# Describing Peripancreatic Collections According to the Revised Atlanta Classification of Acute Pancreatitis

## *An International Interobserver Agreement Study*

*Stefan A. Bouwense, MD, PhD,* Sandra van Brunschot, MD,† Hjalmar C. van Santvoort, MD, PhD,‡ Marc G. Besselink, MD, PhD,‡ Thomas L. Bollen, MD,§ Olaf J. Bakker, MD, PhD,|| Peter A. Banks, MD,¶ Marja A. Boermeester, MD, PhD,‡ Vincent C. Cappendijk, MD, PhD,# Ross Carter, MD,** Richard Charnley, MD, PhD,†† Casper H. van Eijck, MD, PhD,‡‡ Patrick C. Freeny, MD,§§ John J. Hermans, MD, PhD,|||| David M. Hough, MD,¶¶ Colin D. Johnson, MD,## Johan S. Laméris, MD, PhD,*** Markus M. Lerch, MD, PhD,††† Julia Mayerle, MD, PhD,††† Koenraad J. Mortele, MD,‡‡‡ Michael G. Sarr, MD,§§§ Brian Stedman, MD,|||||| Santhi Swaroop Vege, MD,¶¶¶ Jens Werner, MD, PhD,### Marcel G. Dijkgraaf, MD, PhD,**** Hein G. Gooszen, MD, PhD,* and Karen D. Horvath, MD,†††† for the Acute Pancreatitis Interobserver Study Group*

**Objectives:** Severe acute pancreatitis is associated with peripancreatic morphologic changes as seen on imaging. Uniform communication regarding these morphologic findings is crucial for accurate diagnosis and treatment. For the original 1992 Atlanta classification, interobserver agreement is poor. We hypothesized that for the revised Atlanta classification, interobserver agreement will be better.

**Methods:** An international, interobserver agreement study was performed among expert and nonexpert radiologists (n = 14), surgeons (n = 15), and gastroenterologists (n = 8). Representative computed tomographies of all stages of acute pancreatitis were selected from 55 patients and were assessed according to the revised Atlanta classification. The interobserver agreement was calculated among all reviewers and subgroups, that is, expert and nonexpert reviewers; interobserver agreement was defined as poor (≤0.20), fair (0.21–0.40), moderate (0.41–0.60), good (0.61–0.80), or very good (0.81–1.00).

**Results:** Interobserver agreement among all reviewers was good (0.75 [standard deviation, 0.21]) for describing the type of acute pancreatitis and good (0.62 [standard deviation, 0.19]) for the type of peripancreatic collection. Expert radiologists showed the best and nonexpert clinicians the lowest interobserver agreement.

**Conclusions:** Interobserver agreement was good for the revised Atlanta classification, supporting the importance for widespread adaption of this revised classification for clinical and research communications.

**Key Words:** acute pancreatitis, peripancreatic collections, revised Atlanta classification, interobserver agreement

(*Pancreas* 2017;46: 850–857)

From the *Department of OR/Clinical Surgical Research, Radboud university medical center, Nijmegen; Departments of †Gastroenterology and Hepatology, and ‡Surgery, Academic Medical Center, Amsterdam; §Department of Radiology, St. Antonius Hospital, Nieuwegein; ||Department of Surgery, University Medical Center Utrecht, Utrecht, The Netherlands; ¶Department of Gastroenterology, Center for Pancreatic Disease, Brigham and Women's Health Hospital, Harvard Medical School, Boston, MA; #Department of Radiology, Jeroen Bosch Hospital, 's-Hertogenbosch, The Netherlands; **Department of Surgery, Glasgow Royal Infirmary, Glasgow; ††Department of Surgery, Freeman Hospital, Newcastle Upon Tyne, United Kingdom; ‡‡Department of Surgery, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands; §§Department of Radiology, University of Washington Medical Center, Seattle, WA; ||||Department of Radiology, Radboud university medical center, Nijmegen, The Netherlands; ¶¶Department of Radiology, Mayo Clinic, Rochester, MN; ##Department of Surgery, University Hospital Southampton, Hampshire, United Kingdom; ***Department of Radiology, Academic Medical Center, Amsterdam, The Netherlands; †††Department of Medicine A, University Medicine Greifswald, Germany; ‡‡‡Department of Radiology, Beth Israel Deaconess Medical Center, Boston, MA; §§§Department of Surgery, Mayo Clinic, Rochester, MN; ||||||Department of Radiology, University Hospital Southampton, Hampshire, United Kingdom; ¶¶¶Department of Gastroenterology, Mayo Clinic, Rochester, MN; ###Department of Surgery, Ludwig Maximilian University of Munich, Munich, Germany; ****Department of Clinical Research Unit, Academic Medical Center, Amsterdam, The Netherlands; and ††††Department of Surgery, University of Washington Medical Center, Seattle, WA.

Received for publication July 24, 2016; accepted April 21, 2017.

Address correspondence to: Karen D. Horvath, MD, Department of Surgery, University of Washington, 1959 NE Pacific St, Box 365410, Seattle, WA 98195 (e-mail: khorvath@UW.edu).

The authors declare no conflict of interest. All authors have read and approved the article.

Supplemental digital contents are available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.pancreasjournal.com).



DOI: 10.1097/MPA.0000000000000863

Severe acute pancreatitis is often associated with pancreatic and peripancreatic morphologic changes. These changes may consist of peripancreatic edema, peripancreatic necrosis, pancreatic parenchymal necrosis, and different types of peripancreatic collections containing variable amounts of fluid and/or necrosis; these pancreatic and peripancreatic collections may become infected and require intervention. Contrast-enhanced computed tomography (CECT) is used widely to evaluate morphologic changes, which are then correlated with clinical parameters to lead to a disease classification and a treatment plan. Therefore, decisions for treatment are often mainly based on CT findings.[1–3]

The 1992 Atlanta classification was a large step forward in its era.[4] However, with the remarkable advances in imaging technology and patient care, the usefulness of the 1992 Atlanta classification has been challenged.[5] One problem with the 1992 clinically-based approach to the classification was the inability to reproducibly translate morphologic computed tomography (CT) findings into a clear classification resulting in confusion. In fact, when the terms used in the 1992 Atlanta classification were studied for their ability to reliably describe peripancreatic collections on CT, poor interobserver agreement was noted among 5 abdominal radiologists, questioning its clinical usefulness.[6] In the early 2000s, a new set of morphologic terms was developed to describe peripancreatic collections in acute pancreatitis on CT, identified by the acronym "PANCODE," which stands for pancreatic nonenhancement, collection descripition.[7] It showed good to excellent interobserver agreement among internationally recognized experts, such as gastroenterologists, surgeons, and radiologists.[7] Then, in 2012,

the Atlanta classification was revised.[8] The revised Atlanta classification incorporated in part the PANCODE morphologic descriptors and combined these and other better-defined terms with clinical parameters to create a new classification that aimed to facilitate communication among treating physicians and between institutions.[8] Whether the revised Atlanta classification offers reliable interobserver agreement is unknown.

We hypothesized that the interobserver agreement among expert and nonexpert radiologists and clinicians has improved with the revised Atlanta classification. Better classification systems should lead to a more objective and accurate communication among physicians, a more uniform clinical decision-making, and a more accurate research communication, which will help improve patient outcomes. The primary aim of this study was to determine the interobserver agreement of the revised Atlanta classification and to investigate the reliability of translating CT morphology into the terms of the revised Atlanta classification for acute pancreatitis.

## MATERIALS AND METHODS

### Study Population

All abdominal CECTs from patients with predicted severe acute pancreatitis (acute physiology and chronic health evaluation II, >7; Imrie score, >2; C-reactive protein, >150) from 2 Dutch multicenter trials, the PROPATRIA[9] and PANTER trials,[10] were used for this study. All CTs were reviewed by 1 experienced abdominal radiologist (T.L.B.). For every patient, the CT severity index (CTSI; range, 0–10 points) was determined.[3,11,12] Representative CTs of all stages of acute pancreatitis were selected based on the following criteria: use of iodinated contrast material in the pancreatic and/or portal venous phase and availability of a Digital Imaging and Communications in Medicine format (AccuImage Diagnostics Corporation, AccuLite, Version 3.116; San Francisco, Calif) (required for full digital review). Only CTs of patients without an intervention to their peripancreatic collections was used for optimal visualization of their collections.

The design of this study and selection of the CTs was similar to our previous interobserver study.[7]

### Reviewers

Two groups of international reviewers were formed, with an equal distribution of experts and nonexperts to represent clinical care in expert centers and community hospitals. Both experts and nonexperts were subdivided into the following subgroups of reviewers: (1) expert radiologists, (2) expert clinicians, (3) nonexpert radiologists, and (4) nonexpert clinicians (Fig. 1).

"Expert" was defined as a surgeon, gastroenterologist, or radiologist with at least 10 publications on pancreatic diseases, working in a specialized pancreatic diseases unit.

"Nonexpert" was defined as a surgeon, gastroenterologist, or radiologist, with working experience in the gastrointestinal field but no specific scientific interest in pancreatic diseases as evidenced by no publications on this topic.

### Data Collection

Two investigators (S.v.B. and S.A.B.) visited the participating centers and had meetings with the clinicians and radiologists. The study format was fully standardized. First, reviewers were given time to read a sheet with the definitions of pancreatic and extrapancreatic collections, according to the revised Atlanta classification. This was followed by a short PowerPoint presentation explaining the study endpoints, the scoring sheet, and software used. Reviewers were given time to ask additional questions about the reviewing process before they started with the review of the first CT. During the scoring of CTs, the investigators were not present; they were, however, available solely for technical issues but not for explanation or interpretation of the PANCODE terms or definitions of the revised Atlanta classification. During the review process, the definitions of the revised Atlanta classification remained available.

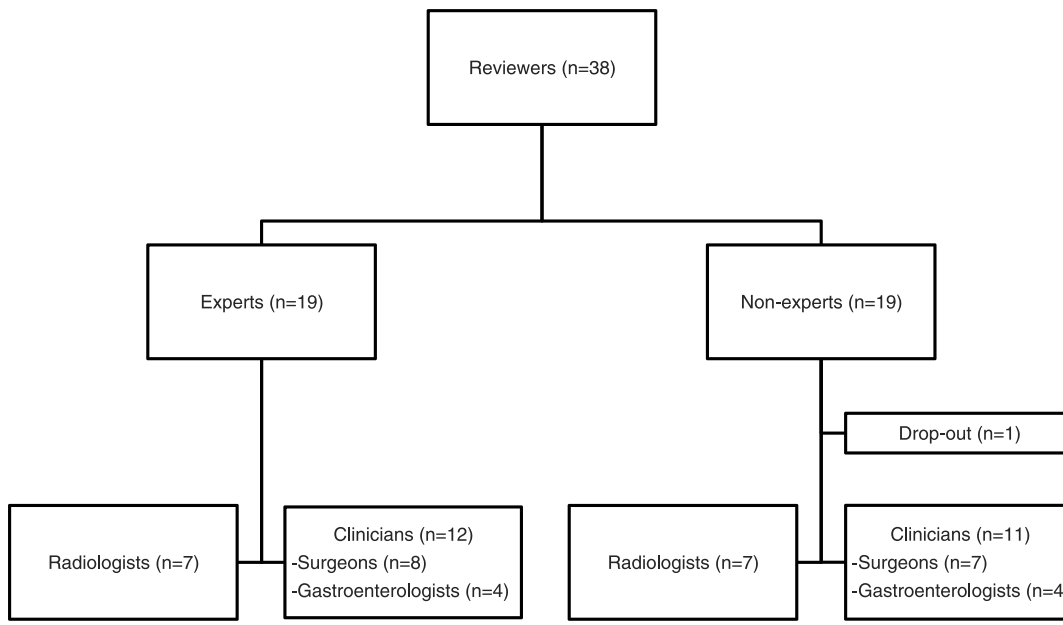


**FIGURE 1.** Distribution of reviewers. The group of experts was an international panel of surgeons, gastroenterologists, and radiologists. All are considered experts in acute pancreatitis as evidenced by at least 10 publications on pancreatic diseases and working in a specialized pancreatic diseases unit. The group of nonexperts was an international panel of surgeons, gastroenterologists, and radiologists. They have working experience in the gastrointestinal field but no special scientific interest in pancreatic diseases as evidenced by no publications on this topic. One nonexpert dropped out because of personal reasons.

Extent of **PA**ncreatic **N**onenhancement?
- Yes
  - <30%
  - 30-50%
  - >50%
- No

Is there a **CO**llection?
- No
- Yes
  - If 'yes', please choose one **DE**scription per question:
    - Relation with pancreas:
      - Intrapancreatic only
      - Intrapancreatic and adjacent to pancreas
      - Only adjacent to pancreas (no parenchymal perfusion defect)
      - Separate
    - Encapsulation:
      - Complete
      - Partial
      - None
    - Content:
      - Homogeneous
      - Heterogeneous (including fat, hermorrhage, loculation/septa or densities higher than fluid)
    - Mass effect (=displacement of adjacent structures: vessels, organs etc.)
      - Yes
      - No
    - Shape:
      - Round or oval
      - Irregular
    - Loculated gas bubbles:
      - Yes
      - No
    - Gas-fluid level:
      - Yes
      - No

**FIGURE 2.** Scoring sheet about the descriptive and morphologic terms to evaluate on CT (PANCODE).

After the instruction, reviewers assessed the CTs and recorded their data on a scoring sheet:

1. First, the reviewers were asked to describe the collection according to the PANCODE criteria.
2. Next, the timing of that particular CT was revealed as time in days after onset of symptoms of acute pancreatitis, with emphasis on whether the CT was obtained less than or greater than 4 weeks after onset of symptoms.
3. The reviewers then completed the second part of the sheet for the same CT and assigned a classification using the revised Atlanta classification definitions.

This process was repeated for all 55 CTs. In case of multiple collections, the reviewer was asked to describe the clinically most relevant collection.

## Scoring Sheets

A scoring sheet was used to mimic the following 2-step process involved in disease classification as performed by a treating clinician: step one, to determine the morphologic classification of CT findings; and step two, to combine morphology with clinical parameters to come up with a classification. The first section of each scoring sheet contained descriptive, morphologic terms (PANCODE) (Fig. 2).[7] The following are the terms of PANCODE that are used to describe the morphological changes: extent of pancreatic nonenhancement, relation with pancreas, encapsulation, content, mass effect, shape, loculated gas bubbles, and gas-fluid level.

The second section of each scoring sheet contained the definitions of the revised Atlanta classification for collections in acute pancreatitis (Fig. 3 and Supplementary Table 1, http://links.lww.com/MPA/A597).[8] First, the type of acute pancreatitis needs to be determined (interstitial edematous pancreatitis, necrotizing pancreatitis, or indeterminate), then the type of peripancreatic collection (acute peripancreatic fluid collection, acute necrotic collection, pancreatic pseudocyst, walled-off necrosis, or indeterminate).

During the reviewing process, the reviewers were not allowed to use any other sources of information than the information provided on the study sheets.

## Data Analysis

No formal sample size was calculated because of the uncertainty about the expected interobserver agreement between groups. We used a convenient sample of CTs and observers, which was similar to one of our previous studies.[7]

For every item on the scoring sheet, the distribution (eg, 20% and 80%) of options (eg, "yes" and "no") within the 55 CTs was assessed for each reviewer individually and described by medians with interquartile ranges (IQRs).

To prevent biased or hypothetic estimates of interobserver agreements, we refrained from frequently used approaches such as "percentage agreement," "$\kappa$," or "prevalence-adjusted bias-adjusted $\kappa$." Instead, we applied the ratio (ratio $\kappa_{max}$) between the observed $\kappa$ and the maximally achievable $\kappa$ given the presence of factors constraining observers in their ability to actually agree or disagree beyond chance ($\kappa_{max}$).[13] Calculation based on the other approaches is provided in Supplementary Table 2, http://links.lww.com/MPA/A597.

Subgroup analysis was performed in the case of normal distribution of individual measurements by analysis of variance with post-hoc Bonferroni analysis for multiple comparisons; otherwise, a Kruskal-Wallis with post-hoc Mann-Whitney *U* analysis was performed. *P* values less than 0.05 were considered statistically significant. The following are the subgroups of reviewers

Determine the type of acute pancreatitis (choose one option):
- Intersitial edematous pancreatitis
- Necrotizing pancreatitis
- Indeterminate

Determine the type of peripancreatic collection (choose one option):
- Acute Fluid Collection (AFC)
- Acute Necrotic Collection (ANC)
- Pancreatic Pseudocyst
- Walled-Off Necrosis (WON)
- Indeterminate

**FIGURE 3.** Scoring sheet about the type of acute pancreatitis and peripancreatic collections as defined by the revised Atlanta classification definitions.

that were compared: (1) expert radiologists, (2) expert clinicians, (3) nonexpert radiologists, and (4) nonexpert clinicians.

An interobserver agreement of 0.81 to 1.00 was defined as very good agreement; 0.61 to 0.80, good agreement; 0.41 to 0.60, moderate agreement; 0.21 to 0.40, fair agreement; and less than 0.20, poor agreement.[14]

## RESULTS

From a total of 248 patients, 55 CTs were included to cover the complete spectrum of morphologic changes in acute pancreatitis, with emphasis on severe disease (ie, presence of peripancreatic collections and parenchymal necrosis). Thirty CTs from 30 consecutive patients who did not undergo drainage and/or operative therapy after their CT were selected (5 patients with a CTSI of 1–2, 5 with a CTSI of 3–4, 5 with a CTSI of 5–6, and 15 with a CTSI of 7–10) and 25 consecutive patients with a CT before intervention (percutaneous drainage or operative interventions) or infected necrosis (irrespective of CTSI).

For this selection of patients, the median time between admission and CT was 18 days (IQR, 9–32 days). Of the 55 patients, 60% had infected necrosis proven by bacterial culture (obtained with fine-needle aspiration or during first intervention). Twenty-five of those patients required operative therapy, and 8 patients underwent solely percutaneous catheter drainage. Nine of these 55 patients died (16%).

Thirty-seven reviewers analyzed the CTs (14 radiologists, 15 surgeons, and 8 gastroenterologists) and were subdivided into the following 4 groups: expert radiologists (n = 7), expert clinicians (n = 12), nonexpert radiologists (n = 7), and nonexpert clinicians (n = 11).

The distribution of the scored options by all reviewers, including subgroups of reviewers, for the revised Atlanta classification definitions is shown in Table 1. The reviewers scored necrotizing pancreatitis (median, 76%) as the type of pancreatitis in most cases and the peripancreatic collections as acute necrotic collections (median, 45%) or walled-off necrosis (median, 25%). The observer could not score the precise type of peripancreatic collections in a median of 7% of cases. One of the CTs that were reviewed is shown in Figure 4.

In Supplementary Tables 3, 4, and 5, http://links.lww.com/MPA/A597, the scored items for the PANCODE terms are mentioned.

### Interobserver Agreement

Among all reviewers, the interobserver agreement for the definitions of the revised Atlanta classification was good for the type of acute pancreatitis and good for the type of peripancreatic collection. Interobserver agreement among the expert radiologists was very good for the type of acute pancreatitis and type of peripancreatic collection. For the nonexpert radiologists, interobserver agreement was good for the type of acute pancreatitis and type of peripancreatic collection. The expert clinicians showed good interobserver agreement for the type of acute pancreatitis and moderate interobserver agreement for type of peripancreatic collection. For the nonexpert clinicians, interobserver agreement was good for the type of acute pancreatitis and moderate for type of peripancreatic collection (Table 2).

### Scored Items for All Subgroups

Among the subgroups, interobserver agreement was good among all subgroups with type of acute pancreatitis showing very good to good interobserver agreement and moderate to very good agreement for the type of peripancreatic collection (Table 2, Figs. 5, 6).

The greatest interobserver agreement among all scored items was found in the expert radiologists groups, followed by the nonexpert radiologists and expert clinicians. Nonexpert clinicians showed the least interobserver agreement of all subgroups. Comparisons between subgroups are shown in Table 3.

## DISCUSSION

Our international, multidisciplinary interobserver agreement study showed good interobserver agreement for type of acute pancreatitis and good interobserver agreement for defining the type of peripancreatic collection, according to the definitions of the revised Atlanta classification. This represents a significant and notable step forward in interobserver agreement over the poor interobserver agreement among clinicians noted for the original 1992 Atlanta classification. The ability to classify patients and to agree on that classification is much better with the revised Atlanta classification and lends strong support for its widespread adaption in both the clinical and research communications.

The clinical course of acute pancreatitis can be mild with little morbidity and very low mortality rates, whereas severe acute pancreatitis is marked by high morbidity and mortality.[8,15] Both disease courses are associated with a variety of morphologic

**TABLE 1.** The Distribution of the Scored Options for the Revised Atlanta Classification Definitions

| | Expert | | Nonexpert | | |
|---|---|---|---|---|---|
| **Term** | **Radiologists (n = 7)** | **Clinicians (n = 12)** | **Radiologists (n = 7)** | **Clinicians (n = 11)** | **Overall (n = 37)** |
| Type of pancreatitis | | | | | |
| Interstitial edematous pancreatitis | 17 (12–23) | 22 (20–25) | 31 (25–36) | 20 (16–24) | 22 (18–26) |
| Necrotizing pancreatitis | 81 (75–88) | 78 (75–80) | 69 (64–75) | 76 (75–80) | 76 (71–81) |
| Indeterminate | 2 (0–2) | 0 (0–0) | 0 (0–2) | 2 (0–7) | 0 (0–2) |
| Type of peripancreatic collection | | | | | |
| Acute fluid collection | 12 (8–16) | 15 (13–20) | 18 (16–25) | 15 (11–18) | 13 (10–20) |
| Acute necrotic collection | 53 (46–62) | 49 (47–53) | 45 (40–47) | 47 (38–51) | 45 (39–54) |
| Pancreatic pseudocyst | 0 (0–2) | 0 (0–0) | 0 (0–4) | 0 (0–0) | 0 (0–2) |
| Walled-off necrosis | 25 (20–32) | 27 (25–29) | 24 (18–28) | 29 (27–31) | 25 (18–29) |
| Indeterminate | 6 (4–11) | 7 (4–9) | 9 (2–13) | 7 (4–16) | 11 (5–16) |

All values are in percentages and are described by medians (IQR).

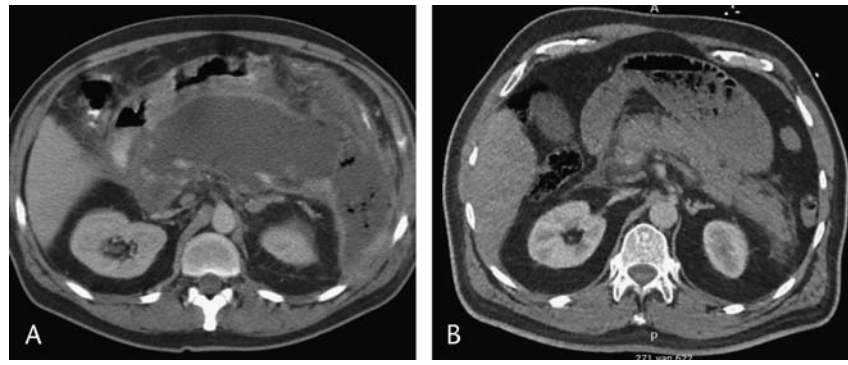


**FIGURE 4.** Two of the 55 CTs reviewed in this interobserver agreement study. A, Most of the reviewers described this CT as >50% pancreatic nonenhancement with a collection that is intrapancreatic and adjacent to the pancreas, which is partially encapsulated, heterogeneous, with mass effect, irregular shaped, with loculated gas bubbles, and without a gas-fluid level. All reviewers defined the type of pancreatitis as necrotizing pancreatitis. The CT was obtained 21 days after onset of disease. Most of the reviewers defined this as necrotizing pancreatitis with a collection defined as an acute necrotic collection. B, Most of the reviewers described this CT as >50% pancreatic nonenhancement with a collection that is adjacent to the pancreas, which has no encapsulation, heterogenous, no mass effect, irregular shaped, with no loculated gas bubbles, and without a gas-fluid level. The CT was obtained at day 10 after onset of disease. Most of the reviewers defined the type of pancreatitis as necrotizing pancreatitis with an acute necrotic collection.

changes in the pancreatic and peripancreatic region, all necessitating their own treatment.[8] One of the pitfalls of the 1992 Atlanta classification was that it was a predominantly clinically based system. With advancements in diagnostic imaging over the past 20 years, many management decisions were made based solely on the evaluations of the CTs. This approach meant that radiologists and clinicians were left with trying to translate the ill-defined CT morphology of the patient with pancreatitis into a clinical classification. The result was poor interobserver agreement, leading to confused classifications in both the clinical and research communications. The value of the revised Atlanta classification is that it combines both morphologic and clinical criteria to create a more useful classification system of pancreatitis. Because correct interpretation of CT morphology combined with the clinical status of the patient determines treatment,[16] it is to be expected that with use of the revised Atlanta classification, patient treatment will be improved.

Our study also touched on generalizability issues. Two other similar interobserver studies were performed before this study, using similar methodology. In the first interobserver study, the CTs of patients with necrotizing pancreatitis were scored with the original 1992 Atlanta classification terms[4] by 5 expert radiologists. Interobserver agreement was poor; the generalizability of the terms was unknown.[6] The second interobserver study showed good to excellent agreement for the PANCODE terms among expert radiologists and clinicians (6); however, because of the absence of nonexpert reviewers, the generalizability was unknown. In the present study, we included international groups of both expert and nonexpert radiologists and clinicians making these results more relevant to daily practice. Interobserver agreement was best among expert radiologists experienced in pancreatic disorders, followed by the nonexpert radiologists and expert clinicians. Even nonexpert clinicians, however, showed good to moderate agreement. This degree of interobserver agreement represents a substantial advance in international understanding of the classification of severe acute pancreatitis. Using the revised Atlanta classification, all physicians can now look at the same information and arrive generally at the same conclusion regarding disease classification for each patient. Beyond facilitating clinical communication, use of the revised classification should also greatly advance research studies.

Much of the decision-making for the revised Atlanta classification hinges on the proper description of CT morphology, which, according to our results showed, is most uniform among radiologists; this is not surprising. The interobserver agreement was significantly better among radiologists than among clinicians (ie, surgeons and gastroenterologists). These data support the need for a multidisciplinary approach to severe acute pancreatitis where radiologists and clinicians work together. Based on our data, the radiologist is essential for the proper description of the CT morphology, and the clinician is needed to translate the type of peripancreatic collection into a patient-specific plan of treatment. Moreover, the expert reviewers as a group did better than the nonexpert reviewers, which suggest that interobserver agreement is better in expert centers. How this result is related to the clinical outcome of patients with acute pancreatitis is uncertain; however, one could argue that the availability of multidisciplinary teams along with better interobserver agreement in assigning a disease classification might improve clinical outcomes. Severe acute pancreatitis occurs with relatively low prevalence, and the patients present with many manifestations in symptoms and CT findings. Within this small group, each patient is unique and when

**TABLE 2.** Interobserver Agreement, Ratio $\kappa_{max}$, Among Reviewers for the Revised Atlanta Classification Definitions

| | Expert | | Nonexpert | | |
|---|---|---|---|---|---|
| **Term** | **Radiologists (n = 7)** | **Clinicians (n = 12)** | **Radiologists (n = 7)** | **Clinicians (n = 11)** | **Overall (n = 37)** |
| Type of pancreatitis | 0.87 (0.11) | 0.79 (0.21) | 0.72 (0.12) | 0.69 (0.27) | 0.75 (0.21) |
| Type of peripancreatic collection | 0.82 (0.10) | 0.59 (0.16) | 0.77 (0.09) | 0.54 (0.20) | 0.62 (0.19) |

Values are ratio $\kappa_{max}$.

Values are expressed as mean (SD).

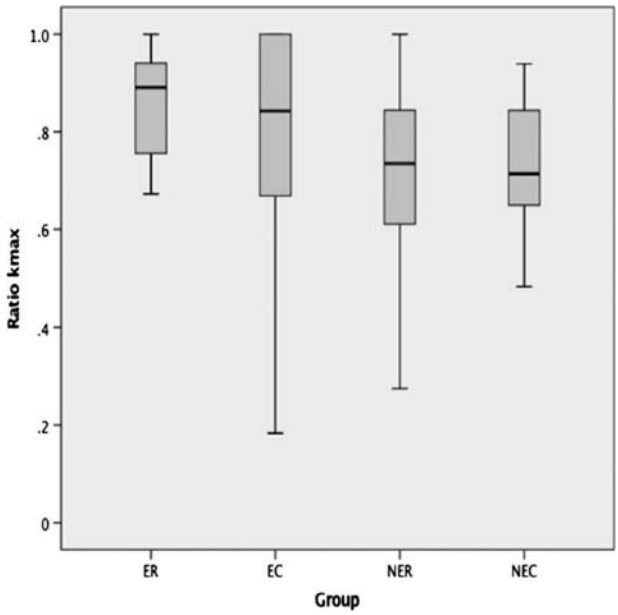


**FIGURE 5.** Boxplot of subgroups of reviewers for the revised Atlanta classification definitions for type of acute pancreatitis. Values are ratio $\kappa_{max}$. ER indicates expert radiologist; EC, expert clinicians; NEC, nonexpert clinicians; NER, nonexpert radiologists.

distributed over many clinicians within a hospital system that does not treat many severely ill patients annually, it can be difficult for a practitioner to gain the experience and perspective needed. We recommend treating patients with severe acute pancreatitis in multidisciplinary teams of expert radiologists and clinicians.[16,17] This is in line with a growing body of literature that has linked increased procedure and/or case volume with improved outcomes

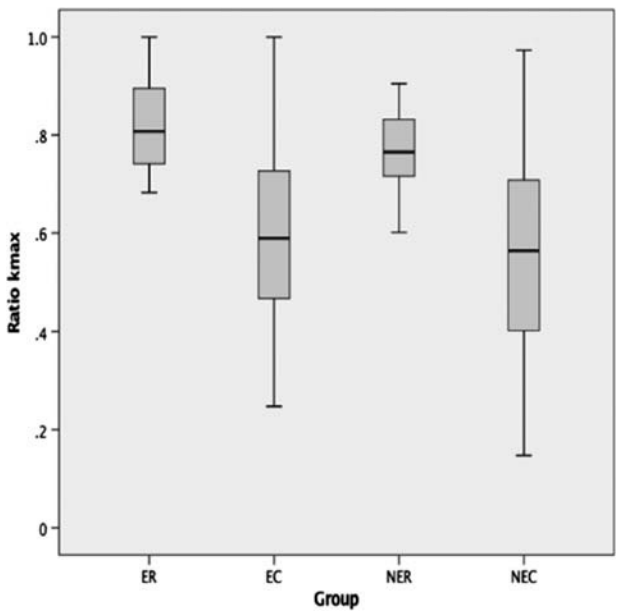


**FIGURE 6.** Boxplot of subgroups of reviewers for the revised Atlanta classification definitions for type of peripancreatic collection. Values are ratio $\kappa_{max}$. ER indicates expert radiologist; EC, expert clinicians; NEC, nonexpert clinicians; NER, nonexpert radiologists.

**TABLE 3.** Significant Differences Between Subgroups of Reviewers for the Revised Atlanta Classification Definitions

| **Type of Pancreatitis** | | | | |
|---|---|---|---|---|
| | **ER** | **EC** | **NER** | **NEC** |
| ER | NA | NS | $P = 0.025$ | $P = 0.025$ |
| EC | NS | NA | NS | NS |
| NER | $P = 0.025$ | NS | NA | NS |
| NEC | $P = 0.025$ | NS | NS | NA |
| **Type of Peripancreatic Collection** | | | | |
| | **ER** | **EC** | **NER** | **NEC** |
| ER | NA | $P < 0.001$ | NS | $P < 0.001$ |
| EC | $P < 0.001$ | NA | $P < 0.001$ | NS |
| NER | NS | $P < 0.001$ | NA | $P < 0.001$ |
| NEC | $P < 0.001$ | NS | $P < 0.001$ | NA |

Significant differences between subgroups of reviewers for ratio $\kappa_{max}$.

ER indicates expert radiologist; EC, expert clinicians; NA, not applicable; NEC, nonexpert clinicians; NER, nonexpert radiologists; NS, not significant.

for patients across a variety of diseases and procedures. Recent evidence suggests that treatment of patients with acute pancreatitis in high volume centers (>120 patients treated a year) results in a decrease in length of intensive care unit stay, length of hospital stay, and mortality.[18]

The revised Atlanta classification is definitely a step forward in the description and management of acute pancreatitis. But this new classification also generates some questions. For instance, does the good interobserver agreement also lead to a better selection of patients for invasive treatment and thereby improved outcomes? Are all terms used in the revised Atlanta classification equally important for clinical decision-making or is treatment dictated by a few descriptive terms, such as content, (extent of) encapsulation, and (presence of) gas bubbles? A known issue with CECT is that it is difficult to distinguish the content (solid vs fluid based) of a peripancreatic collection and has its limitations in determining the extent of encapsulation.[8] In fact, we found only moderate interobserver agreement for the terms content and encapsulation (Supplementary Table 4, http://links.lww.com/MPA/A597). Both terms can be essential in clinical decision-making and for the success of interventions like catheter drainage and necrosectomy. Because magnetic resonance imaging (MRI) and endoscopic ultrasonography can help in better defining the content and encapsulation of a collection,[8] future research should focus on determining the additional value of MRI and endoscopic ultrasonography for describing peripancreatic collections in severe acute pancreatitis. The effect of these diagnostic modalities as adjuncts to interobserver agreement and on the outcome of patients with necrotizing pancreatitis also needs to be evaluated in the light of the revised Atlanta classification.[19,20]

Some limitations have to be taken into account. There are several ways of calculating interobserver agreement. In studies on interobserver agreement for imaging (eg, ultrasonography, CT, or MRI) in benign and malignant pancreatic disorders, multiple approaches have been described, all with advantages and disadvantages.[21–27] Agreement between 2 reviewers can be calculated by Cohen $\kappa$ and in case of multiple reviewers with Fleiss $\kappa$; however, because of a substantial imbalance in distribution for most of the terms, the Cohen and Fleiss $\kappa$ could not be used, because the $\kappa$ statistic is influenced strongly by the prevalence of

the attribute.[28,29] An alternative analysis is the intraclass correlation coefficient; however, this analysis looks at the data as groups instead of paired observations, which we wanted to examine in our study. We used the ratio $\kappa_{max,}$ because it corrects for bias and prevalence, which suits our data best. Another limitation is that the reviewer findings were not compared with a reference standard, for several reasons. First, the aim of this study was not to define whether the reviewers found the "correct" diagnosis; however, if the reviewers described morphologic changes similarly, do we all speak the same language? Second, what is the criterion standard for comparison? One could argue that the diagnosis of the official radiology report of a single radiologist is not, per se, better than the agreement of a group of expert radiologists or expert clinicians. Our decision to use descriptions of the images without reference to any subsequent clinical data reflects the "real world" clinical situation.

In conclusion, present findings highlight a major improvement in interobserver agreement with the 2012 revised Atlanta classification compared with the poor interobserver agreement seen among experts with the 1992 Atlanta classification and lends strong support for complete adoption of the revised Atlanta classification.

## ACUTE PANCREATITIS INTEROBSERVER STUDY GROUP: AUTHORS AND COLLABORATORS

Ali A. Aghdassi (Department of Medicine, University Medicine Greifswald, Greifswald, Germany)

Peter A. Banks (Department of Gastroenterology, Center for Pancreatic Disease, Brigham and Women's Hospital, Harvard Medical School, Boston, MA)

Olaf J. Bakker (Department of Surgery, University Medical Center Utrecht, Utrecht, The Netherlands)

Marc G. Besselink (Department of Surgery, Academic Medical Center, Amsterdam, The Netherlands)

Marja A. Boermeester (Department of Surgery, Academic Medical Center, Amsterdam, The Netherlands)

Thomas L. Bollen (Department of Radiology, St. Antonius Hospital, Nieuwegein, The Netherlands)

Stefan A. Bouwense (Department of OR/Clinical Surgical Research, Radboud university medical center, Nijmegen, The Netherlands)

Sandra van Brunschot (Department of Hepatology, Academic Medical Center, Amsterdam, The Netherlands)

Vincent C. Cappendijk (Department of Radiology, Jeroen Bosch Hospital, 's-Hertogenbosch, The Netherlands)

Ross Carter (Department of Surgery, Glasgow Royal Infirmary, Glasgow, United Kingdom)

Richard Charnley (Department of Surgery, Freeman Hospital, Newcastle upon Tyne, United Kingdom)

Conny C. van Enckevort (Department of Gastroenterology, Canisius-Wilhelmina Hospital, Nijmegen, the Netherlands)

Marcel G. Dijkgraaf (Clinical Research Unit, Academic Medical Center, Amsterdam, The Netherlands)

Casper H. van Eijck (Department of Surgery, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands)

Patrick C. Freeny (Department of Radiology, University of Washington Medical Center, Seattle, WA)

Hein G. Gooszen (Department of OR/Clinical Surgical Research, Radboud university medical center, Nijmegen, The Netherlands)

Robbert J. de Haas (Department of Radiology, Radboud university medical center, Nijmegen, The Netherlands)

John J. Hermans (Department of Radiology, Radboud university medical center, Nijmegen, The Netherlands)

Karin Horsthuis (Department of Radiology, Academic Medical Center, Amsterdam, The Netherlands)

Karen D. Horvath (Department of Surgery, University of Washington Medical Center, Seattle, WA)

Michel F. van der Jagt (Department of Surgery, Radboud university medical center, Nijmegen, The Netherlands)

David M. Hough (Department of Radiology, Mayo Clinic, Rochester, MN)

Colin D. Johnson (Department of Surgery, University Hospital Southampton, Hampshire, United Kingdom)

Niels F. Kok (Department of Surgery, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands)

Desirée H. Koopmanschap (Department of Radiology, Radboud university medical center, Nijmegen, The Netherlands)

Manuel J. Koppe (Department of Surgery, Radboud university medical center, Nijmegen, The Netherlands)

Nanda C. Krak (Department of Radiology, University Medical Center Utrecht, Utrecht, The Netherlands)

Charlotte E. Lane (Department of Radiology, University Hospital Southampton, Hampshire, United Kingdom)

Johan S. Laméris (Departments of Radiology, Academic Medical Center, Amsterdam, The Netherlands)

Jean H. Lee (Department of Radiology, University of Washington Medical Center, Seattle, WA)

Markus M. Lerch (Department of Medicine, University Medicine Greifswald, Greifswald, Germany)

Q. de Lussanet (Department of Radiology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands)

Julia Mayerle (Department of Medicine, University Medicine Greifswald, Greifswald, Germany)

Koenraad J. Mortele (Department of Radiology, Beth Israel Deaconess Medical Center, Boston, MA)

Hjalmar C. van Santvoort (Department of Surgery, Academic Medical Center, Amsterdam, The Netherlands)

Michael G. Sarr (Department of Surgery, Mayo Clinic, Rochester, MN)

Michael D. Saunders (Department of Gastroenterology, University of Washington Medical Center, Seattle, WA)

Brian Stedman (Department of University Hospital Southampton, Hampshire, United Kingdom)

Santhi Swaroop Vege (Department of Gastroenterology, Mayo Clinic, Rochester, MN)

Manon van der Vlugt (Department of Hepatology, Academic Medical Center, Amsterdam, The Netherlands)

Bas van Wageningen (Department of Surgery, Radboud university medical center, Nijmegen, The Netherlands)

Eelco Wassenaar (Department of Surgery, University of Washington Medical Center, Seattle, WA)

Bob J. van Wely (Department of Surgery, Canisius-Wilhelmina Hospital, Nijmegen, the Netherlands)

Jens Werner (Department of Surgery, Ludwig Maximilian University of Munich, Munich, Germany)

Bas P. Wijnhoven (Department of Surgery, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands)

## ACKNOWLEDGMENT

*We would like to thank Dr A.H. Zwinderman, professor of Clinical Epidemiology and Bio-Statistics at the Academic Medical Center of the University of Amsterdam, The Netherlands, for his help with our statistical analysis.*

## REFERENCES

1. Working Party of the British Society of Gastroenterology; Association of Surgeons of Great Britain and Ireland; Pancreatic Society of Great Britain

and Ireland; Association of Upper GI Surgeons of Great Britain and Ireland. UK guidelines for the management of acute pancreatitis. *Gut*. 2005; 54(suppl 3):iii1–iii9.

2. Balthazar EJ, Freeny PC, Vansonnenberg E. Imaging and intervention in acute pancreatitis. *Radiology*. 1994;193:297–306.
3. Banks PA, Freeman ML. Practice guidelines in acute pancreatitis. *Am J Gastroenterol*. 2006;101:2379–2400.
4. Bradley EL 3rd. A clinically based classification system for acute pancreatitis. Summary of the International Symposium on Acute Pancreatitis, Atlanta, GA, September 11 through 13, 1992. *Arch Surg*. 1993;128:586–590.
5. Bollen TL, van Santvoort HC, Besselink MG, et al. The Atlanta Classification of acute pancreatitis revisited. *Br J Surg*. 2008;95:6–21.
6. Besselink MG, van Santvoort HC, Bollen TL, et al. Describing computed tomography findings in acute necrotizing pancreatitis with the Atlanta classification: an interobserver agreement study. *Pancreas*. 2006;33: 331–335.
7. van Santvoort HC, Bollen TL, Besselink MG, et al. Describing peripancreatic collections in severe acute pancreatitis using morphologic terms: an international interobserver agreement study. *Pancreatology*. 2008;8:593–599.
8. Banks PA, Bollen TL, Dervenis C, et al. Classification of acute pancreatitis–2012: revision of the Atlanta classification and definitions by international consensus. *Gut*. 2013;62:102–111.
9. Besselink MG, van Santvoort HC, Buskens E, et al. Probiotic prophylaxis in predicted severe acute pancreatitis: a randomised, double-blind, placebo-controlled trial. *Lancet*. 2008;371:651–659.
10. van Santvoort HC, Besselink MG, Bakker OJ, et al. A step-up approach or open necrosectomy for necrotizing pancreatitis. *N Engl J Med*. 2010;362: 1491–1502.
11. Balthazar EJ, Robinson DL, Megibow AJ, et al. Acute pancreatitis: value of CT in establishing prognosis. *Radiology*. 1990;174:331–336.
12. Forsmark CE, Baillie J; AGA Institute Clinical Practice and Economics Committee, et al. AGA institute technical review on acute pancreatitis. *Gastroenterology*. 2007;132:2022–2044.
13. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85:257–268.
14. Altman DG. *Practical statistics for medical research*. London, UK: Chapman and Hall/CRC; 1991.
15. van Santvoort HC, Bakker OJ, Bollen TL, et al. A conservative and minimally invasive approach to necrotizing pancreatitis improves outcome. *Gastroenterology*. 2011;141:1254–1263.
16. Working Group IAP/APA Acute Pancreatitis Guidelines. IAP/APA evidence-based guidelines for the management of acute pancreatitis. *Pancreatology*. 2013;13(4 Suppl 2):e1–e15.
17. Freeman ML, Werner J, van Santvoort HC, et al. Interventions for necrotizing pancreatitis: summary of a multidisciplinary consensus conference. *Pancreas*. 2012;41:1176–1194.
18. Singla A, Simons J, Li Y, et al. Admission volume determines outcome for patients with acute pancreatitis. *Gastroenterology*. 2009;137: 1995–2001.
19. Morgan DE, Baron TH, Smith JK, et al. Pancreatic fluid collections prior to intervention: evaluation with MR imaging compared with CT and US. *Radiology*. 1997;203:773–778.
20. Bollen TL, Besselink MG, van Santvoort HC, et al. Toward an update of the atlanta classification on acute pancreatitis: review of new and abandoned terms. *Pancreas*. 2007;35:107–113.
21. de Jong K, Nio CY, Mearadji B, et al. Disappointing interobserver agreement among radiologists for a classifying diagnosis of pancreatic cysts using magnetic resonance imaging. *Pancreas*. 2012;41: 278–282.
22. Loizou L, Albiin N, Ansorge C, et al. Computed tomography staging of pancreatic cancer: a validation study addressing interobserver agreement. *Pancreatology*. 2013;13:570–575.
23. Kim JH, Hong SS, Kim YJ, et al. Intraductal papillary mucinous neoplasm of the pancreas: differentiate from chronic pancreatits by MR imaging. *Eur J Radiol*. 2012;81:671–676.
24. Sai JK, Suyama M, Kubokawa Y, et al. Diagnosis of mild chronic pancreatitis (Cambridge classification): comparative study using secretin injection-magnetic resonance cholangiopancreatography and endoscopic retrograde pancreatography. *World J Gastroenterol*. 2008;14: 1218–1221.
25. Sahani DV, Kadavigere R, Blake M, et al. Intraductal papillary mucinous neoplasm of pancreas: multi-detector row CT with 2D curved reformations–correlation with MRCP. *Radiology*. 2006;238: 560–569.
26. Lecesne R, Taourel P, Bret PM, et al. Acute pancreatitis: interobserver agreement and correlation of CT and MR cholangiopancreatography with outcome. *Radiology*. 1999;211:727–735.
27. Kim JH, Eun HW, Park HJ, et al. Diagnostic performance of MRI and EUS in the differentiation of benign from malignant pancreatic cyst and cyst communication with the main duct. *Eur J Radiol*. 2012;81: 2927–2935.
28. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol*. 1990;43:551–558.
29. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol*. 1990;43:543–549.