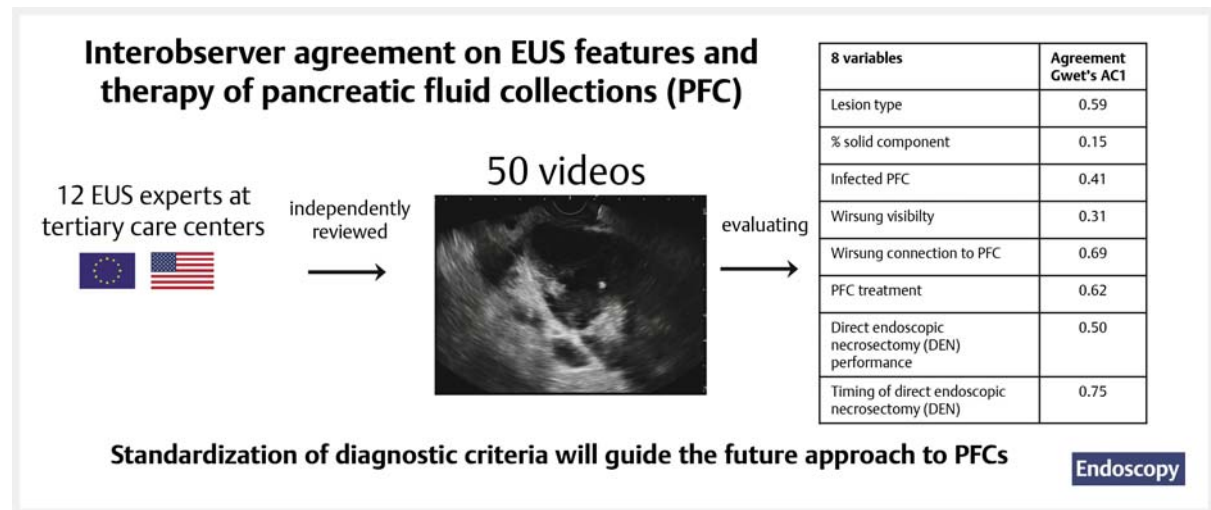


# The endoscopic ultrasound features of pancreatic fluid collections and their impact on therapeutic decisions: an interobserver agreement study

## INFOGRAPHIC



### Authors

Carlo Fabbri<sup>1</sup>, Todd H. Baron<sup>2</sup>, Giulia Gibiino<sup>1</sup> , Paolo Giorgio Arcidiacono<sup>3</sup>, Cecilia Binda<sup>1</sup> , Andrea Anderloni<sup>4</sup>, Gianenrico Rizzatti<sup>5,6</sup>, Manuel Pérez-Miranda<sup>7</sup> , Andrea Lisotti<sup>8</sup> , Loredana Correale<sup>5</sup>, Joan B. Gornals<sup>9</sup> , Iliaria Tarantino<sup>10</sup> , Maria Chiara Petrone<sup>3</sup>, Paolo Cecinato<sup>11</sup> , Pietro Fusaroli<sup>8</sup> , Alberto Larghi<sup>5,6</sup>

### Institutions

- 1 Unit of Gastroenterology and Digestive Endoscopy, Morgagni-Pierantoni Hospital of Forlì and Bufalini Hospital of Cesena, AUSL Romagna, Italy
- 2 Division of Gastroenterology and Hepatology, University of North Carolina, Chapel Hill, North Carolina, USA
- 3 Pancreato-biliary Endoscopy and Endosonography Division, Pancreas Translational and Clinical Research Center, San Raffaele Scientific Institute IRCCS, Milan, Italy
- 4 Digestive Endoscopy Unit, Division of Gastroenterology, Humanitas Research Hospital, Rozzano, Milano, Italy
- 5 Digestive Endoscopy Unit, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy
- 6 CERTT, Center for Endoscopic Research Therapeutics and Training, Catholic University, Rome, Italy
- 7 Department of Gastroenterology and Hepatology, Hospital Universitario Rio Hortega, Valladolid, Spain
- 8 Gastroenterology Unit, Hospital of Imola, Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy
- 9 Endoscopy Unit, Digestive Diseases Department, Hospital Universitari de Bellvitge, Bellvitge Biomedical Research Institute (IDIBELL), University of Barcelona, Barcelona, Spain

- 10 Endoscopy Service, Department of Diagnostic and Therapeutic Services, IRCCS-ISMETT, Palermo, Italy
- 11 Gastroenterology and Digestive Endoscopy Unit, Arcispedale Santa Maria Nuova di Reggio Emilia, AUSL-IRCCS di Reggio Emilia, Reggio Emilia, Italy

submitted 10.1.2021

accepted after revision 8.9.2021

published online 8.9.2021

### Bibliography

Endoscopy 2022; 54: 555–562

DOI 10.1055/a-1640-4365

ISSN 0013-726X

© 2021, Thieme. All rights reserved.

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany



Supplementary material

Supplementary material is available under

<https://doi.org/10.1055/a-1640-4365>

 Scan this QR-Code for the author commentary.



#### Corresponding author

Carlo Fabbri, MD, Unit of Gastroenterology and Digestive Endoscopy, Morgagni-Pierantoni Hospital of Forlì and Bufalini Hospital of Cesena, AUSL Romagna, Italy  
carlo.fabbri@auslromagna.it

#### ABSTRACT

**Background** A validated classification of endoscopic ultrasound (EUS) morphological characteristics and consequent therapeutic intervention(s) in pancreatic and peripancreatic fluid collections (PFCs) is lacking. We performed an interobserver agreement study among expert endosonographers assessing EUS-related PFC features and the therapeutic approaches used.

**Methods** 50 EUS videos of PFCs were independently reviewed by 12 experts and evaluated for PFC type, percentage solid component, presence of infection, recognition of and communication with the main pancreatic duct (MPD), stent choice for drainage, and direct endoscopic necrosectomy (DEN) performance and timing. The Gwet's AC1 coefficient was used to assess interobserver agreement.

**Results** A moderate agreement was found for lesion type (AC1, 0.59), presence of infection (AC1, 0.41), and need for DEN (AC1, 0.50), while fair or poor agreements were stated for percentage solid component (AC1, 0.15) and MPD recognition (AC1, 0.31). Substantial agreement was rated for ability to assess PFC–MPD communication (AC1, 0.69), decision between placing a plastic versus lumen-apposing metal stent (AC1, 0.62), and timing of DEN (AC1, 0.75).

**Conclusions** Interobserver agreement between expert endosonographers regarding morphological features of PFCs appeared suboptimal, while decisions on therapeutic approaches seemed more homogeneous. Studies to achieve standardization of the diagnostic endosonographic criteria and therapeutic approaches to PFCs are warranted.

## Introduction

Pancreatic and peripancreatic fluid collections (PFCs) are amongst the most ominous complications of severe acute pancreatitis [1, 2]. According to the revised Atlanta criteria, PFCs are classified on the basis of the interval time from the episode of acute pancreatitis and the presence or absence of necrotic content in acute fluid collections or acute necrotic collections and in pancreatic pseudocysts (PPCs) or walled-off necrosis (WON) [3]. Acute fluid collections and acute necrotic collections both occur within the first 4 weeks and are characterized by the absence of a mature wall. When spontaneous resolution does not occur, a well-defined capsule with a mature wall develops with formation of a PPC or WON [3]. All these morphological features are usually revealed by computed tomography (CT), which represents the gold standard for their detection [4]. However, CT proves less accurate for defining lesion type, in particular the presence of and percentage necrotic content [5–7].

Endoscopic ultrasound (EUS), which has become the procedure of choice to treat both PPCs and WON [8], can provide additional, and perhaps more accurate, information on the content of PFCs compared with CT. This can be extremely helpful when choosing the most appropriate therapeutic approach, even though a validated classification of PFCs based on EUS findings does not exist.

This lack of an EUS-based PFC morphological classification is responsible for the high variability and lack of therapeutic standardized approaches used by endosonographers to treat these conditions [9]. Differentiation of acute from chronic PFCs is mandatory because the method of drainage and out-

comes following therapeutic intervention differ substantially [10]. In most cases, acute collections do not evolve further and regress with time; only those that develop a mature wall and become symptomatic (i. e. cause multiorgan failure, organ compression, or abdominal compartment syndrome) need to be treated [11]. For chronic PFCs, it is fundamental to distinguish PPCs from WONs and to assess in the latter the amount of necrotic content. This information can guide the choice of stent type (plastic vs. metal), and the probability of needing direct endoscopic necrosectomy (DEN) [12].

To date, no study has focused on the agreement among endosonographers with regard to the EUS criteria to differentiate the PFC types or on the treatment decisions made based on the observed features of the PFCs. To fill in this gap, we performed an interobserver agreement study on the morphological definition of PFCs and on treatment options chosen by endosonographers based on the detected findings.

## Methods

### Observers

We invited experienced endosonographers from 20 tertiary care referral centers with a specific expertise in endoscopic treatment of PFCs to participate. In all these centers more than 30 PFC drainage procedures had been performed yearly over the previous 3–5 years. A total of 10 European and US tertiary care centers agreed to participate in this study. In two centers, two endosonographers participated in the review process. Experienced endosonographers were defined as those who had performed more than 50 PFC drainage procedures, with more than 5 years of continuous activity in EUS.

## Study material

Each participating center provided five high quality videos of PFCs, which were chosen among all procedures performed in patients with PFCs in the year before the study was started, giving a total of 50 videos, and these were shared with all involved endosonographers for review through a dedicated website. The decision on which video to select for review was completely at the discretion of each of the participating centers. The evaluated video was at least 30-seconds long and no still images were provided.

Videos were reviewed by each endosonographer in several sessions over a 2-month period. The characteristics of the patients are shown in ► **Table 1**; all of them had previously signed informed consent for EUS drainage that also granted permission for their examination findings to be used for research purposes. The study protocol was submitted to the Principal Investigator's Research Ethics Board, which determined that this was a quality assurance review that did not constitute human subject research, therefore requirement for approval from the board was waived.

The material was anonymous and in no instances was a patient's identity revealed. Recorded images were not preselected on the basis of quality to avoid bias; endosonographers were blinded regarding the clinical history and any therapeutic procedure(s) performed.

## Evaluation technique

Observers were asked to analyze eight variables for each video and to express their opinion in a database according to the parameters reported in ► **Table 2** and shown in ► **Fig. 1**.

## Statistical analysis

Interobserver agreement for the study descriptors was calculated using the alternative chance-correlated coefficient (AC1) statistic, with 95% CIs [13]. Although the kappa statistic is frequently used to test interobserver agreement, it does have some limitations. In particular, the kappa statistic is affected by the prevalence of the finding under consideration to a similar extent to predictive values being affected by the prevalence of the considered disease. For rare findings, very low values of kappa may not necessarily reflect low rates of overall agreement [14, 15]. The Gwet measure AC1 is supposed to deal with the apparent "paradox" of low agreement values despite a large percentage agreement [13].

Interpretation of the AC1 statistic is similar to the kappa statistic: the AC1 ranges from -1.00 (perfect disagreement) to +1.00 (perfect agreement), with a value of zero indicating a reliability equivalent to chance. Accordingly, the interobserver agreement was classified using criteria established by Landis and Koch: less than chance (<0.00), slight (0.00–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), and almost perfect (>0.80) [16].

We calculated the pairwise AC1 (pair-kappa agreement) for all rater pairs. We then used the arithmetic mean of these estimates to provide an overall index of agreement (i.e. mean agreement). The composite index of agreement for analysis of

► **Table 1** Characteristics of the 50 patients with pancreatic fluid collections from whom videos were taken to assess the interobserver agreement.

Median age (SD), years	62.0 (5.4)
Sex, male/female, n	26/24
Mean body mass index (SD), kg/m <sup>2</sup>	24.4 (4.2)
Smoker, n (%)	22 (44%)
Heavy alcohol intake, n (%)*	19 (38%)
Mean time after onset of acute pancreatitis (SD), weeks	6.0 (1.3)
Pancreatic site, n (%)	
▪ Uncinate process	4 (8%)
▪ Head/neck	17 (34%)
▪ Body	21 (42%)
▪ Tail	8 (16%)

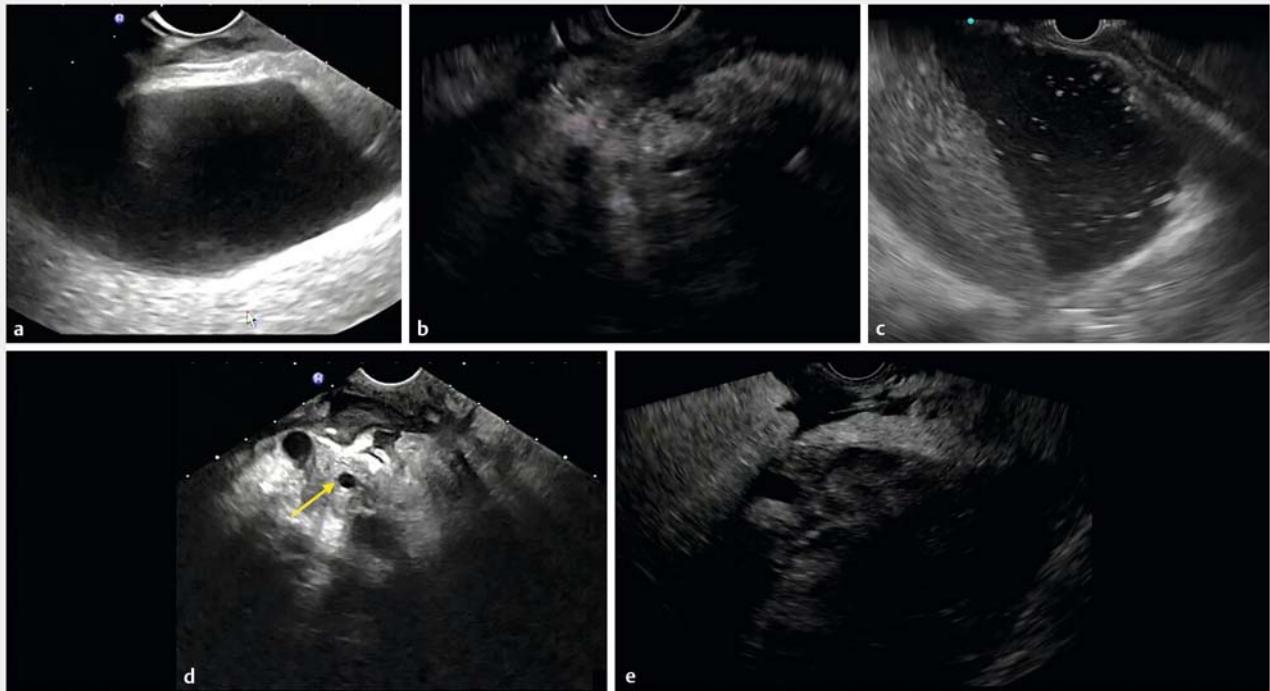
\* For women, >3 alcohol units/day; for men, >4 alcohol units/day.

► **Table 2** Variables analyzed by the 12 endosonographers through review of videos from 50 patients with pancreatic fluid collections (PFCs).

Lesion type: acute fluid collection, acute necrotic collection, pseudocyst, or walled-off necrosis
Evaluation of the percentage solid component: ≤25%, 25%–50%, >50%
Presence of infection demonstrated by recognition of debris in the fluid component of the PFC: Yes or No
Main pancreatic duct recognition: Yes, No, or Not evaluable
Main pancreatic duct communication with the PFC: Yes, No, or Not evaluable
Treatment of choice for the PFC: double-pigtail plastic stent or lumen-apposing metal stent
Performance of direct endoscopic necrosectomy (DEN): Yes or No
Timing of DEN: at the index procedure or after two-three days

multiple raters was also calculated using the extended version of AC1 (kappa) coefficients to statistics for multiple raters (details are given in **Appendix 1 s** with results in **Table 1 s**, see online-only Supplementary material).

Finally, to test for any inter-rater differences in categorical scores, binomial generalized random-effects models (GLMMs) were fitted. These GLMMs included no-fixed effects, and lesion and rater as random effects. The GLMM approach provides an estimate of the agreement repeatability at either the lesion or rater level. We used the bootstrap approach to obtain the 95% CI for the repeatability estimates. For interpretation of repeatability estimates, values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 are indicative of poor, moderate, good, and excellent reliability, respectively.



► **Fig. 1** Images of some evaluated variables: **a,b** type of lesion (**a** pseudocyst; **b** walled-off necrosis); **c** presence of debris as a sign of an infected collection; **d** main pancreatic duct visibility; **e** percentage solid component and indication for necrosectomy.

Details and results of these GLMM models are provided in **Appendix 2s**. All statistical analyses were conducted using R version 3.5.1 (2018–07–02).

## Results

All videos from the 50 PFCs were reviewed by the 12 endosonographers involved in the study, giving a maximum of 600 interpretations (50 PFCs × 12 endosonographers). Not all variables were scored for each patient because it was impossible for the observers to assess them. Results for each of the variables are given in ► **Table 3**.

### Agreement on lesion type

The individual scoring for this variable is shown in ► **Fig. 2a**. A total of 590 observations were available for the analysis. The mean agreement was 71.4% (95%CI 68.3%–74.3%), ranging across all different pairs of raters from 45.0% to 98.0%. The mean AC1 coefficient was 0.59 (95%CI 0.54–0.63), indicating moderate agreement. When the categories for classification were dichotomized as the presence of a PPC versus the absence, acute necrotic collection or WON, the mean agreement was 80.0% (95%CI 78.5%–82.4%) and the AC1 was 0.62 (95%CI 0.59–0.64), indicating moderate to substantial agreement.

### Agreement on the percentage solid component

The individual scoring for this variable is shown in ► **Fig. 2b**. Overall, 306 observations were available for the analysis of percentage solid component. The mean agreement was 42.2% (95%

%CI 37.1%–47.3%), with an AC1 coefficient of 0.15 (95%CI 0.07–0.23), indicating poor agreement.

### Agreement on the assessment of infection

The individual scoring for this variable is shown in **Fig. 1s**. Overall, 585 observations were available for the analysis of PFC infection. The mean agreement was 71.1% (95%CI 57.8%–84.5%), with an AC1 coefficient of 0.41 (95%CI 0.35–0.47), indicating poor to fair agreement.

### Agreement on MPD recognition

The individual scoring for MPD visibility is shown in **Fig. 2s**. Overall, 594 observations were available for analysis of this variable. The mean agreement was 48.8% (95%CI 37.3%–60.4%), with an AC1 coefficient of 0.31 (95%CI 0.19–0.43), indicating that main pancreatic duct recognition was an unreliable sign for the analysis of PFCs.

### Agreement on MPD communication with the PFC

The individual scoring for MPD communication with the PFC is shown in **Fig. 3s**. Overall, 597 observations were available for the analysis of this variable. The mean agreement was 79.7% (95%CI 70.0%–86.8%), with an AC1 of 0.69 (95%CI 0.53–0.76), indicating substantial to moderate agreement.

### Agreement on PFC treatment

The individual scoring for this variable is shown in ► **Fig. 3a**. Overall, 573 observations were available to analyze the type of stent to be used for PFC treatment. The mean agreement was

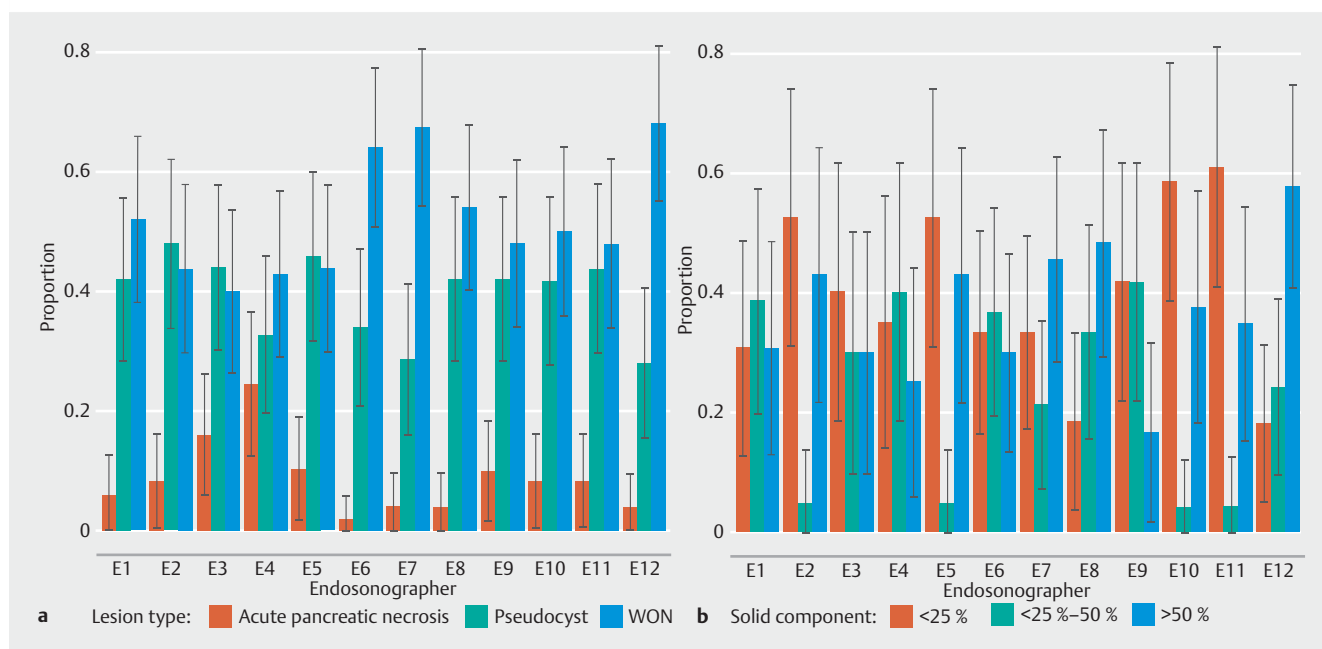
**► Table 3** Interobserver agreement among the 12 endosonographers for the evaluation of the different variables on 50 videos of pancreatic fluid collections as measured by crude agreement and Gwet's AC1.

Study variable	Number of observations	Number of cases	Proportion of agreement (95%CI), %	Gwet's AC1 (95%CI)
Lesion type	590	50	71.4 (68.3–74.3)	0.59 (0.54–0.63)
Percentage solid component <sup>1</sup>	306	41	42.2 (37.1–47.3)	0.15 (0.07–0.23)
Infection	585	50	71.1 (57.8–84.5)	0.41 (0.35–0.47)
EUS evaluation of the wall of the PFC	592	50	62.4 (55.0–69.0)	0.26 (0.18–0.34)
Main pancreatic duct recognition	594	50	48.8 (37.3–60.4)	0.31 (0.19–0.43)
Main pancreatic duct communication with the PFC	597	50	79.7 (70.0–86.8)	0.69 (0.53–0.76)
PFC treatment	573	50	76.4 (74.1–79.0)	0.62 (0.59–0.66)
Performance of necrosectomy <sup>2</sup>	554	50	75.0 (68.7–80.5)	0.50 (0.43–0.58)
Timing of necrosectomy <sup>2</sup>	154	50	75.1 (61.0–88.6)	0.75 (0.62–0.89)

EUS, endoscopic ultrasound; PFC, pancreatic fluid collections.

<sup>1</sup> Evaluated only for lesions classified as walled-off necrosis.

<sup>2</sup> Excluding samples rated as "plastic stents" for treatment purpose.



**► Fig. 2** Individual scores in proportions for each of the 12 endosonographers for assessment of: **a** lesion type; **b** necrotic content. WON, walled off necrosis.

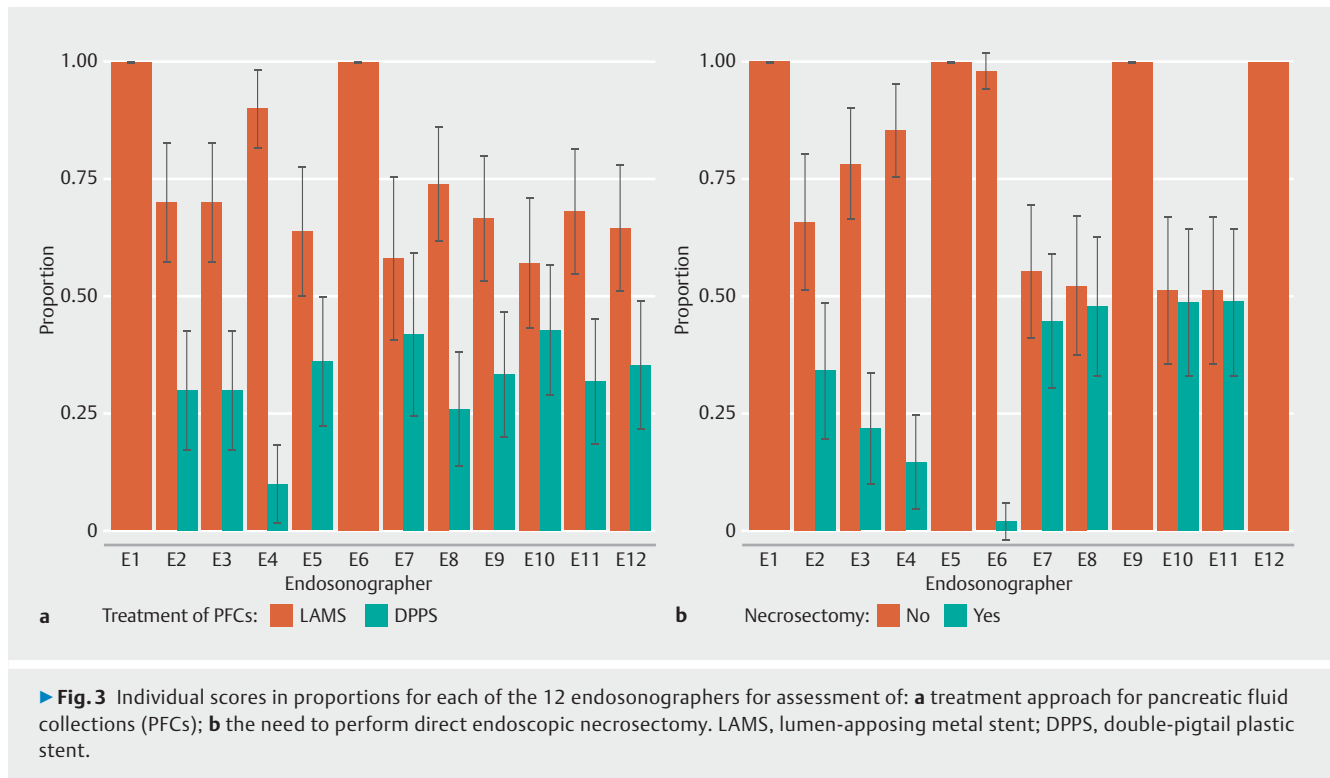
76.4% (95%CI 74.1%–79.0%), with an AC1 of 0.62 (95%CI 0.59–0.66), indicating substantial agreement.

### Agreement on the need for DEN

The individual scoring for this variable is shown in ► **Fig. 3b**. Overall, 554 observations were available for the analysis of the need for DEN. The mean agreement was 75.0% (95%CI 68.7%–80.5%), with an AC1 of 0.50 (95%CI 0.43–0.58), indicating moderate agreement.

### Agreement on the timing of DEN

Because the experts did not score this parameter when DEN was not needed, only 154 observations were available for the analysis of this variable. The rate of observations classified as after 2 or 3 days was 134/154 (87%), ranging from 18% to 100% across the endosonographers, while the remaining 20 observations (13%) were classified as at the index procedure. The mean agreement was 75.1% (95%CI 61.0%–88.6%), with an AC1 of 0.75 (95%CI 0.62–0.89), indicating substantial agreement.



## Discussion

We performed a study aimed at evaluating the interobserver agreement among expert endosonographers on the morphological assessment of PFCs and the treatment options chosen on the basis of the detected findings through review of 50 videos of procedures. Overall, a moderate agreement on distinguishing PFC type, presence of infection, and the need for DEN was found; substantial agreement was reached for the ability to recognize a communication of the PFC with the MPD, the stent type to perform drainage, and the timing for DEN, while poor and fair agreements were achieved for evaluation of the percentage necrotic content and MPD recognition, respectively.

In the early 2000 s, the term “PANCODE” was introduced to describe peripancreatic collections on CT. It was included in the revised Atlanta criteria, which enabled better agreement on classification and on multidisciplinary decision-making for the optimal therapeutic approach [17], according to a previous multidisciplinary interobserver agreement study [3]. The accordance was critically revised by radiologists and then evaluated together with surgeons and gastroenterologists [18]. Overall, a moderate-to-good agreement was observed for the evaluation of the CT criteria for both peripancreatic and pancreatic acute and chronic PFCs in different studies [18–20].

The weakest point of CT scanning is the detection of necrotic debris and its quantification, which can change a treatment decision. This task can be better accomplished by magnetic resonance imaging (MRI), which has been reported to be superior to CT in this setting [21]. In real-life, however, MRI is not usually performed if the patient has already undergone a CT and PFC treatment has been shown to be needed. The assessment of ne-

crotic content relies on the findings at EUS, which is nowadays considered the treatment of choice for symptomatic PFCs, rather than surgical and interventional radiology management [22–25].

The type of PFC, percentage necrotic content, and the presence of infection have a major impact on the decision as to how to perform EUS-guided drainage, even though a clear-cut standardization of the procedure is still lacking [9, 26]. Indeed, in the last few years, the introduction of lumen-apposing metal stents (LAMSs) with a cystotome on the catheter tip, allowing a one-step procedure without the need for accessory exchange, has dramatically revolutionized the approach to PFC treatment [27, 28]. These stents have a large luminal diameter of up to 20 mm, which allows for drainage and the performance of DEN, when needed, and are technically easier to insert compared with double-pigtail plastic stents (DPPSs).

In our study, agreement among the 12 endosonographers for all the above-mentioned variables was only poor or moderate. The moderate agreement on PFC type in the pool of reviewed videos is surprising and might be explained by the presence of PFCs with minimal amounts of necrotic debris, which could be classified as WON by some of observers and as a PPC by others. The poor agreement on the estimation of necrotic content percentage reflects the subjectivity and lack of standardization for the assessment of this parameter, which has a major impact on management decisions. From a Dutch multicenter randomized controlled study, which compared endoscopic and surgical step-up approaches to pancreatic necrosis among 63 patients, in which drainage was accomplished using two 7-Fr DPPSs and one 8.5-Fr nasocystic catheter in the endoscopic arm, the mean rate of necrotic content was less

than 30% in 51% of the patients [25]. However, DEN was deemed necessary in a total of 57% of patients, thereby suggesting the possibility of underestimation of necrotic content in some cases. Conversely, Bang et al. compared DPPSs versus LAMSs for WON treatment [29], and DPPSs proved to be noninferior to LAMSs despite a higher mean necrotic content compared with the LAMS arm (45.3% vs. 40.3%). Studies comparing the assessment of necrotic content by EUS and MRI are needed to define objective EUS criteria for the quantification of necrotic content.

The agreement on MPD visualization was poor, while the evaluation of MPD–PFC communication was rated as substantial only because it was not evaluable in the large majority of cases (81%). By EUS, both parameters are better assessed 6–8 weeks after therapeutic intervention, when the collection has resolved. As reported by Bang et al., EUS identification of disconnected pancreatic duct syndrome is highly correlated with the findings at CT and endoscopic retrograde pancreatography [30]. Its detection is relevant because transmural placement of DPPSs, which remain in place indefinitely, can prevent collection recurrence.

A substantial agreement was found regarding the choice of the stent (DPPSs or LAMSs) to place, a result that does not reflect actual clinical practice, where the selection of one or other stent type is very operator-dependent. Various meta-analyses, mostly based on retrospective studies, have favored LAMSs over DPPSs for both PPCs and WONs [31–33]. However, as discussed before, the only available RCT performed in patients with WON demonstrated non-inferiority of DPPSs over LAMSs, therefore bringing into question the conclusions of the previous meta-analyses. Recommendations from various Endoscopy Societies have also expressed different opinions on this topic. The European Society for Gastrointestinal Endoscopy guidelines do not clarify the stent to be used in patients with WON [34], while the recently published clinical practice guidelines by the American Gastroenterological Association favor the use of LAMSs when necrosis is present [35]. Efforts should also be made to clarify this matter for patients with PPCs where LAMSs might not be cost-effective, favoring less-expensive DPPSs.

Finally, moderate and substantial agreements were scored for the need for and timing of DEN, respectively. The requirement for DEN can be difficult to assess without knowledge of the patient's clinical status. On the other hand, a high necrotic content might predict the need for DEN, even though no study has established the degree of solid debris for which DEN should be performed upfront. Once the need for DEN has been determined, the choice between performing it at the time of the drainage procedure or at a subsequent date becomes a clinical decision. A recent retrospective study on 271 patients with WON reported a reduction in the number of DEN sessions when the first session was performed following the initial drainage procedure, a result that needs to be replicated in a prospective multicenter study before becoming the standard of care [36]. It is possible that agreement/disagreement on both the stent to be placed and the need for and timing of DEN

might have reflected local practice, rather than information obtained from the rated videos.

Our study has several limitations. First, the reviewer findings were not compared with a reference standard. However, our study was not designed to define whether the reviewers were able to reach the “correct” diagnosis, but to evaluate if visual information contained in the EUS videos allowed for a reproducible assessment of PFCs without the aid of other clinical information. This may not reflect real-life conditions but our results, though far from definitive, provide insight into an unexplored field and can be used as the background to stimulate new research.

Second, we did not perform any sample size calculations and the confidence limits for the kappa values for each pair of raters are wide. Nonetheless, our study with 12 observers and 50 cases produced sufficiently narrow CIs around the overall agreement estimates for the detection of clinically relevant variations. Third, our study could have selection bias. It was possible to include only a limited number of patients with PFCs. Therefore, the study population was not able to reflect the entire spectrum of PFC lesions.

Fourth, the reviewers were expert endosonographers from high volume centers, so the generalizability of these results to less experienced endosonographers working in the community is questionable. Fifth, the results from our study also reflect the paradox associated with the kappa statistic, wherein an item or category demonstrates high percentage agreement but a low kappa coefficient. This inherent limitation of kappa is well-established and acknowledged [16]. Therefore, the percentage agreement and AC1 coefficient may be more appropriate measures of reliability.

Finally, we did not perform a “definition phase,” in which consensus criteria were discussed/defined. However, the aim of this work was to evaluate the features of PFCs at EUS that enable the differentiation of different PFC types and the treatment decision made as routinely applied by experienced endosonographers working in independent settings/institutions. For this very reason, it was decided that all endosonographers should evaluate all videos blindly and independently, in the absence of a preformed agreement and consequent bias, with full understanding of the clinical aim of the present study, in particular the fact that they were unaware of the patients' clinical features.

In conclusion, our study suggests that interobserver agreement among experienced endosonographers with regard to the morphological assessment of PFCs is less than satisfactory, while a more homogeneous agreement on therapeutic decisions was found. Future prospective comparative studies will be needed in an attempt to achieve standardization of the diagnostic criteria and therapeutic approaches to these complex clinical entities.

## Competing interests

C. Fabbri has collaborations with Boston Scientific and Steris.

## References

- [1] Working Group IAP/APA Acute Pancreatitis Guidelines. IAP/APA evidence-based guidelines for the management of acute pancreatitis. *Pancreatology* 2013; 13: e1–e15
- [2] Crockett SD, Wani S, Gardner TB et al. American Gastroenterological Association Institute Guideline on initial management of acute pancreatitis. *Gastroenterology* 2018; 154: 1096–1101
- [3] Banks PA, Bollen TL, Dervenis C et al. Classification of acute pancreatitis – 2012; revision of the Atlanta classification and definitions by international consensus. *Gut* 2013; 62: 102–111
- [4] Lankisch PG, Apte M, Banks PA. Acute pancreatitis. *Lancet* 2015; 386: 85–96
- [5] Baron TH, Thaggard WG, Morgan DE et al. Endoscopic therapy for organized pancreatic necrosis. *Gastroenterology* 1996; 111: 755–764
- [6] Kamal A, Singh VK, Akshintala VS et al. CT and MRI assessment of symptomatic organized pancreatic fluid collections and pancreatic duct disruption: an interreader variability study using the revised Atlanta classification 2012. *Abdom Imaging* 2015; 40: 1608–1616
- [7] Sternby H, Verdonk RC, Aguilar G et al. Significant inter-observer variation in the diagnosis of extrapancreatic necrosis and type of pancreatic collections in acute pancreatitis - An international multicenter evaluation of the revised Atlanta classification. *Pancreatology* 2016; 16: 791–797
- [8] Giovannini M. Endoscopic ultrasound-guided drainage of pancreatic fluid collections. *Gastrointest Endosc Clin N Am* 2018; 28: 157–169
- [9] Rimbaş M, Pizzicanella M, Larghi A. EUS-guided treatment of WON using lumen-apposing metal stents: protocol standardisation based on the occurrence of natural healing processes. *Gut* 2019; 68: 1334–1335
- [10] Baron TH, Harewood GC, Morgan DE et al. Outcome differences after endoscopic drainage of pancreatic necrosis, acute pancreatic pseudocysts, and chronic pancreatic pseudocysts. *Gastrointest Endosc* 2002; 56: 7–17
- [11] Tyberg A, Karia K, Gabr M et al. Management of pancreatic fluid collections: A comprehensive review of the literature. *World J Gastroenterol* 2016; 22: 2256–2270
- [12] Rerknimitr R. Endoscopic transmural necrosectomy: timing, indications, and methods. *Clin Endosc* 2020; 53: 49–53
- [13] Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol* 2008; 61: 29–48
- [14] Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990; 43: 543–549
- [15] Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol* 1993; 46: 423–429
- [16] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174
- [17] Yip HC, Teoh AYB. Endoscopic management of peri-pancreatic fluid collections. *Gut Liver* 2017; 11: 604–611
- [18] Bouwense SA, van Brunschot S, van Santvoort HC et al. Describing peripancreatic collections according to the revised Atlanta classification of acute pancreatitis: an international interobserver agreement study. *Pancreas* 2017; 46: 850–857
- [19] Sternby H, Verdonk RC, Aguilar G et al. Significant inter-observer variation in the diagnosis of extrapancreatic necrosis and type of pancreatic collections in acute pancreatitis - An international multicenter evaluation of the revised Atlanta classification. *Pancreatology* 2016; 16: 791–797
- [20] Badat N, Millet I, Corno L et al. Revised Atlanta classification for CT pancreatic and peripancreatic collections in the first month of acute pancreatitis: interobserver agreement. *Eur Radiol* 2019; 29: 2302–2310
- [21] Morgan DE, Baron TH, Smith JK et al. Pancreatic fluid collections prior to intervention: evaluation with MR imaging compared with CT and US. *Radiology* 1997; 203: 773–778
- [22] Varadarajulu S, Bang JY, Sutton BS et al. Equal efficacy of endoscopic and surgical cystogastrostomy for pancreatic pseudocyst drainage in a randomized trial. *Gastroenterology* 2013; 145: 583–590
- [23] Akshintala VS, Saxena P, Zaheer A et al. A comparative evaluation of outcomes of endoscopic versus percutaneous drainage for symptomatic pancreatic pseudocysts. *Gastrointest Endosc* 2014; 79: 921–928
- [24] Bakkar OJ, van Santvoort HC, van Brunschot S et al. Endoscopic transgastric vs surgical necrosectomy for infected necrotizing pancreatitis: a randomized trial. *JAMA* 2012; 307: 1053–1061
- [25] van Brunschot S, van Grinsven J, van Santvoort HC et al. Endoscopic or surgical step-up approach for infected necrotising pancreatitis: a multicentre randomised trial. *Lancet* 2018; 391: 51–58
- [26] Rizzatti G, Rimbaş M, Larghi A. Which stent to use for the management of pancreatic pseudocysts? Time for randomized controlled studies *Endoscopy* 2020; 52: 321
- [27] Bang JY, Varadarajulu S. Lumen-apposing metal stents for endoscopic ultrasonography-guided interventions. *Dig Endosc* 2019; 31: 619–626
- [28] Giovannini M. Endoscopic ultrasound-guided drainage of pancreatic fluid collections. *Gastrointest Endosc Clin N Am* 2018; 28: 157–169
- [29] Bang JY, Navaneethan U, Hasan MK et al. Non-superiority of lumen-apposing metal stents over plastic stents for drainage of walled-off necrosis in a randomised trial. *Gut* 2019; 68: 1200–1209
- [30] Bang JY, Navaneethan U, Hasan MK et al. EUS correlates of disconnected pancreatic duct syndrome in walled-off necrosis. *Endosc Int Open* 2016; 4: E883–E889
- [31] Bazerbachi F, Sawas T, Vargas EJ et al. Metal stents versus plastic stents for the management of pancreatic walled-off necrosis: a systematic review and meta-analysis. *Gastrointest Endosc* 2018; 87: 30–42
- [32] Hammad T, Khan MA, Alastal Y et al. Efficacy and safety of lumen-apposing metal stents in management of pancreatic fluid collections: are they better than plastic stents? A systematic review and meta-analysis *Dig Dis Sci* 2018; 63: 289–301
- [33] Saunders R, Ramesh J, Cicconi S et al. A systematic review and meta-analysis of metal versus plastic stents for drainage of pancreatic fluid collections: metal stents are advantageous. *Surg Endosc* 2019; 33: 1412–1425
- [34] Arvanitakis M, Dumonceau JM, Albert J et al. Endoscopic management of acute necrotizing pancreatitis: European Society of Gastrointestinal Endoscopy (ESGE) evidence-based multidisciplinary guidelines. *Endoscopy* 2018; 50: 524–546
- [35] Baron TH, DiMaio CJ, Wang AY et al. American Gastroenterological Association Clinical Practice Update: Management of pancreatic necrosis. *Gastroenterology* 2020; 158: 67–75
- [36] Yan L, Dargan A, Nieto J et al. Direct endoscopic necrosectomy at the time of transmural stent placement results in earlier resolution of complex walled-off pancreatic necrosis: Results from a large multicenter United States trial. *Endosc Ultrasound* 2019; 8: 172–179