Defining the endoscopic ultrasound features of chronic pancreatitis in Asians: a multicenter validation study

GRAPHICAL ABSTRACT

Assesses inter-rater reliability in recognition of EUS features of chronic pancreatitis

17 EUS videos
11 criteria rated by 21 experts
6 criteria with better inter-rater reliability

Assesses 6 criteria identified in Phase I in a multicenter Asian case-control study

12 centers across Asia participated
132 chronic pancreatitis
152 controls
Absence/presence of 6 criteria assessed
Performance of 6 criteria: AUROC = 0.986

AEG criteria for chronic pancreatitis:
2 or more of the following:
Hyperchoic foci with shadowing
Lobularity with honeycombing
Cysts
Dilated main pancreatic duct
Dilated side branches
Calculi in the main pancreatic duct

Authors
Calvin Jianyi Koh1,2, Sundeep Lakhtakia3, Mitsuhiro Kida4, Cosmas Rinaldi A. Lesmana5, Tiing Leong Ang6, Charles Kieng Fong Vu7, Than Than Aye8, Sun Hwa Park9, Majid A. Almadi10, Charing Ching-Ning Chong11, Raymond Shing Yan Tang12, Xi Wu13, Ida Hilmi14, Vinay Dhir15, Nonthalee Pausawasdi16, Jahangeer Basha17, Zhen-Dong Jin18, Ai Ming Yang13, Anthony Yuen Bun Teoh11, Dong-Wan Seo9, Hsiu-Po Wang18, Khok Yu Ho1,2

Institutions
1 Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
2 Division of Gastroenterology and Hepatology, National University Hospital, Singapore, Singapore
3 Department of Gastroenterology, Asian Institute of Gastroenterology, Hyderabad, India
4 Department of Gastroenterology, Kitasato University Hospital, Sagamihara, Kanagawa, Japan
5 Digestive Disease & GI Oncology Center, Medistra Hospital, Jakarta, Indonesia
6 Department of Gastroenterology and Hepatology, Changi General Hospital, Singapore
7 Department of Gastroenterology and Hepatology, Tan Tock Seng Hospital, Singapore
8 Department of Gastroenterology, Sanpya General Hospital, Thingan Gyun, Myanmar
9 Department of Gastroenterology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea
10 Department of Medicine, King Khalid University Hospital, Riyadh, Saudi Arabia
11 Department of Surgery, Prince of Wales Hospital, The Chinese University of Hong Kong, Shatin, Hong Kong
12 Institute of Digestive Diseases, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China
13 Department of Gastroenterology, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China
14 Division of Gastroenterology, University of Malaya, Kuala Lumpur, Malaysia
15 Institute of Digestive and Liver Care, Global Hospital, Mumbai, India
16 Mahidol University Faculty of Medicine, Siriraj Hospital, Bangkok, Thailand
17 Department of Gastroenterology, The Second Military Medical University, Shanghai, China
18 Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

Original article
Assesses inter-rater reliability in recognition of EUS features of chronic pancreatitis

17 EUS videos
11 criteria rated by 21 experts
6 criteria with better inter-rater reliability

Part 1
Assesses 6 criteria identified in Phase I in a multicenter Asian case-control study

12 centers across Asia participated
132 chronic pancreatitis
152 controls
Absence/presence of 6 criteria assessed
Performance of 6 criteria: AUROC = 0.986

Part 2

AEG criteria for chronic pancreatitis:
2 or more of the following:
Hyperchoic foci with shadowing
Lobularity with honeycombing
Cysts
Dilated main pancreatic duct
Dilated side branches
Calculi in the main pancreatic duct

Koh Calvin Jianyi et al. EUS features of chronic pancreatitis in Asians ... Endoscopy | © 2020. The Author(s). Published online: 2020-07-10
Introduction

Chronic pancreatitis is defined by chronic, progressive pancreatic inflammation and scarring that irreversibly damages the pancreas and results in loss of exocrine and endocrine function [1]. As a condition of heterogeneous etiology [2], varying clinical presentations, and considerable differences in the clinical course, the diagnosis of chronic pancreatitis remains challenging. Various imaging modalities, such as computed tomography (CT), magnetic resonance imaging (MRI), and endoscopic ultrasound (EUS), are available to aid in imaging the structural and morphologic changes in chronic pancreatitis. International consensus guidelines have proposed an algorithmic approach to patients with clinical presentations suspicious for chronic pancreatitis [3].

EUS has particular utility in the diagnosis of chronic pancreatitis [4] as it can detect subtle pancreatic parenchymal and ductal changes, before even external cross-sectional imaging [5], and is considered the most sensitive test to diagnose the early stage of the disease [3]. One of the challenges in the use of EUS in the diagnosis of chronic pancreatitis is that some of the pathognomonic features identified may also be seen in other common conditions, such as obesity and diabetes [6], as well as being related to aging [7]. Furthermore, considerable variation in morphology in chronic pancreatitis limits the specificity of any one endosonographic feature. To address some of the heterogeneity, expert consensus standard criteria have been derived, the most commonly used being the Rosemont criteria [8]. Challenges remain with the utility and performance of these criteria in clinical cohorts [9].

The major drawback of the Rosemont criteria is their considerable interoperator variability [10,11]. In addition, the optimal number of criteria required for a diagnosis of chronic pancreatitis remains unclear [12]. Furthermore, many of the diagnostic criteria were derived solely in Western populations, and some differences in chronic pancreatitis patterns and pathology among Asians have been suggested [13], hence there is a need to review these criteria in a multinational Asian context. It is also important to evaluate and clarify these criteria in a real-world scenario.
Therefore, the aim of this study was to evaluate and prospectively validate EUS criteria for chronic pancreatitis in a multicenter cohort of Asian patients with the aim of deriving a set of criteria that have less interoperator variability and improved clinical utility.

Methods

The study was divided into two parts: the first part was conducted to derive the EUS features of chronic pancreatitis with adequate interoperator agreement; the second was to prospectively evaluate these features in a multicenter cross-sectional study and determine the optimal combination of features for the diagnosis of chronic pancreatitis.

Part I – selecting endosonographic features with good interobserver variability

A set of 17 EUS video clips were retrospectively recorded and edited in standardized format to demonstrate the EUS stations for the evaluation of the pancreas (stomach, duodenal bulb, and second part of the duodenum) and include endosonographic evaluation of the parenchymal and ductal features of the entire pancreas. Fifteen of these videos were obtained from Asian patients who were investigated for suspected chronic pancreatitis, and two were from Asian patients with a normal pancreas. The video clips were chosen for quality and clarity, and were anonymized and made available for online review and scoring.

A panel of experienced endosonographers from across Asia were invited to examine the videos at a live session, being blinded to the patient status and evaluating for the presence or absence of 11 features of chronic pancreatitis, based on literature review: hyperechoic foci with shadowing, non-shadowing hyperechoic foci, strands, lobularity with non-contiguous lobules, lobularity with honeycombing, cysts, dilated main pancreatic ducts (measured on the EUS video), irregular pancreatic duct contour, hyperechoic pancreatic duct wall, dilated pancreatic duct side branches, and calculi in the main pancreatic duct [8, 14–18]. Standardized definitions of these features were further defined for clarity (Table 1) and selected illustrative images were also provided (Fig. 1).

Endosonographers were asked to independently score each video, marking if these features were present or absent. Interobserver agreement in recognizing the various diagnostic EUS features was determined by computing the kappa statistic, as described by Landis and Koch [19]. A perfect interobserver reliability is denoted by a kappa of 1, while values > 0.80 denote excellent reliability, 0.80–0.60 good reliability, 0.59–0.40 fair reliability, and < 0.40 poor reliability. Prior to the study, the participating endosonographers were trained to recognize these features with eight EUS training videos (six of chronic pancreatitis and two of normal pancreas).

For the purpose of this study, the six best performing features with good interobserver agreement and with kappa values of at least 0.6 were then selected to be included in the validation in part II of the study.

Part II – validation of the selected endosonographic features

A multicenter study was conducted prospectively to independently validate the use of these EUS features, with a cross-sectional design [20] being used to enroll subjects, both cases as well as normal controls.

Trained endosonographers at participating centers across Asia were acquainted in the use of a preliminary set of criteria comprising six selected diagnostic EUS features of chronic pancreatitis.

Firstly, eight de-identified EUS training videos (six of chronic pancreatitis and two of normal pancreas) were constructed. Each center study lead then enrolled experienced endosonographers and reviewed the eight selected videos with them to ensure standardization of the endosonographic features.
The check-back process was implemented where the same eight videos (with diagnostic features identified and labelled) was presented back to the endosonographers and competency in recognizing all six diagnostic EUS features was confirmed prior to study participation.

Participating endosonographers then enrolled cases of chronic pancreatitis and normal controls in each of their centers. For each EUS, the presence or absence of the previously derived six diagnostic EUS features was recorded.

As a quality-control measure, all EUS procedural videos were recorded and the de-identified set of EUS videos was subsequently assessed by an expert review committee (from a different study center) for verification of these features.

**Inclusion and exclusion criteria**

**Chronic pancreatitis cases**

To be included as a chronic pancreatitis case for this study, subjects had to have had a definitive diagnosis of chronic pancreatitis satisfying either reference criteria (i), (ii), or (iii) listed below (tabulated in Table 1s, see online-only Supplementary material):
(i) definitive imaging findings characteristic of chronic pancreatitis on CT/magnetic resonance cholangiopancreatography (MRCP), as per the American Pancreatic Association [18]
(ii) endoscopic retrograde cholangiopancreatography (ERCP) findings characteristic of chronic pancreatitis, as per the Japanese Pancreatic Society [17]
(iii) definitive histological findings characteristic of chronic pancreatitis, as per the Japanese Pancreatic Society [17].

Normal pancreas controls
Subjects with normal pancreas based on CT, MRI, and/or ERCP were recruited as controls. These subjects had to have no hepatobiliary disease, including pancreatitis. Other exclusion criteria included diabetes mellitus, a history of heavy alcohol use, family history of pancreatic disease, and history of gallstone or renal stone disease. They underwent EUS because of non-hepatobiliary indications, such as evaluation of subepithelial lesions and mediastinal masses.

An online secure clinical database was established for data coordination. Each study center enrolled subjects on this database, with the presence of each of the six features being recorded, together with patient demographics and characteristics.

Outcome measures
The study outcome measures included the sensitivity, specificity, and accuracy of the use of the six EUS criteria for the diagnosis of chronic pancreatitis in the multicenter Asian population. Both positive predictive value (PPV) and negative predictive value (NPV) were assessed and the performance of the combination of factors was evaluated.

Ethics
The study was endorsed by the individual center’s ethics review committee, as well as the National Health Group Domain Specific Review Board (NHG DSRB 2014/01271-AMD0005) from Singapore.

Statistical methods
For the sample size calculation for Part I of the study, we determined at least 21 experts were needed to each rate at least 17 videos, following methods described for statistical calculations described by Walter et al [21]. For Part II, we estimated for a difference in risk factor prevalence of 0.2 from existing data, we would need to enroll at least 62 cases and 62 controls for a power of 0.8 with a type I error probability of 0.05. Using multiple centers allowed these numbers to be accrued over a shorter period of time.

All data were collated from the individual study centers. Categorical variables were assessed using chi-squared tests. All data were analyzed in SPSS Statistics, version 21.0 (IBM Corp., Armonk, New York, USA).

Results
For Part I, a total of 21 expert endosonographers from Asia participated in the study. Six respondents completed ratings of only some of the videos and were therefore excluded, leaving 15 who rated the 17 EUS imaging videos of the pancreas for the 11 internationally used diagnostic EUS features of chronic pancreatitis, as described in Table 1. The respondents, who comprised experts from Hong Kong, India, Indonesia, Malaysia, Singapore, and Thailand, among others, had a mean experience of 1913 EUS procedures.

The overall interobserver agreement on the presence/absence of the 11 diagnostic EUS features of chronic pancreatitis was below expectation (kappa value 0.54, range 0.14–0.90). The results of the interobserver agreement for the individual features are shown in Table 2. There was no significant difference in interobserver reliability between respondents with more (completed > 1000 procedures) and less (completed < 1000 procedures) experience in the use of EUS (P = 0.56).

Considerable heterogeneity between observers was noted for some of the EUS features, for example hyperechoic duct wall had an interobserver agreement coefficient of 0.135. However, of the 11 EUS features included in the study, they were six diagnostic EUS features that were more consistently recognized by the participating endosonographers, with a combined kappa value for these six features of 0.73, and individual kappa values ranging from 0.60 to 0.90. Two criteria (lobulation with honeycombing and pancreatic duct irregularity) had a kappa of 0.60, but given the prior decision to select six criteria, lobulation with honeycombing was chosen because pancreatic duct irregularity would have overlapped somewhat with dilated main pancreatic duct. These six EUS features with good interobserver agreement were used to form a preliminary set of diagnostic criteria and included: hyperechoic foci with shadowing, lobularity with honeycombing, cysts, dilated main pancreatic duct, dilated side branches, and calculi in the main pancreatic duct.

For Part II, 284 subjects (132 cases, 152 controls) were enrolled from 12 centers. The subjects’ ethnic origins were: Indian 112 (39.4%), Japanese 62 (21.8%), Chinese 42 (14.8%), Malay 4

<p>| Table 2 Interobserver agreement for the individual EUS features. |
|-------------------|-----------------|-------------------|</p>
<table>
<thead>
<tr>
<th>EUS feature</th>
<th>Kappa</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperechoic foci with shadowing</td>
<td>0.64</td>
<td>0.48–0.80</td>
</tr>
<tr>
<td>Non-shadowing hyperechoic foci</td>
<td>0.23</td>
<td>0.07–0.39</td>
</tr>
<tr>
<td>Strands</td>
<td>0.44</td>
<td>0.18–0.71</td>
</tr>
<tr>
<td>Lobularity with non-contiguous lobules</td>
<td>0.34</td>
<td>0.15–0.53</td>
</tr>
<tr>
<td>Lobularity with honeycombing</td>
<td>0.60</td>
<td>0.33–0.86</td>
</tr>
<tr>
<td>Cysts</td>
<td>0.90</td>
<td>0.82–0.99</td>
</tr>
<tr>
<td>Dilated main pancreatic duct</td>
<td>0.61</td>
<td>0.43–0.83</td>
</tr>
<tr>
<td>Irregular pancreatic duct</td>
<td>0.60</td>
<td>0.42–0.77</td>
</tr>
<tr>
<td>Hyperechoic duct wall</td>
<td>0.14</td>
<td>0.01–0.29</td>
</tr>
<tr>
<td>Calculi in the main pancreatic duct</td>
<td>0.82</td>
<td>0.72–0.92</td>
</tr>
</tbody>
</table>
(1.4 %), and other 64 (22.5 %). There were 60.9 % men in this cohort. Age was recorded in bands (<18y, 18 – 30y, 31 – 50y and >50y); the median age group was 51 years and above, with there being participants in all age groups. Compared with the control group, the chronic pancreatitis group tended to be younger, with 55.3 % below 50 years of age compared with 36.2 % below 50 years of age among the controls (P< 0.002). There was also a significant association with male sex among the cases, with 79.5 % of the cases being men compared with 44.7 % in the control group (P<0.001). For the chronic pancreatitis subjects, 66.7 % had a history of ethanol use, 55.6 % smoking, and 53.3 % had additional hepatobiliary diseases.

Most of the case subjects had a diagnosis of chronic pancreatitis based on CT (80.0 %) and/or MRI criteria (57.8 %). The exact distribution of these modalities (CT, MRI, ERCP, or pathology) among the cases is illustrated in Fig. 1s.

The individual performance characteristics of each of the six selected diagnostic EUS features: hyperechoic foci with shadowing, lobularity with honeycombing, cysts, dilated main pancreatic duct, dilated side branches, and calculi in the main pancreatic duct are shown in Table 3. All six features were shown to be useful in distinguishing cases from controls, with accuracy ranging from 63.4 % to 89.1 %. Using the receiver operating curve (ROC), two or more of these six EUS features accurately defined chronic pancreatitis (sensitivity 94.7 %, specificity 98.0 %), with an area under the curve (AUROC) of 0.986 (Fig. 2).

Discussion

Although EUS features have been described in chronic pancreatitis, interoperator agreement remains a challenge for the EUS diagnosis of chronic pancreatitis [20]. There is also a wide variation among operators in the type and number of features required for the diagnosis of chronic pancreatitis. For instance, a questionnaire survey of 21 EUS experts from 10 countries in Asia found a divergence of opinion about the type and number of features required for a diagnosis of chronic pancreatitis [22]. The survey also found a tendency towards overdiagnosis of chronic pancreatitis, based on the currently used Rosemont and International EUS criteria for the diagnosis of chronic pancreatitis. In the survey, nine experts (42.5 %) reported observing at least one abnormal feature in asymptomatic pancreatic EUS examinations, while six experts (28.5 %) reported observing up to three abnormal features during examinations of the pancreas in asymptomatic individuals [22]. It is against this
background of issues that this study aimed to address the EUS diagnostic features of chronic pancreatitis.

To address the issue of interobserver variability in the use of EUS features to diagnose chronic pancreatitis, we asked a group of Asian EUS experts to select features with good interobserver agreement, which were then validated in a multicenter cross-sectional study of Asian patients. This study design differs from previously published studies that either evaluated the interobserver agreement among many endosonographers in a cross-sectional design [8,23,24] or evaluated the utility of the EUS findings in a cohort, but with the interobserver agreement between only two or three endosonographers [10,25].

To address the issue of ease of adoption of these EUS features, we used a two-part process to derive a set of six EUS criteria, for which a cutoff of two or more features corresponded to a sensitivity of 94.7% and a specificity of 98.0% in this cohort. Limiting the existing features to those with high agreement will make the assessment of chronic pancreatitis simple and easy to use, and is likely to improve the overall reproducibility, compared with the Rosemont criteria.

Diagnosing chronic pancreatitis is particularly important for the patient as this disease has implications in terms of quality of life, as well as serious long-term sequelae, such as pancreatic cancer. As chronic pancreatitis is a risk factor for pancreatic adenocarcinoma, there is a role for surveying this cohort of patients. In this regard, EUS may be a helpful surveillance modality, as it is thought to be the most reliable procedure for detecting malignancy.

There are several limitations in this study. Firstly, the cross-sectional design of the study allows a higher proportion of chronic pancreatitis patients to be enrolled than would be found in typical EUS practice. This should not be an issue as the intent of this study was to investigate the diagnostic features of chronic pancreatitis, and not the prevalence of chronic pancreatitis.

Secondly, reflective of real-world experience, the control group was older than the chronic pancreatitis group. It could be argued that the age difference between the case and control groups may skew the interpretation, because age-related ductal and parenchymal changes might mimic those of chronic pancreatitis [7]. In the present study, however, despite the presence of a greater number of older patients in the control group, the AUROC for two or more features remains good at 0.986.

Thirdly, we did not include subjects with early chronic pancreatitis, where the subjects might have had subtle changes demonstrable only on EUS, before being demonstrable as CT or MRI changes [5]. Early chronic pancreatitis could not be addressed in this study as the definition used for the diagnosis of chronic pancreatitis in this cross-sectional study was a combination of compatible clinical presentation and radiological features, which typically diagnoses more established chronic pancreatitis. Histology was not mandatory for the diagnosis of chronic pancreatitis in this study as it is not practical to perform biopsies or surgery for patients with chronic pancreatitis without a clear indication. As there is no single gold standard for chronic pancreatitis, the present study chose to adopt a priori agreed definitions based on either imaging and/or histologic criteria. In any case, this might not be an issue, as the main focus of this study was interobserver agreement amongst the experts and the derivation of the best EUS features for adoption into the clinical practice. We hope that the EUS criteria proposed from this study will form the basis for further prospective evaluation via a vis other imaging or histologic modalities and might have a future role in detecting early chronic pancreatitis.

Finally, it is well known that many imaging features of chronic pancreatitis are not pathognomonic of chronic pancreatitis. Pancreatic cystic lesions, such as pseudocysts, intraductal papillary mucinous neoplasms, or mucinous cystadenomas, may potentially be falsely labelled should there be other concomitant criteria. Furthermore, three of the features described above, dilated pancreatic duct, dilated side branches, and cysts, can also be found with an obstructing pancreatic mass lesion. However, this should not be an issue as the same imaging modality would be expected to pick up the mass lesion.

One other point of note is that EUS is thought to be more sensitive than CT and MRI in chronic pancreatitis, particularly in detecting early morphologic changes. However, these changes might also correspondingly have high interindividual variability. For instance, in our study, lobularity without shadowing had a low kappa of 0.22, but this finding could potentially precede lobularity with shadowing. On the other hand, having “standard” criteria that EUS operators disagree on would be a superfluous exercise. This reflects the clinical equipoise that many criteria face in balancing the need for a sensitive test and a reproducible one.

Patient selection remains a methodological limitation of this cross-sectional study, as it remains to be seen if these criteria hold true in a separate unselected population. However, these patients were drawn from the same patient population as we would expect an Asian EUS center to encounter.

This study has numerous strengths. First, a relatively large number of patients were enrolled prospectively from multiple centers in Asia. Second, to ensure stringent standardization of the criteria used, we conducted upfront standardization training of the experts, followed by recruitment of patients into this study. Third, a high number of endosonographers (n=17) took part in the agreement, compared with other reported studies that have typically had between two or three observers participating in the study. Fourth, all EUS videos were recorded and later audited by an independent committee to ensure consistency with the reported features.

While there is some suggestion of racial heterogeneity in chronic pancreatitis [13,26], there is a paucity of data to support the hypothesis that the endosonographic appearances between Asians and Caucasians differ significantly. Nevertheless, it is noted that differences in ethnic origin exist between other reported interobserver agreements [10,25] and those in this study. This could be related to differences in genetic risk factors, for example differences in the population prevalence of SPINK or CFTR mutations [27,28].

In summary, this is the first multicenter Asian EUS study to provide a set of EUS features diagnostic of chronic pancreatitis that are relatively objective, reproducible, and simple to use. It
is anticipated that these features will be helpful in clinical practice and further research.

Acknowledgments
This study was supported by an unrestricted grant from the Asian EUS Group.

Competing interests
The authors declare that they have no conflict of interest.

References