



Published in final edited form as:

*Abdom Radiol (NY)*. 2020 May ; 45(5): 1481–1487. doi:10.1007/s00261-020-02521-7.

## Inter-observer Variability of Radiologists for Cambridge Classification of Chronic Pancreatitis Using CT and MRCP - Results from a Large Multi-Center Study

**Temel Tirkes, M.D.\* [Associate Professor Radiology],**

Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN USA

**Zarine K. Shah, M.D. [Assistant Professor of Radiology],**

Department of Radiology, Ohio State University Wexner Medical Center, Columbus, OH USA

**Naoki Takahashi, M.D.,**

Department of Radiology, Mayo Clinic, Rochester, MN USA

**Joseph R. Grajo, M.D. [Assistant Professor of Radiology],**

University of Florida College of Medicine, Gainesville, FL USA

**Stephanie T. Chang, M.D.,**

Department of Radiology and Division of Body MRI, Stanford University School of Medicine, Stanford, CA USA

**Ashley M. Wachsman, M.D.,**

Department of Radiology, Cedars-Sinai Medical Center, Los Angeles, CA USA

**Kareem Mawad, MD, MA,**

Department of Radiology, South San Francisco Medical Center, The Permanente Medical Group, South San Francisco, CA USA

**Carlos A Farinas, MD,**

Department of Radiology, Baylor College of Medicine, Houston, TX, USA

**Liang Li, Ph.D. [Professor],**

Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX USA

**Savitri N. Appana, MS,**

Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

---

Terms of use and reuse: academic research for non-commercial purposes, see here for full terms. <https://www.springer.com/aam-terms-v1>

\* *Corresponding author* Temel Tirkes, M.D., F.A.C.R., Indiana University School of Medicine, Department of Radiology and Imaging Sciences, 550 N. University Blvd. Suite 0663, Indianapolis, IN, 46202, Tel: (317) 944-1837, Fax: (317) 944-1848, [atirkes@iu.edu](mailto:atirkes@iu.edu).

Conflicts of Interest

None.

**Publisher's Disclaimer:** This Author Accepted Manuscript is a PDF file of an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

**Darwin L. Conwell, M.D. [Professor of Medicine],**

Department of Medicine, Division of Gastroenterology, Hepatology & Nutrition, Ohio State University Wexner Medical Center, Columbus, OH, USA

**Dhiraj Yadav, M.D. MPH [Professor of Medicine],**

Department of Medicine, Division of Gastroenterology, Hepatology & Nutrition, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

**Anil K. Dasyam, M.D. [Associate Professor of Radiology]**

Department of Radiology, University of Pittsburgh Medical Center, Pittsburgh, PA USA

**On behalf of the Consortium for the Study of Chronic Pancreatitis, Diabetes, and Pancreatic Cancer****Abstract**

**Purpose**—Determine inter-observer variability among radiologists in assigning Cambridge Classification (CC) of chronic pancreatitis (CP) based on magnetic resonance imaging (MRI)/magnetic resonance cholangiopancreatography (MRCP) and contrast-enhanced CT (CECT).

**Methods**—Among 422 eligible subjects enrolled into the PROCEED study between 6/2017 and 8/2018, 39 were selected randomly for this study (chronic abdominal pain (n=8; CC of 0), suspected CP (n=22; CC of 0, 1 or 2) or definite CP (n=9; CC of 3 or 4). Each imaging was scored by the local radiologist (LRs) and three of five central radiologists (CRs) at other consortium sites. The CRs were blinded to clinical data and site information of the participants. We compared the CC score assigned by the LR with the consensus CC score assigned by the CRs. The weighted kappa statistic (K) was used to estimate the inter-observer agreement.

**Results**—For the majority of subjects (34/39), the group assignment by LR agreed with the consensus composite CT/MRCP score by the CRs (concordance ranging from 75% to 89% depending on cohort group). There was moderate agreement (63% and 67% agreed, respectively) between CRs and LRs in both the CT score (weighted Kappa [95% CI] = 0.56 [0.34, 0.78]; p-value = 0.57) and the MR score (weighted Kappa [95% CI] = 0.68 [0.49, 0.86]; p-value = 0.72). The composite CT/MR score showed moderate agreement (weighted Kappa [95% CI] = 0.62 [0.43, 0.81]; p-value = 0.80).

**Conclusion**—There is a high degree of concordance among radiologists for assignment of CC using MRI and CT.

**Keywords**

Chronic Pancreatitis; Inter-observer Variability; Computerized Tomography; Magnetic Resonance Cholangiopancreatography

**Introduction**

Chronic pancreatitis (CP) is a chronic inflammatory disorder of the pancreas with abdominal pain as the predominant clinical symptom. Other manifestations of CP include acute pancreatitis, progressive pancreatic exocrine and endocrine insufficiency, as well as long-term complications such as pancreatic cancer [1–3]. Advanced CP is characterized by any

combination of calcifications, parenchymal atrophy, and fibrosis which in turn can lead to distortion or irregular dilation of the pancreatic duct (PD)[4]. Changes of advanced CP can be identified easily on imaging, but diagnosing early CP is challenging [5–7]. The understanding of the pathogenesis of CP is incomplete and many classifications have been proposed [8–11]. More recently, a mechanistic definition accepted by major international societies has been proposed, which recognizes the different stages during the evolution of CP and the role of genetic, environmental and other factors[7,12].

Prospective Evaluation of Chronic Pancreatitis for Epidemiologic and Translational Studies (PROCEED) is an ongoing multi-institutional prospective longitudinal cohort study being performed by the Consortium for the study of Chronic Pancreatitis, Diabetes, and Pancreatic Cancer (CPDPC)[13]. In addition to establishing a research cohort, the objectives of PROCEED are to define the risk of progression from suspected to definite CP, development of exocrine and or endocrine dysfunction, test the predictive capability of biomarkers of diagnosis and prognosis of CP, and to develop a platform to conduct translational, mechanistic and genetic studies. Subjects in PROCEED are enrolled into three sub cohorts, namely controls, suspected CP and definite CP to represent the spectrum of the natural history of CP[12]. Controls are subdivided into those with no pancreatic disease (Green I) and those with chronic abdominal pain of suspected pancreatic origin (Green II), while suspected CP sub cohort consists of three groups, including indeterminate CP (Yellow I), acute pancreatitis (Yellow II) and recurrent acute pancreatitis (Yellow III). Subjects with definite CP are assigned to the Red group. Except for no pancreas disease controls, the assignment of subjects in each of other groups is based on clinical history and specific findings on CT scan and/or MRI/MRCP (Table 1). Imaging studies for study participants are reviewed by a designated radiologist at each participating clinical center[13].

PROCEED investigators chose to use the Cambridge classification (CC) as the primary criteria for assigning subjects into specific groups (Table 2). CC, initially proposed over 3 decades ago to define the severity of CP on endoscopic retrograde cholangiopancreatography (ERCP), ultrasound and CT scan, has later been adapted for MRCP [14,8,1]. CC has been endorsed by the practice guidelines of the American Pancreatic Association[1]. The rationale and implications for using CC at baseline assessment and during follow-up in PROCEED have been published previously[15].

Since no prior study has systematically used CC to assign subjects into subgroups and follow them to determine disease progression similar to PROCEED, the primary aim of this study was to determine the validity of group assignment of subjects enrolled into PROCEED. Moreover, to the best of our knowledge, interobserver variability of CC by CT scan and MRI/MRCP has not been validated previously. Therefore, a secondary aim of our study was to evaluate intra- and inter-observer variability of radiologists for CC of CP using CT and MRCP.

## Materials and Methods

The PROCEED study was approved by the Institutional Review Boards of the participating institutions and all subjects provided an informed consent prior to enrollment. Detailed

methodology of the PROCEED study ([ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT03099850) ID NCT03099850) has been recently published[13]. Of 422 participants enrolled in PROCEED between June 2017 and August 2018, 39 subjects were randomly sampled from four categories of disease severity without reference to their data on the case report forms. There are four pre-defined cohort groups in PROCEED: chronic upper abdominal pain of suspected pancreatic origin or Green II (n = 8; CC = 0), indeterminate CP with no prior acute pancreatitis or Yellow I (n = 4; CC = 1 or 2), acute pancreatitis (Yellow II) or recurrent acute pancreatitis (Yellow III) (n = 18; CC = 0, 1, or 2), and definite CP or Red groups (n = 9; CC = 3 or 4) (Table 1). Anonymized images from CT scan and secretin-MRCP examination of each participant used for PROCEED group assignment were anonymized and uploaded to a HIPAA secure website for distribution to the consortium radiologists participating in this study. Each image was reviewed by three of five subspecialty trained abdominal radiologists (called Central Radiologists, or CR) from 5 different participating clinical centers under CPDPC. The radiologists at clinical center who contributed to the PROCEED cohort assignment at enrollment are called Local Radiologists, or LR.

### Imaging Technique

CT scans for all subjects were performed with intravenous contrast. Contiguous 5 mm or thinner axial sections were displayed. All subjects underwent contrast-enhanced pancreatic protocol MRI and all, but 7 patients underwent MRCP, with or without secretin, with standard imaging parameters (Table 3). The CT and MRI machines, type and amount of intravenous contrast administered for these exams varied based on institutional preferences.

### Image Analysis

The CRs were blinded to any clinical or demographic information as the images were completely anonymized. They interpreted the images using their preferred DICOM viewer on an independent workstation and recorded their findings to a secure database maintained by the CPDPC (<https://cpdpc.mdanderson.org>). Each imaging study was scored by 3 CRs, who were randomly selected from the participating 5 CRs. Each CR independently scored the CT scans and MRCP exams into one of three CC scoring categories: 0 (normal imaging appearance of pancreas), 1–2 (equivocal or findings of mild CP), and 3–4 (definitive imaging findings of CP by CC). An exception was that a scoring category of 3–4 was also assigned to images with definite pancreatic calcifications, independent of the score per CC. For each CT scan, the consensus CT score from the three CRs was considered the majority vote among the three scoring categories: 0, 1–2, and 3–4. In case the scores from the three CRs were: 0, 1–2 and 3–4, with no majority vote, we used the middle range, 1–2, as the consensus score. All CTs were decided by majority vote. The consensus MRI/MRCP score was defined similarly for each MRI/MRCP exam. Lack of majority vote was seen for only 1 MR exam for which consensus score of 1–2 was adopted. For each subject, the higher category between the consensus CT score and the consensus MRI/MRCP score was considered the consensus composite score of both CT and MRCP. One of the CT scans was unavailable for review due to technical issues and was not included in the analysis.

## Statistical Analyses

For CT images, we compared the consensus CT score by the CR and the Cambridge CT score by the LR in a 3 by 3 table for the 38 of the 39 subjects for whom CT was available for review. Similar 3 by 3 tables were produced for the consensus MRI/MRCP score, and the consensus composite score. Weighted Kappa coefficients are calculated to quantify agreement. For each of the four groups (Green II, Yellow I, Yellow II-III, Red), we calculated the proportion of patients whose consensus composite scores by the CR agreed with the enrollment criteria for imaging for that group, e.g. for Green II CC for both CT and MR was scored 0 by the LR as well as consensus of CRs. The weighted kappa statistic (K) was used to estimate the intra- and inter-observer agreement as following: poor (<0.4), moderate (0.4–0.6), good (0.6–0.8), or excellent (>0.8). SAS v9.4 (SAS Institute, Inc. Cary, NC) was used for statistical analysis.

## Results

Median age of the patients was 49 years (range 20–67 years); there were 18 males and 21 females. For the majority of the patients, the actual cohort assignment by the LR agrees with the consensus composite CT/MR score by the CRs, with concordance ranging from 75% to 88.9% (Table 4). Interestingly, lower concordance was seen only in the Indeterminate CP group, where the total number of subjects was 4. For all other groups, the concordance was excellent, i.e. >80%. Among the composite CC scores assigned by the LRs to the subjects included in the study, a score of 1 or 2 was the most common at 41% (n=16 of 39) followed by a score of 0 at 36% (n=14). Among the composite CC scores assigned by the CRs, scores of 0 and scores of 1 or 2 were equally common at 36% (n=14 each). There was moderate agreement (63% and 67% agreed, respectively) between CRs and LRs in both the CT score (weighted Kappa [95% CI] = 0.56 [0.34, 0.78]; p-value = 0.57) and the MR score (weighted Kappa [95% CI] = 0.68 [0.49, 0.86]; p-value = 0.72) (Table 5A and Table 5B). The composite CT/MR score showed moderate agreement (weighted Kappa [95% CI] = 0.62 [0.43, 0.81]; p-value = 0.80) (Table 5C). Agreement for each of the studies as well as the composite scores was good to excellent at the extremes, i.e. for scores of 3–4 or 0, but moderate for scores of 1–2.

## Discussion

We found high degree of concordance between the LR and CRs for assignment of subjects using CC of CP into PROCEED. Moreover, we also noted at least moderate agreement between the CRs and LR for CC of CP for both CT and MRI/MRCP. These data are an important step towards standardization of imaging reports for CP which has implications for research as well as clinical care of patients.

PROCEED is the first longitudinal cohort study of CP in the United States. Due to a lack of consensus regarding the diagnostic features of CP, PROCEED investigators chose CC on cross-sectional imaging as the primary criteria to objectively classify patients into different groups representing the spectrum of CP during its evolution. After initial enrollment, subjects will be followed longitudinally for progression based on imaging findings which will be correlated with clinical symptomatology (e.g. acute pancreatitis, diabetes, etc.) and

other factors (e.g. demographics, risk factors, etc.) that may play a role in determining the natural history of CP. Therefore, accurate classification of subjects into sub cohort groups is critical to make valid inference of the accumulated data. Therefore, our results demonstrating a high degree of concordance for cohort assignment between the LR and the consensus interpretation among the CRs is reassuring. In only one of the four groups, indeterminate CP, the concordance was <80% - however, the sample size of this group was too small to draw definitive conclusions. Since the primary outcome of PROCEED is transition to definite CP (CC 3 or 4) for which the concordance was high (89%), we do not feel that a lower concordance for CC 1–2 category will affect the primary results of the study. Moreover, whether and how often morphological changes in patients with CC 1–2 category evolve over time is an important clinical question that will be answered by secondary analyses of PROCEED data.

In addition to CC, PROCEED investigators will also capture granular data on parenchymal and ductal features as well as pancreas function on CT and MR images. This includes findings such as T1 signal intensity, duodenal filling after secretin, pancreatic ductal compliance with secretin, etc.[15]. We believe that such analyses as well as novel features on MR will address the limitations of CC and help in developing newer measures to characterize and standardize CT and MR changes aiding in diagnosis and follow-up of CP[16,17].

Ours is the first study to evaluate interobserver variability of CC on CT and MR. There was overall good to excellent agreement between LRs and CRs for assignment of subjects into PROCEED cohorts using CC based composite CT/MR scores. The high degree of concordance observed in our study between LRs and CRs for assignment of subjects into the definite CP (Red) and the Green (Chronic abdominal pain) categories is due to the lack of ambiguity in interpretation of either advanced features of CP such as pancreatic ductal obstruction, filling defects and severe irregularity or a normal appearing pancreas on imaging. Inclusion of pancreatic calcifications (not described in CC as one of the criteria for assigning patients into the definite CP category) may also have accounted for the high degree of concordance in the Red (Definite CP) group. A somewhat lower degree of concordance between LRs and CRs for assignment of subjects to indeterminate CP group (Yellow I) based on CC scores of 1 or 2, is likely related to ambiguous and subjective nature of some of the features included in CC such as “slight gland enlargement”, “heterogeneous parenchyma” and “abnormal main duct” (Figure 1).

Given limitations of CC, there is a need for a newer and more comprehensive, standardized and validated classification system for CP, the capitalizing on the recent advances in cross-sectional imaging[18]. The necessity for such a classification was endorsed by the international guidelines for imaging and severity scoring of CP[19]. The PROCEED study is adopting the reporting standards proposed by the CPDPC CP on cross-sectional imaging[15]. We believe that data from the PROCEED will be a step towards developing a new cross-sectional imaging-based classification system for CP.

Our study has a few limitations. The study population included is relatively small, especially in the Indeterminate CP group. Therefore, similar studies by other groups and in a larger

number of patients would be helpful to determine external validity of our observations. A recent small study reported on interobserver variability among radiologists for interpretation on CT findings of CP[20]. Although the study did not specifically evaluate CC, findings of this study show moderate agreement for CT scan findings and are reassuring. The CT and MRI technique was not uniform in all patients as the imaging was performed in different hospitals. However, all patients had contrast-enhanced CT and MRI exams using widely established imaging parameters.

## Conclusion

We found a high degree of concordance for group assignment of cross-sectional imaging-based Cambridge scores for CP by the LR and CRs in the PROCEED study. There was at least moderate agreement among radiologists for CC on CT and MR. These observations provide validity for the use of CC as the primary criteria for phenotyping subjects in the PROCEED study. It is anticipated that primary and secondary analyses of data from the study will add new knowledge to our understanding of morphologic evolution of CP.

## Acknowledgments

### Source of Funding

Research reported in this publication was supported by National Cancer Institute and National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under award numbers R01DK116963 (MINIMAP), U01DK108323 (IU), U01DK108306 (UPMC), U01DK108327 (OSU), U01DK108314 (CSMC), DKP3041301 (UCLA), U01DK108300 (Stanford) and U01DK108288 (Mayo). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

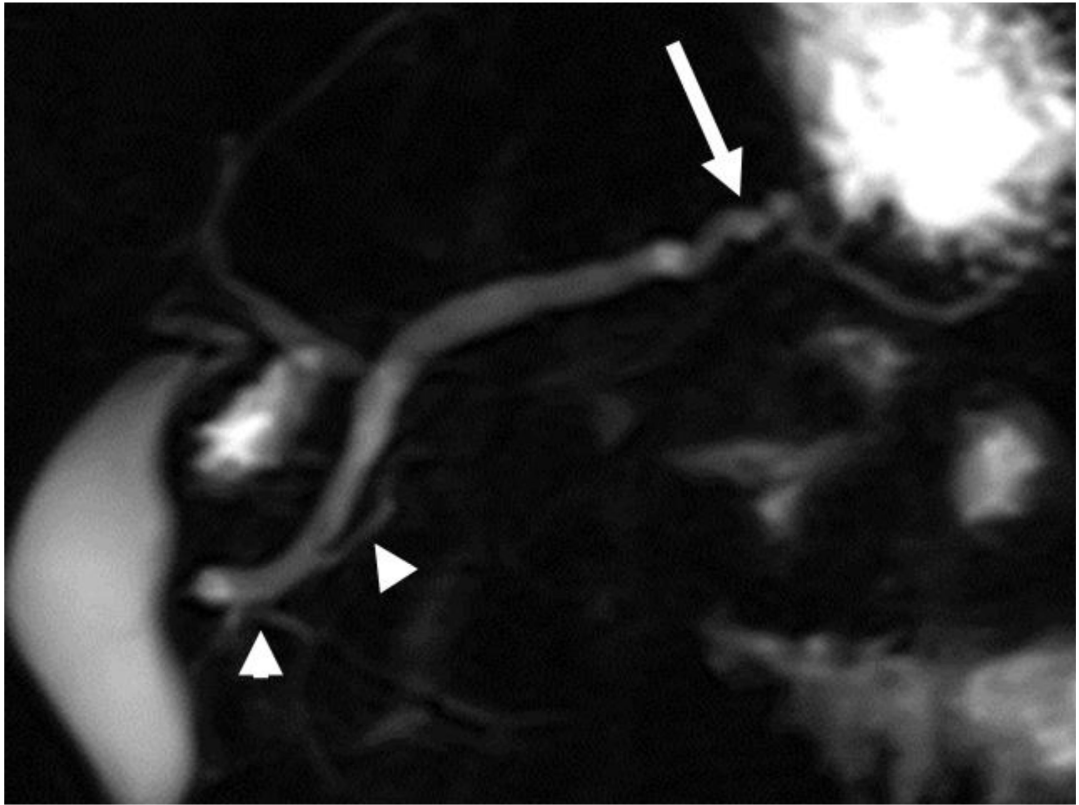
## References

1. Conwell DL, Lee LS, Yadav D, Longnecker DS, Miller FH, Morteale KJ, Levy MJ, Kwon R, Lieb JG, Stevens T (2014) American Pancreatic Association practice guidelines in chronic pancreatitis: evidence-based report on diagnostic guidelines. *Pancreas* 43 (8):1143–1162 [PubMed: 25333398]
2. Braganza JM, Lee SH, McCloy RF, McMahon MJ (2011) Chronic pancreatitis. *The Lancet* 377 (9772):1184–1197
3. Majumder S, Chari ST (2016) Chronic pancreatitis. *The Lancet* 387 (10031):1957–1966
4. Klöppel G, Detlefsen S, Feyerabend B (2004) Fibrosis of the pancreas: the initial tissue damage and the resulting pattern. *Virchows Archiv* 445 (1):1–8 [PubMed: 15138818]
5. Sainani NI, Kadiyala V, Morteale K, Lee L, Suleiman S, Rosenblum J, Wang W, Banks PA, Conwell DL (2015) Evaluation of Qualitative Magnetic Resonance Imaging Features for Diagnosis of Chronic Pancreatitis. *Pancreas* 44 (8):1280–1289 [PubMed: 26465953]
6. Choueiri NE, Balci NC, Alkaade S, Burton FR (2010) Advanced imaging of chronic pancreatitis. *Current gastroenterology reports* 12 (2):114–120 [PubMed: 20424983]
7. Whitcomb DC, Shimosegawa T, Chari ST, Forsmark CE, Frulloni L, Garg P, Hegyi P, Hirooka Y, Irisawa A, Ishikawa T (2018) International consensus statements on early chronic Pancreatitis. Recommendations from the working group for the international consensus guidelines for chronic pancreatitis in collaboration with The International Association of Pancreatology, American Pancreatic Association, Japan Pancreas Society, PancreasFest Working Group and European Pancreatic Club. *Pancreatology*
8. Schneider A, Löhr JM, Singer MV (2007) The M-ANNHEIM classification of chronic pancreatitis: introduction of a unifying classification system based on a review of previous classifications of the disease. *Journal of gastroenterology* 42 (2):101–119 [PubMed: 17351799]

9. Catalano MF, Sahai A, Levy M, Romagnuolo J, Wiersema M, Brugge W, Freeman M, Yamao K, Canto M, Hernandez LV (2009) EUS-based criteria for the diagnosis of chronic pancreatitis: the Rosemont classification. *Gastrointestinal endoscopy* 69 (7):1251–1261 [PubMed: 19243769]
10. Shimosegawa T, Kataoka K, Kamisawa T, Miyakawa H, Ohara H, Ito T, Naruse S, Sata N, Suda K, Hirota M (2010) The revised Japanese clinical diagnostic criteria for chronic pancreatitis. *Journal of gastroenterology* 45 (6):584–591 [PubMed: 20422433]
11. Ammann RW (1997) A clinically based classification system for alcoholic chronic pancreatitis: summary of an international workshop on chronic pancreatitis. *Pancreas* 14 (3):215–221 [PubMed: 9094150]
12. Whitcomb DC, Frulloni L, Garg P, Greer JB, Schneider A, Yadav D, Shimosegawa T (2016) Chronic pancreatitis: an international draft consensus proposal for a new mechanistic definition. *Pancreatology* 16 (2):218–224 [PubMed: 26924663]
13. Yadav D, Park WG, Fogel EL, Li L, Chari ST, Feng Z, Fisher WE, Forsmark CE, Jeon CY, Habtezion A (2018) PROspective Evaluation of Chronic Pancreatitis for EpidEmiologic and Translational StudIes: Rationale and Study Design for PROCEED From the Consortium for the Study of Chronic Pancreatitis, Diabetes, and Pancreatic Cancer. *Pancreas* 47 (10):1229–1238 [PubMed: 30325862]
14. Sarner M, Cotton P (1984) Classification of pancreatitis. *Gut* 25 (7):756–759 [PubMed: 6735257]
15. Tirkes T, Shah ZK, Takahashi N, Grajo JR, Chang ST, Venkatesh SK, Conwell DL, Fogel EL, Park W, Topazian M, Yadav D, Dasyam AK, For the Consortium for the Study of Chronic Pancreatitis D, Cancer P (2019) Reporting Standards for Chronic Pancreatitis by Using CT, MRI, and MR Cholangiopancreatography: The Consortium for the Study of Chronic Pancreatitis, Diabetes, and Pancreatic Cancer. *Radiology* 290 (1):207–215. doi:10.1148/radiol.2018181353 [PubMed: 30325281]
16. Tirkes T, Yadav D, Conwell DL, Territo PR, Zhao X, Venkatesh SK, Kolipaka A, Li L, Pisegna JR, Pandol SJ, Park WG, Topazian M, Serrano J, Fogel EL, Consortium for the Study of Chronic Pancreatitis D, Pancreatic C (2019) Magnetic resonance imaging as a non-invasive method for the assessment of pancreatic fibrosis (MINIMAP): a comprehensive study design from the consortium for the study of chronic pancreatitis, diabetes, and pancreatic cancer. *Abdom Radiol (NY)* 44 (8):2809–2821. doi:10.1007/s00261-019-02049-5 [PubMed: 31089778]
17. Dasyam AK, Vipperla K, Slivka A, Gong T, Papachristou GI, Whitcomb DC, Yadav D (2019) Computed tomography based scoring system in a prospectively ascertained cohort of patients with chronic pancreatitis. *Pancreatology* 19 (8):1027–1033 [PubMed: 31630919]
18. Dasyam AK, Shah ZK, Tirkes T, Dasyam N, Borhani AA (2019) Cross-sectional imaging-based severity scoring of chronic pancreatitis: why it is necessary and how it can be done. *Abdominal Radiology*:1–11 [PubMed: 29967984]
19. Frøkjær JB, Akisik F, Farooq A, Akpınar B, Dasyam A, Drewes AM, Haldorsen IS, Morana G, Neoptolemos JP, Olesen SS (2018) Guidelines for the Diagnostic Cross Sectional Imaging and Severity Scoring of Chronic Pancreatitis. *Pancreatology* 18 (7):764–773 [PubMed: 30177434]
20. Razeq AAKA, Elfar E, Abubacker S (2019) Interobserver agreement of computed tomography reporting standards for chronic pancreatitis. *Abdominal Radiology*:1–7 [PubMed: 29967984]



(a)



(b)



**Figure 1.**

2D thick-slab MRCP image (*a*) obtained 2 minutes after secretin administration demonstrates smooth contour of a prominent main pancreatic duct, which is tortuous (white arrow) in the region of pancreatic body. Note 2 abnormal branch ducts in the pancreatic head/neck (arrowheads). An axial image (*b*) from contrast-enhanced CT scan of the same patient demonstrates the prominent and slightly tortuous pancreatic duct (white arrows) in the body of pancreas. The apparent contour irregularity of PD from the tortuosity likely resulted in 2 of the 3 CRs categorizing the CT as grade 3 per CC with composite CT/MR score of 3, discordant from the corresponding composite CT/MR score of 1 by the LR.

**Table 1.** Inclusion criteria for clinical symptomatology and imaging findings for PROCEED groups relevant to this analysis

	Chronic Abdominal pain of suspected pancreatic origin (Green II) <sup>(***)</sup>	Indeterminate CP (Yellow I) <sup>(***)</sup>	AP (Yellow II) or RAP (Yellow III) <sup>(***)</sup>	Definite CP (Red) <sup>(***)</sup>
Clinical symptoms	Chronic abdominal pain	Chronic abdominal pain	Clinical symptoms yes or no	Clinical symptoms yes or no
Cambridge score on CT	0	1 or 2	0, 1 or 2	3 or 4 or pancreatic calcifications
Cambridge score on MR	0	1 or 2	0, 1 or 2	3 or 4 or pancreatic calcifications
Composite Cambridge score on CT or MR <sup>*</sup>	0	1 or 2	0, 1 or 2	3 or 4 or pancreatic calcifications

AP= Acute pancreatitis, RAP= Recurrent acute pancreatitis, CP=Chronic pancreatitis

<sup>(\*)</sup> Higher score between CT or MR

<sup>(\*\*\*)</sup> Subjects in Green II and Yellow I-III are required to have both CT and MR at enrollment; subjects in Red group are not required to have both CT and MR if one of the imaging studies shows calcifications. Detailed inclusion and exclusion criteria for PROCEED groups are published in Yadav, Dhiraj, et al. "Prospective Evaluation of Chronic Pancreatitis for Epidemiologic and Translational Studies: Rationale and Study Design for PROCEED From the Consortium for the Study of Chronic Pancreatitis, Diabetes, and Pancreatic Cancer." *Pancreas* 47.10 (2018): 1229-1238.

**Table 2:**

**CT/MRI/MRCP Grading of CP based on Cambridge classification**

	<b>Equivocal (I)</b>	<b>Mild (II)</b>	<b>Moderate (III)</b>	<b>Marked (IV)</b>
<b>CT/MRI Scan</b>	One of the following: main duct enlarged (2–4 mm), slight gland enlargement, heterogeneous parenchyma, small cavities (<10mm), irregular ducts, focal pancreatitis, increased echogenicity of main duct wall, irregular head/body contour	>=2 of the following: enlarged main duct (2–4mm), gland enlargement, heterogeneous parenchyma, small cavities (<10mm), irregular ducts, focal acute pancreatitis, increased echogenicity of main duct wall, irregular head/body contour	Cannot distinguish from mild	Moderate changes plus >=1 of the following: large cavities (>10mm), gland enlargement, intraductal filling defects/calculi, duct obstruction, stricture or gross irregularity
<b>MRCP</b>	<3 abnormal side branch changes	>=3 abnormal side branch changes	Abnormal main duct; >3 abnormal side branches	Moderate changes plus 1 of the following: obstruction, filling defects, severe irregularity/dilatation of main duct

Adapted from Conwell, Darwin L., et al. "American Pancreatic Association Practice Guidelines in Chronic Pancreatitis: evidence-based report on diagnostic guidelines." *Pancreas* 43.8 (2014): 1143.

**Table 3.**

Pancreatic protocol MRI parameters

T2-weighted images (Fast-spin-echo (FSE), turbo-spin-echo (TSE) or a variant: axial, without fat suppression
T2-weighted images with fat suppression: axial, without fat suppression
Axial T1-weighted, 3-D, 2-point Dixon gradient echo sequence
Post-contrast T1 weighted images: T1-weighted 3-D gradient-echo with fat suppression in pre-contrast, arterial, portal venous and 5-minute delayed phases
MRCP examination: <ul style="list-style-type: none"> <li>• 2D MRCP (thick slab), 40 mm thick, 8 paracoronal projections.</li> <li>• 3D MRCP (thin slab): 2–3 mm, respiratory synchronized 3-D TSE sequence.</li> <li>• Secretin-stimulated 2D (thick slab) MRCP: coronal single shot FSE sequence or a variant, repeated every 30 seconds for 10 minutes after intravenous infusion of Secretin.</li> </ul>

Agreement between the actual cohort assignment and the consensus composite CT/MR score by the Central Radiologists

**Table 4.**

Patient cohort	Cambridge score by the enrollment criteria	Number of participants	Number (%) of patients with consensus composite CT/MR score satisfying enrollment criteria
<b>Green II</b> (Chronic Abdominal pain of suspected pancreatic origin)	0	8	7 (87.5%)
<b>Yellow I</b> (Indeterminate CP)	1,2	4	3 (75.0%)
<b>Yellow II and III</b> AP (Yellow II) or RAP (Yellow III)	0,1,2	18	16 (88.9%)
<b>Red</b> (Definite CP)	3,4	9	8 (88.9%)
<b>Total</b>	-	39	34 (87.2%)

AP= Acute pancreatitis. RAP= Recurrent acute pancreatitis. CP=Chronic pancreatitis

**Table 5A:** Agreement between Central Radiologists (CR) and Local Radiologists (LR) on the CT score.

Frequency	Consensus CT score by CR			Total
	0	1-2	3-4	
0	15	6	1	22
1-2	3	3	2	8
3-4	0	2	6	8
<b>Total</b>	18	11	9	38 <sup>(*)</sup>

<sup>(\*)</sup> One CT image was removed due to failure to review for technical issues

Agreement between Central Radiologists (CR) and Local Radiologists (LR) on the MR score

Table 5B:

Frequency	Consensus MR score by CR			Total
	0	1-2	3-4	
0	11	5	0	16
MR score by LR 1-2	7	8	0	15
3-4	0	1	7	8
Total	18	14	7	39



**Table 5C:** Agreement between Central Radiologists (CR) and Local Radiologists (LR) on the composite CT/MR score

Frequency	Consensus composite CT/MR score by CR			Total
	0	1-2	3-4	
0	9	5	0	14
Composite CT/MR score by LR	1-2	8	3	16
	3-4	1	8	9
Total	14	14	11	39