OPEN

Digital Image Analysis of the Proliferation Markers Ki67 and Phosphohistone H3 in Gastroenteropancreatic Neuroendocrine Neoplasms: Accuracy of Grading Compared With Routine Manual Hot Spot Evaluation of the Ki67 Index

Dordi Lea, MD,†‡ Einar G. Gudlaugsson, MD, PhD,* Ivar Skaland, PhD,* Melinda Lillesand, MSc,* Kjetil Søreide, MD, PhD, FACS, FRCSEd,†‡§ and Jon A. Søreide, MD, PhD, FACS, FISS†§*

Abstract: Gastroenteropancreatic neuroendocrine neoplasms (GEP-NENs) are rare epithelial neoplasms. Grading is based on mitotic activity or the percentage of Ki67-positive cells in a hot spot. Routine methods have poor intraobserver and interobserver consistency, and objective measurements are lacking. This study aimed to evaluate digital image analysis (DIA) as an objective assessment of proliferation markers in GEP-NENs. A consecutive cohort of patients with automated DIA measurement of Ki67 (DIA Ki67) and phosphohistone H3 (DIA PHH3) on immunohistochemical slides was analyzed using Visiopharm image analysis software (Hoersholm, Denmark). The results were compared with the Ki67 index from routine pathology reports (pathology Ki67). The study included 159 patients (57% males). The median pathology Ki67 was 2.0% and DIA Ki67 was 4.1%. The interclass correlation coefficient of the DIA Ki67 compared with the pathology Ki67 showed an excellent agreement of 0.96 [95% confidence interval (CI): 0.94-0.96]. The observed kappa value was 0.86 (95% CI: 0.81-0.91) when comparing grades based on the same methods. PHH3 was measured in 145 (91.2%) cases. The observed kappa value was 0.74. (95% CI: 0.65-0.83) when comparing grade based on the DIA PHH3 and the pathology Ki67. The DIA Ki67 shows excellent agreement with the pathology Ki67. The DIA PHH3 measurements were more varied and cannot replace other methods for grading GEP-NENs.

Key Words: neuroendocrine tumor, neuroendocrine carcinoma, proliferation, digital image analysis, immunohistochemistry

(*Appl Immunohistochem Mol Morphol* 2021;29:499–505)

Gastroenteropancreatic neuroendocrine neoplasms (GEP-NENs) comprise a heterogeneous group of rare, benign, or malignant epithelial tumors (carcinoids) originating from the pancreas (PNENs) or gastrointestinal tract (GI-NETs). The reported annual incidence varies between 2.39 and 5.83 per 100,000 inhabitants according to international literature,[1,2] with an estimated prevalence of 35 per 100,000 because of the long survival times. The 5-year survival rates vary between 40% and 100% and are associated with the primary tumor site, tumor grade, and stage of disease at the time of diagnosis.[3–5] Moreover, GEP-NENs classified as functional tumors (which secrete hormones or peptides to cause clinical symptoms or syndromes) show a different biological behavior from those classified as nonfunctional GEP-NENs,[6] and tumor behavior is also associated with the histopathologic pattern, including the features of an adenocarcinoma.[7] The diagnostic criteria of neuroendocrine tumors are based on morphology and the positive staining of the neuroendocrine markers synaptophysin and/or chromogranin A by immunohistochemistry (IHC).[8]

According to the World Health Organization (WHO) criteria, the grading of GEP-NENs is based on the evaluation of mitotic activity, either by counting mitosis, the so-called "mitotic activity index" (MAI), on hematoxylin and eosin (HE)-stained slides or by calculating the percentage of Ki67-positive cells in a hot spot (Table 1).[9,10] The highest grade should apply if any discordance between the MAI and Ki67 index assessment occurs.[10] The Ki67 index predicts prognosis better than MAI.[11]

Received for publication November 23, 2020; accepted February 22, 2021.

From the Departments of *Pathology; §Gastrointestinal Surgery; ‡Gastrointestinal Translational Research Unit, Molecular Laboratory, Hillevåg, Stavanger University Hospital, Stavanger; and †Department of Clinical Medicine, University of Bergen, Bergen, Norway.

This study was financially supported in part by an unrestricted research grant provided by the CarciNor, which is the Norwegian patient advocacy association for patients with neuroendocrine tumors.

The authors declare no conflict of interest.

Reprints: Dordi Lea, MD, Department of Pathology, Stavanger University Hospital, Pb 8100, Stavanger N-4068, Norway (e-mail: dordilea@gmail.com).

TABLE 1. WHO Grading of Neuroendocrine Tumors (WHO 2019)

	Grade	Mitotic Activity, Per 2 mm^2*	Ki67%*
NET Grade 1	Low	1	< 3
NET Grade 2	Intermediate	2-20	3-20
NET Grade 3	High	> 20	> 20
LCNEC	High†	> 20	> 20
SCNEC	High†	> 20	> 20
MiNEN	Variable	Variable	Variable

*Mitotic rates are expressed as the number of mitoses/2 mm^2 as determined by counting in 50 fields of 0.2 mm^2 (ie, in a total area of 10 mm^2); the Ki67 proliferation index value is determined by counting at least 500 cells in the regions of highest labeling (hot spots), which are identified at scanning magnification.

†Poorly differentiated NECs are not formally graded but are considered high-grade by definition.

LCNEC indicates large-cell neuroendocrine carcinoma; MiNEN, mixed neuroendocrine–non-neuroendocrine neoplasm; NEC, neuroendocrine carcinoma; NET, neuroendocrine tumor; SCNEC, small-cell neuroendocrine carcinoma; WHO, World Health Organization.

One of the challenges related to the current routine grading procedure is the time-consuming counting of > 500 cells by a pathologist, which may lead to an eyeball estimation as a short-cut in a busy routine practice. Moreover, the identification of "hot spots" in a section may be difficult,[12] which may partly explain the reported poor intraobserver and interobserver reliability.[13] Consequently, objective measurements, including digital image analysis (DIA), are warranted for accurate grade reporting.

The heterogeneous biological behavior of GEP-NENs has encouraged a search for better prognostic markers. Phosphohistone H3 (PHH3) has been identified as a promising marker for the prediction of disease-free survival and disease-specific survival in PNENs.[14,15] In contrast to Ki67, which is present in cell nuclei in the G1, S, G2, and M phases of the cell cycle, PHH3 stains mitotic cells (M phase). Thus, with PHH3, mitotic activity can be specifically determined;[16] therefore, PHH3 has been suggested as an alternative to the Ki67 index in PNENs.[17] Furthermore, counting MAI is challenging because apoptotic figures can be misidentified as mitotic figures.[17,18] PHH3 does not stain apoptotic cells and can therefore be a better biomarker for mitosis than MAI. Several studies have shown good concordance between the number of mitoses and PHH3.[19,20] Accordingly, PHH3 is regarded as promising for the assessment of grading in GEP-NENs in general.

In this study, we evaluated and compared the DIA Ki67 with the routine procedure for the Ki67 index assessment (pathology Ki67). In addition, we applied DIA for PHH3 assessment (DIA PHH3) to explore possible associations between PHH3 and the Ki67 index to evaluate the potential advantages or challenges of PHH3 as a proliferation marker in routine practice.

MATERIALS AND METHODS

Ethical Approval

All procedures performed in studies involving human participants were per the ethical standards of the institutional and/or National Research Committee and the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This project was approved by the Regional Ethics Committee of the Western Health Authority (REK 2016/1622). Patients still alive have signed a written consent form to participate.

Materials

We identified all consecutive patients diagnosed with GEP-NENs and treated at Stavanger University Hospital from 2003 to 2013. The hospital serves as the only hospital for a well-defined Norwegian population of ~380,000 people.[2] Of 204 consecutive patients during that period, 35 declined to participate. In addition, 1 patient was excluded because of the possibility of primary pulmonary neuroendocrine carcinoma, and 9 patients were excluded because of a lack of tissue for analysis. Thus, 159 (77.9%) patients were included.

Archived formalin-fixed paraffin-embedded tissue from the hospital's diagnostic biobank was obtained. The current WHO criteria for neuroendocrine tumors[10] were used, and pathologic tumor-node-metastasis (TNM) staging was performed according to the American Joint Committee on Cancer (AJCC) 8th edition.[21] For Ki67, 500 to 2000 tumor cells were assessed in hot spots by microscopic evaluation.[21] The Ki67 index was retrieved from the original routine pathology reports, and cases without available information were re-evaluated to complete pertinent information on all patients, as reported in our previous study.[2]

All tumors were confirmed as NENs by positive IHC staining for synaptophysin and/or chromogranin A. Neuroendocrine carcinoma was included as grade 3 (high grade). Mixed neuroendocrine–non-neuroendocrine neoplasms were excluded. The sample included 63 (39.6%) biopsies and 96 (60.4%) surgical specimens. MAI was not evaluated because of the high number of biopsies with an area <10 mm^2.

Methods

IHC

Ki67. The MIB-1 clone (Dako, Glostrup, Denmark) was used for routine staining at the Department of Pathology. The method has had minor changes in processing from 2003 to 2013, but the MIB-1 clone has been the same. The MIB-1 clone is validated for GEP-NEN grading.[22,23]

PHH3. Antigen retrieval and antibody dilution were optimized before study onset. Paraffin sections adjacent to the HE sections were cut into 2-µm-thick sections and mounted on Superfrost Plus slides. The slides were incubated at 60°C for 1 hour and stained using a Dako Omnis immunostainer. PHH3 (nr. 06-570; Merck Group, Darmstadt, Germany) was used at a dilution of 1:5000. The Dako EnVision Flex+ (Dako GV80011-2) detection system was used in line with the recommendations of the manufacturer.

DIA

IHC staining of Ki67 was already performed, and the available archived sections were retrieved from the hospital's archive. HE sections, Ki67, and PHH3 were scanned at ×40 magnification using a Leica SCN400 slide

scanner (Leica Biosystems, Wetzlar, Germany) and uploaded to the image analysis software (Visiopharm, Hoersholm, Denmark). The person responsible for the DIA evaluation (D.L.) was blinded to the previously reported routine Ki67 index results and other parameters when DIA was performed. Patients were excluded from the DIA if there was insufficient tumor material for analysis (< 500 tumor cells). If a biopsy material and not a surgical specimen was used for the original routine Ki67 evaluation, DIA was also performed on the biopsy sample, given that sufficient materials were available.

For Ki67, the percentage of positive tumor cells was measured in the hot spot of the tumor in an area that includes 500 to 2000 tumor cells with the Visiopharm program (Fig. 1). The software program identified positive nuclei (label 1) and negative nuclei (label 2) within a manually selected area named the region of interest (ROI). Stroma and stromal cells were excluded from the ROI by the software program so that only the tumor cells were evaluated, similar to our DIA Ki67 method described for breast cancer.[12] The percentage of tumor cells was calculated. If the hot spot was ill-defined on a slide, the measurement was repeated in different areas, and the area with the highest positivity was counted. The program measured the areas of positive and negative nuclei, and based on the size of the nuclei in each tumor, an estimate of tumor cells was calculated. A percentage was calculated from ~2000 tumor cells. If there were <2000 cells but > 500 cells, a percentage was given based on the available cells in the section. This value was compared with the pathology Ki67 (regarded as the gold standard).

The number of PHH3-positive cells was calculated in 4 different ROIs of 2 mm^2 (ROI: 1 to 4) (Fig. 2). The different ROIs were chosen subjectively from the visual identification of areas with the highest number of PHH3-positive cells. The number of PHH3-positive cells within each ROI was calculated. If there was insufficient material for 4 ROIs, fewer ROIs were chosen for measurement. For PHH3, Visiopharm was programmed to detect IHC-stained mitotic cells as described by others.[24] Cells in all 4 substages of mitosis, prophase, metaphase, anaphase, and telophase, were regarded as positive.[25] Objects smaller than mitotic figures were removed by a size filter. The remaining objects were dilated to fuse the chromatin structures of cells in anaphase or telophase into 1 object. All remaining objects were counted, and the number of objects per 2 mm^2 was calculated. The counted objects were encircled in the original image, allowing a visual inspection of the counted mitotic cells. Apart from the manual selection of the 4 different ROIs, the DIA procedure and all calculations were fully automated.

Statistics

IBM SPSS Statistics for Windows, Version 26.0 (IBM Corporation, Armonk, NY) was used for statistical calculations. The quadratically weighted kappa was used to measure the agreement between ordinal variables.[26] The interclass correlation coefficient was used to measure the agreement between continuous variables (single rater, absolute agreement). All agreement estimates are presented with 95% confidence intervals (CIs). Values <0.50, between 0.50 and 0.75, between 0.75 and 0.90, and > 0.90 indicated poor, moderate, good, and excellent reliability, respectively.[27] To plot the difference in pathology Ki67 and DIA Ki67 measurement against average value, we used a Bland-Altman plot.[28]

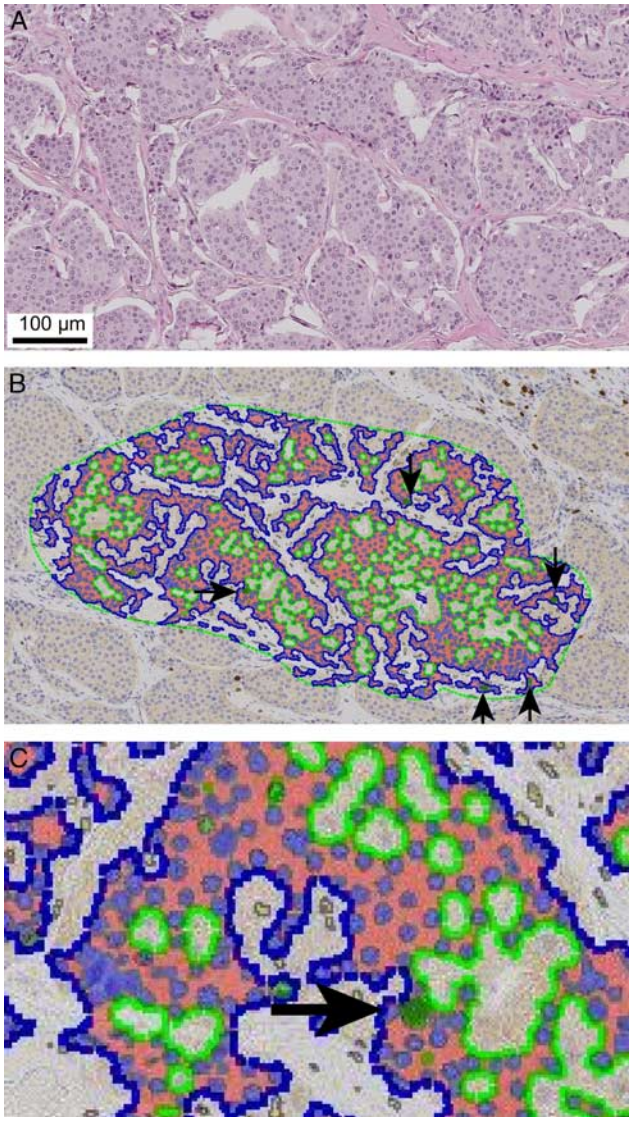

FIGURE 1. A, Hematoxylin and eosin staining of a grade 2 neuroendocrine tumor. B and C Immunohistochemical staining of Ki67 with digital image analyses performed on the same tumor. Black arrows point to some of the positive cells.

RESULTS

The clinicopathologic characteristics of the patients are shown in Table 2. The distribution of grading based on the DIA Ki67, DIA PHH3, and pathology Ki67 are shown in Figure 3. This figure shows that different methods influence grading, with more grade 2 tumors and fewer grade 1 tumors with the methods based on DIA than with the pathology Ki67. The median pathology Ki67 was 2.0% (range: 1.0% to 100%). The median DIA Ki67 value was 4.1% (range: 0.0%

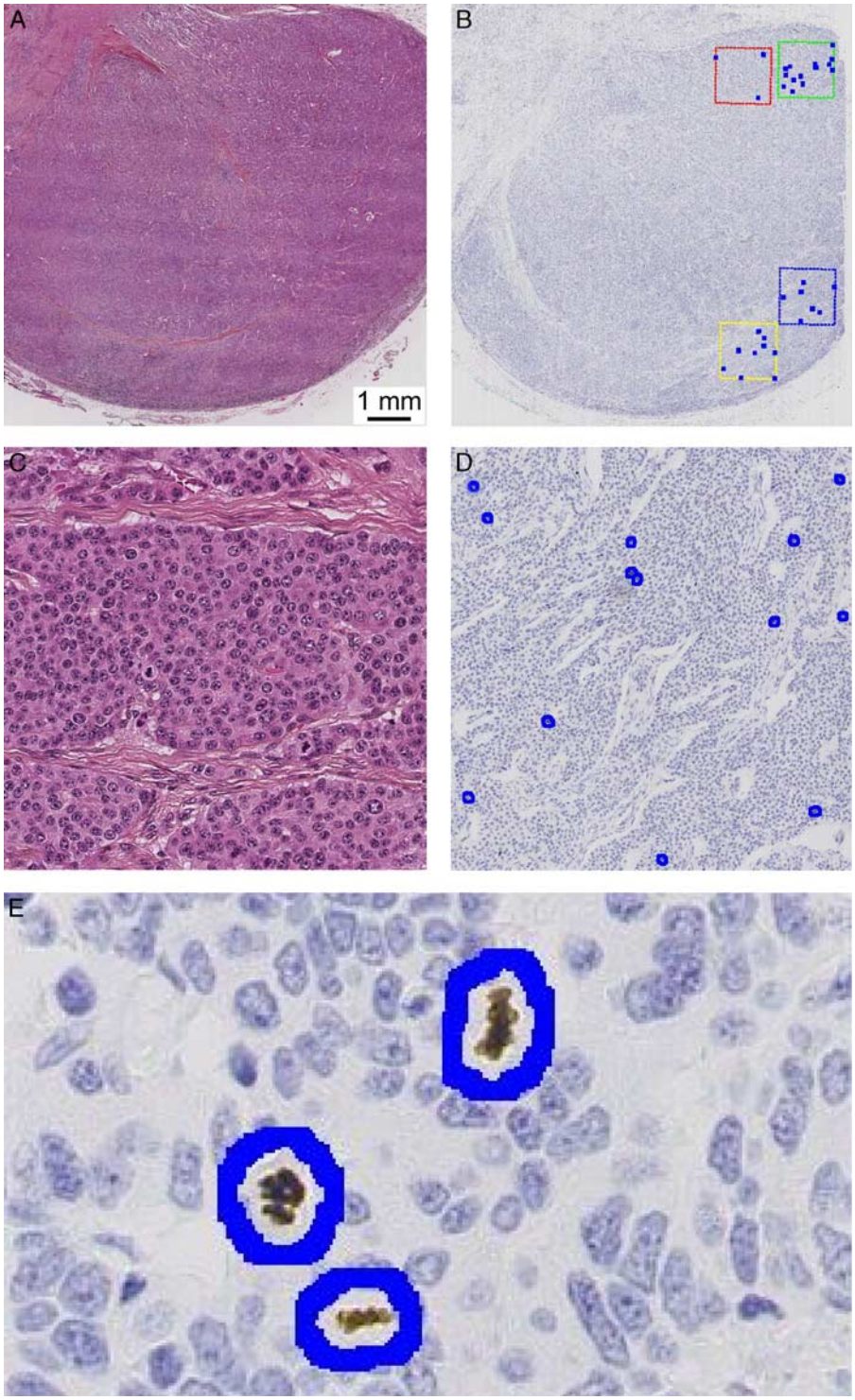

FIGURE 2. A, Hematoxylin and eosin staining of a grade 2 neuroendocrine tumor. B, Digital image measurement of immunohistochemically stained phosphohistone H3 in 4 different regions of interest. C, Close-up image of hematoxylin and eosin staining with marking of mitosis. D and E, Close-up image of the measurement of phosphohistone H3 in the regions of interest with the highest mitotic activity.

to 99.9%). The interclass correlation coefficient of the DIA Ki67 and pathology Ki67 showed an excellent agreement of 0.96 (95% CI: 0.94-0.98). Figure 4A shows a scatter plot of the distribution of the DIA Ki67 and pathology Ki67. The observed kappa between grading based on the DIA Ki67 and the pathology Ki67 was 0.86 (95% CI: 0.81-0.91).

TABLE 2. Demographics and Clinical Characteristics of the Patients

Characteristics	Value
Age, median (range), y	61.8 (12.5-94.2)
Sex, n (%)	
Male	91 (57.2)
Female	68 (42.8)
Tumor size, median* (range), cm	1.7 (0.1-13.8)
Localization of tumor, n (%)	
esophagus	2 (1.3)
Stomach	9 (5.7)
Duodenum	5 (3.1)
Small intestine	54 (34.0)
Meckel	1 (0.6)
Appendix	31 (19.5)
Pancreas	21 (13.2)
Colon	15 (9.4)
Rectum	14 (8.8)
Metastasis liver/unknown primary	7 (4.4)
WHO grade, n (%)	
1	85 (53.5)
2	38 (23.9)
3	36 (22.6)
T classification*, n (%)	
T1	51 (37.5)
T2	21 (15.4)
T3	50 (36.8)
T4	14 (10.3)
N classification*, n (%)	
N0	69 (44.5)
N1	86 (55.5)
M classification*, n (%)	
M0	95 (61.3)
M1	60 (38.7)
AJCC stage*, n (%)	
I	51 (32.4)
II	12 (7.6)
III	34 (21.6)
IV	60 (38.2)

*Numbers may not add up because of missing data: n = 18 for tumor size, n = 23 for T-stage, n = 4 for N-stage and M-stage and n = 2 for AJCC stage.

AJCC indicates American Joint Committee on Cancer; WHO, World Health Organization.

A Bland-Altman plot was created to visualize the difference in agreement against average value of pathology Ki67 and DIA Ki67 (Fig. 5).

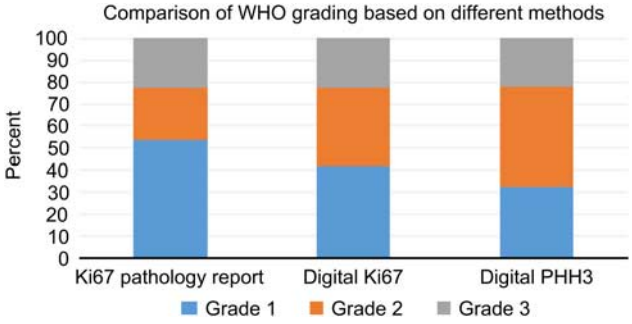

FIGURE 3. Comparison of World Health Organization (WHO) grading based on different methods for proliferation measurement. PHH3 indicates phosphohistone H3. full color online

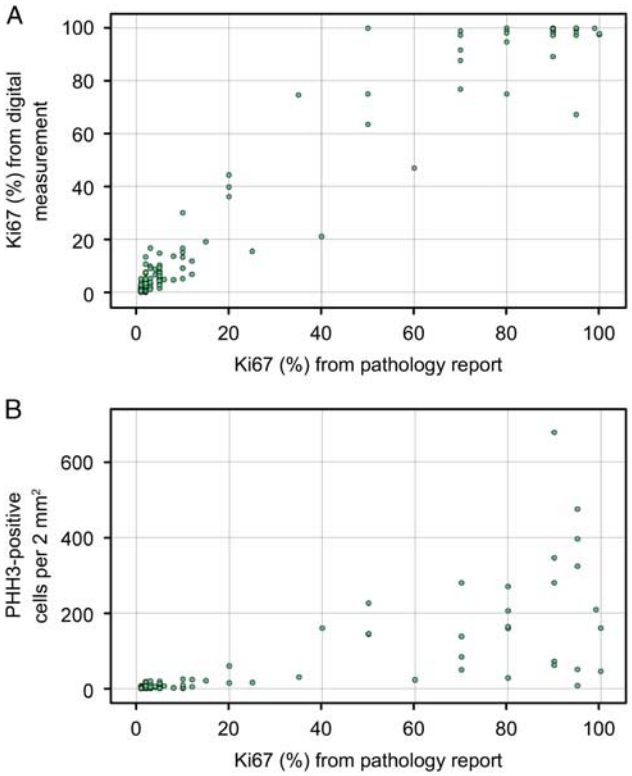

FIGURE 4. A, Scatter plot showing the correlation between the Ki67 index from the pathology report and Ki67 based on digital image analysis. B, Scatter plot showing the correlation between the Ki67 index from the pathology report and digital measurement of the number of phosphohistone H3 (PHH3)-positive cells per 2 mm^2. full color online

The agreement between grading based on the DIA Ki67 and the pathology Ki67 is shown in Table 3. None of the grade 3 tumors was graded as grade 1 and vice versa. Cases with less correlation between the methods comprised mostly grade 1 and grade 2 tumors, and the percentage was mostly measured higher (ie, > 3%) by DIA than with the pathology Ki67. Among 26 cases with a discrepancy between grades 1 and 2, Ki67 values of 2% to

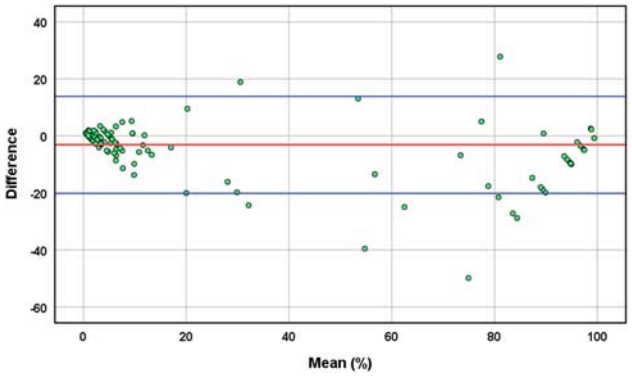

FIGURE 5. A Bland-Altman plot showing the difference Ki67 index against mean value from the pathology report and Ki67 based on digital image analysis. full color online

TABLE 3. Agreement Between Grading of Tumors Based on Digital Image Analysis of Ki67 and Grading From Routine Pathology Reports

	Grade (DIA Ki67) n, (%)			
	1	2	3	Total, n (%)
Grade (routine)				
1	63 (39.6)	22 (13.8)	0	85 (53.5)
2	4 (2.5)	33 (20.8)	1 (0.6)	38 (23.9)
3	0	1 (0.6)	35 (22.0)	36 (22.6)
Total	67 (42.1)	56 (35.2)	36 (22.6)	159

DIA indicates digital image analysis.

4% were found in 24 (92.3%) of the cases by either DIA Ki67 or pathology Ki67.

PHH3 was measured in 145 (145/159 = 91.2%) of the patients. Fourteen patients were excluded because of a lack of available tumor material for analysis. A median of 3 mitoses (range: 0 to 678 cells/mm^2) was observed. Figure 4B shows a scatter plot of the distribution of DIA PHH3 and the pathology Ki67. The observed kappa between grading based on the DIA PHH3 and pathology Ki67 was 0.742. (95% CI: 0.65-0.83). The agreement between grading based on DIA PHH3 and the pathology Ki67 is shown in Table 4. For the DIA of PHH3, there were no grade 3 tumors that were graded as grade 1. Moreover, 35 NENs were upgraded from grade 1 to grade 2 based on the DIA PHH3 measurement compared with pathology Ki67. In 20 (20/35 = 57.1%) of these tumors, the PHH3 value was 2 or 3.

DISCUSSION

The current study explored several comparative measures for grading by the use of DIA and alternative markers for proliferation. Grading based on the DIA Ki67 showed good reliability compared with grading based on the pathology Ki67. The use of DIA PHH3 for grading did not show similar results, with a higher number of NENs migrating between grades, especially grade 2, as a consequence. Whether this represents true grade migration or just artefacts from variation in scores remains unproven.

This study confirms an excellent agreement between the DIA Ki67 and the pathology Ki67, a finding that is supported by other studies.[13,29] DIA can improve the reliability and reproducibility of grading in routine practice.[30] Thus, likely the pathology departments with this method or similar DIA available, can use DIA Ki67 as a part of their routine diagnostics. Digital pathology has during recent years, been introduced at more pathology departments, and the numbers are increasing.[31,32] Manual counting is often difficult because of the high cellularity commonly encountered in these tumors.[30] One of the benefits of our DIA method of Ki67 is the automatic separation of stroma and stromal cells from tumor cells. This does not apply to all DIAs, which may partly explain the previously reported poor concordance between the DIA and manual analysis of Ki67 found by others.[33]

TABLE 4. Agreement Between Grading of Tumors Based on Digital Image Analysis of Phosphohistone H3 and Grading from Routine Pathology Reports

	Grade (DIA PHH3) n, (%)			
	1	2	3	Total n, (%)
Grade (routine)				
1	43 (29.7)	35 (24.1)	0	78 (53.8)
2	4 (2.8)	28 (19.3)	4 (2.8)	36 (24.8)
3	0	3 (2.1)	28 (19.3)	31 (21.4)
Total	47 (32.4)	66 (45.5)	32 (22.1)	145

DIA indicates digital image analysis; PHH3, phosphohistone H3.

This study included a consecutive series from a population-representative cohort of GEP-NEN patients in a well-defined region of Norway. Compared with other studies of GEP-NENs, the number of included patients with grades 2 and 3 was high, and the distribution of different grades was more even.[13,17,30,33–35] A limitation of our study is that almost 40% of the tumor samples were from biopsies. However, this is in line with current routine practice at many centers since surgical specimens are not achievable for all patients, especially patients with advanced disease.

Grading based on the DIA PHH3 agreed less well with the WHO grade based on the pathology Ki67. This is in line with the results of others.[15,33,36] Accordingly, we believe the DIA and use of PHH3 are not supported for routine use or as an alternative in GEP-NEN grading. Some of the cases with low Ki67 had high PHH3. Others have reported a specific value for PHH3 in PNENs,[14] but we could not confirm or refute this based on the limited number of PNENs.

For both DIA Ki67 and DIA PHH3, the proportion of grade 2 tumors was higher than the routinely reported grading. This illustrates the difficulty in separating grade 1 and grade 2 tumors.[33,37] A factor that might explain this finding is that the DIA method makes it easier to identify a hot spot than manual evaluation in a busy routine practice. With DIA, measurements can be repeated or several areas can be measured if the hot spot is difficult to identify. The size of the hot spot (500 to 2000 cells) might also influence the grading. A study of G1 and G2 PNENs found that DIA overestimated the Ki67 index compared with manual evaluation. The difference was reduced by increasing the size of the hot spot.[38] Kroneman et al[37] found that eyeball estimations of Ki67 tended to downgrade more NETs to grade 1 than evaluations by DIA. These studies support our findings. In many cases with a discrepancy, the value of Ki67 or PHH3 was near the cut-off level for the grading criteria. Discussions are ongoing regarding which method should be regarded as the gold standard.[13,34,35] For breast cancers, there is a debate about whether the manual analysis of Ki67 should be the gold standard since several studies have shown superior prognostic information from DIA.[39,40]

In conclusion, standardized DIA Ki67 scoring gave similar results as subjective scoring but may be a time-saving supplementary tool in surgical pathology[29] that can improve the poor intraobserver and interobserver reliability with manual evaluation methods.[41] DIA PHH3 do not agree so well with routine grading and is not recommended as an alternative to MAI or Ki67 in routine practice.

ACKNOWLEDGMENTS

The authors thank statistician Ingvild Dalen, PhD, Stavanger University Hospital, Norway, for providing help with the statistical methods.

REFERENCES

1. Merola E, Rinzivillo M, Cicchese N, et al. Digestive neuroendocrine neoplasms: a 2016 overview. *Dig Liver Dis*. 2016;48:829–835.
2. Sandvik OM, Søreide K, Gudlaugsson E, et al. Epidemiology and classification of gastroenteropancreatic neuroendocrine neoplasms using current coding criteria. *Br J Surg*. 2016;103:226–232.
3. Modlin IM, Lye KD, Kidd M. A 5-decade analysis of 13,715 carcinoid tumors. *Cancer*. 2003;97:934–959.
4. Maggard MA, O'Connell JB, Ko CY. Updated population-based review of carcinoid tumors. *Ann Surg*. 2004;240:117–122.
5. Søreide JA, van Heerden JA, Thompson GB, et al. Gastrointestinal carcinoid tumors: long-term prognosis for surgically treated patients. *World J Surg*. 2000;24:1431–1436.
6. Lee MS, O'Neil BH. Summary of emerging personalized medicine in neuroendocrine tumors: are we on track? *J Gastrointest Oncol*. 2016; 7:804–818.
7. Yao JC, Hassan M, Phan A, et al. One hundred years after "carcinoid": epidemiology of and prognostic factors for neuroendocrine tumors in 35,825 cases in the United States. *J Clin Oncol*. 2008;26:3063–3072.
8. Bellizzi AM. Immunohistochemistry in the diagnosis and classification of neuroendocrine neoplasms: what can brown do for you? *Hum Pathol*. 2020;96:8–33.
9. Nagtegaal ID, Odze RD, Klimstra D, et al. The 2019 WHO classification of tumours of the digestive system. *Histopathology*. 2020;76:182–188.
10. WHO Classification of Tumours. *Digestive System Tumours*, 5 ed. Vol. 1. Lyon, France: IARC Publications; 2019.
11. 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.
12. Gudlaugsson E, Skaland I, Janssen EA, et al. Comparison of the effect of different techniques for measurement of Ki67 proliferation on reproducibility and prognosis prediction accuracy in breast cancer. *Histopathology*. 2012;61:1134–1144.
13. 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.
14. Villani V, Mahadevan KK, Ligorio M, et al. Phosphorylated Histone H3 (PHH3) is a superior proliferation marker for prognosis of pancreatic neuroendocrine tumors. *Ann Surg Oncol*. 2016;23 (suppl 5):609–617.
15. Ozturk Sari S, Taskin OC, Gundogdu G, et al. The impact of phosphohistone-H3-assisted mitotic count and Ki67 Score in the determination of tumor grade and prediction of distant metastasis in well-differentiated pancreatic neuroendocrine tumors. *Endocr Pathol*. 2016;27:162–170.
16. Sun A, Zhou W, Lunceford J, et al. Level of phosphohistone H3 among various types of human cancers. *BMJ Open*. 2012;2:e001071.
17. Tracht J, Zhang K, Peker D. Grading and prognostication of neuroendocrine tumors of the pancreas: a comparison study of Ki67 and PHH3. *J Histochem Cytochem*. 2017;65:399–405.
18. Skaland I, Janssen EA, Gudlaugsson E, et al. Validating the prognostic value of proliferation measured by phosphohistone H3 (PPH3) in invasive lymph node-negative breast cancer patients less than 71 years of age. *Breast Cancer Res Treat*. 2009;114:39–45.
19. Bossard C, Jarry A, Colombeix C, et al. Phosphohistone H3 labelling for histoprognostic grading of breast adenocarcinomas and computer-assisted determination of mitotic index. *J Clin Pathol*. 2006; 59:706–710.
20. Ribalta T, McCutcheon IE, Aldape KD, et al. The mitosis-specific antibody anti-phosphohistone-H3 (PHH3) facilitates rapid reliable grading of meningiomas according to WHO 2000 criteria. *Am J Surg Pathol*. 2004;28:1532–1536.
21. American Joint Committee on Cancer Staging Manual, 8th ed. Cham, Switzerland: Springer International Publishing AG; 2017.
22. Gentil Perret A, Mosnier JF, Buono JP, et al. The relationship between MIB-1 proliferation index and outcome in pancreatic neuroendocrine tumors. *Am J Clin Pathol*. 1998;109:286–293.
23. Klöppel G, Couvelard A, Perren A, et al. ENETS Consensus Guidelines for the Standards of Care in Neuroendocrine Tumors: towards a standardized approach to the diagnosis of gastroenteropancreatic neuroendocrine tumors and their prognostic stratification. *Neuroendocrinology*. 2009;90:162–166.
24. Skaland I, Janssen EA, Gudlaugsson E, et al. Phosphohistone H3 expression has much stronger prognostic value than classical prognosticators in invasive lymph node-negative breast cancer patients less than 55 years of age. *Mod Pathol*. 2007;20:1307–1315.
25. Dumars C, Foubert F, Touchefeu Y, et al. Can PPH3 be helpful to assess the discordant grade in primary and metastatic enteropancreatic neuroendocrine tumors? *Endocrine*. 2016;53:395–401.
26. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull*. 1968;70:213–220.
27. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016; 15:155–163.
28. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet*. 1995;346:1085–1087.
29. Satturwar SP, Pantanowitz JL, Manko CD, et al. Ki-67 proliferation index in neuroendocrine tumors: can augmented reality microscopy with image analysis improve scoring? *Cancer Cytopathol*. 2020;128: 535–544.
30. 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.
31. Tschuchnig ME, Oostingh GJ, Gadermayr M. Generative adversarial networks in digital pathology: a survey on trends and future potential. *Patterns (N Y)*. 2020;1:100089.
32. Jahn SW, Plass M, Moinfar F. Digital pathology: advantages, limitations and emerging perspectives. *J Clin Med*. 2020;9:3697.
33. Hacking SM, Sajjan S, Lee L, et al. Potential pitfalls in diagnostic digital image analysis: experience with Ki-67 and PHH3 in gastrointestinal neuroendocrine tumors. *Pathol Res Pract*. 2020;216:152753.
34. Dhall D, Mertens R, Bresee C, et al. Ki-67 proliferative index predicts progression-free survival of patients with well-differentiated ileal neuroendocrine tumors. *Hum Pathol*. 2012;43:489–495.
35. 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.
36. Kim MJ, Kwon MJ, Kang HS, et al. Identification of phosphohistone H3 cutoff values corresponding to original WHO grades but distinguishable in well-differentiated gastrointestinal neuroendocrine tumors. *Biomed Res Int*. 2018;2018:1013640.
37. 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.
38. Owens R, Gilmore E, Bingham V, et al. Comparison of different anti-Ki67 antibody clones and hotspot sizes for assessing proliferative index and grading in pancreatic neuroendocrine tumours using manual and image analysis. *Histopathology*. 2020;77:646–658.
39. Stålhammar G, Robertson S, Wedlund L, et al. Digital image analysis of Ki67 in hot spots is superior to both manual Ki67 and mitotic counts in breast cancer. *Histopathology*. 2018;72:974–989.
40. Klauschen F, Wienert S, Schmitt WD, et al. Standardized Ki67 diagnostics using automated scoring—clinical validation in the GeparTrio Breast Cancer Study. *Clin Cancer Res*. 2015;21: 3651–3657.
41. Trikalinos NA, Chatterjee D, Lee J, et al. Accuracy of grading in pancreatic neuroendocrine neoplasms and effect on survival estimates: an institutional experience. *Ann Surg Oncol*. 2020;27:3542–3550.

