ORIGINAL ARTICLE

Conventional versus Rosemont endoscopic ultrasound criteria for chronic pancreatitis: Comparing interobserver reliability and intertest agreement

Bruce Kalmin MD, Brenda Hoffman MD FASGE, Robert Hawes MD FASGE, Joseph Romagnuolo MD MSc FASGE

B Kalmin, B Hoffman, R Hawes, J Romagnuolo. Conventional versus Rosemont endoscopic ultrasound criteria for chronic pancreatitis: Comparing interobserver reliability and intertest agreement. Can J Gastroenterol 2011;25(5):261-264.

BACKGROUND: The Rosemont criteria (RC) were recently proposed by expert consensus to standardize endoscopic ultrasound (EUS) features and thresholds for diagnosing chronic pancreatitis (CP); however, they are cumbersome and are not validated.
OBJECTIVE: To determine interobserver agreement between RC and conventional criteria (CC), and to assess intertest agreement in the diagnosis of CP.
METHODS: Thirty-six consecutive patients who underwent EUS for abdominal pain or pancreatitis were retrospectively reviewed. Anonymized images were independently chosen as best representations of the pancreatic body and reviewed by three experts who recorded the presence of CC and RC features. Agreement (proportion and kappa statistic) between CC and RC was calculated. Interobserver agreement within the CC and RC was assessed. Secondary comparisons with endoscopic retrograde cholangiopancreatography were made where available.
RESULTS: Using CC, 60 readings (83.3%) were negative for CP, while 12 readings (16.7%) were positive. Using RC, 59 readings (81.9%) were negative for CP, while 13 readings (18.1%) were positive. The weighted kappa for interobserver agreement for CC (four categories: normal/low probability, indeterminate, high probability or calcific) was 0.50, with 80.0% overall agreement, versus 0.27 and 68.1% for the four RC categories (normal, indeterminate, suggestive of and consistent with). Agreement on a positive diagnosis with CC was 86.1% (P=0.38 [McNemar's exact test]), with a kappa of 0.47; for RC, agreement was lower at 80.6% (P=0.016 [McNemar's exact test]), with a kappa of 0.38. For patients who underwent endoscopic retrograde cholangiopancreatography (n=12), false-negative and false-positive rates between CC and RC did not appear to be different.
CONCLUSIONS: The RC do not appear to achieve the goals of improving accuracy and interobserver agreement for diagnosing CP.

Key Words: Accuracy; Chronic pancreatitis; Criteria; ERCP; EUS; Interobserver reliability

La classification habituelle et la classification de Rosement pour l'échoendoscopie en cas de pancréatite chronique : comparer la fiabilité entre observateurs et l'entente entre les tests

HISTORIQUE : Un consensus d'experts a récemment proposé la classification de Rosemont (CR) pour normaliser les caractéristiques et les seuils de l'échoendoscopie (ÉES) afin de diagnostiquer la pancréatite chronique (PC). Cette classification est toutefois fastidieuse et non validée.
OBJECTIF : Déterminer l'entente entre observateurs au sujet de la CR et de la classification habituelle (CH) et évaluer l'entente entre les tests pour diagnostiquer la PC.
MÉTHODOLOGIE : Les chercheurs ont procédé à l'analyse rétrospective de 36 patients consécutifs qui ont subi une ÉES en raison de douleurs abdominales ou de pancréatite. Ils ont choisi des images anonymisées de manière indépendante, qui constituaient les meilleures représentations du corps pancréatique, lesquelles ont été évaluées par trois experts qui ont consigné les caractéristiques de la CH et de la CR. Les chercheurs ont ensuite calculé l'entente (proportion et statistique Kappa) entre la CH et la CR. Ils ont évalué l'entente entre observateurs pour ce qui est de la CH et de la CR. Lorsque c'était possible, ils ont effectué des comparaisons secondaires avec la cholangiopancréatographie rétrograde endoscopique.
RÉSULTATS : Au moyen de la CH, 60 lectures (83,3 %) excluaient la PC, tandis que 12 (16,7 %) étaient positives. Au moyen de la CR, 59 lectures (81,9 %) excluaient la PC, tandis que 13 (18,1 %) étaient positives. L'indice pondéré de Kappa de l'entente entre observateurs à l'égard de la CH (quatre catégories : normal/faible probabilité, indéterminé, probabilité élevée ou calcifiée) était de 0,50, avec 80,0 % d'entente globale, par rapport à 0,27 et à 68,1 % des catégories de CR (normal, indéterminé, évocateur et compatible). L'entente à l'égard d'un diagnostic positif au moyen de la CH s'élevait à 86,1 % (P=0,38 [test exact de McNemar]), selon un indice pondéré de Kappa de 0,47. Au moyen de la CR, l'entente était inférieure à 80,6 % (P=0,016 [test exact de McNemar]), selon un indice pondéré de Kappa de 0,38. Pour ce qui est des patients qui avaient subi une cholangiopancréatographie rétrograde endoscopique (n=12), les taux de faux-négatifs et de faux-positifs entre la CH et les CR semblaient similaires.
CONCLUSIONS : La CR ne semble pas parvenir à améliorer l'exactitude et l'entente entre observateurs pour diagnostiquer la PC.

Chronic pancreatitis (CP) can be a difficult diagnosis to make, especially when calcifications and duct dilation are absent. Endoscopic ultrasound (EUS), however, appears to be a sensitive and specific test for this disease (1). Various endosonographic features have been used to determine whether a patient has CP. However, while most endosonographers agree on the final diagnosis, the kappa values for individual criteria are modest at best (2). Although most endosonographers look for nine criteria, they consider calcifications, or fulfilling five or more of the remaining eight criteria, to be a 'high-probability' diagnosis, while fulfilling three to four criteria is considered to be 'indeterminate'. There is heterogeneity in the literature regarding the criteria to look for and the thresholds for diagnosis (3-9), with thresholds for the number of criteria required varying from one to six, and the denominator of criteria sought varying from five to 12 (10). A common frustration among endosonographers is that the concept of 'counting criteria' assumes that the criteria (eg, duct dilation versus strands) have equal weight, which is probably not accurate.

The 'gold' or reference standard for CP has traditionally been endoscopic retrograde cholangiopancreatography (ERCP), pancreatic function testing or histology, although it is known that ERCP and pancreatic function testing miss a significant proportion of early CP. Studies have shown that finding fewer than three criteria effectively excludes moderate and severe CP on ERCP. Asymptomatic control patients may have some features of CP, especially if they drink alcohol, and it is not clear whether these represent false positives or subclinical disease (11). One study comparing EUS with surgical pathology (12) revealed that three or more EUS criteria provided the best balance of sensitivity (83%) and specificity (80%), while five or more was very specific; another similar study (13) found that four or more criteria was the best threshold value.

In April 2007, an international consensus panel met in Rosemont, Illinois (USA), in an attempt to improve the reliability of EUS for CP, and included one of the authors of the present article (JR). This meeting resulted in the formulation of the so-called 'Rosemont criteria' (RC) for

TABLE 1
Endoscopic ultrasound criteria for the diagnosis of chronic pancreatitis

Conventional criteria	Rosemont criteria
Parenchymal criteria	**Major criteria A**
Hyperechoic foci	Hyperechoic foci (>2 mm in length/width with
Hyperechoic strands	shadowing)
Hypoechoic lobules, foci or areas	Major duct calculi (echogenic structure[s] within the MPD with acoustic shadowing)
Cyst	
	Major criteria B
	Lobularity (≥3 contiguous lobules = 'honeycombing')
Duct criteria	
Irregular duct contour	**Minor criteria**
Visible side branches	Cyst (anechoic, round/elliptical with or without septations)*
Hyperechoic duct margin	Dilated duct (≥3.5 mm in body or >1.5 mm in tail)*
Dilated main duct	Irregular duct contour (uneven or irregular outline and ectatic course)
Stone	Dilated side branch (>3 tubular anechoic structures each measuring ≥1 mm in width, budding from the MPD)*
	Hyperechoic duct wall (echogenic, distinct structure >50% of entire MPD in the body and tail)
	Hyperechoic strands (≥3 mm in at least 2 different directions with respect to the imaged plane)
	Hyperechoic foci (>2 mm in length/width that are nonshadowing)*
	Lobularity (>5 mm, noncontiguous lobules)

If any of these minor criteria are present, the patient cannot be classified as 'normal'. MPD Main pancreatic duct. Data from references 14 and 15

TABLE 2
Classification of patients based on endoscopic criteria

Conventional criteria	Rosemont criteria
Normal (or low probability)	**Consistent with**
0–2 criteria	2 major A
	1 major A + 1 major B
Indeterminate or intermediate probability	1 major A + ≥3 minor
3–4 criteria	
	Suggestive of
High probability	Major A + <3 minor
5–9 criteria	Major B + ≥3 minor
	≥5 minor, no major
	Indeterminate
	Major B alone + <3 minor
	Normal
	<3 minor, no major

Data from references 14 and 16

or patient demographics. An electronic radial or a linear Olympus echoendoscope (Olympus, USA) was used, with 47.2% of procedures being performed with the latter device (range 41.7% to 54.2% for the three reviewers' cases). Each reviewer was instructed to record the presence or absence of the features used to diagnose CP using CC and the newly proposed RC. Each patient's images were evaluated by two randomly chosen reviewers (72 reviews). Each reviewer examined 24 cases; one-half of those cases were re-examined by one of the other two endosonographers, and one-half by the other. Thus, the 36 patients were effectively divided into 3×12 patient groups – each group was reviewed by a two-physician pairing of the three endosonographers. None of the images were believed by any of the endosonographers to be of unacceptable quality for interpretation.

Using the CC, patients were classified as 'normal' (less than three of nine criteria), 'borderline' (three to four criteria), 'high probability' (five or more criteria) or 'calcific'. Using the RC, patients were classified as 'normal', 'indeterminate', 'suggestive of' or 'consistent with' (Table 2). Studies were then dichotomized such that patients who were classified as normal/borderline under the CC or normal/indeterminate under the RC, were separated from those classified as high-probability/calcific under the CC or suggestive of/consistent with under the RC. For the purposes of the study, the latter grouping was considered to be a positive test for CP. If an ERCP had been performed, the Cambridge criteria of that pancreatogram were noted.

Statistical analysis and sample size
Interobserver variability was determined for each of the two systems using weighted kappa statistics (weighted because defining high probability versus calcific is not as significant a clinical disagreement as defining normal versus high probability, for example) to assess agreement beyond chance. Because variability may differ among different pairings of physicians, an exact test for heterogeneity was used to check for significant variation in agreement using Stata version 7.0 (Stata, USA). To assess how often the RC and CC arrived at different conclusions, inter-test agreement between the two classification systems regarding the diagnosis of CP in individual patients was determined by comparing the number of patients diagnosed with CP (positive test) using CC versus using RC, calculating both proportions of agreement and kappa statistics. For the subset of patients who underwent ERCP, the EUS findings were compared with the ERCP as a secondary exploratory analysis to help explain any disagreements between CC and RC that may have occurred.

Assuming a point estimate of 0.8 for kappa, and a 15% rate of positive EUS examinations (for CP), a 95% CI for the kappa of ±0.1 would be achieved with 36 patients.

the EUS diagnosis of CP, which were recently published (14). This classification system attempted to standardize and more explicitly define the endosonographic features and thresholds for the diagnosis of chronic pancreatitis, and the grouping of criteria into major and minor importance categories (Tables 1 and 2) (14,15).

However, these newly proposed criteria have never been validated or proven to have better interobserver agreement than the previous criteria. Furthermore, the final classification system is rather complex compared with the current system. In the present study, we attempted to determine how the conventional criteria (CC) compare with the RC in terms of interobserver agreement, how often one disagrees with the other (and with ERCP in a small subgroup), to preliminarily assess whether this set of new criteria is worthy of further study and/or ready for 'prime-time' regarding inclusion in endoscopy report-writing software and/or research protocols.

METHODS

Patients
Consecutive patients who underwent an EUS between September 17, and November 19, 2008, at the Medical University of South Carolina (MUSC [South Carolina, USA]) for the indication of abdominal pain or 'pancreatitis' were considered for inclusion in the present retrospective study. Patients with pancreatic masses were excluded. The study was reviewed by the institutional review board of MUSC.

Categorizations
Endosonographic images from each of the EUS studies were independently chosen by one of the authors (BK) as the best representations of the body of the pancreas, and were subsequently anonymized and independently reviewed again by three experienced quarternary-care endosonographers (more than 1000 pancreatic cases in total and more than 100 pancreatic cases/year), without access to the study indication

TABLE 3
Summary of interobserver agreement between conventional criteria and Rosemont criteria

	Conventional criteria	Rosemont criteria
Weighted kappa (for 4-category assessment)	0.50 (P<0.001 for agreement beyond chance) 80.0% overall agreement	0.27 (P=0.01 for agreement beyond chance) 68.1% overall agreement
Kappa for a positive test*	0.47 (P=0.002 for agreement beyond chance) 86.1% overall agreement	0.38 (P=0.002 for agreement beyond chance) 80.6% overall agreement†
Kappa for a normal test	0.50 (P=0.01 for agreement beyond chance) 75.0% overall agreement	0.18 (P=0.13 for agreement beyond chance‡) 58.3% overall agreement

*Positive test refers to any test that was not 'normal'/'low probability' and not 'indeterminate'; †McNemar's exact test P<0.05 (significant); ‡No significant agreement beyond what is expected by chance

TABLE 4
Cambridge classification of chronic pancreatitis using endoscopic retrograde cholangiopancreatography

Class	Definition
0 – Normal	Visualization of entire duct system with uniform filling of side branches without acinar opacification, with a normal main duct and normal side branches
1 – Equivocal	Normal main duct
	1–3 abnormal side branches
2 – Mild	Normal main duct
	>3 abnormal side branches
3 – Moderate	Dilated main duct with irregularity
	>3 abnormal side branches
	Small cysts (<10 mm)
4 – Marked or severe	Large cysts (>10 mm)
	Gross irregularity of main pancreatic duct
	Intraductal calculus or calculi
	Stricture(s)
	Obstruction with severe dilation

Data from reference 12

RESULTS

A total of 36 patients met the inclusion criteria for the present study. Of these, 53% had only abdominal pain as an indication, while 47% had 'pancreatitis' suspected by an imaging abnormality, unexplained hyperlipasemia or recurrent acute pancreatitis. The mean age of the patients was 50 years and 44% were men. Significant alcohol use was noted in 11% of patients. Calcifications were found on EUS in 5% of patients, and 11% had main pancreatic duct dilation (4 mm or greater in the body). There was a total of 72 readings of these 36 studies.

Intertest agreement

Using CC, 60 readings (83.3%) were negative for CP, while 12 readings (16.7%) were positive. Using RC, 59 readings (81.9%) were negative for CP, while 13 readings (18.1%) were positive. Three readings (4.2%) were positive using the RC but negative with CC, while one reading (1.4%) was negative on RC but positive on CC. Weighted kappas for agreement beyond chance for the three reviewers were 0.58, 0.62 and 0.82, respectively, comparing each reviewer's score (from 1 to 4) using CC versus the score they assigned using RC (P<0.001 for all).

Interobserver agreement

Interobserver agreement for CC regarding the four categories (normal/low probability, indeterminate, high-probability or calcific) was associated with a weighted kappa score of 0.50 (P<0.001 for agreement beyond chance) and 80.0% overall agreement (Table 3). For RC, the weighted kappa was 0.27 (P=0.01 for agreement beyond chance) among the four categories (normal, indeterminate, suggestive of and consistent with), and agreement was 68.1%.

Agreement regarding a positive diagnosis with CC was 86.1% (P=0.38 [McNemar's exact test]), with a kappa of 0.47 (P=0.002); for the RC, agreement was 80.6% (ie, a higher rate of disagreement [McNemar's exact test P=0.016 for significant difference among reviews]), with a kappa of 0.38 (P=0.002). Note that kappas are lower here because they measured agreement beyond chance, and when there are only two possible answers (as opposed to the four-category analysis above), chance agreement is more likely and, therefore, agreement beyond chance is more difficult to achieve. When the definition of 'abnormal' or 'positive' is changed to any case not reported as 'normal', the agreement (75%) and kappa (0.50) (P=0.01 for agreement beyond chance) are lower for both CC and for RC (agreement 58.3%; kappa 0.18; P=0.13 for agreement beyond chance), but the agreement for CC remains higher for CC than for RC.

Each of the three endosonographers agreed with the other endosonographer assigned to be the second reviewer (three possible pairings of 12 patients per reviewer pairing) that the EUS was positive

for CP at similar rates: 75.0%, 83.3% and 83.3%, respectively (P=1.0 [exact test for heterogeneity]) using the CC criteria (average agreement 86.1%). For the RC, again, each endosonographer agreed with the other endosonographer that the test was positive at similar rates: 75.0%, 75.0% and 83.3% of cases, respectively (average agreement 86.1%; P=1.0 [exact test for heterogeneity]).

Subgroup comparison of CC and RC with ERCP

There were 12 patients who underwent ERCP in addition to their EUS. The number of cases that had been assigned to each reviewer ranged from eight, nine and seven among the three reviewers, respectively. Of the 12 ERCP patients, 10 (83.3%) had normal pancreatograms (Cambridge class 0 or 1) and two (16.7%) had evidence of CP (Cambridge 2 ['mild'] or higher); the latter ERCP-positive rate, stratified by each of the three reviewers, was 1 of 8, 2 of 9 and 1 of 7, respectively. False positive (CC or RC classification that was positive [not normal/indeterminate], but the Cambridge class was 0 or 1 [normal or equivocal]) rates for each reviewer appeared similar for CC versus RC: the CC rates were 1 of 1, 2 of 2 and 0 of 1 versus 1 of 1, 3 of 4 and 0 of 1 with RC. False negative (Cambridge classification abnormal [not normal or equivocal], but the CC or RC classification was negative [normal or indeterminate]) rates, were also similar for CC and RC: with CC, the rates were 1 of 7, 2 of 7 and 0 of 6 versus 1 of 7, 1 of 5 and 0 of 6 for RC (Table 4).

DISCUSSION

The purpose of the present study was to compare the CC – as used at our institution – with the newer proposed RC. Creation of the newer RC carried the hope of accomplishing the following: higher accuracy (implying that RC would disagree with the old criteria in a clinically significant number of patients and, when it did, RC would be more likely to be correct); better interobserver agreement; the descriptors used ('consistent with', 'suggestive of', etc) would be more well-received by referring doctors than the more conventional descriptors 'high probability', 'indeterminate' or 'intermediate probability', etc). The former two aims were assessed in the present study.

However, we found that the majority of patients had the same diagnosis regardless of the criteria used. Roughly speaking, considering the way the categories are defined in the RC, one can see that calcific pancreatitis (and the most obvious forms of high-probability CP under the CC) would fall under the RC's 'consistent with'; most high-probability noncalcific pancreatitis in CC falls under 'suggestive of' the indeterminate/intermediate probability of the CC criteria matches well

with the RC 'indeterminate' category; and low probability and normal in CC correspond to 'normal' in RC. Therefore, it is not surprising that the diagnoses did not change at a substantial rate; it is doubtful, given the above, that one set of criteria will prove to have a substantially different diagnostic performance compared with the other, and may not be worth the work and expense of repeating validation studies on this new set of criteria. The exploratory comparison with ERCP in the subgroup of our cohort that had ERCP results available (to investigate whether disagreements between RC and CC were more often in favour of RC being correct or vice versa) did not show a trend for the false-positive or false-negative results being any different. The CC have already been extensively studied in comparison with ERCP and pancreatic function tests, clinical and radiological follow-up, and to histology.

The interobserver agreement of CC has been studied previously (2) and was found to be modest at best, but it is not clear whether using the more complex algorithm proposed in the RC can reduce the problem of interobserver disagreement. In fact, the measures of agreement beyond chance were lower for RC than for CC. It is not clear why this occurred because as the RC were developed, more explicit definitions were specifically required in an attempt to improve the agreement among endosonographers. One proposed explanation is that although the specific criteria are spelled out in greater detail in the RC, every subcriterion (defining the main criterion) that is added likely adds a layer of agreement challenge – eg, instead of having to agree on whether lobules are present, one has to now agree on how many and whether they are contiguous; or on what percentage of the duct has a hyperechoic wall – this added level of factors requiring agreement may trump the specificity with which the criteria were defined.

There are limitations to the present study. First, it was retrospective in nature and used still images. Video recordings of the EUS would have been better; however, for comparisons of interobserver reliability, the limitation of the still images would have been balanced on both sides of the comparison. Second, the small sample size limited subgroup and exploratory analyses with ERCP, and resulted in wide CIs.

An important observation made qualitatively in the course of the current study is that using the newer RC appeared to be very time-consuming and required multiple quantitative measurements. It also involved a relatively complex algorithm to stratify the multiple categories (Table 2); practically, the algorithm would need to be posted in the EUS room and one would need to consult it frequently to produce a report. This may preclude the majority of endosonographers from using the RC, and subsequently create inconsistencies in the scoring systems used in practice. We did not formally compare the time burden of each scoring system, but again, the extra time might not be supported by its proposed potential benefits, especially because those benefits, at least preliminarily, do not appear to exist.

CONCLUSION

While the goal of accurately and consistently diagnosing early noncalcific CP is desirable, it is not clear whether using a new nonvalidated set of criteria will achieve this. In the present study, there was no clear advantage of using the RC over the CC, but there were some time- and complexity-related disadvantages noted. The RC do not have a clear advantage in terms of interobserver reliability, which was a major impetus for developing the RC and, in fact, appears to be worse – there was a significant rate of disagreement on the final diagnosis (P=0.02), a lower weighted kappa for agreement beyond chance on rating the individual categories (0.50 versus 0.27) and no evidence of agreement beyond chance for a 'normal' examination (P=0.13). Preliminarily, the RC do not appear to more accurately diagnose patients

with CP over the CC and, in our exploratory subanalysis of patients who also underwent ERCP, it did not correspond better with ERCP than CC. Therefore, at the present time, evidence supporting the preferential use of RC over CC in research studies involving EUS or in endoscopic reports in general practice appears to be lacking. Further study would be required of the RC's diagnostic performance and reliability before it could be recommended in these capacities. Funded by the American Society of Gastrointestinal Endoscopy, our group has recently begun a prospective, randomized, multicentre trial measuring the interobserver reliability of CC and RC, and comparing CC and RC results with ERCP.

REFERENCES

1. Chong AK, Romagnuolo J, Lewin D, et al. Diagnosis of chronic pancreatitis with endoscopic ultrasound: A comparison with histopathology. Gastrointest Endosc 2005;61:AB77. (Abst)
2. Wallace MB, Hawes RH, Durkalski V, et al. The reliability of EUS for the diagnosis of chronic pancreatitis: Interobserver agreement among experienced endosonographers. Gastrointest Endosc 2001;53:294-9.
3. Kahl S, Glasbrenner B, Leodolter A, et al. EUS in the diagnosis of early chronic pancreatitis: A prospective follow-up study. Gastrointest Endosc 2002;55:507-11.
4. Hollerbach S, Klamann A, Topalidis T, et al. Endoscopic ultrasonography (EUS) and fine-needle aspiration (FNA) cytology for diagnosis of chronic pancreatitis. Endoscopy 2001;33:824-31.
5. Hastier P, Buckley MJ, Francois E, et al. A prospective study of pancreatic diseases in patients with alcoholic cirrhosis: Comparative diagnostic value of ERCP and EUS and long-term significance of isolated parenchymal abnormalities. Gastrointest Endosc 1999;49:705-9.
6. Catalano MF, Lahoti S, Geenen JE, et al. Prospective evaluation of endoscopic ultrasonography, endoscopic retrograde pancreatography, and secretin test in the diagnosis of chronic pancreatitis. Gastrointest Endosc 1998;48:11-7.
7. Sahai AV, Zimmerman M, Aabakken L, et al. 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.
8. Buscail L, Escourrou J, Moreau J, et al. Endoscopic ultrasonography in chronic pancreatitis: A comparative prospective study with conventional ultrasonography, computed tomography, and ERCP. Pancreas 1995;10:251-7.
9. Wiersema MJ, Hawes RH, Lehman G, et al. Prospective evaluation of endoscopic ultrasonography and endoscopic retrograde cholangiopancreatography in patients with chronic abdominal pain of suspected pancreatic origin. Endoscopy 1993;25:555-64.
10. Romagnuolo J. EUS in inflammatory diseases of the pancreas. In: Hawes R, Fockens P, eds. Endosonography, 2nd edn. London: Elsevier 2010:127-47.
11. Sahai AV, Mishra G, Penman ID, et al. EUS to detect evidence of pancreatic disease in patients with persistent or nonspecific dyspepsia. Gastrointest Endosc 2000;52:153-9.
12. Chong A, Romagnuolo J. Gender-related changes in the pancreas by EUS. Gastrointest Endosc 2005;62:475. (Lett)
13. Varadarajulu S, Eltoum I, Tamhane A, Eloubeidi MA. Histopathologic correlates of noncalcific pancreatitis by EUS: A prospective tissue characterization study. Gastrointest Endosc 2007;66:501-9.
14. Catalano MF, Sahai A, Levy M, et al. EUS-based criteria for the diagnosis of chronic pancreatitis: The Rosemont classification. Gastrointest Endosc 2009;69:1251-61.
15. Hernandez LV, Sahai A, Brugge WR, et al. Standardized weighted criteria for EUS features of chronic pancreatitis: The Rosemont Classification. Gastrointest Endosc 2008;67:AB96. (Abst)