THIEME
OPEN
ACCESS

Interobserver agreement among cytopathologists in the evaluation of pancreatic endoscopic ultrasound-guided fine needle aspiration cytology specimens*

Authors   Rawad Mounzer[1], Roy Yen[1], Carrie Marshall[2], Sharon Sams[2], Sanjana Mehrotra[2], Mohamed Sherif Said[3], Joshua C. Obuch[1], Brian Brauer[1], Augustin Attwell[1], Norio Fukami[1], Raj Shah[1], Stuart Amateau[1], Matthew Hall[1], Lindsay Hosford[1], Robert Wilson[1], Amit Rastogi[4], Sachin Wani[1]

Institutions   Institutions are listed at the end of article.

Corresponding author
Sachin Wani, MD
Division of Gastroenterology
and Hepatology
University of Colorado Anschutz
Medical Center
Mail Stop F735
1635 Aurora Court
Rm 2.031
Aurora
CO 80045
USA
Fax: +1-720-848-2749
sachinwani10@yahoo.com

Background and aims: Endoscopic ultrasound with fine needle aspiration (EUS-FNA) has become the standard of care in the evaluation of solid pancreatic lesions. Limited data exist on interobserver agreement (IOA) among cytopathologists in assessing solid pancreatic EUS-FNA specimens. This study aimed to evaluate IOA among cytopathologists in assessing EUS-FNA cytology specimens of solid pancreatic lesions using a novel standardized scoring system and to assess individual clinical and cytologic predictors of IOA.

Methods: Consecutive patients who underwent EUS-FNA of solid pancreatic lesions at a tertiary care referral center were included. EUS-FNA slides were evaluated by four blinded cytopathologists using a standardized scoring system that assessed final cytologic diagnosis and quantitative (number of nucleated/diagnostic cells) and qualitative (bloodiness, inflammation/necrosis, contamination, artifact) cytologic parameters. Final clinical diagnosis was based on final cytology, surgical pathology, or 1-year clinical follow-up. IOA was calculated using multi-rater kappa (κ) statistics. Bivariate analyses were performed comparing cases with and without uniform agreement among the cytopathologists followed by logistic regression with backward elimination to model likelihood of uniform agreement.

Results: Ninety-nine patients were included (49% males, mean age 64 years, mean lesion size 26 mm). IOA for final diagnosis was moderate (κ= 0.45, 95% confidence interval (CI) 0.4–0.49) with minimal improvement when combining suspicious and malignant diagnoses (κ=0.54, 95%CI 0.49–0.6). The weighted kappa value for overall diagnosis was 0.65 (95%CI 0.54–0.76). IOA was slight to fair (κ=0.04–0.32) for individual cytologic parameters. A final clinical diagnosis of malignancy was the most significant predictor of agreement [OR 3.99 (CI 1.52–10.49)].

Conclusions: Interobserver agreement among cytopathologists for pancreatic EUS-FNA specimens is moderate-substantial for the final cytologic diagnosis. The final clinical diagnosis of malignancy was the strongest predictor of agreement. These results have significant implications for patient management and need to be validated in future trials.

Introduction
▼

Pancreatic cancer is currently the fourth leading cause of cancer-related deaths in both men and women in the United States with an estimated 40,560 deaths annually [1]. Between 2010 and 2030, this disease is projected to be amongst cancers with the highest increases in incidence with an estimated increase of 55% [2]. Despite vast efforts to stem this disease, mortality rates have remained fairly unchanged with 5-year survival rates increasing from 3.1% to only 6.9% over the past three decades [3, 4].

With the advent and increased availability of endoscopic ultrasound (EUS), our ability to detect and diagnose pancreatic cancer has greatly improved. This has largely been driven by the high sensitivity and specificity (>90%) of EUS in detecting small pancreatic lesions (<2 cm), its capability of tissue acquisition by fine needle aspiration (FNA), and the excellent safety profile of this modality [5–10]. EUS-FNA has also been shown to have a high sensitivity and specificity in the diagnosis of solid pancreatic neoplasms, reported to be 85% (95% confidence interval [CI], 84–86%) and 98% (95%CI, 97–99%) respectively in a recent meta-analysis [11–13].

Despite the superior performance characteristics of EUS-FNA in the diagnosis of solid malignant neoplasms, interobserver agreement amongst cy-

* Results of this study were presented in part at Digestive Disease Week 2015, Washington, DC, USA.

topathologists in the evaluation of FNA samples of these lesions has yet to be extensively and rigorously evaluated. Currently available data are limited to the evaluation of pancreatic neuroendocrine tumors and the histologic grading of pancreatic cancer on fine needle biopsy samples [14,15]. Although standardized nomenclature for pancreaticobiliary cytology has recently been published by the Papanicolaou Society of Cytopathology [16], there currently exist no standardized criteria for the evaluation of EUS-FNA sample adequacy that incorporates various factors that can impact the overall diagnostic impression. Such factors include the presence of gastrointestinal contaminants and amount of blood or non-diagnostic cells that may obscure visualization of diagnostic tissue [17–19].

An accurate diagnosis of solid pancreatic lesions is essential to the appropriate and timely administration of therapy. Failure to achieve an accurate diagnosis can lead to patient anxiety, a delay in treatment or inappropriate treatment with potentially poor clinical outcomes. Interobserver variability amongst pathologists in evaluating pathology specimens has been extensively studied in other areas including Barrett's esophagus. Studies identifying poor interobserver reproducibility in diagnosing the degree of dysplasia and early malignancy in this disease process have led to the modification of current guidelines, which now require confirmation of the pathologic diagnosis by an expert pathologist [20–22]. Identifying such variability in the assessment of EUS-FNA samples can have similar implications on the handling of these specimens and ultimately impact patient care.

The primary aim of this study was to assess interobserver agreement among cytopathologists in evaluating EUS-FNA cytology specimens of solid pancreatic lesions for overall diagnosis and individual specimen-related quantitative and qualitative parameters by utilizing a novel standardized scoring system. The secondary aim was to evaluate the individual clinical and cytologic parameters impacting interobserver agreement.

Methods
▼
Study setting
This study was conducted at a tertiary care referral center. Approval for the study was obtained from the Institutional Review Board and Human Research Protection Office at the University of Colorado Anschutz Medical Center.

Study population
Consecutive patients who underwent EUS-FNA of solid pancreatic lesions from August 2011 to August 2012 were identified and included in the study. Patient demographics (age, sex), clinical history of acute or chronic pancreatitis, and presenting symptoms (weight loss, jaundice) were collected by chart review. EUS reports were also reviewed to extract data with regard to lesion size, location, echogenicity, and sampling technique (including needle gauge and number of passes). All procedures were performed by experienced endosonographers, each having performed >500 cases. The final cytologic and clinical diagnoses were also recorded. Patients were considered to have a final clinical diagnosis of malignancy based on a final cytologic diagnosis of malignancy or surgical pathology revealing malignancy. Benign disease was based on a final cytologic read of no evidence of malignancy and at least 1 year of clinical follow-up during which the patient was not subsequently diagnosed with a pancreatic malignancy. Patients who did not undergo surgical resection or had less than 1 year of clinical follow-up were excluded from the study.

EUS-FNA sampling
EUS-FNA samples were obtained using a 19-, 22–or 25-guage EUS-FNA needle (Echotip Ultra HD endoscopic ultrasound needle, Cook Medical, Winston Salem, NC, or Expect endoscopic ultrasound needle, Boston Scientific, Natick, MA, USA). Needle size selection was based on the preference of the endosonographer. All slides were prepared by an experienced cytotechnologist. Tissue samples were flushed out of the needle onto a glass slide using a 10 mL air-filled syringe with the remaining tissue rinsed with normal saline into 10% buffered formalin for cell block preparation. Typically, two or three slides (one alcohol-fixed Papanicolaou-stained slide and one or two air-dried slides stained with modified Giemsa stain (DiffQuik)) were made for each needle pass. A cytopathologist was present on-site during each EUS procedure and evaluated air-dried Giemsa stained slide samples to assess whether the pancreatic tissue sample obtained was adequate for cytologic evaluation. Tissue samples were then further processed within the pathology department before definitive cytologic evaluation.

Scoring tool
All EUS-FNA slides included in the study were de-identified and evaluated by four blinded cytopathologists (three experienced Board Certified cytopathologists, one cytopathology fellow)

Table 1 Standardized scoring tool to assess individual EUS-FNA slides* and final cytologic diagnosis based on predefined quantity and quality measures.

	Score		
	1	**2**	**3**
Quantitative measures			
No. nucleated cells/slide	Few: <25 cells	Moderate: 25–500 cells	Numerous: >500 cells
No. diagnostic cells/slide	Few: <25 cells	Moderate: 25–500 cells	Numerous: >500 cells
Qualitative measures			
Blood	Absent or non-obscuring	Mild obscuring (<25% lesional cells affected)	Extensive obscuring (>25% lesional cells affected)
Inflammation and necrosis	Absent or non-obscuring	Mild obscuring (<25% lesional cells affected)	Extensive obscuring (>25% lesional cells affected)
Gastrointestinal contaminant	Minimal/none	5–25% of nucleated cells	>25% of nucleated cells
Preparation/staining Artifacts	None	Minimal artifact (<25% lesional cells affected)	Extensive artifact (>25% lesional cells affected)

* Typically, one alcohol-fixed Papanicolaou-stained slide and one air-dried modified Giemsa-stained slide were made for each pass.

using a previously described modified scoring system that had been reviewed, modified, and standardized amongst the cytopathologists (◐ **Table 1**) [17 – 19]. This scoring tool required the cytopathologists to broadly quantify both the number of nucleated cells (i.e. inflammatory cells, gastrointestinal contaminant as well as cells from the targeted lesion) and the number of diagnostic cells on each slide. Diagnostic cells were defined as cells of pancreatic origin (acinar, ductal, or islet cells) that would allow for a cytologic diagnosis. Cells with significant cytologic atypia or frank features of malignancy would be considered diagnostic (suggestive of malignancy) as would clusters of acinar cells with increased interstitial fibrosis (suggestive of chronic pancreatitis). The amount of blood referred specifically to clots that had formed within the biopsy needle and entrapped nucleated cells. Preparation and staining artifacts included overstained or understained slides, very thick smears, and bubbles under the coverslip.

The reviewing cytopathologists were blinded to all clinical information including tumor size, location within the pancreas, patient symptoms, radiologic findings, and type of mucosa (gastric or duodenal) that was traversed to obtain the FNA sample. The scoring system assessed specimen-related quantitative factors (number of nucleated cells and number of diagnostic cells per slide) and qualitative factors (amount of blood, inflammation and necrosis, gastrointestinal contamination, and quality of preparation/staining). All EUS-FNA slides from all sampling passes of each lesion were reviewed by all cytopathologists and each slide was individually scored on a scale of 1 – 3 for each of the abovementioned quantitative and qualitative parameters using the described scoring system. Additionally, for every case, an overall score was provided for each graded parameter (typically the most common score given to the individual slides) and for the final cytologic diagnosis. The "overall" cytologic diagnosis was the most definitive diagnosis reached after reviewing all slides and a diagnosis of malignancy was recorded even if only a single slide from the total case had been scored as malignant. The evaluation of cell block specimens was left to the discretion of the cytopathologists. The cell block material (histology) was not graded in the manner of the aspirate slides as the purpose of this study was to evaluate for cytologic features. Final cytologic diagnoses were categorized as: insufficient, benign, atypical, suspicious for malignancy, or malignant. Neoplasms such as neuroendocrine tumors and solid-pseudopapillary neoplasms were categorized as malignant for the purposes of this study. The final clinical diagnosis was then defined as benign or malignant based on the abovementioned clinical data.

Statistical analysis

Collected data were incorporated into a datasheet using Microsoft Excel for Windows 2007 (Microsoft Corp, Redmond, WA, USA) and then coded for analysis. The statistical analysis was performed by a senior outcomes researcher (M.H.). Values for mean, median, range, and standard deviation were calculated using Microsoft Excel for Windows. Interobserver variability was calculated using multi-rater kappa (κ) statistics with a 95 % confidence interval (CI). Landis and Koch definitions were used to evaluate the strength of rater agreement and were categorized as: slight (0 – 0.20); fair (0.21 – 0.40); moderate (0.41 – 0.60); substantial (0.61 – 0.80); almost perfect (0.81 – 1.00) [23]. For agreement on the overall diagnosis, weighted kappa values were computed that used weights (i.e. Cicchetti-Allison weights) to quantify the relative difference between ordinal categories. Bivariate analyses were also performed to compare cases with and without uniform

agreement followed by logistic regression with backward elimination to model the likelihood of uniform agreement. Independent predictors of unanimous agreement that were evaluated using bivariate analysis included patient demographics, clinical parameters, EUS findings, and cytologic parameters. Test statistics with $P < 0.05$ were considered significant. SAS version 9.4 (SAS Institute, Cary, NC, United States) was used for data analysis. Assuming an overall κ of 0.8 amongst four cytopathologists, a prevalence of malignancy of 0.6, and a lower limit of 95 % CI ≥ 0.7, a sample size of at least 72 cases was required.

Results
▼

A total of 99 patients who underwent EUS-FNA for further evaluation of solid pancreatic lesions were identified. Patient characteristics are presented in ◐ **Table 2**. The mean patient age was 64 years (SD 13) and 49 % of patients were males. The median lesion size on EUS was 26 mm with the distribution of lesion location being: 56 % head/uncinate, 8 % neck, 21 % body, and 15 % tail. A 22-gauge needle was used for EUS-FNA in 78 % of the cases. The median number of passes performed during EUS was 3 (range 1 – 8). At the time of presentation, 44 % of patients had experienced weight loss and 28 % had developed jaundice. Amongst these patients, 16 % had developed a recent episode of acute pancreatitis and 7 % had a known clinical diagnosis of chronic pancreatitis. The final cytologic diagnosis was that of malignancy in 60 % of cases, suspicious for malignancy in 4 %, atypical in 5 %, benign in 28 %, and inadequate in 3 %. All patients who had a final cytologic

Table 2 Patient demographics, clinical parameters, EUS findings and final cytologic diagnosis.

Patient demographics (n = 99)	
Age, mean (SD), years	64 (13)
Gender, male, n (%)	48 (48.5)
EUS parameters	
Lesion size, median (range), mm	26 (4 – 53)
Lesion location, n (%)	
Head/uncinate	55 (56)
Neck	8 (8)
Body	21 (21)
Tail	15 (15)
Lesion echogenicity, n (%)	
Hypoechoic	77 (78)
Hyperechoic	1 (1)
Isoechoic	2 (2)
Mixed	11 (11)
Anechoic	8 (8)
Chronic pancreatitis, yes, n (%)	10 (10)
Needle gauge, median (range)	22 (19 – 25)
Needle passes, median (range)	3 (1 – 8)
Clinical parameters	
Weight loss, yes, n (%)	44 (44)
Jaundice, yes, n (%)	28 (28)
Acute pancreatitis, yes, n (%)	16 (16)
Chronic pancreatitis, yes, n (%)	7 (7)
Final cytologic diagnosis	
Inadequate	3 (3)
Benign	28 (28)
Atypical cells	5 (5)
Suspicious for malignancy	4 (4)
Malignant	59 (60)

diagnosis of "inadequate" had a benign final clinical diagnosis. Of the five patients who had a final cytologic diagnosis of "atypical cells", two patients had a final clinical diagnosis of intraductal papillary mucinous neoplasia (IPMN), one patient was found to have adenocarcinoma, and two patients had a benign final clinical diagnosis. All four patients with a final cytologic diagnosis of "suspicious for malignancy" were found to have malignant disease as their final clinical diagnosis. Of the 99 patients in this study, 24 patients ultimately underwent surgical resection. Three (12.5%) of the patients who had surgery were found to have benign disease and 21 (87.5%) were found to have malignancy on surgical pathology.

Interobserver agreement

Evaluation of interobserver agreement among the cytopathologists for the overall cytologic diagnosis revealed only moderate agreement with κ=0.45 (95%CI 0.40–0.49). Interobserver agreement for each of the individual parameters of the standardized tool was also assessed (○ Table 3). Only slight overall agreement among cytopathologists was found for several qualitative parameters including amount of blood, amount of gastrointestinal contaminant, and overall quality of slide preparation with κ values of 0.14 (95%CI 0.08–0.20), 0.14 (95%CI 0.08–0.20), and 0.04 (95%CI –0.04 to 0.11), respectively. There was fair agreement over the degree of inflammatory and/or necrotic tissue seen on each slide with κ=0.21 (95%CI 0.14–0.28). Fair agreement was also noted for quantitative parameters including number of nucleated cells and number of diagnostic cells per slide with κ values of 0.31 (95%CI 0.24–0.37) and 0.32 (95%CI 0.26–0.37), respectively.

Subgroup analysis

Weighted kappa statistics resulted in an improvement in interobserver agreement [κ=0.65 (95%CI 0.49–0.60), standard error 0.05]. In addition, only marginal improvement in interobserver agreement was found in a subanalysis that combined the categories of suspicious for malignancy and malignant [κ=0.54 (95%CI 0.49–0.60)] in evaluating overall cytologic diagnosis (○ Table 3). There was no significant change in κ values for final cytologic diagnosis and individual quantitative and qualitative measures in an analysis that excluded individual cytopathologists (○ Table 4). Additionally, an assessment of interobserver agreement for each of the parameters during the evaluation of individual slides revealed poor to fair agreement among the cytopathologists (κ=–0.04 to 0.58) (data not shown).

Predictors of agreement

Amongst the 99 patients evaluated, there was unanimous agreement among all cytopathologists with regard to the final cytologic diagnosis in 48 patients. Patients with unanimous agreement were found to be more likely to present with jaundice (P=0.001), have a larger lesion size (P=0.04), a greater number of nucleated and diagnostic cells per slide (P<0.001), and a lower amount of gastrointestinal contamination (P=0.03) on EUS-FNA samples. Additionally, patients with the final cytologic (P=0.02) and final clinical (P=0.003) diagnosis of malignancy were more likely to have unanimous agreement among all cytopathologists (○ Table 5). On multivariable analysis, the only predictor for uniform agreement among cytopathologists was a final clinical diagnosis of malignancy [OR 3.99 (CI 1.52–10.49)]. When the accuracy of the individual cytopathologists was calculated using the final clinical diagnosis as the gold standard and excluding specimens that were deemed inadequate, the accuracy of the four cytopathologists was: 94.1%, 92.6%, 90.3%, and 90.8%.

Discussion

▼

In recent years, EUS-FNA has come to play a key role in the diagnosis and staging of pancreatic cancer due to its accuracy in localizing these lesions and safety profile for acquiring tissue samples [3, 24, 25]. Despite the broad utilization of this technique, data regarding to interobserver agreement among cytopathologists evaluating EUS-FNA samples of solid pancreatic lesions remain very limited. This lies in stark contrast to the extensive literature evaluating interobserver agreement among pathologists in other gastrointestinal diseases such as Barrett's esophagus, colon polyps, and inflammatory bowel disease [26–32]. In this study, we evaluated interobserver agreement among cytopathologists

Parameter	Kappa (95%CI)	Standard Error	Strength of agreement
Final cytologic diagnosis			
Overall diagnosis	0.45 (0.40–0.49)	0.02	Moderate
Overall diagnosis combining suspicious and malignant	0.54 (0.49–0.60)	0.03	Moderate
Quantity measures			
Number of nucleated cells/slide	0.31 (0.24–0.37)	0.03	Fair
Number of diagnostic cells/slide	0.32 (0.26–0.37)	0.03	Fair
Quality measures			
Amount of blood	0.14 (0.08–0.20)	0.03	Slight
Degree of inflammation/necrosis	0.21 (0.14–0.28)	0.04	Fair
Amount of gastrointestinal contaminants	0.14 (0.08–0.20)	0.03	Slight
Quality of slide preparation/staining	0.04 (–0.04 to 0.11)	0.04	Slight

Table 3 Interobserver agreement with strength of agreement among cytopathologists – overall kappa values for individual quantity and quality measures and final cytologic diagnosis.

Based on Landis and Koch definitions, strength of rater agreement was categorized as: 0–0.2, slight; 0.21–0.4, fair; 0.41–0.6, moderate; 0.61–0.8, substantial; 0.81–1, almost perfect.

Table 4 Interobserver agreement with strength of agreement among cytopathologists – overall kappa values for individual quantity and quality measures and final cytologic diagnosis with the exclusion of individual cytopathologists (CyP).

Parameter	Kappa (95%CI)				
	Overall	Excluding CyP1	Excluding CyP2	Excluding CyP3	Excluding CyP4
Quantity measures					
Number of nucleated cells/slide	0.31 (0.24 – 0.37)	0.25 (0.16 – 0.34)	0.48 (0.38 – 0.57)	0.23 (0.15 – 0.32)	0.25 (0.16 – 0.34)
Number of diagnostic cells/slide	0.32 (0.26 – 0.37)	0.27 (0.19 – 0.35)	0.34 (0.26 – 0.42)	0.38 (0.3 – 0.46)	0.26 (0.18 – 0.34)
Quality measures					
Amount of blood	0.14 (0.08 – 0.20)	0.12 (0.03 – 0.21)	0.16 (0.08 – 0.25)	0.13 (0.05 – 0.21)	0.09 (0.01 – 0.17)
Degree of inflammation/necrosis	0.21 (0.14 – 0.28)	0.13 (0.03 – 0.23)	0.3 (0.2 – 0.39)	0.21 (0.12 – 0.31)	0.14 (0.05 – 0.24)
Amount of gastrointestinal contaminants	0.14 (0.08 – 0.20)	0.02 (−0.07 – 0.11)	0.08 (−0.01 – 0.16)	0.25 (0.17 – 0.34)	0.15 (0.05 – 0.24)
Quality of slide preparation/staining	0.04 (−0.04 – 0.11)	0.02 (−0.09 – 0.13)	0.02 (−0.09 – 0.13)	0.05 (−0.06 – 0.16)	0 (−0.11 – 0.11)
Final cytology diagnosis					
Overall diagnosis	0.45 (0.40 – 0.49)	0.47 (0.4 – 0.54)	0.42 (0.35 – 0.49)	0.5 (0.43 – 0.56)	0.39 (0.32 – 0.45)
Overall diagnosis combining suspicious and malignant	0.54 (0.49 – 0.60)	0.53 (0.45 – 0.61)	0.55 (0.47 – 0.62)	0.58 (0.51 – 0.66)	0.51 (0.44 – 0.59)

in the assessment of EUS-FNA samples of solid pancreatic lesions utilizing a novel standardized scoring system. Quantitative and qualitative sample characteristics in addition to clinical parameters were also evaluated for their impact on interobserver agreement.

Results of this study show that the interobserver agreement amongst cytopathologists was moderate ($\kappa = 0.45$) for the final cytologic diagnosis. There was marginal improvement in the level of agreement ($\kappa = 0.54$) when both suspicious and malignant categories were grouped together and further improvement in the level of agreement using weighted kappa statistics ($\kappa = 0.65$). This may partially be driven by the fact that the cytopathologists were blinded to patient clinical information. These findings have significant clinical implications, however, as the final cytologic diagnosis ultimately directs patient management. A diagnosis of malignancy would potentially prompt both surgical and oncological consultation whereas a non-malignant diagnosis will likely result in patient observation or repeat sampling if considerable clinical suspicion for an underlying malignancy persists.

Several factors that were significantly associated with unanimous agreement amongst cytopathologists were also identified. These included a clinical presentation of jaundice, lesion size, number of diagnostic and nucleated cells per slide, and degree of gastrointestinal contamination. The final cytologic and clinical diagnoses were also found to be significantly associated with improved agreement. The highest level of agreement amongst all parameters of the standardized scoring tool was the number of diagnostic cells per slide ($\kappa = 0.32$). Overall, there appeared to be better agreement on quantitative measures in comparison to the qualitative measures of the tool. This is potentially due to the greater degree of objectivity that can be applied in assessing and reporting quantitative parameters. It is reassuring to note that the single most significant predictor on multivariate analysis for unanimous agreement among cytopathologists was the final clinical diagnosis of malignancy.

As shown in recently published studies and studies currently still in press, the cytologic interpretation of aspirates from solid pancreatic lesions is fraught with difficulties including poorly sampled lesions due to technical factors such as the accessibility of the lesion, the endosonographer's skill and expertise in procuring adequately cellular and representative samples, the availability of on-site cytopathology evaluation and skill in preparing optimal smears, all of which may influence the cellularity of the lesion or the quality of the cytologic preparation [33 – 37]. Cases with low cellularity or poor technical quality of the cytologic preparations are more likely to receive less accurate diagnoses (i.e. diagnoses short of negative or positive, or indeterminate diagnoses) than cases of adequate cellularity and good quality. The second category of factors that may impact the rate of indeterminate diagnosis are lesion-related variables, such as very well-differentiated tumors with very little cytologic atypia, tumors showing extensive desmoplasia, necrosis, and cystic change, all of which may impact the cellularity of the specimen. Additionally, lesions occurring in the setting of pancreatitis can pose diagnostic challenges.

In this study, a 22 G needle was used for tissue acquisition in the majority of cases (77.7%) with the second most frequently used needle being 25 G (20.2%). This is in concordance with current practices where a 22 G needle is preferred amongst endosonographers for the sampling of pancreatic masses [33]. There is however, a trend towards increasing use of the 25 G needle for FNA which has been supported by a large meta-analysis showing increased sensitivity and comparable specificity of the 25 G needle in diagnosing malignancy in the pancreas [33, 38, 39].

At the present time, data on interobserver agreement of EUS-FNA samples of pancreatic lesions remain scant. In a recent study by Larghi et al. [15], EUS-guided fine needle biopsy samples from patients who had undergone surgical resection of pancreatic adenocarcinoma were evaluated by four experienced pathologists for tumor grade. The total interobserver agreement among the pathologists was found to be fair ($\kappa = 0.27$; 95%CI 0.14 – 0.38) in that study with kappa values for tumor grade ranging from 0.09 to 0.41. Interobserver agreement for the Ki-67 labeling index in pancreatic neuroendocrine tumors has also been evaluated in a study by Weyand et al. [14]. EUS-FNA samples and their corresponding surgically resected specimens were assessed for tu-

	No uniform agreement (n = 51)	Uniform agreement (n = 48)	P value
Demographics			
Age, median (range), y	63 (55 – 73)	66.5 (56.5 – 74)	0.41
Gender, male, n (%)	27 (52.9)	21 (43.8)	0.36
EUS parameters			
Lesion size, median (range), mm	23 (14,31)	30 (19,34)	0.04
Location, n (%)			
Head/uncinate	26 (51.0)	29 (60.4)	0.11
Neck	5 (9.8)	3 (6.3)	
Body	15 (29.4)	6 (12.5)	
Tail	5 (9.8)	10 (20.8)	
Echogenicity, hypoechoic, n (%)	34 (66.7)	43 (89.6)	0.07
Chronic pancreatitis, yes, n (%)	5 (9.8)	5 (10.4)	0.92
Needle gauge, median (range)	22 (22 – 22)	22 (22 – 25)	0.13
Needle passes, median (range)	4 (2 – 5)	3 (2 – 5.5)	0.60
Clinical parameters			
Weight loss, yes	19 (37.3)	25 (52.1)	0.27
Jaundice, yes	7 (13.7)	21 (43.8)	0.001
Acute pancreatitis, yes	9 (17.6)	7 (14.6)	0.52
Chronic pancreatitis, yes	3 (5.9)	4 (8.3)	0.70
Cytologic parameters[1]			
No.	2 (2 – 2.5)	3 (2.5 – 3)	<0.001
No. of diagnostic cells, median (range)	1.5 (1 – 2)	2.5 (2 – 3)	<0.001
Blood, median (range)	2 (1 – 2)	1.5 (1 – 2)	0.34
Inflammation/necrosis, median (range)	1 (1 – 1.5)	1 (1 – 1.5)	0.53
Gastrointestinal contaminants, median (range)	1.5 (1 – 2)	1 (1 – 1.5)	0.002
Preparation/staining, median (range)	1 (1 – 1)	1 (1 – 1)	0.42
Final cytology read			
Inadequate, n (%)	2 (3.9)	1 (2.1)	0.02
Benign, n (%)	21 (41.2)	7 (14.6)	
Atypical, n (%)	3 (5.9)	2 (4.2)	
Suspicious, n (%)	3 (5.9)	1 (2.1)	
Malignant, n (%)	22 (43.1)	37 (77.1)	
Final diagnosis[2]			
Benign, n (%)	24 (47.1)	9 (18.8)	0.003
Malignant, n (%)	27 (52.9)	39 (81.3)	

Table 5 Demographic, EUS and clinical parameters associated with IOV for EUS-FNA samples with no uniform and uniform agreement between cytopathologists.

[1] Median scores of all four cytopathologists were used for individual quantity and quality measures.
[2] Final diagnosis was based on final cytologic diagnosis, surgical pathology or clinical follow-up of at least 1 year.

mor grade using the WHO grading system. Very good interobserver agreement was found between the two cytopathologists evaluating both EUS-FNA samples and surgical specimens. Comparison of tumor grade obtained from surgical specimens with that from EUS-FNA samples, however, revealed discrepancies in tumor grading with cytology found to more likely underestimate the tumor grade. Moreover, in an earlier study by Larghi et al. [40], preoperative EUS-FNA sampling of pancreatic neuroendocrine tumors was found to be safe and highly accurate in diagnosing pancreatic neuroendocrine tumors and determining Ki-67 proliferation indices. That study, however, was limited by a small sample size with only 12 patients eventually undergoing surgery [40].

There are several limitations in this study that are worth mentioning. The use of cell block was not consistent amongst all cytopathologists and was left to their discretion in individual cases. This is reflective of the use of cell block in our practice. The purpose of this study was to assess interobserver agreement with regard to cytologic features. Additionally, all cytopathologists were blinded to clinical data when evaluating the FNA specimens. Although clinical information is usually presented in the setting of on-site cytopathology evaluation, and helps guide the cytopa-

thologist's evaluation of specimens, these data were not available to them in our study. The impact of this on the study results, however, is limited given that all four cytopathologists who participated in the study did not have access to the clinical data. The true impact of clinical data provided to the cytopathologists (tumor size, location within the pancreas, patient symptoms, and radiologic findings) on interobserver agreement needs to be addressed in future studies. Given the standardized nature of slide preparation at our institution, this study was unable to assess for a correlation between the number of slides prepared per pass and interobserver agreement. Finally, data with regard to the EUS approach in tissue acquisition (i. e. transduodenal vs. transgastric) were not collected. We were therefore unable to evaluate the impact of the sampling approach on the specimen quality and accuracy of diagnosis. This would be worth evaluating in future studies.

Evaluating interobserver agreement among cytopathologists reading EUS-FNA samples of pancreatic lesions can similarly have significant implications on patient care and management, particularly when evaluating for underlying malignancy. Furthermore, standardization in the evaluation of specimen quality and determining predictors of agreement can help reduce varia-

bility and provide uniformity in the reporting on these samples. In this study, we evaluated interobserver agreement among cytopathologists in the assessment of EUS-FNA samples of solid pancreatic lesions utilizing a standardized scoring system. Based on these data, interobserver agreement remains modest at best when assessing the overall cytologic diagnosis. While certain parameters with regard to specimen adequacy appear to affect agreement (number of nucleated and diagnostic cells per slide and the amount of gastrointestinal contamination), the final clinical diagnosis of malignancy was the only predictor associated with unanimous agreement among all cytopathologists. Enhancing tissue acquisition techniques to improve the yield of diagnostic cells while decreasing gastrointestinal contamination is likely to improve interobserver agreement. Formal training of cytopathologists on the utilization of a standardized tool for assessing specimen adequacy and diagnostic criteria may also contribute to better outcomes. Large multicenter studies are required to validate the proposed scoring tool and the results of this study.

Competing interests: None.

Institutions
[1] Division of Gastroenterology and Hepatology, University of Colorado Anschutz Medical Center, Aurora, CO, USA
[2] Department of Pathology, University of Colorado Anschutz Medical Center, Aurora, CO, USA
[3] Department of Pathology, Denver Health Medical Center, Denver, CO, USA
[4] Division of Gastroenterology, University of Kansas School of Medicine and Veterans Affairs Medical Center, Kansas City, MO, USA

Acknowledgments
▼

References
1 *Siegel RL*, *Miller KD*, *Jemal A*. Cancer statistics, 2015. CA Cancer J Clin 2015; 65: 5–29
2 *Smith BD*, *Smith GL*, *Hurria A* et al. Future of cancer incidence in the United States: burdens upon an aging, changing nation. J Clin Oncol 2009; 27: 2758–2765
3 *Tempero MA*, *Malafa MP*, *Behrman SW* et al. Pancreatic adenocarcinoma, version 2.2014: featured updates to the NCCN guidelines. J Natl Compr Canc Netw 2014; 12: 1083–1093
4 *Sun H*, *Ma H*, *Hong G* et al. Survival improvement in patients with pancreatic cancer by decade: a period analysis of the SEER database, 1981–2010. Sci Rep 2014; 4: 6747
5 *Rosch T*, *Lorenz R*, *Braig C* et al. Endoscopic ultrasound in pancreatic tumor diagnosis. Gastrointest Endosc 1991; 37: 347–352
6 *Owens DJ*, *Savides TJ*. Endoscopic ultrasound staging and novel therapeutics for pancreatic cancer. Surg Oncol Clin N Am 2010; 19: 255–266
7 *Eloubeidi MA*, *Jhala D*, *Chhieng DC* et al. Yield of endoscopic ultrasound-guided fine-needle aspiration biopsy in patients with suspected pancreatic carcinoma. Cancer 2003; 99: 285–292
8 *Eloubeidi MA*, *Tamhane A*. Prospective assessment of diagnostic utility and complications of endoscopic ultrasound-guided fine needle aspiration. Results from a newly developed academic endoscopic ultrasound program. Dig Dis 2008; 26: 356–363
9 *Adler DG*, *Jacobson BC*, *Davila RE* et al. ASGE guideline: complications of EUS. Gastrointest Endosc 2005; 61: 8–12
10 *Ngamruengphong S*, *Swanson KM*, *Shah ND* et al. Preoperative endoscopic ultrasound-guided fine needle aspiration does not impair survival of patients with resected pancreatic cancer. Gut 2015; 64: 1105–1110
11 *Hewitt MJ*, *McPhail MJ*, *Possamai L* et al. EUS-guided FNA for diagnosis of solid pancreatic neoplasms: a meta-analysis. Gastrointest Endosc 2012; 75: 319–331
12 *Fisher L*, *Segarajasingam DS*, *Stewart C* et al. Endoscopic ultrasound guided fine needle aspiration of solid pancreatic lesions: Performance and outcomes. J Gastroenterol Hepatol 2009; 24: 90–96
13 *Shin HJ*, *Lahoti S*, *Sneige N*. Endoscopic ultrasound-guided fine-needle aspiration in 179 cases: the M.D. Anderson Cancer Center experience. Cancer 2002; 96: 174–180
14 *Weynand B*, *Borbath I*, *Bernard V* 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: 389–395
15 *Larghi A*, *Correale L*, *Ricci R* et al. Interobserver agreement and accuracy of preoperative endoscopic ultrasound-guided biopsy for histological grading of pancreatic cancer. Endoscopy 2015; 47: 308–314
16 *Pitman MB*, *Centeno BA*, *Ali SZ* et al. Standardized terminology and nomenclature for pancreatobiliary cytology: the Papanicolaou Society of Cytopathology guidelines. Diagn Cytopathol 2014; 42: 338–350
17 *Rastogi A*, *Wani S*, *Gupta N* et al. A prospective, single-blind, randomized, controlled trial of EUS-guided FNA with and without a stylet. Gastrointest Endosc 2011; 74: 58–64
18 *Wani S*, *Gupta N*, *Gaddam S* et al. A comparative study of endoscopic ultrasound guided fine needle aspiration with and without a stylet. Dig Dis Sci 2011; 56: 2409–2414
19 *Wani S*, *Early D*, *Kunkel J* et al. Diagnostic yield of malignancy during EUS-guided FNA of solid lesions with and without a stylet: a prospective, single blind, randomized, controlled trial. Gastrointest Endosc 2012; 76: 328–335
20 *Downs-Kelly E*, *Mendelin JE*, *Bennett AE* et al. Poor interobserver agreement in the distinction of high-grade dysplasia and adenocarcinoma in pretreatment Barrett's esophagus biopsies. Am J Gastroenterol 2008; 103: 2333–2340; quiz 2341
21 *Skacel M*, *Petras RE*, *Gramlich TL* et al. The diagnosis of low-grade dysplasia in Barrett's esophagus and its implications for disease progression. Am J Gastroenterol 2000; 95: 3383–3387
22 *Wang KK*, *Sampliner RE*. Practice Parameters Committee of the American College of Gastroenterology. Updated guidelines 2008 for the diagnosis, surveillance and therapy of Barrett's esophagus. Am J Gastroenterol 2008; 103: 788–797
23 *Landis JR*, *Koch GG*. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159–174
24 *Gonzalo-Marin J*, *Vila JJ*, *Perez-Miranda M*. Role of endoscopic ultrasound in the diagnosis of pancreatic cancer. World J Gastrointest Oncol 2014; 6: 360–368
25 *Chang KJ*, *Nguyen P*, *Erickson RA* et al. The clinical utility of endoscopic ultrasound-guided fine-needle aspiration in the diagnosis and staging of pancreatic carcinoma. Gastrointest Endosc 1997; 45: 387–393
26 *Wani S*, *Mathur SC*, *Curvers WL* et al. Greater interobserver agreement by endoscopic mucosal resection than biopsy samples in Barrett's dysplasia. Clin Gastroenterol Hepatol 2010; 8: 783–788
27 *Kaye PV*, *Haider SA*, *Ilyas M* et al. Barrett's dysplasia and the Vienna classification: reproducibility, prediction of progression and impact of consensus reporting and p53 immunohistochemistry. Histopathology 2009; 54: 699–712
28 *Montgomery E*, *Bronner MP*, *Goldblum JR* et al. Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. Hum Pathol 2001; 32: 368–378
29 *Kerkhof M*, *van Dekken H*, *Steyerberg EW* et al. Grading of dysplasia in Barrett's oesophagus: substantial interobserver variation between general and gastrointestinal pathologists. Histopathology 2007; 50: 920–927
30 *Tanaka M*, *Masuda T*, *Yao T* et al. Observer variation of diagnoses based on simple biopsy criteria differentiating among Crohn's disease, ulcerative colitis, and other forms of colitis. J Gastroenterol Hepatol 2001; 16: 1368–1372
31 *Geboes K*, *Riddell R*, *Ost A* et al. A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. Gut 2000; 47: 404–409
32 *Osmond A*, *Li-Chang H*, *Kirsch R* et al. Interobserver variability in assessing dysplasia and architecture in colorectal adenomas: a multicentre Canadian study. J Clin Pathol 2014; 67: 781–786
33 *Karadsheh Z*, *Al-Haddad M*. Endoscopic ultrasound-guided fine-needle aspiration needles: which one and in what situation? Gastrointest Endosc Clin N Am 2014; 24: 57–69
34 *Chen J*, *Yang R*, *Lu Y* et al. Diagnostic accuracy of endoscopic ultrasound-guided fine-needle aspiration for solid pancreatic lesion: a systematic review. J Cancer Res Clin Oncol 2012; 138: 1433–1441

35 *Lee LS, Nieto J, Watson RR* et al. Randomized noninferiority trial comparing diagnostic yield of cytopathologist-guided versus 7 passes for EUS-FNA of pancreatic masses. Dig Endosc 2015: [Epub ahead of print] DOI 10.1111/den.12594

36 *Savides TJ, Donohue M, Hunt G* et al. EUS-guided FNA diagnostic yield of malignancy in solid pancreatic masses: a benchmark for quality performance measurement. Gastrointest Endosc 2007; 66: 277–282

37 *Bergeron JP, Perry KD, Houser PM* et al. Endoscopic ultrasound-guided pancreatic fine-needle aspiration: potential pitfalls in one institution's experience of 1212 procedures. Cancer Cytopathol 2015; 123: 98–107

38 *Madhoun MF, Wani SB, Rastogi A* et al. The diagnostic accuracy of 22-gauge and 25-gauge needles in endoscopic ultrasound-guided fine needle aspiration of solid pancreatic lesions: a meta-analysis. Endoscopy 2013; 45: 86–92

39 *Wani S, Muthusamy VR, Komanduri S.* EUS-guided tissue acquisition: an evidence-based approach (with videos). Gastrointest Endosc 2014; 80: 939–959 e7

40 *Larghi A, Capurso G, Carnuccio A* 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: 570–577