UNIVERSITA' DEGLI STUDI DI VERONA

DEPARTMENT OF

General and Pancreatic Surgery Unit

GRADUATE SCHOOL OF

Health and Life Sciences

DOCTORAL PROGRAM IN

Inflammation, Immunity and Cancer

Cycle / year (1° year of attendance): XXXI

TITLE OF THE DOCTORAL THESIS

Preoperative fine-needle aspiration of pancreatic neuroendocrine tumors: a reliable tool to assess diagnosis and grading - A prospective single-center analysis of 100 cases

S.S.D. MED/18

Coordinator: Prof.ssa Gabriela Constantin

Signature _____

Tutor: Prof. Claudio Bassi

Signature _____

Doctoral Student: Dott. Salvatore Paiella

Signature _____

This thesis would never have been possible without the precious research of Dr. Luca Landoni in the field of the pancreatic neuroendocrine tumors of the pancreas.

INDEX

Introduction

Pancreatic neuroendocrine tumors (PanNENs) are rare malignancies with a good prognosis, especially when compared to their exocrine counterpart. PanNENs are divided into functioning (F-PanNENs) and non-functioning (NF-PanNENs) subtypes, based on hormone production. They are characterized by low mitotic rates, determined using histopathology[1]. The biology of these tumors is complex[2] and there are currently no reliable markers of biological behaviour. One of the most adopted tool in this setting is the calculation of the Ki-67 index. The World Health Organization (WHO)[3] and the European Neuroendocrine Tumor Society (ENETS) [4] Guidelines rely on the Ki-67 index to discriminate three categories of tumors (G1 for of Ki-67 index values < 2%, G2 for values between 3% and 20%, G3 for values > 20%). While surgery is recommended for F-PanNENs, for resectable non-metastatic NF-PanNENs, surgery should only be considered when the tumor is over 20 mm in diameter in cases with symptoms, with G2 or G3 tumors, for smaller tumors when non-surgical management is contraindicated, or according to patient's wishes. Instead, for tumors smaller than 20 mm and graded as G1 or low G2, surveillance is sufficient. Several clinical studies have demonstrated that the histological Ki-67 index value is a valid indicator of the tumor's biology [5-8]. The ability to obtain a cytological Ki-67 index value at the time of diagnosis using fine-needle aspiration (FNA) is highly clinically valuable and in need of further investigation.

Considering that pancreatic surgery is still burdened by high rates of morbidity and mortality [9], tailored treatment based on a proper analysis of the biology of the tumor might improve patient outcome. For example, small NF-PanNENs with high preoperative Ki-67 index values may benefit from surgery regardless of dimensions; however, high Ki-67 index values might cause the surgeon to perform a standard pancreatic resection (over a parenchyma-sparing one) or a regional lymphadenectomy.

It remains unclear whether the preoperative Ki-67 index value accurately represents the postoperative Ki-67 index value of the resected tumor, as some studies describe a good correlation while others do not[10-18]. Recently, Weiss et al. reported the lack of a correlation between the preoperative FNA results and the final

histology, for both Ki-67 and grading, with an average difference in the former of 5.9%. The authors state that preoperative FNA leads to under-grading the tumor and that this consequently results in undertreatment[19]. Intratumor cellular heterogeneity is a confounding factor that might be responsible for the dissonance between the pre- and postoperative analyses[13, 20, 21]. Furthermore, obtaining an adequate number of cells to determine a reliable Ki-67 index value using is challenging using FNA.

In this study, we address multiple issues. First, we establish the concordance rate between cytological and histological Ki-67 index values (cKi-67 and hKi-67) and grading (cG and hG). Second, we explore the possible differences between these rates when percutaneous ultrasound-guided (pUS) or endoscopic ultrasound-guided (eUS) FNA is employed. Finally, we assessed the diagnostic rate of FNA.

Materials and Methods

Patient selection

Patients who underwent pancreatic resection from January 2011 to May 2017 were selected from the prospectively maintained electronic database of the General and Pancreatic Surgery Unit, Pancreas Institute, University of Verona Hospital Trust. Only patients with preoperative FNAs for a presumed PanNEN and postoperative histological analyses were included. Then, according to the issues addressed, multiple sub-cohorts were identified. The primary group included patients who submitted to resection after a preoperative presumed diagnosis of PanNENs. In this group, cKi-67 vs. hKi-67 and cG vs. hG were compared.

The second and the third sub-cohort groups were determined according to the technique used (pUS vs. eUS) and the diagnostic rate for each group was added to the main endpoint.

To ensure a homogeneous cohort, the following cases were excluded from this study: metastatic cases; multifocal cases; PanNENs that either at the preoperative FNA or upon final histology had a mixed neoplastic component (e.g. combinations of PanNENs and acinar cell carcinoma, mixed-adenoneuroendocrine carcinoma, or pancreatic adenocarcinoma); and PanNENs treated preoperatively with neoadjuvant therapy. The Ki-67 assessment was performed according to the

WHO 2017 Guidelines [3]. To ensure accurate grading, at least 500 cells per cell block were counted. In our experience, cases without sufficient tumor cells in their cell blocks are rare and those were not considered for this study. A triple immunocytochemical analysis with synaptophysin, chromogranin, and Ki-67 was performed to determine the diagnostically-relevant neuroendocrine nature of the samples.

From 2015, each case was previously discussed at a dedicated and institutionalized multidisciplinary meeting.

Population characteristics

Demographic, clinical, radiological, cytological, and histological data were obtained. PanNENs were classified according to the WHO criteria, including data regarding the differentiation status and Ki-67 index values. The institutional preoperative and surgical management of PanNENs has been previously described[22]. Patients submitted either to pUS-FNA or eUS-FNA for diagnosis, staging, and grading according to the specialist's prescription, and, considering the site, the dimensions and the shape of the tumor.

pUS-FNA

Previous cross-sectional imaging was reviewed before FNA by the radiologist that would have been performed the FNA. Peripancreatic vessels were examined using Doppler US to determine the safest approach. Prior to the sampling procedure, routine sterile preparation of the abdominal wall was performed and a local anesthetic was injected at the chosen entry point. US-FNAs were performed by an experienced radiologist using a Sequoia 512 system (Acuson/Siemens Medical Solutions, Mountain View, CA, USA). Convex multi-frequency probes with a lateral guidance kit and 20 G or 21 G modified Menghini-type aspiration needles were used. A rapid on-site evaluation (ROSE) was performed after each procedure by experienced cytopathologists. The cytologic sample was immediately smeared onto a glass slide, fixed with 95% ethanol and stained using the modified Papanicolaou method.

eUS-FNA

This procedure was performed by two experienced endoscopists with caseloads greater than 350 eUS-FNA per year. Either a standard (EchoTip Ultra, Cook Medical, Limerick, Ireland) or, more frequently, a side-fenestrated (Echotip ProCore, Cook Medical) 25 G needle was used. The institutional-specific techniques are described in a previous publication [23]. The ROSE was not available after EUS-FNA due to logistic reasons.

Statistical analysis

Differences between groups were evaluated using the chi-square test. Fisher's exact test was used when appropriate. The independent two-tailed t-test was used to compare the means of continuous numerical data. When Ki-67 was reported as < 1% then the value of zero was used. When Ki-67 was reported as a range, the greatest value was used. The correlation between pre- and postoperative Ki-67 index values and grading was calculated using Spearman's rho and/or Cohen's kappa. A rho value of 0.00 - 0.19 was considered to be very weak, 0.20 - 0.39 was weak, 0.40 - 0.59 was moderate, 0.60 - 0.79 was strong, and 0.80 - 1.00 was very strong. A kappa value of 0.20 was considered to be poor, 0.21 - 0.40 was fair, 0.41 - 0.60 was moderate, 0.61 - 0.80 was good, and 0.81 - 1.00 was very good.

A Bland-Altman analysis was used to assess the correlation between cKi-67 and hKi-67, where the difference between the measured values against their means was represented [24]. The Bland-Altman analysis calculates the mean difference between 2 methods of measurement (the "bias") and 95% LOA as the mean difference (1.96 SD). The Bland-Altman analysis calculates the mean difference between two methods of measurement (i.e. the "bias") and 95% LOA as the mean difference (1.96 SD). The 95% CI for the limits were also determined. In this visual method, the smaller is the range between the two limits, the better is the agreement [25].

The difference in Ki-67 index values (ΔKi-67) was calculated by subtracting the preoperative value from the postoperative one. The ΔKi-67 that were lower than -2% or greater than 2% were arbitrarily considered to be outliers and any possible variables were investigated using the chi-square test. All statistical tests were two-sided, and P-values < 0.05 were considered statistically significant. These statistical

analyses were performed using SPSS (IBM, Armonk, NY, USA) and MedCalc (MedCalc Software, Ostend, Belgium).

Results

The final study population consisted of 100 patients. Eighty-seven PanNENs were non-functioning and 13 tumors were functioning (12 insulinomas and one gastrinoma). eUS-FNA and pUS-FNA were performed in 85 and 15 cases respectively. Table 1 contains the demographic, clinical, pathological, surgical and radiological features of the study population. FNA was reached a diagnosis in 85 patients. No periprocedural complications were reported.Eighty-five FNA were performed via eUS, the remaining 15 via pUS. No periprocedural complications were reported.

Table 1. Demographic, clinical and surgical features of the study population

Age (SD)	54 ± 12 years
Sex (M/F)	48/52
Symptoms, (yes, n, %) Abdominal pain (any kind of) Symptoms of insulinoma/gastrinoma Jaundice Acute pancreatitis	30 (30%) 15 (50%) 13 (43.4%) 1 (3.3%) 1 (3.3%)
Site Head Body Tail	 46 (46%) 33 (33%) 20 (20%)
Syndromic cases (MEN1)	3 (3%)
Tumor diameter (mm, mean; SD) ≤/> 20 mm (n, %)	25.5 ± 19.1 48/51 (48/51%)
eUS-FNA / pUS-FNA	85/15 (85/15%)
cG G1 G2 G3	63 (74.1%) 47 (74.6%) 19 (20.6%) 2 (4.8%)
Surgery DPS PD E DPSP MP TPS	 35 34 12 8 7 4

eUS-FNA: Endoultrasound-guided fine-needle aspiration; pUS-FNA:

percutaneous-guided; MEN1: Multiple endocrine neoplasia type 1; DPS: distal pancreatectomy with splenectomy; PD: pancreatico-duodenectomy; E: enucleation; DPSP: distal pancreatectomy with spleen preservation; Middle pancreatectomy; TP: total pancreatectomy; TPS: total pancreatectomy with splenectomy

Diagnostic rate, concordance rates for diagnosis

In 84 of 100 presumed cases of PanNEN, a diagnosis was reached preoperatively. Non-diagnostic cases were biopsied using eUS and pUS in 15 (15/85, 17.6%) and one (1/15, 6.6%) of the cases. Table 2 presents the pathologic data and concordance rates for diagnosis considering the final histology. The overall concordance rate for diagnosis was 97.6%. NF-PanNENs, the most common diagnosis, represented 65 of 76 cases (85.5%). The concordance rate for G3 tumors was 33.3%. Indeed, in one case only the cytology was accurate, while the remaining two cases were revealed to actually be G1 PanNENs with a Ki-67 index value of 2% upon postoperative histological analysis. Notably, in these two cases, the cytology was obtained using a percutaneous approach.

Table 2. Pathologic data

Diagnosis from FNA	**N (%)**	**Final histology**
cG	63 (74.1%)	
G1	47 (74.6%)	-
G2	13 (20.6%)	-
G3	3 (4.8%)	-
Non-diagnostic	-	16 (16%)
NF-PanNENs	-	11 (68%)
MANEC	-	2 (12.5%)
PanNENs G3	-	1 (6.5%)
NEC with PC	-	1 (6.5%)
cPanNEN	-	1 (6.5%)
Diagnostic	84 (84%)	All confirmed
NF-PanNEN	68 (81%)	
Insulinoma	12 (14.2%)	
Gastrinoma	1 (1.2%)	
cPanNEN	3 (3.5%)	
Nodal status (N+)		
Overall	23 (23%)	
G1	9 (39.1%, 19.1% of G1)	

G2	11 (47.9%, 52.6% of G2)
G3	3 (13%, 100% of G3)

NF-PanNEN: non-functioning pancreatic neuroendocrine tumor; MANEC: Mixed Adenoneuroendocrine Carcinoma; NEC: neuroendocrine carcinoma; cPanNEN: cystic PanNEN

Concordance rates for grading

The cG was available for 63 of 84 diagnostic samples (75%) and, therefore, this was the study population for the analysis of the reliability of cG. Assessment of hG was available for 98 of 100 patients. Discordance was observed in 17 of 63 (26.9%) cases. The mean tumor diameter of the discordant cases was 35 ± 32 mm and of the concordant cases was 26 ± 16 mm ($p = 0.001$). The overall sensitivity for G was 76.2%, whereas for G1, G2, and G3 it was 76.6%, 84.6% and 33.3% respectively (Table 3).

Table 3. Concordance rate for grading

	hG1	**hG2**	**hG3**	**Total**
cG1	**36**	11	0	47
cG2	2	**11**	0	13
cG3	2	0	**1**	3
Total	25	19	2	**63**

The overall level of agreement between the pre- and postoperative grading was moderate (Cohen's $k = 0.455$, 95% CI from 0.219 to 0.691, $p < 0.001$). When considering PanNEN less than 20 mm (n = 26, 41%) and greater than 20 mm (n = 37, 59%), the agreement was moderate ($k = 0.438$, 95% CI from 0.070 to 0.79518, $p = 0.014$ and $k = 0.450$, 95% CI from 0.141 to 0.760, $p < 0.001$). When considering the whole cohort, Spearman's rho indicated a moderate positive agreement between the pre- and postoperative grading ($r_s = 0.430$, 95% CI from 0.204 to 0.613, $p < 0.001$). The agreement was moderate for both PanNEN less than 20 mm ($r_s = 0.472$, 95% CI from 0.121 to 0.718, $p = 0.013$) and greater than 20 mm ($r_s = 0.430$, 95% CI from 0.124 to 0.661, $p = 0.030$).

With the endoscopic approach, the agreement was good (n = 54, $r_s = 0.610$, 95% CI from 0.409 to 0.755, $p < 0.001$). However, with the percutaneous approach,

the agreement was negative but not statistically significant (n = 9, r_s = -0.253, 95% CI from -0.785 to 0.494, p = 0.511).

Concordance rates for Ki-67

The mean index values of cKi-67 and hKi-67 were 4.35 ± 9.5% and 5.26 ± 12% respectively (p = 0.334). The average ΔKi-67 was -0.7 ± 7.5%.

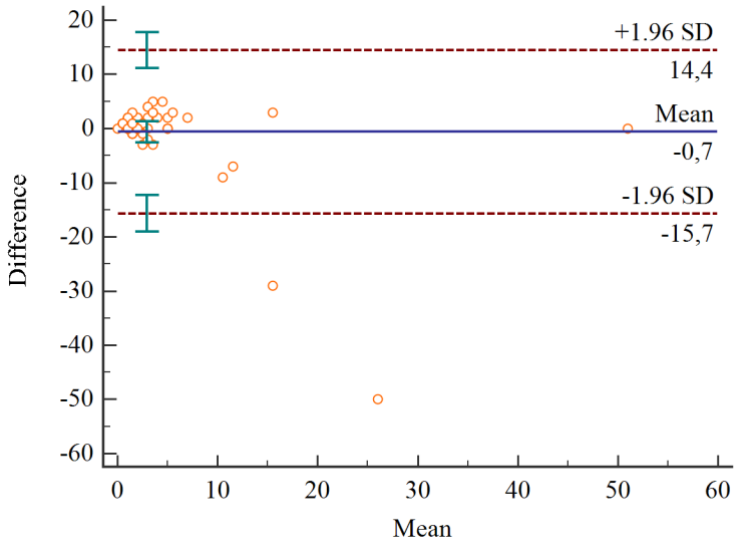
Spearman's rho was good for the overall population (r_s = 0.615, 95% CI from 0.434 to 0.749, p < 0.001). When analyzing tumors less and greater than 20 mm, the agreement was very good and moderate respectively (r_s = 0.862, 95% CI from 0.7180 to 0.9359, p < 0.001; r_s = 0.596, 95% CI from 0.329 to 0.775, p = 0.0002).

Using the endoscopic approach, the agreement was moderate (n = 54, r_s = 0.558, 95% CI from 0.342 to 0.718, p < 0.001) while the percutaneous one was negative but not statistically significant (n = 9, r_s = -0.217, 95% CI from -0.770 to 0.552, p=0.534).

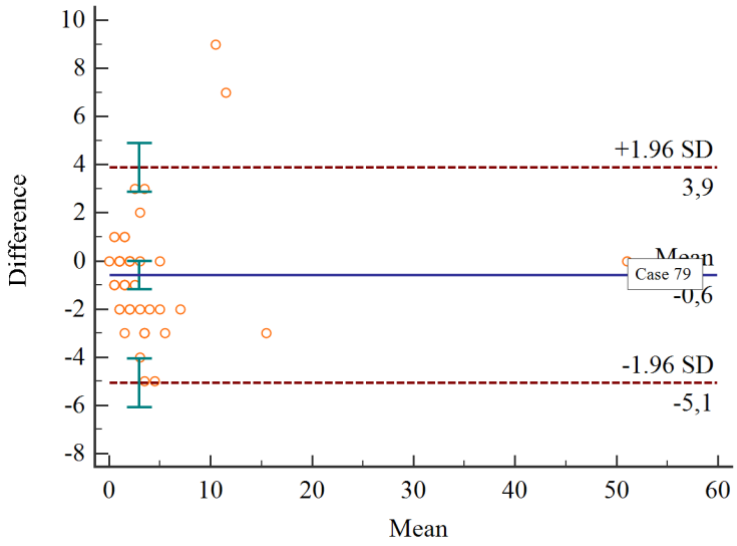
The Bland-Altman plot (Figure 1) shows an agreement between cKi-67 and hKi-67 assessment, as 95% of the data points lie within ± 1.96 SD of the limits of agreement[26]. Due to the skewness of the data (Figure 1A), we removed two outliers from the analysis and the limits of agreement reduced considerably (Figure 1B).

Figure 1. Bland-Altman plot. A. Overall cohort. B. Subgroup analysis with exclusion of the outliers to correct the skewness of the data.

A



B



Outliers of ΔKi-67

Twenty-six (39.4%) outliers were identified. The majority of these cases were found using eUS (n = 18, 69.2%) and the remaining using pUS (n = 8, 30.8%). In the univariate analysis, the pUS approach was more frequently associated with outliers than the eUS one (87.5% vs. 31.6%, p = 0.004). The mean diameter of the outliers was 36.4 ± 30.2 mm, while the consistent cases were 23.5 ± 12.2 mm in diameter (p = 0.001). The other variables considered (sex, age [≤ / > 50 years], symptoms [yes/no], tumor dimensions [≤ / > 20 mm], tumor site [head, neck, tail]) were not statistically significant.

Discussion

Owing to the development of cross-sectional imaging and endoscopic examination, PanNENs have been increasingly detected and diagnosed over the past few decades[27]. An ensuing issue is whether all of these lesions need to be treated aggressively with surgery or whether they can be managed with surveillance. Currently, this choice is decided based on several features, including the tumor's dimensions, site, metabolic activity, presumed biological behavior, and the patient's wishes and comorbidities. The biological behavior is assessed using the Ki-67 index value and grading determined by FNA. In particular, in the era of

personalized medicine, the assessment of grading accuracy is important to inform tailored treatment strategies (e.g. resection vs. surveillance or non-surgical treatments; standard *vs.* parenchyma-sparing resection). However, the literature contains heterogeneous reports of the concordance rate for PanNEN grading, ranging from 69% to 90% [28]. In this study, we investigated whether cKi-67 and cG accurately represents hKi-67 and hG to establish the diagnostic concordance rates of FNA.

We found a very good diagnostic rate of 84% with all diagnoses being confirmed by final histology. These numbers are in line with previous findings[17, 29-31] and they highlight the reliability of the FNA-based diagnosis of PanNENs, apart from the cases with complex histology (mixed forms).

In regards to the grading concordance, we found an acceptable overall rate of 72.1%, which is a bit lower than previous reports[28, 32-34]. This rate was higher for G2 tumors (> 80%) and lower for G3 (33.3%). The overall agreement was moderate and we did not find any difference using the dimensional cut-off of 20 mm. Of note, 11 of 47 (23.4%) cG1 PanNENs were found to actually be hG2 and 2 of 13 (15.3%) cG2 were actually hG1. Two FNA-determined G3 PanNENs were revealed to actually be G1 upon histological analysis and these were all biopsied percutaneously. They were performed at the beginning of our policy to biopsy PanNENs during the therapeutic work-up and at that time the eUS was not available. For G3 tumors, we might speculate that the ROSE was not informative and that additional immunocytochemical evaluation would have been helpful for these complex cases. However, it is likely that the heterogeneous architecture of a G3 tumor itself acted as a confounding factor. The agreement between the grading was good for eUS and negative for pUS, though this was not statistically significant. We hypothesize that this negative agreement from pUS is an "artifact" as the tumors biopsied using pUS were bigger and histologically complex and the patients were well-selected. Nevertheless, the small number of pUS-FNA cases mitigates our considerations. Taken together, the overall agreement between cG and hG was moderate and under-grading (n = 11) was more frequent than over-grading (n = 4). These findings have to somehow be considered at the time of therapeutic decision in order to select the optimal treatment course.

In terms of concordance between Ki-67 index values, we discovered a small underestimation of cKi-67 compared to hKi-67 (mean difference of -0.7%). The overall correlation was good and very good (r_s = 0.862) for tumors smaller than 20 mm. This has already been described by other authors who found a negative correlation between the concordance rate for Ki-67 and the tumor dimensions, where the bigger the tumor, the higher the cellular heterogeneity, and the higher the risk of misestimation[13, 32, 35].

From the analysis of the outliers of ΔKi-67 (n = 26, 39.3%), we found that pUS was significantly more associated with this result than eUS. These outlier cases were stastistically significant bigger tumors with, presumably, higher tissue heterogeneity.

Multiple lessons can be gleaned from the results of the present investigation, which represents the biggest cohort study published to date comparing the cytology and histology-based determinations of PanNENs. First, FNA is an effective diagnostic tool when facing a suspected PanNEN. It enables a high rate of diagnosis, excellent diagnostic concordance, and negligible to no side effects. Second, the grading assessment was satisfactory as it was accurate in about 75% of cases; however, under-grading is possible and must be considered. This finding might affect the therapeutic approach in peculiar cases where the optimal therapeutic approach is not clear. For example, small PanNENs with a doubtful metabolic activity (e.g. Ki-67 3-4%) and unclear lymphadenopathies may be classified as G1 with FNA and later be revealed as G2 at final histology and, thus, may benefit from surgery; or G2 tumors with potential vascular involvement may benefit from non-surgical oncological treatment as these are “true” G2 tumors at histology). Third, we found a very high agreement for Ki-67 index values when considering tumors less than 20 mm. This finding might help the therapeutic management of this sub-cohort of PanNENs where the therapy might be controversial (surgery vs. follow-up). Fourth, bigger PanNENs are more prone to misdiagnosis and have the highest levels of ΔKi-67. Fifth, in our institution, eUS-FNA, even without a ROSE, reaches higher levels of concordance than pUS-FNA; therefore, this approach is preferable for ensuring an accurate diagnosis.

This study also has some limitations that must be considered as well. We did not perform a survival analysis comparing the data from FNA and histology because we believed it was beyond the scope of this study. The pathologists were not blinded and the same pathologists might have performed both the cytology and histology for a single case. The choice to subject the patients to eUS or pUS was at the specialist's discretion and this might have generated a selection bias. Moreover, ROSE was not available for eUS as it was at pUS and this might confound the comparison of the two techniques. Not all PanNENs who underwent surgery underwent FNA and this choice was at the surgeon's discretion, which might have generated a selection bias. Finally, a review of the discordant and non-diagnostic cases that could've identified possible influencing factors was not performed.

In conclusion, in cases of suspected PanNEN, FNA is a reliable tool for assessing the tumor's nature, especially using eUS. Preoperative cytological approaches may under-grade compared to the postoperative histology analyses. Nevertheless, we recommend this preoperative procedure for reaching a diagnosis in unclear cases, especially for small tumors, in order to classify subtypes, tailor treatments, and improve patient outcome.

References

1. Kloppel G. Tumour biology and histopathology of neuroendocrine tumours. Best Pract Res Clin Endocrinol Metab. 2007;21:15-31.
2. Scarpa A, Chang DK, Nones K, Corbo V, Patch AM, Bailey P, et al. Whole-genome landscape of pancreatic neuroendocrine tumours. Nature. 2017;543:65-71.
3. Who Classification of Tumours of Endocrine Organs. 4th ed2018.
4. Falconi M, Eriksson B, Kaltsas G, Bartsch DK, Capdevila J, Caplin M, et al. ENETS Consensus Guidelines Update for the Management of Patients with Functional Pancreatic Neuroendocrine Tumors and Non-Functional

Pancreatic Neuroendocrine Tumors. Neuroendocrinology. 2016;103:153-71.

5. Bahra M, Jacob D, Pascher A, Plockinger U, Kristiansen G, Neuhaus P, et al. Surgical strategies and predictors of outcome for malignant neuroendocrine tumors of the pancreas. J Gastroenterol Hepatol. 2007;22:930-5.

6. Ekeblad S, Skogseid B, Dunder K, Oberg K, Eriksson B. Prognostic factors and survival in 324 patients with pancreatic endocrine tumor treated at a single institution. Clin Cancer Res. 2008;14:7798-803.

7. La Rosa S, Klersy C, Uccella S, Dainese L, Albarello L, Sonzogni A, et al. Improved histologic and clinicopathologic criteria for prognostic evaluation of pancreatic endocrine tumors. Hum Pathol. 2009;40:30-40.

8. Pape UF, Berndt U, Muller-Nordhorn J, Bohmig M, Roll S, Koch M, et al. Prognostic factors of long-term outcome in gastroenteropancreatic neuroendocrine tumours. Endocr Relat Cancer. 2008;15:1083-97.

9. De Pastena M, Paiella S, Marchegiani G, Malleo G, Ciprani D, Gasparini C, et al. Postoperative infections represent a major determinant of outcome after pancreaticoduodenectomy: Results from a high-volume center. Surgery. 2017;162:792-801.

10. Khashab MA, Yong E, Lennon AM, Shin EJ, Amateau S, Hruban RH, et al. EUS is still superior to multidetector computerized tomography for detection of pancreatic neuroendocrine tumors. Gastrointest Endosc. 2011;73:691-6.

11. Ardengh JC, de Paulo GA, Ferrari AP. EUS-guided FNA in the diagnosis of pancreatic neuroendocrine tumors before surgery. Gastrointest Endosc. 2004;60:378-84.

12. Horwhat JD, Paulson EK, McGrath K, Branch MS, Baillie J, Tyler D, et al. A randomized comparison of EUS-guided FNA versus CT or US-guided FNA for the evaluation of pancreatic mass lesions. Gastrointest Endosc. 2006;63:966-75.

13. Hasegawa T, Yamao K, Hijioka S, Bhatia V, Mizuno N, Hara K, et al. Evaluation of Ki-67 index in EUS-FNA specimens for the assessment of malignancy risk in pancreatic neuroendocrine tumors. Endoscopy. 2014;46:32-8.

14. Hosoda W, Takagi T, Mizuno N, Shimizu Y, Sano T, Yamao K, et al. Diagnostic approach to pancreatic tumors with the specimens of endoscopic ultrasound-guided fine needle aspiration. Pathol Int. 2010;60:358-64.

15. Li J, Lin JP, Shi LH, Wang WJ, Li AQ, Si JM, et al. How reliable is the Ki-67 cytological index in grading pancreatic neuroendocrine tumors? A meta-analysis. J Dig Dis. 2016;17:95-103.

16. Laskiewicz L, Jamshed S, Gong Y, Ainechi S, LaFemina J, Wang X. The diagnostic value of FNA biopsy in grading pancreatic neuroendocrine tumors. Cancer Cytopathol. 2018;126:170-8.

17. Hooper K, Mukhtar F, Li S, Eltoum IA. Diagnostic error assessment and associated harm of endoscopic ultrasound-guided fine-needle aspiration of neuroendocrine neoplasms of the pancreas. Cancer Cytopathol. 2013;121:653-60.

18. Weynand B, Borbath I, Bernard V, Sempoux C, Gigot JF, Hubert C, et al. Pancreatic neuroendocrine tumour grading on endoscopic ultrasound-guided fine needle aspiration: high reproducibility and inter-observer agreement of the Ki-67 labelling index. Cytopathology. 2014;25:389-95.

19. Weiss VL, Kiernan C, Wright J, Merchant NB, Coogan AC, Shi C. Fine-needle aspiration–based grading of pancreatic neuroendocrine neoplasms using Ki-67: is accurate WHO grading possible on cytologic material? Journal of the American Society of Cytopathology. 2018;7:154-9.

20. 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;35:853-60.

21. Liszka Ł. Tissue heterogeneity contributes to suboptimal precision of WHO 2010 scoring criteria for Ki67 labeling index in a subset of neuroendocrine neoplasms of the pancreas. Polish Journal of Pathology. 2017;67:318-31.

22. Landoni L, Marchegiani G, Pollini T, Cingarlini S, D'Onofrio M, Capelli P, et al. The Evolution of Surgical Strategies for Pancreatic Neuroendocrine Tumors (Pan-NENs): Time-trend and Outcome Analysis From 587 Consecutive Resections at a High-volume Institution. Ann Surg. 2017.

23. Crinò SF, Conti Bellocchi MC, Bernardoni L, Manfrin E, Parisi A, Amodio A, et al. Diagnostic yield of EUS-FNA of small (≤15 mm) solid pancreatic lesions using a 25-gauge needle. Hepatobiliary & Pancreatic Diseases International. 2018;17:70-4.

24. Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986;327:307-10.

25. Myles PS, Cui J. I. Using the Bland–Altman method to measure agreement with repeated measures. British Journal of Anaesthesia. 2007;99:309-11.

26. Kalra A. Decoding the Bland–Altman Plot: Basic Review. J Pract Cardiovasc Sci. 2017;3:36-8.

27. Sugimoto M, Takagi T, Hikichi T, Suzuki R, Watanabe K, Nakamura J, et al. Efficacy of endoscopic ultrasonography-guided fine needle aspiration for pancreatic neuroendocrine tumor grading. World J Gastroenterol. 2015;21:8118-24.

28. Zilli A, Arcidiacono PG, Conte D, Massironi S. Clinical impact of endoscopic ultrasonography on the management of neuroendocrine tumors: lights and shadows. Dig Liver Dis. 2018;50:6-14.

29. Larghi A, Capurso G, Carnuccio A, Ricci R, Alfieri S, Galasso D, et al. Ki-67 grading of nonfunctioning pancreatic neuroendocrine tumors on histologic samples obtained by EUS-guided fine-needle tissue acquisition: a prospective study. Gastrointest Endosc. 2012;76:570-7.

30. Manta R, Nardi E, Pagano N, Ricci C, Sica M, Castellani D, et al. Pre-operative Diagnosis of Pancreatic Neuroendocrine Tumors with Endoscopic Ultrasonography and Computed Tomography in a Large Series. J Gastrointestin Liver Dis. 2016;25:317-21.

31. Puli SR, Kalva N, Bechtold ML, Pamulaparthy SR, Cashman MD, Estes NC, et al. Diagnostic accuracy of endoscopic ultrasound in pancreatic

neuroendocrine tumors: a systematic review and meta analysis. World J Gastroenterol. 2013;19:3678-84.

32. Unno J, Kanno A, Masamune A, Kasajima A, Fujishima F, Ishida K, et al. The usefulness of endoscopic ultrasound-guided fine-needle aspiration for the diagnosis of pancreatic neuroendocrine tumors based on the World Health Organization classification. Scand J Gastroenterol. 2014;49:1367-74.

33. Piani C, Franchi GM, Cappelletti C, Scavini M, Albarello L, Zerbi A, et al. Cytological Ki-67 in pancreatic endocrine tumours: an opportunity for pre-operative grading. Endocr Relat Cancer. 2008;15:175-81.

34. Farrell JM, Pang JC, Kim GE, Tabatabai ZL. Pancreatic neuroendocrine tumors: accurate grading with Ki-67 index on fine-needle aspiration specimens using the WHO 2010/ENETS criteria. Cancer Cytopathol. 2014;122:770-8.

35. Fujimori N, Osoegawa T, Lee L, Tachibana Y, Aso A, Kubo H, et al. Efficacy of endoscopic ultrasonography and endoscopic ultrasonography-guided fine-needle aspiration for the diagnosis and grading of pancreatic neuroendocrine tumors. Scand J Gastroenterol. 2016;51:245-52.