and 3%, which has led to a variety of interpretations (or no explanation of methodology at all) in the literature.^{2,5–11} More recently, the North American Neuroendocrine Tumor Society and others have recommended that grade 1 tumors be designated as those with Ki67 proliferative index of $< 3\%$.^{12–15} At this time, the general consensus is to use the $< 3\%$ cutoff for grade 1 tumors, and this clarification will also be reflected in the forthcoming edition of the World Health Organization Classification of Tumors of the Digestive System (as per Dr Ralph Hruban, personal communication). However, given that the majority of well-differentiated gastrointestinal neuroendocrine tumors have a Ki67 proliferative index in the 0–5% range, this improvement does not obviate the need for accurate and precise assignment of Ki67 proliferative index to discriminate between grade 1 and grade 2 tumors.

Synaptophysin is a glycoprotein widely expressed in normal neuroendocrine cells of the gastrointestinal tract as well as gastrointestinal neuroendocrine neoplasms. We hypothesized that simultaneous visualization of tumor by synaptophysin expression and proliferating nuclei by Ki67 expression would facilitate more consistent interpretation of gastrointestinal neuroendocrine tumor proliferative index and tumor grade by aiding in the distinction of tumor from non-tumor. To test this idea, we developed a synaptophysin-Ki67 double stain using a commercially available immunohistochemistry kit and conducted a two-part interobserver study, first using synaptophysin-Ki67 double-stained slides and then, after a washout period, using Ki67-only stained slides (along with routine hematoxylin- and eosin-stained slides). The primary aim of the study was to assess interobserver agreement with regard to Ki67 proliferative index and tumor grade. The secondary aim was to evaluate the impact of individual observer, specimen type (ie, biopsy vs resection), and staining method (ie, double stain vs Ki67-only stain) on proliferative index as compared to a gold

standard assessment to identify possible sources of bias.

Materials and methods

Cases

This study was approved by the Institutional Review Board at the University of California Davis Medical Center. Well-differentiated gastrointestinal neuroendocrine tumors diagnosed at the University of California Davis Medical Center from 1 January 2008 to 1 August 2014 were identified through a search of the pathology database. Of the 95 eligible cases, 53 of the most recent cases were selected based on availability of slides and paraffin blocks. Three of these cases were used for optimization of the immunohistochemical staining protocol: one randomly selected tumor from year 2009, one randomly selected tumor from year 2011, and one randomly selected tumor from year 2013 (allowing evaluation of staining quality on tissue blocks stored for different lengths of time). The 50 cases used in the Ki67 quantitative analysis derived from various gastrointestinal sites (8 stomach, 13 small bowel, 5 appendix, 3 colon, 16 rectum, 5 pancreas) and represented both biopsies ($n=29$) and resections ($n=21$).

Immunohistochemistry

The synaptophysin-Ki67 double immunostain was prepared using the Dako Envision Flex, High pH Mini Kit (Dako, Carpinteria, CA, USA) with the G2 system alkaline phosphatase rabbit/mouse Permanent Red Kit and Dako antibodies: synaptophysin (clone DAK-SYNAP, IR660 FLEX RTU Synaptophysin) and Ki67 (clone MIB-1, IR626 FLEX RTU Ki-67). This kit is designed for simultaneous detection of two antigens using horseradish peroxidase with DAB as one chromogen and alkaline phosphatase with Permanent Red as the second chromogen. The DAB chromogen was chosen for detection of Ki67, and the Permanent Red chromogen was chosen for detection of synaptophysin. The initial double stain protocol was designed by the local Dako representative and was deemed adequate by one of the pathologists (KM) based on the intensity and specificity of cytoplasmic synaptophysin and nuclear Ki67 staining, as well as the absence of background staining on the three test sections (Figure 1). Thus, no further optimization was performed. Briefly, 4 μm paraffin-embedded sections were deparaffinized and rehydrated. Antigen retrieval was performed using the Dako PT Link instrument at 97 °C for 20 min. Sections were subjected to Ki67 antibody first (20 min) and horseradish peroxidase (20 min), followed by synaptophysin antibody (20 min), G2 rabbit/mouse secondary antibody (20 min), G2 alkaline phosphatase (20 min), G2 Permanent Red

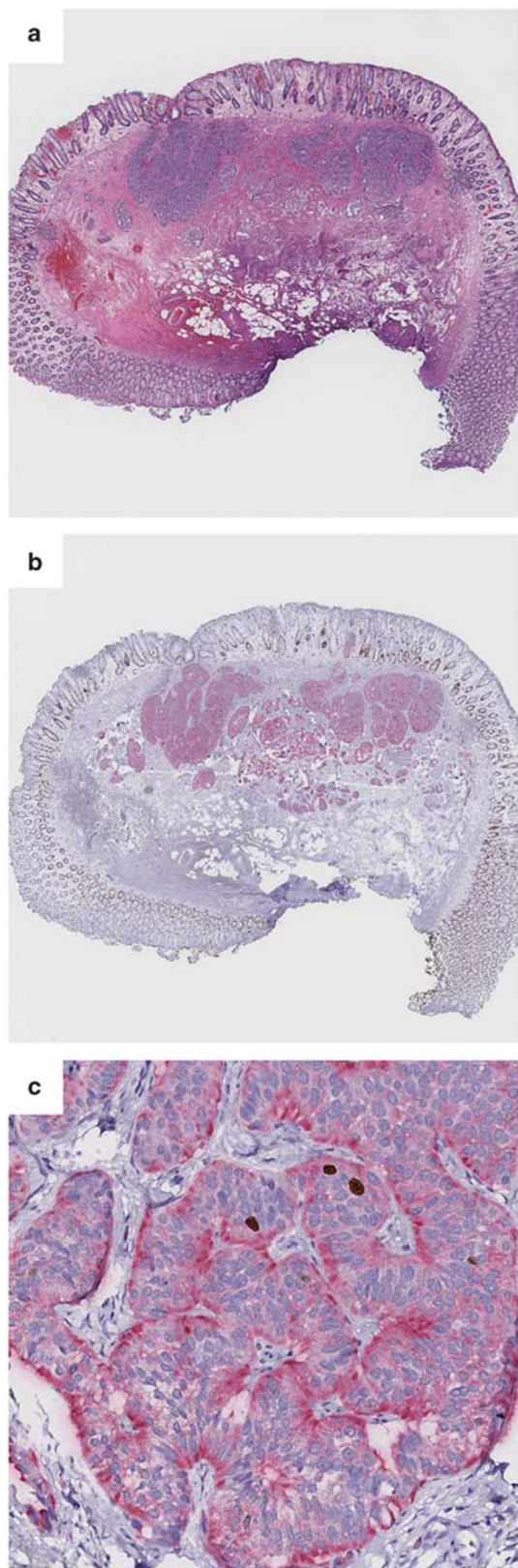


Figure 1 Validation of the synaptophysin-Ki67 double stain, validation case no. 3. (a) Hematoxylin and eosin stain. (b) Synaptophysin-Ki67 double stain. (c) Synaptophysin-Ki67 double stain. All images digitally scanned at x20 magnification.

(10 min), and hematoxylin (2 min) (detailed protocol available upon request). Consecutive tissue sections of the same 50 gastrointestinal neuroendocrine tumors were used to generate Ki67-only stained slides and hematoxylin- and eosin-stained slides. The Ki67-only stained slides were prepared as per the manufacturer's protocol (Dako).

Quantification of Ki67 Proliferative Index

As a first step, three gastrointestinal pathologists (KM, KO, DG) individually assessed the Ki67 proliferative index of the 50 gastrointestinal neuroendocrine tumors using the synaptophysin-Ki67 double-stained slides. Each pathologist was instructed to select the tumor area of highest Ki67 positivity (ie, hotspot) and simultaneously count the number of tumor cells demonstrating nuclear immunoreactivity for Ki67 and the total number of tumor cells (range of 500–2000 tumor cells). Immunopositivity for Ki67 was defined as any brown staining tumor nucleus, regardless of intensity of staining (as described by David Klimstra, personal communication, and Young *et al*¹⁰). All counting was performed by visual assessment using Olympus BX41 microscopes outfitted with x10 (field number 22) oculars, without the use of a graticule. Pathologists each chose the magnification at which counting was performed, and this varied from case to case. These parameters were deliberately chosen to simulate as much as possible typical working conditions. Pathologists tallied tumor cells using a free downloadable smartphone counting application of their choice, such as Cell Counter (<https://play.google.com/store/apps/details?id=edu.sjsu.cellcounter>). Additionally, each pathologist was asked to make a note of any tumor for which Ki67 proliferative index appeared more challenging to assess and also comment on the potential source of difficulty (eg, obscuring inflammation, entrapped non-neoplastic glands).

To minimize the impact of recall bias, a waiting period (ranging from 1 week to 1 month) was instituted. The three pathologists then individually reviewed the set of Ki67-only stained slides (of the same 50 tumors), along with the corresponding hematoxylin- and eosin-stained slides. All other conditions for review of the Ki67-only stained slides were the same as for the double-stained slides. At the completion of the Ki67-only slide review, the pathologists were asked to answer a few written questions regarding ease of use of each Ki67 staining method (ie, double stain, Ki67-only stain), preference of staining method, and the perceived benefits or shortcomings of the synaptophysin-Ki67 double stain.

As a gold standard, one pathologist (KM) independently measured the Ki67 proliferative index of each tumor by selecting and photographing the tumor area of highest Ki67 activity (on Ki67-only stained slides), printing the image on a color printer,

and manually ticking off each immunopositive and immunonegative tumor cell with a pencil until at least 500 neoplastic cells were counted (ie, manual counting method). For imaging, the Infinity 2 camera and Infinity Analyze software (Lumenera Corporation, Ottawa, ON, Canada) were used along with the HP Laserjet 6940 color printer (Hewlett Packard, Palo Alto, CA, USA).

Assessment of grade was determined using the World Health Organization 2010 criteria: grade 1 = Ki67 proliferative index $\leq 2\%$; grade 2 = 3–20%; grade 3 = $>20\%$. Numerical values between 2 and 3% were rounded to the nearest whole number (eg, 2.4% was rounded to 2%; 2.5% was rounded to 3%), as described previously.^{2,6} This rounding method was also the practice of the participating pathologists at the time of the study. Of note, the current North American Neuroendocrine Tumor Society guidelines¹⁴ and the upcoming edition of the World Health Organization Classification of Tumors of the Digestive System (as per Dr Ralph Hruban, personal communication) consider the Ki67 cutoff for grade 1 gastrointestinal neuroendocrine tumors to be $<3\%$. Using this higher threshold increased overall grade concordance for both staining methods, but did not impact the observed variation in Ki67 proliferative index (as described by intraclass correlation).

Statistical Analysis

Intraclass correlations with 95% confidence intervals were calculated to assess inter-rater agreement on Ki67 proliferative index for each staining method (ie, synaptophysin-Ki67 double stain, Ki67-only stain) using the method of Hankinson *et al*,¹⁶ with the SAS package from Hertzmark and Spiegelman.¹⁷ Fisher's Z transformation was used to obtain *P*-values for testing the hypothesis that intraclass correlation = 0. A multirater version of the κ -statistic, as described by Landis and Koch¹⁸ and implemented in the SAS Macro Magree¹⁹ was calculated to assess inter-rater agreement for tumor grade by each staining method.

Friedman's rank test for blocked data was used to compare the relative scoring of Ki67 proliferative index across pathologists to test the null hypothesis that no individual pathologist was systematically scoring higher or lower than the others.

Repeated-measures analysis of variance was used to assess factors influencing the difference between proliferative index (log transformed) as rated by the three pathologists and the gold standard. Factors considered included pathologist, staining method (Ki67-only vs synaptophysin-Ki67 double stain), and type of specimen (biopsy vs resection). Finally, paired *t*-tests were used to compare pathologist score with gold standard score for individual pathologists in specific subgroups, as contrasts to highlight significant findings of the repeated-measures analysis.

Results

Synaptophysin-Ki67 Double Stain

All 50 tumors were diffusely positive for synaptophysin, and all tumors demonstrated internal positive controls for Ki67 (eg, endothelial cells, lymphocytes). Although the intensity of synaptophysin staining from case to case ranged from moderate to strong (2–3+, using a scale analogous to the Allred score for estrogen receptor expression in breast carcinoma²⁰), in each case the borders of the tumor in relation to background stroma were sharp, allowing clear distinction of tumor from non-tumor. No significant intratumoral variation in synaptophysin staining was observed.

Gold Standard

The average number of tumor nuclei counted per tumor for the gold standard assessment was 814 (range 500–1620). The mean Ki67 proliferative index of all tumors was 2.7%, with 62% (31 of 50 tumors) in the 1–5% range and 22% (11 of 50) in the 2–3% range. Per gold standard, the Ki67 proliferative index of tumors ranged from 0.4 to 16.8%. Of these, 39 tumors were grade 1; 11 were grade 2; and none was grade 3. Overall, all three pathologists agreed with the gold standard in 32 of 50 cases when using the Ki67-only stain and in 34 of 50 cases when using the synaptophysin-Ki67 double stain.

Interobserver Agreement

Individual pathologist scores (proliferative index, grade) for each tumor are included in Supplementary Data Tables 1 and 2. Ki67 index as determined by the pathologists ranged from 0 to 29.6%. Pathologist no. 1 counted exactly 500 tumor nuclei for each case, whereas pathologist no. 2 counted an average of 868 tumor nuclei per case (range 501–2205) and pathologist no. 3 counted an average of 1068 tumor nuclei per case (range 500–2751).

Interobserver agreement for Ki67 proliferative index and tumor grade was higher between pathologists when using the synaptophysin-Ki67 double stain. For Ki67 proliferative index, intraclass correlation between pathologists was 0.51 (95% confidence interval 0.35–0.66) when using the Ki67-only stain and 0.79 (95% confidence interval 0.69–0.86) when using the synaptophysin-Ki67 double stain (Table 1). Grade concordance among pathologists was present in 33 of 50 cases ($\kappa = 0.39$, $P < 0.001$, fair agreement) when using the Ki67-only stain and in 37 of 50 cases ($\kappa = 0.58$, $P < 0.001$, moderate agreement) when using the synaptophysin-Ki67 double stain (Table 2). Using a strict $<3\%$ cutoff for grade 1 gastrointestinal neuroendocrine tumors, grade concordance among pathologists was present in 35 of 50 cases with the

Table 1 Intraclass correlations for Ki67 proliferative index of 50 well-differentiated gastrointestinal neuroendocrine tumors using the Ki67-only and synaptophysin-Ki67 slide sets (all three pathologists)

Staining method	Intraclass correlation (95% confidence interval)	P-value ^a
Ki67-only	0.51 (0.35–0.66)	< 0.001
Synaptophysin-Ki67	0.79 (0.69–0.86)	< 0.001

^aTwo-sided vs H₀:r=0.**Table 2** Multirater κ values for grade of 50 well-differentiated gastrointestinal neuroendocrine tumors using the Ki67-only and synaptophysin-Ki67 slide sets (all three pathologists)

Staining method	κ-value	P-value
Ki67-only	0.39	< 0.001
Synaptophysin-Ki67	0.58	< 0.001

Ki67-only stain and in 38 of 50 cases with the synaptophysin-Ki67 stain.

Analysis of individual pathologists' scores revealed that pathologist no. 1 tended to have the highest Ki67 proliferative index among the pathologists, a finding that was statistically significant for both staining methods ($P < 0.001$ by Friedman test).

Subgroup analysis of pathologists' scores as compared with the gold standard revealed that pathologist no. 1 tended to overscore biopsy cases using the Ki67-only stain ($P = 0.02$), but had no other systematic differences compared with the gold standard, whereas pathologist nos 2 and 3 tended to underscore cases compared with the gold standard regardless of staining method or specimen type (P -values ≤ 0.05) (Table 3).

Discordant Cases

Twenty cases were discordant among the pathologists using either the Ki67-only or synaptophysin-Ki67 staining method. Of those, 10 cases were discordant for both staining methods, seven cases were discordant only with the Ki67-only method, and three cases were discordant only with the synaptophysin-Ki67 method. The most common cause of grade discordance for either staining method was higher grade reported by pathologist no. 1, accounting for 53% of discrepancies.

Pathologist Feedback

Feedback from pathologists with regard to the synaptophysin-Ki67 double stain was positive overall, and the double stain was the stain of choice for assessment of Ki67 proliferative index. Pathologists cited improved confidence when counting tumor nuclei in areas containing a significant number of obscuring non-neoplastic cells as the primary reason

Table 3 Analysis of Ki67 proliferative index by pathologist, staining method, and specimen type—compared with gold standard (biopsy and/or resection, as relevant)

Variable	Estimate ^a	Standard error	P-value
Pathologist no. 1			
Ki67-only stain			
Biopsies and resections	0.224	0.099	0.030
Biopsies only	0.317	0.127	0.020
Resections only	0.096	0.157	0.550
Synaptophysin-Ki67 stain			
Biopsies and resections	-0.018	0.123	0.880
Biopsies only	0.121	0.127	0.350
Resections only	-0.230	0.239	0.350
Pathologist no. 2			
Ki67-only stain			
Biopsies and resections	-0.408	0.097	0.0001
Biopsies only	-0.242	0.101	0.020
Resections only	-0.637	0.174	0.002
Synaptophysin-Ki67 stain			
Biopsies and resections	-0.382	0.099	0.0003
Biopsies only	-0.428	0.130	0.003
Resections only	-0.319	0.156	0.050
Pathologist no. 3			
Ki67-only stain			
Biopsies and resections	-0.562	0.078	< 0.0001
Biopsies only	-0.626	0.100	< 0.0001
Resections only	-0.473	0.123	0.001
Synaptophysin-Ki67 stain			
Biopsies and resections	-0.378	0.087	< 0.0001
Biopsies only	-0.342	0.109	0.004
Resections only	-0.434	0.156	0.008

^aEstimate is the mean of the difference of proliferative index on a log scale between pathologists and gold standard (pathologist-gold standard).

for their preference of the synaptophysin-Ki67 double stain (Figure 2).

Pathologists were also asked to note those cases with which they experienced particular difficulty evaluating the Ki67 proliferative index. In total, 17 different cases were identified by the pathologists as being challenging to assess due to obscuring non-neoplastic cells (Supplementary Table 3). Of these, six cases were discrepant using the Ki67-only slides, whereas only two cases (a subset of the six) were discrepant using the synaptophysin-Ki67-stained slides. In four of the 17 cases, two or more comments were made by pathologists regarding challenges in interpretation related to obscuring non-neoplastic cells. Of these, three of four tumors showed grade discordance with the Ki67-only stain. In contrast, all four tumors were grade concordant with the synaptophysin-Ki67 double stain (Figure 3).

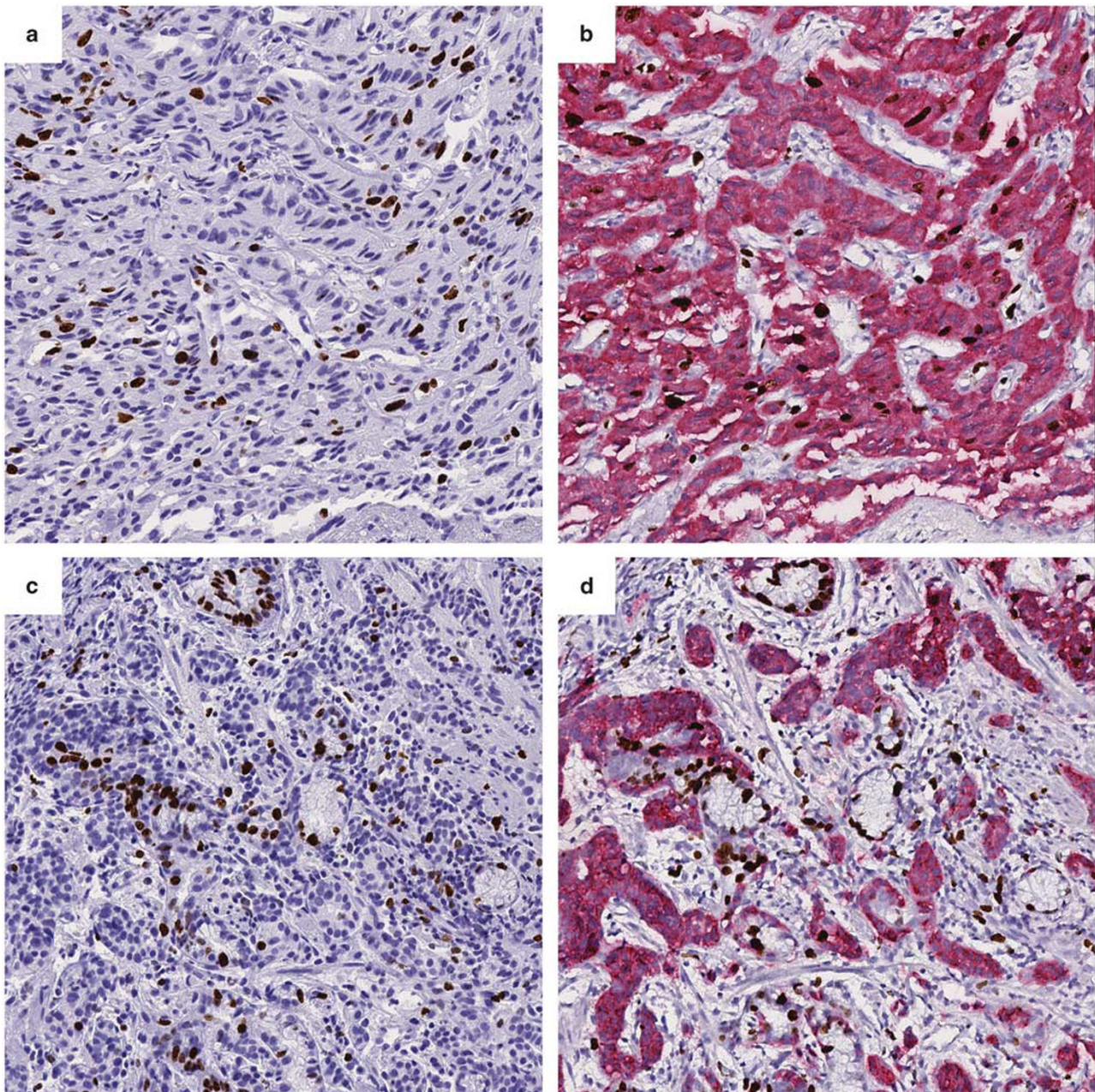


Figure 2 The synaptophysin-Ki67 double stain enhances distinction of tumor from non-tumor, particularly in the presence of proliferating intratumoral vessels (a and b, test case no. 17), entrapped non-neoplastic glands (c and d, test case no. 22), and crush artifact (e and f, test case no. 29). Left panel, Ki67-only stain; right panel, synaptophysin-Ki67 double stain. All images digitally scanned at x20 magnification.

Discussion

It is now well established that performing a Ki67 count is necessary for all gastrointestinal neuroendocrine tumors, as endorsed by the World Health Organization, European Neuroendocrine Tumor Society, American Joint Committee on Cancer, College of American Pathologists, and North American Neuroendocrine Tumor Society. This mandate is based on a large body of evidence demonstrating that the Ki67 proliferative index holds prognostic significance. However, implicit within this mandate

is a requirement that the count be accurate and reproducible. Precision is a key issue in the usefulness of numerical cutoffs for grading, and thus identifying sources of interobserver variation and developing strategies to overcome them is essential. Also critical is the need to balance improvements in precision with consideration of laboratories' financial resources (particularly across an international community) and their impact on pathologists' workload and turnaround time.

Obscuring non-tumor Ki67-positive nuclei are a well-known source of error in deriving the Ki67

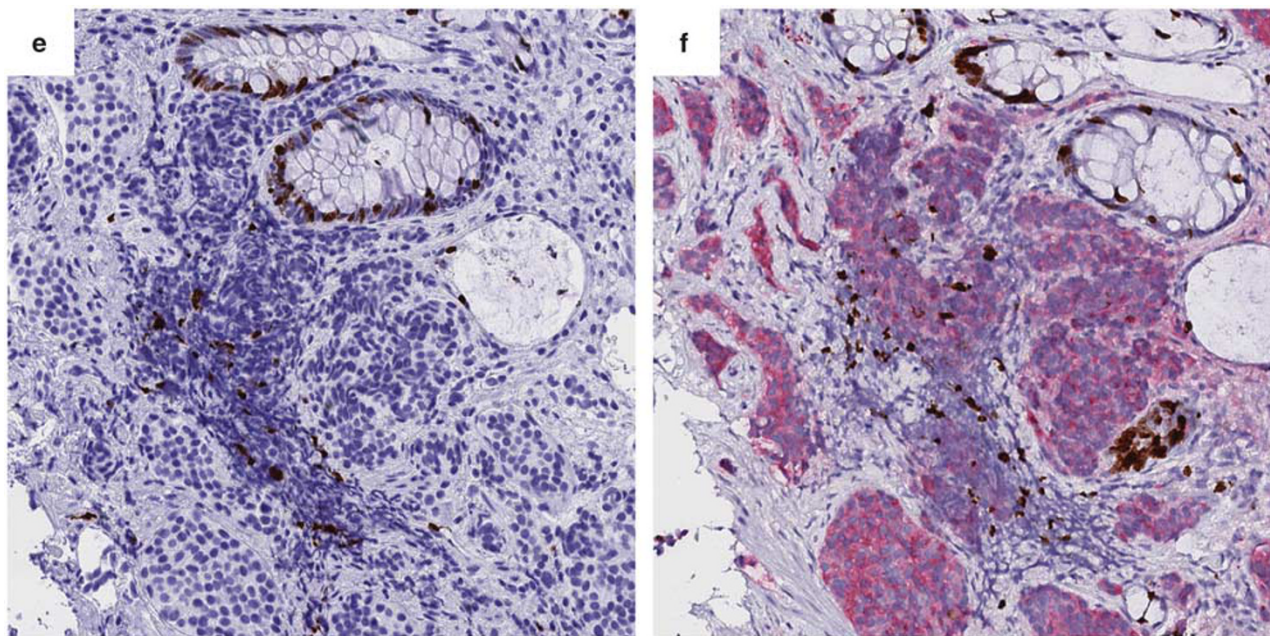


Figure 2 Continued.

proliferative index of gastrointestinal neuroendocrine tumors.^{5,7,21} Unlike automated image analysis and manual image analysis that focus on simplifying the arduous task of counting 500–2000 tumor cells, the synaptophysin-Ki67 double stain is the first technique specifically directed at overcoming human error related to distinguishing gastrointestinal neuroendocrine tumor from non-tumor. Importantly, errors related to mis-assignment of Ki67-positive nuclei to tumor are mathematically far more deleterious to precision than are accidentally including a few more (or missing a few) Ki67-negative tumor nuclei. While there are some automated image analysis programs that will consider nuclear size and shape to avoid counting lymphocytes and/or stromal cells, there is some disagreement regarding how successful these programs are.^{5,21} In addition, such programs require substantial capital investment and considerable technical expertise. More recently, free online image analysis tools using photocopied tumors have become available.^{22,23} These software programs decrease capital costs but suffer from the same limitations as other automated tools with regard to distinguishing tumor from non-tumor.

The synaptophysin-Ki67 double stain was developed with minimal technical expertise (above and beyond what is necessary in a lab performing routine automated immunohistochemistry) and required a minimal number of validation steps. Furthermore, as per the manufacturer, the antibody vials need not be dedicated to the double stain technique. That is, the same vial of Ki67 antibody routinely used as a single immunostain (eg, for the evaluation of breast and hematolymphoid malignancies) can be used for the double stain, albeit not simultaneously. Similarly,

the same vial of synaptophysin antibody used as a general marker of neuroendocrine differentiation can be used for both single and double chromogen applications. Thus, the main added cost is the double stain kit itself. This price varies according to institutional contract, and our local representative quoted the following prices: US\$2847.95 for the Envision Flex Mini Kit, High pH, 125–190 tests (Dako; K802321) and US\$1184.50 for Envision G2 System/AP, Rabbit/Mouse Permanent Red, 50 tests (Dako; K535511). Certainly, this is an upfront cost that would need to be considered in the utilization of this methodology.

On the surface, the interobserver agreement (as demonstrated by intraclass correlations and κ -statistics) might appear relatively low in the current study as compared with other interobserver studies on Ki67 scoring.^{7,10,24} However, in these other studies, hotspot selection was taken out of the equation by having pathologists review predetermined fields or tissue microarray samples,^{4,7,10,24,25} both of which have essentially a single field for counting Ki67-positive nuclei. This format, while easing the task of the study's pathologists, does not reflect typical working conditions, and thus limits its applicability to general practice. Emphasizing the importance of this, Adsay noted, 'In our experience, almost half of the cases are placed in a different World Health Organization grade on the basis of the area chosen for counting.'⁵ In our study, interobserver agreement for Ki67-only stained slides was similar to that seen by Reid *et al*²¹ using their 'manual eye count'—a method equivalent to that used in the current study. It is also worth noting that a subset of studies focuses solely on pancreatic neuroendocrine tumors^{8,21,26}—

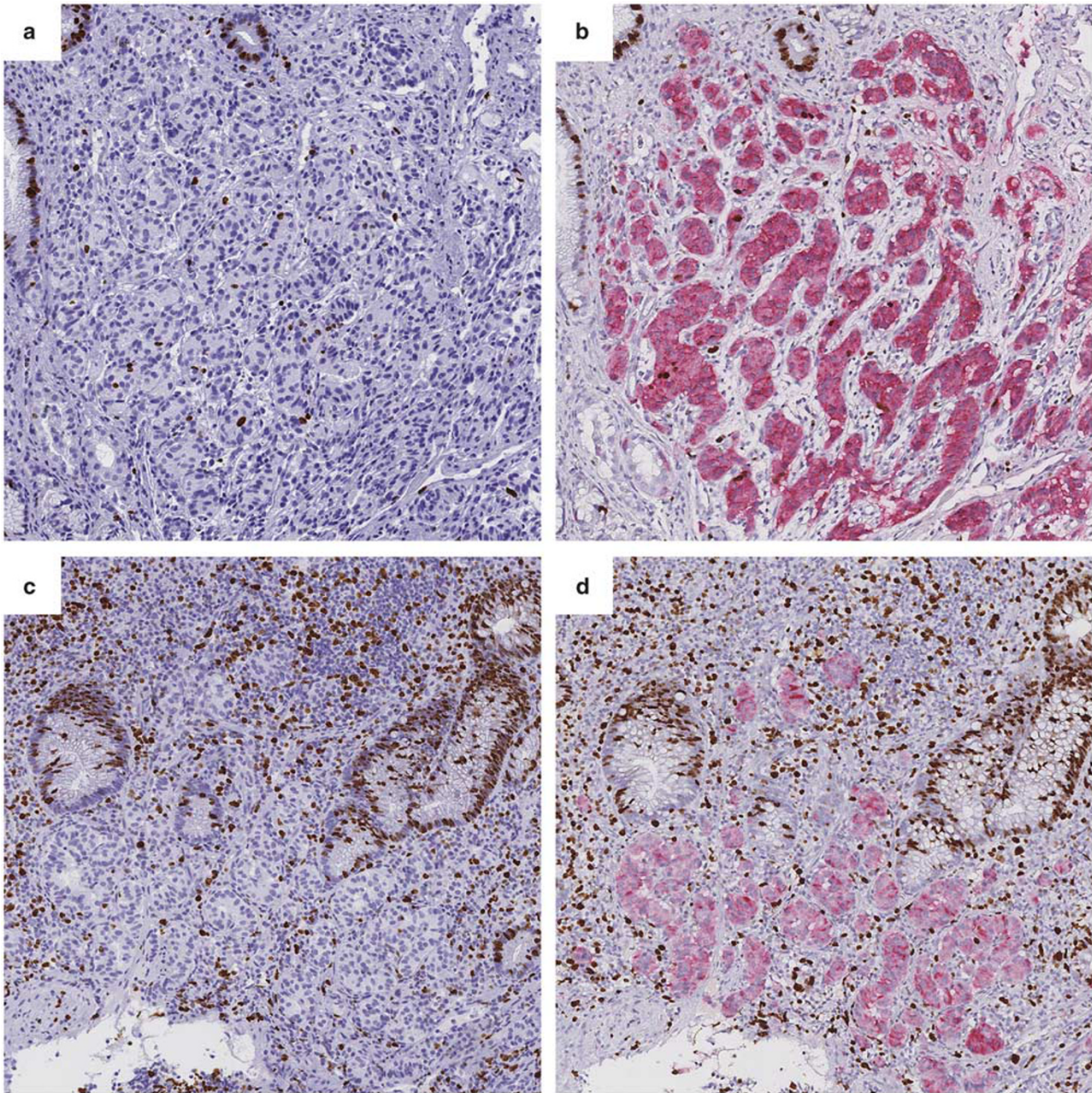


Figure 3 Three cases considered challenging with regard to assessment of Ki67 proliferative index. (a and b) Test case no. 2: This case was grade discordant when the pathologists reviewed the Ki67-only stained slide but grade concordant on the synaptophysin-Ki67 slide. (c and d) Test case no. 12: This case was grade discordant on the Ki67-only stained slide and grade concordant on the synaptophysin-Ki67 slide. (e and f) Test case no. 41. This case was considered challenging but was nonetheless grade concordant on both the Ki67- and synaptophysin-Ki67-stained sections. Left panel, Ki67-only stain; right panel, synaptophysin-Ki67 double stain. All images digitally scanned at x20 magnification.

which are morphologically more homogeneous with few peritumoral stromal lymphocytes or entrapped non-neoplastic glands. Thus, these studies may not capture the intricacies related to assessing Ki67 proliferative index of gastrointestinal neuroendocrine tumors in general.

At the same time, our approach that deliberately left many more variables in the test system was also a limitation of the study. A more focused approach

using images of the same area of tumor on both the Ki67-only and synaptophysin-Ki67-stained slides would have assessed more specifically the value of the double stain method in distinguishing tumor from non-tumor. Additionally, by allowing pathologists to count within a range of 500–2000 tumor cells, we found there was a wide variation in the number of tumor cells counted. In particular, pathologist no. 1 counted 300–500 fewer tumor cells

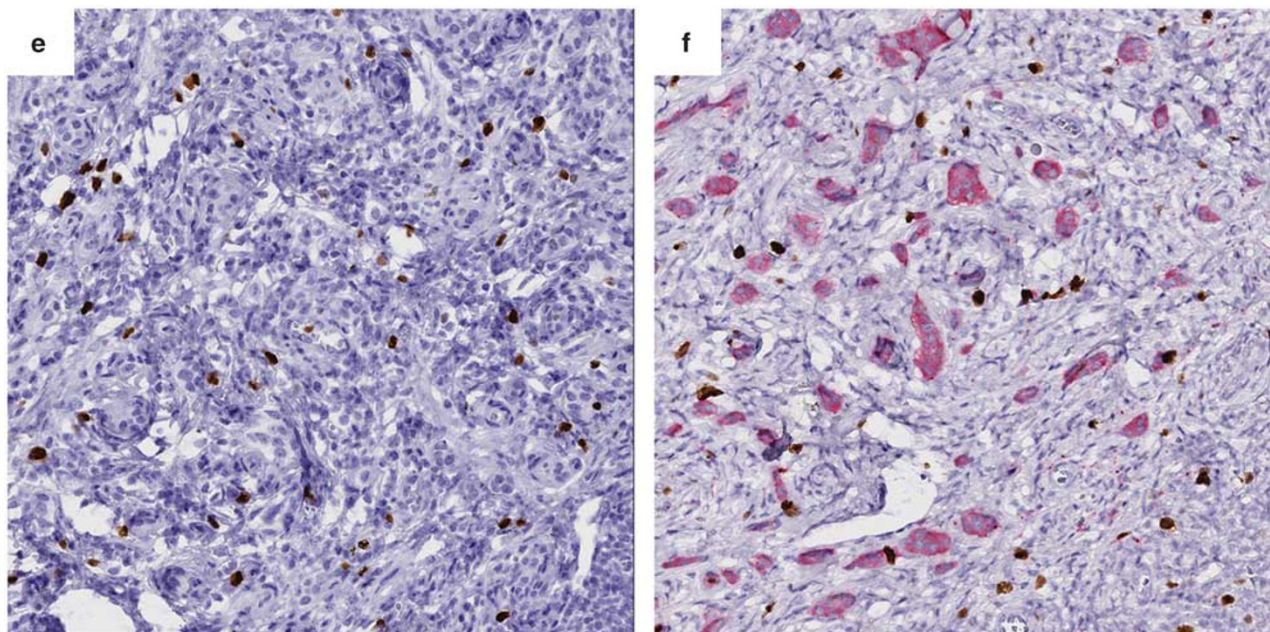


Figure 3 Continued.

on average than the other two pathologists. This correlated with generally higher grade designations, a finding previously observed by others.⁴ Notably, higher grade designations by pathologist no. 1 were the most common cause of grade discordance among the pathologists, accounting for a majority of the grade discrepancies for both staining methods. Thus, counting method alone may have a significant impact on grade.

Although we were not expecting to find any significant differences in pathologists' scores in relation to the gold standard, the fact that pathologist no. 1 tended to overscore biopsies (but not resections) was intriguing. It seems plausible that because the choice of hotspot is more limited in biopsies, more frequent counting in areas containing obscuring non-tumor Ki67-positive cells may be necessary, a situation in which there is increased risk of overcounting. In addition, both pathologist nos 2 and 3 underscored all tumors regardless of staining method compared with the gold standard; one possibility is that eye counting overcounts the number of Ki67-negative tumor nuclei compared with the manual counting method (ie, ticking off tumor nuclei on a printed image). Of note, we chose the eye counting method for this study because it reflects our current practice; however, the findings suggest a re-evaluation of this protocol is warranted.

In conclusion, this study provides preliminary evidence that the synaptophysin-Ki67 double stain improves interobserver agreement in the grading of well-differentiated gastrointestinal neuroendocrine tumors. Although improved grade concordance was restricted to a small number of cases, the overall increase in agreement on Ki67 proliferative index (as described by intraclass correlation) is significant in

that it is a measure entirely independent of grade thresholds. Furthermore, that grade concordance specifically improved in the subset of cases deemed challenging by the pathologists emphasizes the value of the synaptophysin-Ki67 stain for its intended purpose and suggests that its greatest utility is in this pathologist-defined set of cases.

Acknowledgments

We thank Dr Lee-Way Jin for critical review of this manuscript and Ramona Clarke for technical assistance. Support for this project was provided by intradepartmental faculty funding (KM). The technical component of the work performed at the University of California Davis was supported in part by the UC Davis Comprehensive Cancer Center Support Grant awarded by the National Cancer Institute (NCI P30CA093373).

Disclosure/conflict of interest

The authors declare no conflict of interest.

References

- 1 Rindi G, Arnold R, Bosman FT, *et al*. Nomenclature and classification of neuroendocrine neoplasms of the digestive system. In: Bosman FT, Carneiro C, Hruban RH, *et al* (eds). *WHO Classification of Tumours of the Digestive System. World Health Organization Classification of Tumours*, 4th edn. International Agency for Research on Cancer: Lyon, France, 2010, pp 13–14.

- 2 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;37:671–1677.
- 3 Khan MS, Luong TV, Watkins J, *et al*. A comparison of Ki-67 and mitotic count as prognostic markers for metastatic pancreatic and midgut neuroendocrine neoplasms. *Br J Cancer* 2013;108838–1845.
- 4 van Velthuysen ML, Groen EJ, Sanders J, *et al*. Reliability of proliferation assessment by Ki-67 expression in neuroendocrine neoplasms: eyeballing or image analysis?. *Neuroendocrinology* 2014;10088–292.
- 5 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.
- 6 Remes SM, Tuominen VJ, Helin H, *et al*. Grading of neuroendocrine tumors with Ki-67 requires high-quality assessment practices. *Am J Surg Pathol* 2012;36:1359–1363.
- 7 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.
- 8 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.
- 9 Nadler A, Cukier M, Rowsell C, *et al*. Ki-67 is a reliable pathological grading marker for neuroendocrine tumors. *Virchows Arch* 2013;462:501–505.
- 10 Young HT, Carr NJ, Green B, *et al*. Accuracy of visual assessments of proliferation indices in gastroenteropancreatic neuroendocrine tumours. *J Clin Pathol* 2013;66:700–704.
- 11 Blank A, Wehweck L, Marinoni I, *et al*. Interlaboratory variability of MIB1 staining in well-differentiated pancreatic neuroendocrine tumors. *Virchows Arch* 2015;467:543–550.
- 12 Klimstra DS, Modlin IR, Coppola D, *et al*. The pathologic classification of neuroendocrine tumors: a review of nomenclature, grading, and staging systems. *Pancreas* 2010;39:707–712.
- 13 Klimstra DS. Pathology reporting of neuroendocrine tumors: essential elements for accurate diagnosis, classification, and staging. *Semin Oncol* 2013;40:23–36.
- 14 Kunz PL, Reidy-Lagunes D, Anthony LB, *et al*. Consensus guidelines for the management and treatment of neuroendocrine tumors. *Pancreas* 2013;42:557–577.
- 15 Reid MD, Balci S, Saka B *et al*. Neuroendocrine tumors of the pancreas: current concepts and controversies. *Endocr Pathol* 2014;25:65–79.
- 16 Hankinson SE, Manson JE, Spiegelman D, *et al*. Reproducibility of plasma hormone levels in postmenopausal women over a 2–3-year period. *Cancer Epidemiol Biomarkers Prev* 1995;4:649–654.
- 17 Hertzmark E, Spiegelman D. SAS ICC9 Macro 2010. Available at: <https://cdn1.sph.harvard.edu/wp-content/uploads/sites/271/2012/09/icc9.pdf> (20 April 2016).
- 18 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174.
- 19 SAS. Magree Macro V1.0. Available at: <https://github.com/friendly/SAS-macros/blob/master/magree.sas> (20 April 2016).
- 20 Harvey JM, Clark GM, Osborne CK, *et al*. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. *J Clin Oncol* 1999;17:1474–1481.
- 21 Reid 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;28:686–694.
- 22 Immunoratio. Available at: <http://153.1.200.58:8080/immunoratio/> (20 April 2016).
- 23 Image-J. Available at: <http://rsbweb.nih.gov/ij/features.html> (20 April 2016).
- 24 Polley MY, Leung SC, Gao D, *et al*. An international study to increase concordance in Ki67 scoring. *Mod Pathol* 2015;28:778–786.
- 25 Mengel M, von Wasielewski R, Wiese B, *et al*. Interlaboratory and inter-observer reproducibility of immunohistochemical assessment of the Ki-67 labelling index in a large multi-centre trial. *J Pathol* 2002;198:292–299.
- 26 Kroneman TN, Voss JS, Lohse CM, *et al*. Comparison of three Ki-67 index quantification methods and clinical significance in pancreatic neuroendocrine tumors. *Endocr Pathol* 2015;26:255–262.

Supplementary Information accompanies the paper on Modern Pathology website (<http://www.nature.com/modpathol>)