Ultrasound in Med. & Biol., Vol. 43, No. 4, pp. 735–743, 2017

0301-5629

http://dx.doi.org/10.1016/j.ultrasmedbio.2016.11.020

● *Original Contribution*

DIAGNOSTIC ACCURACY OF TRANSABDOMINAL ULTRASOUND IN CHRONIC PANCREATITIS

Trond Engjom,*† Dag A. Sangnes,‡ Roald F. Havre,†‡ Friedemann Erchinger,*§ Khanh Do-Cong Pham,‡ Ingfrid S. Haldorsen,*¶ Odd H. Gilja,*† and Georg Dimcevski*‡

*Department of Clinical Medicine, University of Bergen, Bergen, Norway; †National Centre for Ultrasound in Gastroenterology, Department of Medicine, Haukeland University Hospital, Bergen, Norway; ‡Department of Medicine, Haukeland University Hospital, Bergen, Norway; §Department of Medicine, Voss Hospital, Voss, Norway; and ¶Department of Radiology, Haukeland University Hospital, Bergen, Norway

(*Received* 5 *September* 2016; *revised* 21 *November* 2016; *in final form* 27 *November* 2016)

Abstract—The performance of transabdominal ultrasound (US) in chronic pancreatitis (CP) following the advances in US technology made during recent decades has not been explored. Our aim in this prospective study was to evaluate the diagnostic accuracy of modern abdominal US compared with the Mayo score in CP. One hundred thirty-four patients referred for suspected CP were included in the study. Fifty-four patients were assigned the diagnosis CP. After inclusion, transabdominal US was performed. Ductal features (calculi, dilations and caliber variations, side-branch dilations and hyper-echoic duct wall margins) and parenchymal features (calcifications, cysts, hyper-echoic foci, stranding, lobulation and honeycombing) were recorded. Features were counted and scored according to a weighting system defined at the international consensus meeting in Rosemont, Illinois (Rosemont score). Diagnostic performance indices (95% confidence interval) of US were calculated: The unweighted count of features had a sensitivity of 0.69 (0.54–0.80) and specificity of 0.97 (0.90–1). The Rosemont score had a sensitivity of 0.81 (0.69–0.91) and specificity of 0.97 (0.90–1). Exocrine pancreatic failure was most pronounced in Rosemont groups I and II ($p < 0.001$). We conclude that using both unweighted and weighted scores, the diagnostic accuracy of modern transabdominal US is good. The extent of pancreatic changes detected by the method is correlated with exocrine pancreatic function. (E-mail: Trond.engjom@helse-bergen.no)

Key Words: **Chronic pancreatitis, Transabdominal Ultrasound, Mayo score, Diagnostic accuracy, Sensitivity, Specificity, Exocrine pancreatic function.**

INTRODUCTION

Modern guidelines and reviews on chronic pancreatitis (CP) include transabdominal ultrasound (US) as a relevant first-line imaging method in the evaluation of the pancreas (Conwell et al. 2014; Drewes et al. 2015; Erchinger et al. 2011; Forsmark 2013; Martinez et al. 2013; Mayerle et al. 2013). No prospective studies have reported the diagnostic performance using the most recent US technology. Endoscopic US (EUS) with characterization of ductal and parenchymal changes with or without the aid of weighted scores like the "Rosemont score" (Catalano et al. 2009) is presently the gold standard for sonographic imaging of the pancreas in CP. The Rosemont score was established through an international consensus meeting in Rosemont, Illinois (April 13–14, 2007). The result of the conference was agreement on definitions and weighting for five parenchymal and five ductal features of CP. The diagnostic quality of EUS compared with other imaging modalities like computed tomography (CT) and magnetic resonance imaging (MRI) is regarded as good (Catalano et al. 1998; Kalmin et al. 2011; Luetmer et al. 1989; Manfredi et al. 2000; Pungpapong et al. 2007; Wiersema and Wiersema, 1995). The severity of pathologic EUS changes correlates to histopathologic

Address correspondence to: Trond Engjom, Haukeland University Hospital, 5021 Bergen, Norway. E-mail: Trond.engjom@helse-bergen.no

Conflict of interest disclosure: All authors received salary from Haukeland University Hospital. T.E., G.D., O.H.G. and I.H. have received salary from the University of Bergen. R.F.H. and K.D.C.P. have received a salary from Helse Vest, the regional hospital consortium. T.E., R.F.H., F.E. and K.D.C.P. have received limited Norwegian Gastroenterology Association grants related to the actual research protocol.

findings and extent of exocrine dysfunction (Albashir et al. 2010; Catalano et al. 1998; Chong et al. 2007; Kalmin et al. 2011; Luetmer et al. 1989; Manfredi et al. 2000; Pungpapong et al. 2007; Wiersema and Wiersema, 1995). However, low inter-rater agreement, especially for the minor EUS-detected features in the Rosemont score, has been reported (Del et al. 2012; Kalmin et al. 2011; Stevens et al. 2010). The invasive character of the EUS examination and the long learning curve for operators are drawbacks of the method.

Transabdominal US is limited by longer sound wave distance to the pancreas compared with EUS. Bowel gas, obesity and individual patient factors reduce the quality of US images. Previous studies have explored the visualization of morphologic changes in CP by transabdominal US and reported varying diagnostic accuracy with a sensitivity of 70% to 80% (Bolondi et al. 1989; Foley et al. 1980; Gebel et al. 1985; Ikeda et al. 1994; Martinez-Noguera and D'Onofrio 2007). One large study reported sensitivities for US >85% for the features calcifications, pancreatic duct dilations and cysts, compared with CT (Ikeda et al. 1994). In recent decades, there has been an overall advance in US technology (Dimcevski et al. 2013; Whitsett 2009). Re-evaluation using modern technology is warranted (Conwell et al. 2014). The advances in software and screen technology and the introduction of harmonic imaging reduce noise and improve image resolution in all US systems. Modern US probes with dynamic frequencies, better depth-focusing technology and better high-frequency probes also add significantly to noise reduction and improved resolution. Especially in the characterization of calcifications, where the evaluation of shadowing is highly significant, overall reduction of random noise features is important. In the detection of minor changes in pancreatic ducts and parenchyma, we postulate that improved image resolution probably increases the sensitivity of the method compared with earlier studies.

In this prospective observational cohort study, our aim was to evaluate the diagnostic accuracy of features detected by a high-end transabdominal US scanner compared with the CP diagnosis defined by a diagnostic score combining clinical and imaging features, the Mayo score (Layer et al. 1994). Furthermore, US findings were evaluated according to criteria from the Rosemont consensus to address possible improvements in diagnostic accuracy from a weighted score.

METHODS

Participants

One hundred forty-one eligible participants were recruited among patients referred to our outpatient clinic with suspected CP. Reasons for referral were presenting symptoms or classic CP characteristics based on previous diagnostic imaging. Patients who did not fulfill the protocol for an adequate Mayo score were not included. We excluded patients for whom US visualization of the pancreas was insufficient because of obesity, repeated overlying bowel air or other factors. The final number of patients was 124. Patient characteristics are summarized in Table 1.

Diagnostic standards

Relevant clinical data and imaging reports from EUS (n = 74), pancreatic CT (n = 111) and/or pancreatic MRI (n = 22) were retrieved from the medical records or obtained as necessary for the clinical workup. All EUS examinations were performed according to a modern standard using the linear EG-3870 UTK or radial EG-3670 URK scope from Pentax Medical (Pentax Europe, Hamburg, Germany) by two experienced operators whose focus is pancreatic EUS. CT and MRI scans were performed according to standardized and similar protocols at different hospital scanners as part of the routine diagnostic workup. All CT protocols included intravenous contrast.

Mayo score

On the basis of clinical information, endocrine and exocrine failure, and findings on CT, MRI and EUS, patients were diagnosed according to a modified Mayo score (also named Layer score) (Erchinger et al. 2013; Layer et al. 1994) (Table 2). Patients with a Mayo score ≥4 were diagnosed with CP.

Sonographic examination

After overnight fasting, patients were examined with US while in the supine or right lateral position, with the

Table 1. Demographic data, laboratory results and exocrine pancreatic function

	Chronic pancreatitis (n = 54)	Others (n = 70)	*p*
Age	59 (24–79)*	55 (16–79)	
Sex, female/male	29/25†	33/37	
Body mass index	22.7 (20.1–24.4)	24.2 (21.1–26.7)	<0.05
Smokers, including ex-smokers	45.3%	27.3%	<0.05
Alcohol units/week	0 (0–1)	1 (0–3)	<0.05
HbA1c, %	5.8 (5.4–6.6)	5.5 (5.1–5.7)	<0.05
Fecal elastase, μg/g	151 (16–458)	376 (189–508)	<0.05
Duodenal bicarbonate, mmol/L	71 (38–101)	109 (90–122)	<0.05

* Minimum–maximum.
† Median (interquartile range) unless otherwise stated. Age (min–max).

Table 2. Mayo score: Diagnostic score for chronic pancreatitis modified from Layer et al. (1994) *

Pancreatic calcifications or typical histologic findings	4 points
Moderate or marked morphologic changes on ultrasonography, computed tomography or EST	3 points
Definite morphologic changes on magnetic resonance imaging	3 points
Reduced exocrine pancreatic function by EST or fecal elastase 1 level	2 points
History of acute pancreatitis or upper abdominal pain	2 points
Diabetes mellitus or impaired glucose tolerance test	1 point

EST = endoscopic.
* The diagnosis requires 4 points.

transverse or oblique epigastric probe in the lateral/posterior left subcostal position. Each examination was performed by a single operator. The skilled operators (T.E., F.E., G.D., 5–25 y of experience in pancreas scanning) knew the reason for referral, but were blinded to clinical data and results from earlier examinations. A GE Logic E9 scanner with a 1- to 5-MHz curvilinear probe was used. Whenever possible with respect to image depth and quality, the examination was supplemented by a 9-MHz linear probe (GE Medical Systems and Primary Care Diagnostics, Milwaukee, WI, USA). The default abdomen configuration of the scanner was used to acquire the images (CRA Probe—frequency: 4.0 MHz [CRA Probe] and 9.0 MHz [linear probe], frame rate: 15–22 f/s, dynamic range: 34, varying depth of scanning).

Complete US scanning of the pancreas was performed. The ductal and parenchymal features (Fig. 1) were recorded on a standardized form. The visibility of pancreatic head, body and tail was graded from 1 to 4 (1 = good, 2 = adequate, 3 = poor, 4 = not visible). The data were acquired from the segments of the pancreas with the best visualization.

Rosemont score

The weighted Rosemont score for EUS-detected features of CP comprises five parenchymal and five ductal features (Catalano et al. 2009). These features with definitions and weighting are described in Table 3. We applied the definitions and weighting unchanged from the Rosemont consensus on the detected transabdominal US features.

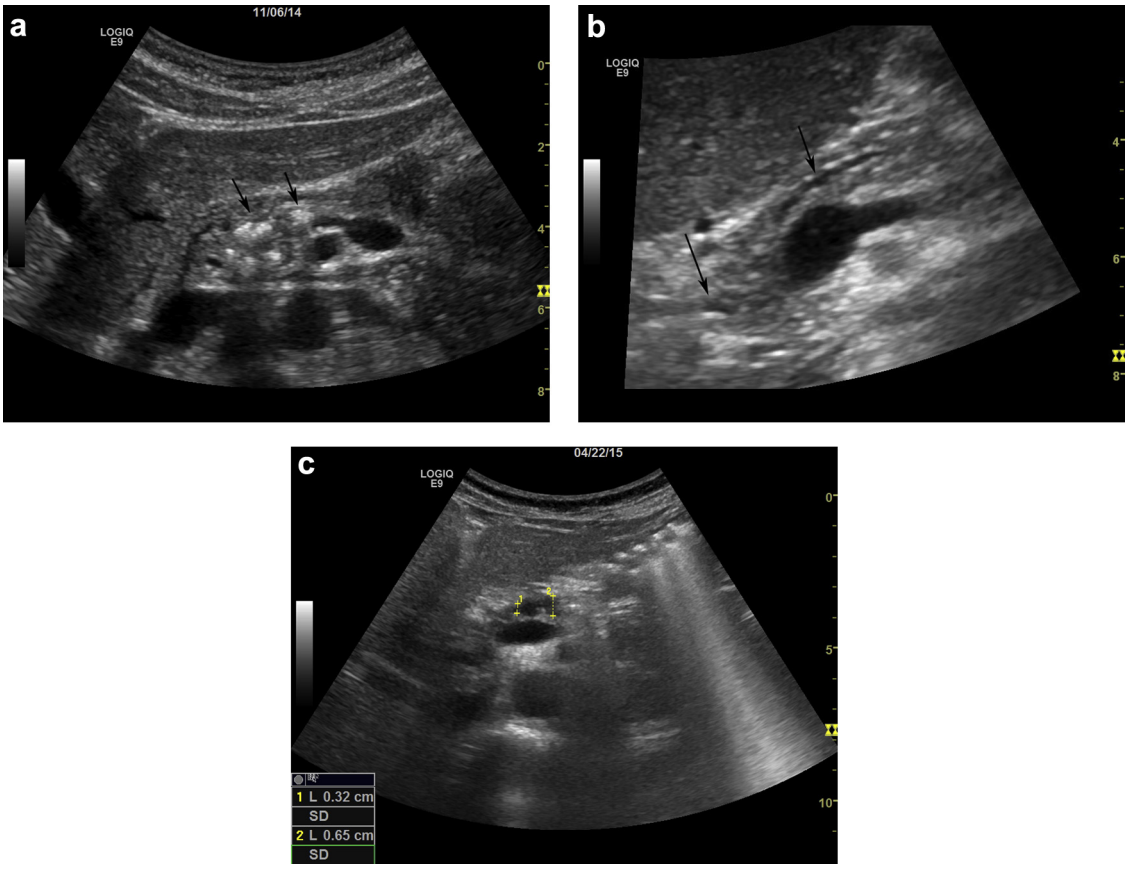

Fig. 1. Images revealing different ultrasound-detected features of chronic pancreatitis. (a) Ultrasound image of several large, shadowing calcifications in the head and body of the pancreas. (b) Ultrasound image revealing a large, shadowing calcification in the main pancreatic duct (*left arrow*), with dilated duct in pancreatic body (*right arrow*). (c) Ultrasound image revealing chronic dilation and caliber variations in the main pancreatic duct in the body of the pancreas.

Table 3. Parenchymal and ductal features of chronic pancreatitis in the Rosemont criteria

Rank	Definition	Major	Minor
Parenchymal features			
1. Hyper-echoic foci with shadowing	Echogenic structures >2 mm in length and width that produce a shadow; at least 3 are needed to be a marker of CP	Major A	
2. Lobularity			
A. Without honeycombing	Well circumscribed >5-mm structures with rims hyper-echoic relative to central areas; at least 3 lobules in the body or tail; not to be considered in the pancreatic head		Yes Major B
B. With honeycombing	When at least 3 of the lobules are contiguous, the feature is termed *honeycombing*		
3. Hyper-echoic foci without shadowing	Echogenic structures >2 mm in length/width; no shadow		Yes
4. Cysts	Anechoic, rounded/elliptic structures >2 mm in short axis		Yes
5. Stranding	Hyper-echoic lines >3 mm, seen in at least 2 image planes; at least 3 strands necessary; should be evaluated in body and tail and ventral pancreas		Yes
Ductal features			
1. MPD calculi	Echogenic structures with acoustic shadowing within MPD; can be considered in all segments	Major A	
2. Irregular MPD contour	A main duct that is uneven and ectatic in its course; should be assessed only from the pancreatic body and tail		Yes
3. Dilated side branches	Presence of 3 or more tubular anechoic structures each measuring >1 mm in width, communicating with the MPD; assessed only from the pancreatic body and tail		Yes
4. MPD dilation	MPD diameter >3.5 mm within the pancreatic head/body or >1.5 mm within the tail (no consensus)		Yes
5. Hyperechoic MPD margin	A relatively hyper-echoic duct wall found in >50% of the entire MPD in the body and tail; when imaged in a parallel or perpendicular orientation, both proximal and distal MPD borders must be hyper-echoic		Yes

Rosemont score for parenchymal and ductal features of chronic pancreatitis	
I. Consistent with CP	A. 1 major A feature + ≥3 minor features B. 1 major A feature + 1 major B feature C. 2 major A features
II. Suggestive of CP	A. 1 major A feature + <3 minor features B. 1 major B feature + ≥3 minor features C. ≥5 minor features (any)
III. Indeterminate for CP	A. 3 to 4 minor features, no major features B. Major B feature alone or with 3 minor features
IV. Normal	≤2 minor features, no major features

CP = chronic pancreatitis, MPD = main pancreatic duct.
Source: Modified from Catalano et al. (2009).

Classic scoring of sonographically detected features

The classic EUS assessment of sonographic features in CP is performed by counting the features detected in each patient. Traditionally, a cutoff of three to five features has been considered for the diagnosis of CP (Catalano et al. 1998; Iglesias-Garcia et al. 2015; Sahai et al. 1998; Stevens et al. 2010; Wiersema et al. 1993). We also performed an unweighted counting of the US-detected features. In our study, we chose to use three features as the cutoff for calculation of accuracy. Others have described the registration of nine EUS features (Stevens et al. 2010). For simplicity, we chose to use all 10 Rosemont EUS features as the basis for the counting.

Exocrine function testing

Pancreatic exocrine function was evaluated using a timed, secretin-stimulated endoscopic short test described elsewhere (Erchinger et al. 2013; Tjora et al. 2013). Patients were offered conscious sedation with intravenous midazolam during the test procedure. The peak bicarbonate concentration was measured; the cutoff for exocrine failure was defined as <80 mmol/L (Conwell et al. 2003; Erchinger et al. 2013; Stevens et al. 2008).

Fecal elastase 1 was analyzed with a commercial monoclonal analysis kit (ScheBo Biotech, Giessen, Germany). A fecal elastase 1 level <200 μg/mg was considered pathologic (Loser et al. 1996). Patients with both tests under threshold were considered exocrine insufficient according to the Mayo score.

Statistical analysis

Statistics were calculated using SPSS Version 22 (IBM, Armonk, NY, USA) and SigmaPlot 11 (Systat Software, San Jose, CA, USA). Normal distribution of the samples was tested with the Kolmogorov–Smirnov

test. Results are given as the median (95% confidence interval [CI] or interquartile [IQ] range). Comparisons between groups were made using Student's *t*-test or the Mann–Whitney *U*-test as appropriate. The level of statistical significance was set at 5%. Accuracy was calculated from receiver operating characteristic (ROC) curves. Inter-rater agreement for the detection of calcifications and cysts by US compared with CT was calculated as Cohen's κ. Agreement was defined according to Landis and Koch (1977): 0 = no agreement, 0–0.20 = slight agreement, 0.21–0.40 = fair agreement, 0.41–0.60 = moderate agreement, 0.61–0.80 = substantial agreement and 0.81–1 = almost perfect agreement. Clinically relevant agreement usually requires a value >0.5 (Altman 1997).

Ethics

The study was conducted in accordance with the Helsinki Declaration (World Medical Association General Assembly 2015) and received institutional review board approval from the Regional Committee for Ethics in Medical and Biologic Research, Western Norway (REK–Vest, Registration No. 2010/2857-7). The study is registered as a clinical trial: ClinicalTrials.gov Identifier: NCT01059669. All patients signed an informed consent. The protocol adheres to the STARD statement (Bossuyt et al. 2015).

RESULTS

Participants

Patient enrollment was conducted during the period January 2011 to June 2016. One hundred forty-one eligible patients were evaluated. Seven patients did not fulfill the protocol for an adequate Mayo score and were not included. Ten patients (7%) who had insufficient US visualization of the pancreas were excluded. A few patients were included despite missing values for one of the exocrine tests or missing demographics.

Accordingly, we included 124 patients with acceptable-quality US. Patient inclusion and exclusion are displayed in Figure 2. CP was diagnosed in 54 patients with a Mayo score ≥4. Seventy patients with varying diagnoses (recurrent acute pancreatitis, functional dyspepsia, bile stone disease or other cause of abdominal pain) were assigned a Mayo score <4. These patients were included as a control group. The Rosemont score from US categorized the patients into four groups: (I) consistent with CP (n = 31), (II) suggestive of CP (n = 14), (III) indeterminate for CP (n = 10) and (IV) normal pancreas (n = 69). The CP patients had lower body mass indexes and higher median HbA1c values; a larger percentage of CP patients was smokers and had reduced exocrine pancreatic function compared with the non-CP group. Median alcohol intake at the time of inclusion was lower in the CP than in the non-CP group. There were no differences between the groups with respect to age or sex. Other than some discomfort accompanying the endoscopic examinations, no adverse events were reported during any of the examination procedures.

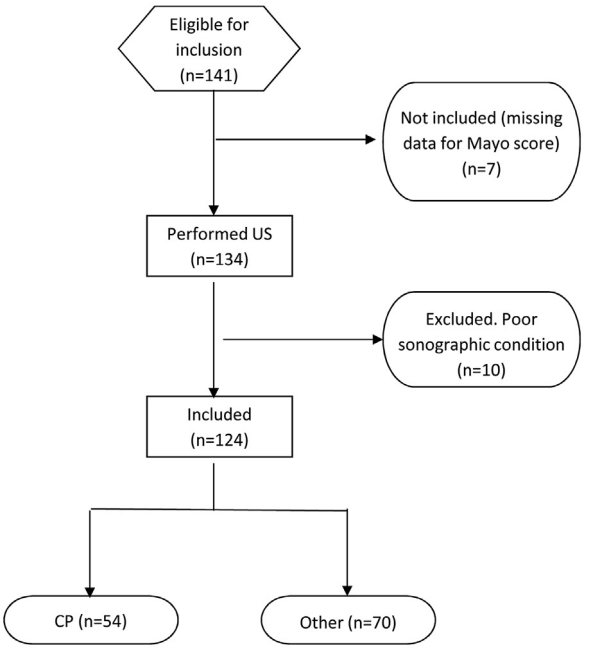

Fig. 2. Flowchart of patient enrollment. CP = chronic pancreatitis, US = ultrasound.

Visualization

Of the 134 total eligible patients who fulfilled the Mayo score, the whole pancreas could be visualized in 82 patients (61%) (visualization score ≤2 in all 3 segments), and sufficient visualization for inclusion was achieved in 124 patients (92%). The pancreatic tail was the part of the pancreas most frequently incompletely visualized (34 patients, 25%). In 10 patients, visualization of the entire pancreas was inadequate to determine a US score.

Visualization of calcifications and cysts by US

We calculated inter-rater agreement between US and CT for calcifications and cysts. The calculation was performed on the subgroup of 111 patients for whom CT scans were available. US and CT detection of calcifications were in almost perfect agreement (κ = 0.91). Agreement between US detection and CT detection of cysts was substantial (g = 0.69).

Diagnostic accuracy of transabdominal US

We calculated ROC curves for the two US scoring methods (Rosemont and classic scores) against the

diagnosis CP by Mayo score (Fig. 3). Cutoffs for the sonographic diagnosis were defined as a classic score of ≥3 and Rosemont categories I and II. Both scoring systems yielded very high accuracies of ≥0.95 (Table 4). There was somewhat better sensitivity and specificity for US in the Rosemont score compared with the classic score for the suggested cutoffs, but the difference between the areas under the ROC curves did not reach statistical significance (Fig. 3). We performed a subanalysis on the group with minimal change CP represented by Mayo scores 0–6 (n = 90). In this group we calculated a sensitivity of 0.55 (0.32–0.77) and a specificity of 0.99 (0.92–1) for the Rosemont score.

Rosemont US category and exocrine function

When we divided the patients into groups with prominent US changes (US Rosemont categories I and II) and less prominent US changes (US Rosemont categories III and IV), we found that the groups with the most prominent changes also had the most severe exocrine failure ($p < 0.001$) and the highest prevalence of exocrine failure ($p < 0.001$) (Fig. 4).

DISCUSSION

To our knowledge, this is the only prospective clinical report evaluating the use of transabdominal US compared with a relevant diagnostic standard for CP during the last 20 y. We applied two different scoring systems developed for EUS to the pancreatic US findings in a population of well-characterized patients. Our study had three main findings: First, we found that pancreatic US has good diagnostic accuracy for diagnosing CP in both scoring systems. We also found that disease severity represented by grade of exocrine failure was correlated to severity of sonographic changes in the CP group. The agreement between transabdominal US and CT in the detection of calcifications and cysts was excellent.

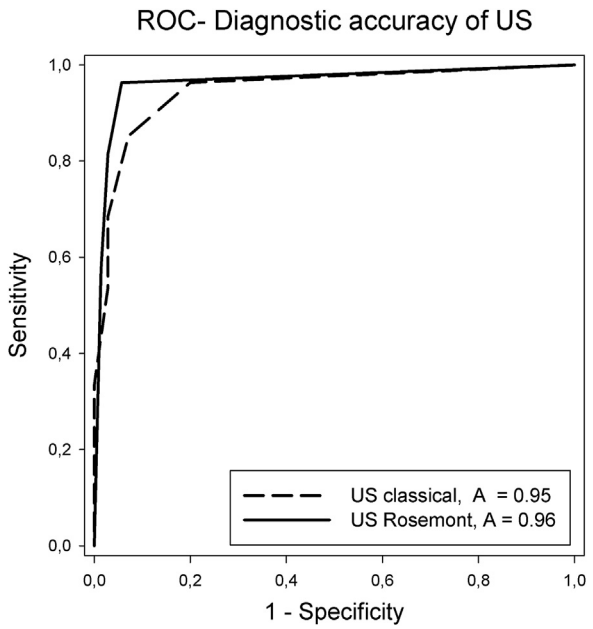

Fig. 3. Receiver operating characteristic curve of diagnostic accuracy of the Rosemont and classic scores of transabdominal ultrasound (US) features. Variances for the area under the curve (*A*) are listed in Table 4.

Knowledge of the diagnostic performance of transabdominal US in CP is based on old studies (Bolondi et al. 1989; Foley et al. 1980; Gebel et al. 1985; Ikeda et al. 1994). The German clinical practice guideline for CP (Mayerle et al. 2013) describes the specificity of US as good (70%–97%), whereas the method has pitfalls with respect to sensitivity (60%–81%). Reports evaluating transabdominal US of the pancreas in CP after the substantial improvement in US and imaging technology during the last two decades have been requested (Conwell et al. 2014). Our results confirm that dedicated transabdominal US of the pancreas has good specificity under adequate scanning conditions and that the overall accuracy has improved compared with most of the existing studies. The advantage of the Rosemont score over the traditional counting of EUS features has not been determined in EUS (Del et al. 2012; Kalmin et al. 2011; Stevens et al. 2010). We found that by applying the principles from the Rosemont consensus to transabdominal US findings we achieved better sensitivity for our suggested cutoffs, but the accuracy did not significantly differ. The most reliable and major features in the Rosemont score detected by US are calcifications. This feature is scored with the highest weight in the Rosemont score, whereas some minor features have poorer detection rate with US and are consequently weighted lower.

This study adds to our knowledge on the overall diagnostic accuracy of US features in CP, with use of a modern US scanner, and explores the advantage of weighted scores in the evaluation of findings obtained with this modality.

Methodologic considerations

The Rosemont score was developed to fit the performance of EUS. The score is not validated for US. Direct

Table 4. Accuracy of unweighted (classic) and Rosemont scores for the diagnosis of chronic pancreatitis

US scores	Sensitivity	Specificity	Cutoff	Accuracy
Classical score	0.69 (0.54–0.80)*	0.97 (0.90–1)	≥3	0.95 (0.91–0.99)
Rosemont score	0.81 (0.69–0.91)	0.97 (0.90–1)	≤2	0.97 (0.93–1)

* Median (95% confidence interval).

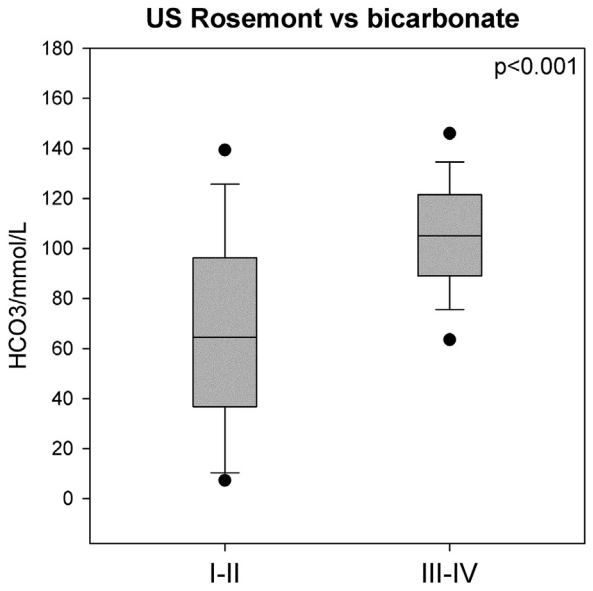

Fig. 4. Boxplot of peak duodenal bicarbonate concentration of secretin-stimulated duodenal juice for Rosemont ultrasound groups I and II versus groups III and IV. *Boxes* represent medians and quartiles. *Error bars* above and below the box indicate the 90th and 10th percentiles, respectively. *Dots* indicate outliers outside the 90% confidence interval.

adaption to transabdominal US with lower scan frequencies and much longer sound wave distance to examined structures may be difficult. The limitations in inter-rater agreement for defined sonographic features of CP described earlier for EUS may also be a limitation for transabdominal US. The operators discussed and agreed on general definitions of the individual features before starting the study. We did not analyze inter-rater agreement.

The detail in EUS is superior to that in most US examinations. The depth and probe frequency may interfere with the interpretation of US features. However, under perfect sonographic conditions in non-obese individuals and with the additional use of transabdominal high-frequency probes, we were able to obtain high-quality images of the whole or a sufficient part of the pancreas. In many cases we failed to obtain a complete US image of the pancreas from body to tail. This may reduce sensitivity in cases of focal disease. For some features, the frequency of the transducer may play a role. Interface echoes appear as the US traverses echo-rich structures and layers. Ductal walls reflect a more prominent echo at lower US frequency than at higher US frequency. This may have implications for the interpretation of ductal structures. Because of interface echo differences, the cutoff values of EUS for main ductal dilation may be inappropriate for US measures. However, in the present study, we did not alter this definition. We strictly applied the "leading edge to leading edge" measurement principle (Erchinger et al. 2011) to minimize the effect of transducer frequency variation.

For visualization of calcifications and cysts with unspecified location, we found that US performed well compared with CT. On the other hand, differentiation between intra- and extra-ductal locations of the calcifications and the detection of minor features such as stranding, minor duct dilations and duct wall hyperechogenicity are challenging in US. We postulate that because of the higher spatial resolution of EUS, the detection of these weaker features is more relevant for the EUS modality. This may explain why the weighted score, in which major features are given a high weighting, performs better on US.

Diagnostic standards for CP are challenging, and a new definition has been proposed (Whitcomb et al. 2016). The old Cambridge criteria (Sarner and Cotton, 1984) are based on endoscopic retrograde cholangiopancreatography (ERCP) and, thus, not practically applicable as a modern standard because ERCP is no longer a frequently used diagnostic modality. The most widespread standard is the M-ANNHEIM criteria (Schneider et al. 2007). This comprehensive classification system comprises several subsystems for diagnostic definition, risk factor classification, disease staging and multimodal image evaluation. For research, the system has drawbacks in its complexity.

We chose to use the simple diagnostic MAYO score, originally developed by Peter Layer (Layer et al. 1994). This score combines imaging, histology, symptoms and features of exocrine and endocrine failure in a simple numerical score. The score was originally developed for calcifying alcoholic pancreatitis, and is not validated in all groups of CP. For both research and clinical purposes, this score has the advantage of a clear numerical diagnostic definition. The score probably has limitations in defining minimal-change CP.

Supplementary methods such as contrast-enhanced ultrasonography and elastography have been developed during the last decade. Small studies (Azemoto et al. 2015; Llamoza-Torres et al. 2016; Uchida et al. 2009; Yashima et al. 2012) have evaluated the performance of these modalities applied by external US in patients with CP, but at present no clinical consensus regarding CP has been arrived at with respect to these methods. The evaluation of such modalities was outside the scope of this study.

Study limitations

Studies on US and EUS are operator dependent. The examinations were performed by operators blinded to earlier history and radiologic imaging. Blinding to patient appearance and communication during the procedure was not feasible. Blinding biases may exist.

Patients with poor sonographic imaging of the pancreas were excluded. Exclusions caused by obesity or bowel gas may introduce a selection bias.

In characterization of the patients, different image modalities (CT, MRI, EUS) were chosen as the best clinical approach in the diagnostic workup process. The use of different modalities may introduce variations in classifications. We argue that this reflects a realistic image of practical CP diagnostics.

The direct application of unadjusted EUS scoring systems may not be completely feasible for the US-detected features. To our knowledge, there are no validation studies on transcutaneous US using the Rosemont score.

CONCLUSIONS

We have reported a clinically relevant evaluation of the accuracy of transabdominal US of the pancreas using modern US technology. We found that the modality has good diagnostic accuracy and that the extent of sonographic changes is reflected by the grade of exocrine failure. Some minor features were difficult to visualize by the transabdominal approach. Adjustment of weighted scores to fit the performance of transabdominal US in practical CP diagnostics should be explored in future studies. Our results support the place of US as a simple, radiation-free and non-invasive first-line modality in the clinical workup of patients with CP.

Acknowledgments—This study was supported by MedViz (http://medviz.uib.no/), an interdisciplinary research cluster from Haukeland University Hospital, University of Bergen and Christian Michelsen Research AS. We thank our laboratory technicians Liv Aasmul and Aud Sissel Hjartholm for analyses of duodenal juice and gastrointestinal unit nurses for aid during endoscopic examinations.

REFERENCES

Albashir S, Bronner MP, Parsi MA, Walsh RM, Stevens T. Endoscopic ultrasound, secretin endoscopic pancreatic function test, and histology: Correlation in chronic pancreatitis. Am J Gastroenterol 2010; 105:2498–2503.

Altman DG. Practical statistics for medical Research. London: Chapman & Hall; 1997. ch 14.3.

Azemoto N, Kumagi T, Yokota T, Hirooka M, Kuroda T, Koizumi M, Ohno Y, Yamanishi H, Abe M, Onji M, Hiasa Y. Utility of contrast-enhanced transabdominal ultrasonography to diagnose early chronic pancreatitis. Biomed Res Int 2015;2015:393124.

Bolondi L, Li BS, Gaiani S, Barbara L. Sonography of chronic pancreatitis. Radiol Clin North Am 1989;27:815–833.

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA, Cohen JF, STARD. An updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527.

Catalano MF, Lahoti S, Geenen JE, Hogan WJ. Prospective evaluation of endoscopic ultrasonography, endoscopic retrograde pancreatography, and secretin test in the diagnosis of chronic pancreatitis. Gastrointest Endosc 1998;48:11–17.

Catalano MF, Sahai A, Levy M, Romagnuolo J, Wiersema M, Brugge W, Freeman M, Yamao K, Canto M, Hernandez LV. EUS-based criteria for the diagnosis of chronic pancreatitis: The Rosemont classification. Gastrointest Endosc 2009;69:1251–1261.

Chong AK, Hawes RH, Hoffman BJ, Adams DB, Lewin DN, Romagnuolo J. Diagnostic performance of EUS for chronic pancreatitis: A comparison with histopathology. Gastrointest Endosc 2007; 65:808–814.

Conwell DL, Lee LS, Yadav D, Longnecker DS, Miller FH, Mortele KJ, Levy MJ, Kwon R, Lieb JG, Stevens T, Toskes PP, Gardner TB, Gelrud A, Wu BU, Forsmark CE, Vege SS. American Pancreatic Association practice guidelines in chronic pancreatitis: Evidence-based report on diagnostic guidelines. Pancreas 2014;43:1143–1162.

Conwell DL, Zuccaro G Jr, Vargo JJ, Trolli PA, Vanlente F, Obuchowski N, Dumot JA, O'Laughlin C. An endoscopic pancreatic function test with synthetic porcine secretin for the evaluation of chronic abdominal pain and suspected chronic pancreatitis. Gastrointest Endosc 2003;57:37–40.

Del PD, Poves E, Tabernero S, Beceiro I, Moral I, Villafruela M, Sanz C, Borrego G. Conventional versus Rosemont endoscopic ultrasound criteria for chronic pancreatitis: Interobserver agreement in same day back-to-back procedures. Pancreatology 2012;12:284–287.

Dimcevski G, Erchinger FG, Havre R, Gilja OH. Ultrasonography in diagnosing chronic pancreatitis: New aspects. World J Gastroenterol 2013;19:7247–7257.

Drewes AM, Frokjaer JB, Jørgensen MT, Knoop F, Mortensen MB, Schaffalitzky de Muckadell O, Nøjgaard C, Olesen S, Phillipsen E, Rasmussen H. Diagnostik og behandling af kronisk pankreatit. Aalborg: Dansk Selskab for Gastroenterologi og Hepatologi; 2015.

Erchinger F, Dimcevski G, Engjom T, Gilja O. Transabdominal ultrasound of the pancreas: Basic and new aspects. Imaging Med 2011; 3:411–422.

Erchinger F, Engjom T, Tjora E, Hoem D, Hausken T, Gilja OH, Dimcevski G. Quantification of pancreatic function using a clinically feasible short endoscopic secretin test. Pancreas 2013;42: 1101–1106.

Foley WD, Stewart ET, Lawson TL, Geenan J, Loguidice J, Maher L, Unger GF. Computed tomography, ultrasonography, and endoscopic retrograde cholangiopancreatography in the diagnosis of pancreatic disease: A comparative study. Gastrointest Radiol 1980;5:29–35.

Forsmark CE. Management of chronic pancreatitis. Gastroenterology 2013;144:1282–1291.

Gebel M, Stiehl M, Freise J. [Value of sonographic imaging of the pancreatic duct for the diagnosis of chronic pancreatitis and pancreatic cancer compared to ERCP]. Ultraschall Med 1985;6:127–130.

Iglesias-Garcia J, Larino-Noia J, Lindkvist B, Dominguez-Munoz JE. Endoscopic ultrasound in the diagnosis of chronic pancreatitis. Rev Esp Enferm Dig 2015;107:221–228.

Ikeda M, Sato T, Morozumi A, Fujino MA, Yoda Y, Ochiai M, Kobayashi K. Morphologic changes in the pancreas detected by screening ultrasonography in a mass survey, with special reference to main duct dilatation, cyst formation, and calcification. Pancreas 1994;9:508–512.

Kalmin B, Hoffman B, Hawes R, Romagnuolo J. Conventional versus Rosemont endoscopic ultrasound criteria for chronic pancreatitis: Comparing interobserver reliability and intertest agreement. Can J Gastroenterol 2011;25:261–264.

Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.

Layer P, Yamamoto H, Kalthoff L, Clain JE, Bakken LJ, DiMagno EP. The different courses of early- and late-onset idiopathic and alcoholic chronic pancreatitis. Gastroenterology 1994; 107:1481–1487.

Llamoza-Torres CJ, Fuentes-Pardo M, Alvarez-Higueras FJ, Alberca-de-Las-Parras F, Carballo-Alvarez F. Usefulness of percutaneous elastography by acoustic radiation force impulse for the non-invasive diagnosis of chronic pancreatitis. Rev Esp Enferm Dig 2016;108:450–456.

Loser C, Mollgaard A, Folsch UR. Faecal elastase 1: A novel, highly sensitive, and specific tubeless pancreatic function test. Gut 1996; 39:580–586.

Luetmer PH, Stephens DH, Ward EM. Chronic pancreatitis: Reassessment with current CT. Radiology 1989;171:353–357.

Manfredi R, Costamagna G, Brizi MG, Maresca G, Vecchioli A, Colagrande C, Marano P. Severe chronic pancreatitis versus suspected pancreatic disease: Dynamic MR cholangiopancreatography after secretin stimulation. Radiology 2000;214:849–855.

Martinez J, Abad-Gonzalez A, Aparicio JR, Aparisi L, Boadas J, Boix E, de las HG, Dominguez-Munoz E, Farre A, Fernandez-Cruz L, Gomez L, Iglesias-Garcia J, Garcia-Malpartida K, Guarner L, Larino-Noia J, Lluis F, Lopez A, Molero X, Moreno-Perez O, Navarro S, Palazon JM, Perez-Mateo M, Sabater L, Sastre Y, Vaquero EC, De-Madaria E. [Recommendations of the Spanish Pancreatic Club on the diagnosis and treatment of chronic pancreatitis: Part 1 (diagnosis)]. Gastroenterol Hepatol 2013;36:326–339.

Martinez-Noguera A, D'Onofrio M. Ultrasonography of the pancreas: 1. Conventional imaging. Abdom Imaging 2007;32:136–149.

Mayerle J, Hoffmeister A, Werner J, Witt H, Lerch MM, Mossner J. Chronic pancreatitis: Definition, etiology, investigation and treatment. Dtsch Arztebl Int 2013;110:387–393.

Pungpapong S, Wallace MB, Woodward TA, Noh KW, Raimondo M. Accuracy of endoscopic ultrasonography and magnetic resonance cholangiopancreatography for the diagnosis of chronic pancreatitis: A prospective comparison study. J Clin Gastroenterol 2007;41:88–93.

Sahai AV, Zimmerman M, Aabakken L, Tarnasky PR, Cunningham JT, van VA, Hawes RH, Hoffman BJ. Prospective assessment of the ability of endoscopic ultrasound to diagnose, exclude, or establish the severity of chronic pancreatitis found by endoscopic retrograde cholangiopancreatography. Gastrointest Endosc 1998;48:18–25.

Sarner M, Cotton PB. Classification of pancreatitis. Gut 1984;25:756–759.

Schneider A, Lohr JM, Singer MV. The M-ANNHEIM classification of chronic pancreatitis: Introduction of a unifying classification system based on a review of previous classifications of the disease. J Gastroenterol 2007;42:101–119.

Stevens T, Conwell DL, Zuccaro G Jr, Van LF, Lopez R, Purich E, Fein S. A prospective crossover study comparing secretin-stimulated endoscopic and Dreiling tube pancreatic function testing in patients evaluated for chronic pancreatitis. Gastrointest Endosc 2008;67:458–466.

Stevens T, Lopez R, Adler DG, Al-Haddad MA, Conway J, Dewitt JM, Forsmark CE, Kahaleh M, Lee LS, Levy MJ, Mishra G, Piraka CR, Papachristou GI, Shah RJ, Topazian MD, Vargo JJ, Vela SA. Multicenter comparison of the interobserver agreement of standard EUS scoring and Rosemont classification scoring for diagnosis of chronic pancreatitis. Gastrointest Endosc 2010;71:519–526.

Tjora E, Wathle G, Engjom T, Erchinger F, Molven A, Aksnes L, Haldorsen IS, Dimcevski G, Njolstad PR, Raeder H. Severe pancreatic dysfunction but compensated nutritional status in monogenic pancreatic disease caused by carboxyl-ester lipase mutations. Pancreas 2013;42:1078–1084.

Uchida H, Hirooka Y, Itoh A, Kawashima H, Hara K, Nonogaki K, Kasugai T, Ohno E, Ohmiya N, Niwa Y, Katano Y, Ishigami M, Goto H. Feasibility of tissue elastography using transcutaneous ultrasonography for the diagnosis of pancreatic diseases. Pancreas 2009;38:17–22.

Whitcomb DC, Frulloni L, Garg P, Greer JB, Schneider A, Yadav D, Shimosegawa T. Chronic pancreatitis: An international draft consensus proposal for a new mechanistic definition. Pancreatology 2016;16:218–224.

Whitsett MC. Ultrasound imaging and advances in system features. Ultrasound Clin 2009;4:391–401.

Wiersema MJ, Hawes RH, Lehman GA, Kochman ML, Sherman S, Kopecky KK. Prospective evaluation of endoscopic ultrasonography and endoscopic retrograde cholangiopancreatography in patients with chronic abdominal pain of suspected pancreatic origin. Endoscopy 1993;25:555–564.

Wiersema MJ, Wiersema LM. Endosonography of the pancreas: normal variation versus changes of early chronic pancreatitis. Gastrointest Endosc Clin North Am 1995;5:487–496.

World Medical Association General Assembly. World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. Available at: http://www.wma.net/en/30publications/10 policies/b3/; 2015. Accessed September 2, 2016.

Yashima Y, Sasahira N, Isayama H, Kogure H, Ikeda H, Hirano K, Mizuno S, Yagioka H, Kawakubo K, Sasaki T, Nakai Y, Tada M, Yoshida H, Omata M, Koike K. Acoustic radiation force impulse elastography for noninvasive assessment of chronic pancreatitis. J Gastroenterol 2012;47:427–432.