Acute Pancreatitis. Biomarkers and radiology assessment

Sternby, Hanna

2017

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Acute Pancreatitis
Biomarkers and radiology assessment

Hanna Sternby

Lund University

DOCTORAL DISSERTATION
by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at Lilla aulan, MFC, Skåne University Hospital Malmö
October 6, 2017, at 1:15 pm.

Faculty opponent
Associate Professor Claes Jönsson
Sahlgrenska University Hospital, Göteborg University
Abstract
Background: Acute pancreatitis (AP) is a common and potentially severe disease. Early identification of severity grade is crucial for the outcome of the patient. For study approaches and comparison of inter-institutional data the AP patients need to be uniformly classified.

The aims of this thesis were to investigate the early stratification capacity of biomarkers in AP and to assess the morphological criteria of the revised Atlanta classification (RAC).

Methods: The studies are based on two cohorts. The first cohort encompasses 285 AP patients from six European centers. All patients had at least one CT performed during the first three months after onset of disease. In total 388 CTs were scored by six local radiologists and one central expert radiologist according to the morphological criteria of the RAC. The results were compared using interobserver agreement levels.

The second cohort was enrolled at Skåne University Hospital Malmö. Blood samples from 232 AP patients were collected upon admission and daily as long as the pancreatic amylase was elevated. Exact time for onset of pain was requested at inclusion. From the serum samples selected biomarkers were analysed and compared with severity outcome of the patients.

Results: In general the interobserver agreement between the local radiologists and the central expert radiologist was good regarding clinically important radiological features. Two areas of interpretation inconsistencies were identified: presence of extrapancreatic necrosis and necrotic debris of peripancreatic or pancreatic collections.

In biomarker analysis we found that cut-off levels for severe disease established by previous studies did not reach sufficient stratification capacity in our cohort. However, a combination of IL-6 and CRP (with cut-off levels 23.6 pg/ml and 57 mg/L respectively) demonstrated good potential in differentiating mild from non-mild (moderately severe and severe AP according to the RAC) disease. When investigating the temporal development of biomarkers we found that the mean values of IL-6 and IL-1β increased significantly in the severe group between 0-24 and 25-48 hours after onset of disease. Additionally, differences in mean values (i.e. delta-values) varied significantly between the mild and severe group for IL-6, IL-1β and IL-10.

Conclusions: Our results indicate that radiologists are unfamiliar with some of the new morphological categories of the RAC, potentially resulting in inconsistent reporting of necrotising pancreatitis. Among prognostic biomarkers in AP, in general CRP, IL-6, IL-1β and IL-10 demonstrated superior stratification capacity in our cohort.

Key words: Acute pancreatitis, severity grade, biomarkers, morphology, classification, radiology assessment
Acute Pancreatitis

Biomarkers and radiology assessment

Hanna Sternby
We can’t plan life. All we can do is be available for it

Lauryn Hill
Content

List of papers ................................................................................................................. 9
Abbreviations ................................................................................................................. 11
Introduction ................................................................................................................... 13
  History ......................................................................................................................... 13
  Anatomy and physiology .......................................................................................... 13
Acute pancreatitis .......................................................................................................... 15
  Aetiologies .................................................................................................................. 15
  Pathogenesis ............................................................................................................... 16
  Local inflammation ..................................................................................................... 16
  Systemic inflammation ............................................................................................... 19
  Diagnosis ..................................................................................................................... 21
  Classification .............................................................................................................. 21
  Imaging ......................................................................................................................... 24
  Severity prediction ..................................................................................................... 26
  Management of the patient ....................................................................................... 29
Aims .................................................................................................................................. 31
Materials and Methods ................................................................................................. 33
  Paper I .......................................................................................................................... 33
    Study design and cohort ........................................................................................... 33
    Statistics .................................................................................................................... 34
  Paper II-IV ................................................................................................................... 34
    Study design and cohort ........................................................................................... 34
    Biomarkers ............................................................................................................... 35
    Statistics .................................................................................................................... 36
Results ............................................................................................................................. 37
  Paper I .......................................................................................................................... 37
    Patients and CECTs ................................................................................................. 37
    Interobserver agreement ......................................................................................... 37
  Paper II .......................................................................................................................... 38
    Patients ....................................................................................................................... 38
    Biomarkers and cut-off levels ................................................................................. 39
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper III</td>
<td>40</td>
</tr>
<tr>
<td>Patients</td>
<td>40</td>
</tr>
<tr>
<td>Biomarkers</td>
<td>40</td>
</tr>
<tr>
<td>Paper IV</td>
<td>41</td>
</tr>
<tr>
<td>Patients</td>
<td>41</td>
</tr>
<tr>
<td>Biomarkers</td>
<td>41</td>
</tr>
<tr>
<td>Discussion</td>
<td>43</td>
</tr>
<tr>
<td>Acute pancreatitis</td>
<td>43</td>
</tr>
<tr>
<td>Radiological assessment</td>
<td>43</td>
</tr>
<tr>
<td>Biomarkers</td>
<td>45</td>
</tr>
<tr>
<td>Methodological considerations</td>
<td>47</td>
</tr>
<tr>
<td>Conclusions</td>
<td>51</td>
</tr>
<tr>
<td>Future perspectives</td>
<td>53</td>
</tr>
<tr>
<td>Populärvetenskaplig sammanfattning</td>
<td>55</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>59</td>
</tr>
<tr>
<td>References</td>
<td>61</td>
</tr>
</tbody>
</table>
This thesis is based on the following original papers, which will be referred to in the text by their Roman numerals


IV. Sternby H, Hartman H, Johansen D, Thorlacius H, Regnér S. The initial development of interleukins differs significantly between mild and severe acute pancreatitis. *Submitted*

*Published papers were reproduced with permission from Elsevier (Pancreatology) and Karger (European Surgical Research)*
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Atlanta classification of 1992</td>
</tr>
<tr>
<td>ANC</td>
<td>acute necrotic collection</td>
</tr>
<tr>
<td>AP</td>
<td>acute pancreatitis</td>
</tr>
<tr>
<td>APACHE</td>
<td>acute physiology and chronic health evaluation</td>
</tr>
<tr>
<td>APFC</td>
<td>acute peripancreatic fluid collection</td>
</tr>
<tr>
<td>AUC</td>
<td>area under curve</td>
</tr>
<tr>
<td>BUN</td>
<td>blood urea nitrogen</td>
</tr>
<tr>
<td>CRP</td>
<td>c-reactive protein</td>
</tr>
<tr>
<td>CT</td>
<td>computed tomography</td>
</tr>
<tr>
<td>CECT</td>
<td>contrast-enhanced computed tomography</td>
</tr>
<tr>
<td>DBC</td>
<td>determinant based classification on acute pancreatitis</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>ERCP</td>
<td>endoscopic retrograde cholangiopancreatography</td>
</tr>
<tr>
<td>EUS</td>
<td>endoscopic ultrasound</td>
</tr>
<tr>
<td>HAPS</td>
<td>harmless acute pancreatitis score</td>
</tr>
<tr>
<td>Hb</td>
<td>hemoglobin</td>
</tr>
<tr>
<td>ICAM</td>
<td>intercellular adhesion molecules</td>
</tr>
<tr>
<td>ICU</td>
<td>intensive care unit</td>
</tr>
<tr>
<td>IEP</td>
<td>interstitial edematous pancreatitis</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>interferon gamma</td>
</tr>
<tr>
<td>IL</td>
<td>interleukin</td>
</tr>
<tr>
<td>MCP-1</td>
<td>monocyte chemoattractant protein-1</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>MODS</td>
<td>multiple organ dysfunction syndrome</td>
</tr>
<tr>
<td>MRCP</td>
<td>magnetic resonance cholangiopancreatography</td>
</tr>
<tr>
<td>MRI</td>
<td>magnetic resonance imaging</td>
</tr>
<tr>
<td>NFκB</td>
<td>nuclear factor-κB</td>
</tr>
<tr>
<td>NP</td>
<td>necrotising pancreatitis</td>
</tr>
<tr>
<td>PAF</td>
<td>platelet activating factor</td>
</tr>
<tr>
<td>RAC</td>
<td>revised Atlanta classification on acute pancreatitis</td>
</tr>
<tr>
<td>ROC</td>
<td>receiver operating characteristic</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>SIRS</td>
<td>systemic inflammatory reaction syndrome</td>
</tr>
<tr>
<td>SOFA</td>
<td>sepsis-related organ failure assessment</td>
</tr>
<tr>
<td>TNF-α</td>
<td>tumor necrosis factor alpha</td>
</tr>
<tr>
<td>WON</td>
<td>walled-off necrosis</td>
</tr>
</tbody>
</table>
Introduction

The pancreas has discouraged researchers and clinicians for more than a century. With its anatomical inaccessibility, complex physiology and ill-understood pathobiology this organ humbles anyone who dare to confront it.

History

Pan kreas (all flesh) is claimed to be recognised by the Greeks around 300 years BC. Initially regarded as a protective organ either for the stomach or the large adjacent vessels its functions remained unfamiliar for many years. In 1642 Johann Georg Wirsung described the main duct which also holds his name, but it was not until the 19th century that any further understanding of the gland was established. By this time the digestive enzymes were identified, and in 1869 the medical student Paul Langerhans published his revolutionary findings, later named the islands of Langerhans. Major advances during the 20th century were the discovery by Banting and coworkers on how to extract insulin and the description of the exocrine pancreatic cell by George Paladan, for which they both received the Nobel Prize. By the turn of the 20th century acute pancreatitis was recognised as a separate individual disease, and not just a part of general intraabdominal necrosis related to any severe state [1].

Anatomy and physiology

The pancreas is situated ventrally to the second lumbar vertebra in the retroperitoneum, thus holding a central position in the upper abdomen. The location implies close proximity to a number of important structures such as large vessels and the duodenum (Figure 1) but also organs like the spleen, the stomach and the transverse colon [2]. In acute pancreatitis, the position of the gland plays an important role for possible complications of the disease.

The adult pancreas weights around 100 gram and is divided into head, neck, body and tail, with the head constituting the largest mass. The pancreatic duct (duct of Wirsung) passes from the tail to the head where it unifies with the common bile
duct (Figure 1). Normally they jointly enter the duodenum at the papilla Vateri, however cases with an accessory pancreatic duct also exist (Figure 1)[2].

![Diagram of the pancreas and periampullary region](image)

**Figure 1**

Anatomy of the pancreas and periampullary region. Reprinted with courtesy of dr Daniel Ansari[3], illustration by Anders Flood ©

Although most famous for its endocrine functions, the distribution of pancreatic cells are approximately 85 % exocrine and 2 % endocrine. The endocrine cells, gathered together in the islets of Langerhans, produce and secrete peptides (mainly insulin and glucagon) into the bloodstream. The exocrine pancreas is composed of ductal cells forming the ductal system and groups of several hundred linked acinar cells which, together with duct endings, form so called acinis. Ductal and acini secretion is regulated by both hormonal input and neural stimulation in response to food. Ductal cells secrete a bicarbonate-rich fluid serving several tasks, including transportation of digestive enzymes from the acinis to the duodenum and neutralization of gastric acid for optimal pH in the duodenum in the context of digestion. About two liters of pancreatic juice is secreted into the intestinal lumen daily[3].
The acinar cells are the functional units of the exocrine pancreas where digestive (proteolytic, lipolytic and amylolytic) enzymes are synthesized, stored and secreted. Large amounts of pro-enzymes (trypsinogen, procarboxypeptidase and chymotrypsinogen) are produced by and kept in the endoplasmatic reticulum systems. The endoplasmatic reticulum additionally accommodates calcium, which is a regulator of enzyme release into the ductal system. Due to their aggressive nature, proteases are secreted as inactive pro-enzymes (zymogens), packed within zymogen granules together with secretory trypsin inhibitors. Through exocytosis, the membrane-bound compartments merge with the membrane of the acinar cell and the inactive pro-enzymes are released into the lumen of the pancreatic duct. Isotonic NaCl-rich fluid from the acinus and bicarbonate from the ductal cells inhibit the initiation of autodigestion before the precursors reach the duodenum. As food enters the duodenum the enzyme cholecystokinin is released which in turn initiates the exocytosis. In the duodenum trypsinogen is activated by enterokinase from the duodenal mucosa to trypsin, setting off the activation of other peptidases. The intestinal mucosa itself is protected from the proteases by protease inhibitors [2,4].

Acute pancreatitis

Acute pancreatitis (AP) is an inflammatory disorder of the pancreas. Depending on demographic variations the annual incidence ranges between 13-73 per 100 000 inhabitants, and numbers are reported to be increasing worldwide [5, 6]. In 2012, Peery et al reported AP to be the third most common disease for hospital admission among gastroenterology disorders in the United States. Additionally it was demonstrated to be the fifth highest cause of in-hospital mortalities, in second place regarding total hospital stay and the largest contributor to increasing hospital costs [7].

Aetiologies

The dominating aetiologies of AP within western populations are gallstones and alcohol, together accounting for 70-80% of the cases in most studies [5, 8]. Biliary disease is more common in women and elderly whereas alcohol misuse appears more often among men [5]. In 10-30% of the cases the aetiology remains unknown. Less common causes are pancreatic duct obstruction (tumours and other strictures of the common bile duct) endoscopic retrograde cholangiopancreatograh (ERCP), other iatrogenic causes, hypercalcemia, hyperlipidemia, drugs (thiazides, azathioprine, tetracyklin), pancreatic trauma,
infections and parasites, autoimmune diseases, cystic fibrosis etc \[2, 9\]. In Malmö, Sweden, around 55% of the patients suffer from acute biliary pancreatitis, whereas alcohol misuse causes 15% of the cases with AP.

**Pathogenesis**

Despite decades of research the complex pathophysiological pathways of AP are still poorly understood. Both precise initial events and subsequent reactions are still subjects of debate \[8\].

It is generally recognized that dysfunction in basalmembranes of acinar cells and subsequent defects in pancreatic secretion is, regardless of aetiology, the catalyst of the disease \[10\]. Subsequent events are likely a combination of multiple pathways. Secretion blockage leads to an abnormal increase in intracellular calcium level which prevents the regular zymogen exocytosis \[11\]. As the regular inhibitory system thus becomes reduced the precursor containing zymogens converge with lysosomes holding cathepsin B which has the capacity to convert premature trypsinogen to trypsin. The result is an inappropriate intraacinar activation of trypsin, which has been considered a key event in AP \[12, 13\]. Until recently the trypsin-mediated course has been the dominating theory on the pathogenesis behind AP hedström \[14\]. However, another early event in AP is the activation of intraacinar nuclear factor- κB (NFκB). Increasing evidence suggest that NFκB is not only a crucial pathway in AP but also trypsin-independent with sufficient capacity to induce the inflammatory response autonomously \[15, 16, 17\]. An additional event that might be of importance is the increased intraacinar calcium levels which themselves lead to acinar apoptosis.

Regardless of the different theories, trypsin converts other digestive pro-enzymes leading to the process of autodigestion of the acinus \[18, 19\]. Destruction of cell membranes and tissue by the activated proteases (trypsin, chymotrypsin and elastase) results in pancreatic damage with oedema and necrosis.

**Local inflammation**

In immediate response to pancreatic cell injury, a local inflammatory reaction is initiated including vasodilatation, increased permeability of vessels and infiltration of leukocytes (primarily neutrophils and monocytes) into the pancreatic tissue. The complex process of leukocyte recruitment is since long considered central for the determination of disease severity \[20\]. Local pancreatic tissue damage results in the unleashing of cytokines and free radicals directly from the destructed gland, yielding changes in the microvascular system as well as leukocyte activation and
enrolment. The aggregation of leukocytes in the pancreas conduce further release of various promotors of inflammation. Subsequently a cycle of gradually enhanced production of inflammatory mediators is accomplished\(^ {\text{[21, 22, 23]}}\).

**Pro-inflammatory cytokines**

Cytokines are low molecular weight proteins able to interact with various target cells. Such events result in both extraction and release of other cytokines as well as the self-amplification of their own production. Although the inflammatory pathways of AP are unclarified multiple pro- and anti-inflammatory mediators, playing important roles for both local as systemic reactions, have been identified\(^ {\text{[19]}}\). From a temporal perspective both the acinar cell injury and the accumulation of leukocytes stimulate the release of interleukin-1β (IL-1β) and tumor necrosis factor-α (TNF-α) followed by production and leakage of interleukin-6 (IL-6) and interleukin-8 (IL-8)\(^ {\text{[24, 25]}}\). As they correlate with degree of inflammation these biomarkers have been repeatedly investigated as markers for sepsis and systemic complications\(^ {\text{[26]}}\).

TNF-α is mainly secreted by macrophages and measurable as early an hour after onset of AP, with a reported distinct rise over the next six hours\(^ {\text{[23]}}\). In the context of AP increased levels of TNF-α has been found to correlate with severity grade, probably by increasing vascular permeability and upgrading of the inflammatory cytokine and chemokine response\(^ {\text{[23, 27, 28]}}\). However, TNF-α has also proven to be an unsteady marker of severe disease since it is promptly purged from the blood by the liver and thus difficult to detect in the clinical setting\(^ {\text{[27]}}\).

Like TNF-α, interleukin-1 (IL-1) is derived from macrophages. IL-1 exists in two forms IL-1β and interleukin-1α (IL-1α). Being a potent inflammatory mediator, IL-1β has the capacity to induce severe disease in animal models and possesses a central role in the systemic inflammatory reaction. IL-1β works synergistically with TNF-α and additionally stimulates the production of cytokines like IL-6 and IL-8\(^ {\text{[27, 29, 30]}}\).

As a response to TNF-α and IL-1β stimulation, IL-6 is derived from various cell types\(^ {\text{[19]}}\). Being the principal mediator of the acute phase response, IL-6 rises six hours after onset of AP and peaks 24-48 hours before CRP\(^ {\text{[21, 31]}}\). IL-6 additionally affects both the immune (B- and T-cells) and coagulation system. Multiple studies have demonstrated its discriminative properties in AP where increased levels have been associated with both organ failure and systemic complications\(^ {\text{[32, 33, 34]}}\).
Chemokines

Chemokines, a subfamily of the cytokines, are key components in the immune system possessing strong capacity for recruitment, migration and activation of leukocytes \[23, 35\]. Both IL-8 and monocyte chemoattractant protein-1 (MCP-1) belong to the chemokine family. Upon stimulation by TNF-α, IL-8 is synthesized by a large variety of cells and has a strong neutrophil attractant capacity causing release of, among others, the tissue dilapidator elastase. It is associated with severe AP, in particular the presence of respiratory complications, and can be detected from 24 hours after onset of disease \[21, 36, 37\]. The chemokine MCP-1 both recruits and activates monocytes in the pancreatic tissue. Upregulated MCP-1 expression has been found in both experimental models and human tissue implicating a role of this marker in the course of AP \[38\]. Additionally, levels of MCP-1 have been correlated with disease severity \[39, 40\].

Anti-inflammatory cytokines

As pro-inflammatory mediators are released a compensatory anti-inflammatory reaction is initiated. Among identified anti-inflammatory interleukins IL-1 receptor antagonist (IL-1ra), interleukin-4 (IL-4) and interleukin-10 (IL-10) are the most investigated. Levels of IL-1ra have demonstrated to be significantly more increased in patients with severe AP as it moderates the effects of IL-1. It has also been found to decrease severity in experimental AP \[31, 41\]. IL-4 inhibits the production of cytokines in general \[42\]. Interferon-γ (IFN-γ) is released by IL-1, however levels of IFN-γ have been found to be suppressed in severe AP, indicating that IFN-γ has anti-inflammatory characteristics \[42, 43, 44\]. The most well-known anti-inflammatory cytokine is IL-10 which reduces levels and functions of proinflammatory mediators like IL-1β, TNF-α, IL-6 and IL-8 and also correlates with grade of severity \[45, 46, 47\]. The levels of IL-10 elevate within 24 hours after onset of disease and, in the patients with severe AP, remain increased during the first week \[36, 48\].

Numerous additional cytokines with possibly interesting roles in AP have been investigated, however they will not be further discussed in this thesis.

Additional inflammatory mediators

Apart from cytokines other inflammatory mediators play important roles in the pathogenic course of AP.

IL-1 stimulates the release of platelet activating factor (PAF) \[42\]. PAF is a proinflammatory phospholipid originating from the vascular endothelium. As well
as being a potent vasodilator capable to increase capillary permeability, it also has the potential to activate and recruit neutrophils and amplify the production of IL-1β, IL-6 and TNF-α. Increased levels of PAF have been measured in AP \cite{49, 50}. It has thus been suggested that specific PAF-antagonists could reduce the inflammatory response \cite{51, 52}. However in a large multicentre phase III study the antagonist (lexipafant) did not demonstrate improved outcomes in patients with severe disease \cite{53}.

IL-1, IL-8 and TNF-α induce the expression and up-regulation of intercellular adhesion molecules (ICAMs) on the surface of endothelial cells. The process facilitates the adhesion of leukocytes making ICAMs play an important role in the process of leukocyte migration in AP \cite{54, 55}.

**Pancreatic necrosis**

Impaired microcirculation with hypotension, tissue hypoxia and cellular damage subsequently results in necrosis of the pancreas \cite{56, 57}. In the mild disease cell destruction generally occurs through controlled apoptosis whereas in severe AP the tissue damage and necrosis develops unrestrainedly \cite{58}. Repeated bursts of cytokine release not only result in capillary leakage but also mucosal damage and increased intestinal permeability. The causal factor of infected necrosis is suggested to be translocation of bacteria due to ischemia of the bowel and impaired gut barrier function \cite{59}. It normally takes about one week into the course of the disease before infectious complications emerge. Evolution of infected tissue creates a second burst of cytokine and chemokine release, leading to clinical deterioration of the patient and increased risk of organ failure and mortality \cite{60, 61, 62}.

**Systemic inflammation**

Apart from acting locally, cytokines and inflammatory mediators propagate via the vena portae into the systemic circulation. The result is a general distortion of normal physiology and initiation of a systemic inflammatory response syndrome (SIRS). By up-regulation and systemic propagation of cytokines the vascular endothelium is activated leading to capillary leakage and migration of leukocytes into all tissues. Additionally oxygen radicals and proteases are released which bring on parenchymal and endothelial cell damage causing deterioration in microcirculation and subsequent tissue oxygen deficiency \cite{26}. Enhanced vascular leakage, peripheral leukopenia and secondary oedema increase the risk of organ degradation and failure. The state of coagulopathy and fibrinolysis with elevated levels of d-dimer are reported as important features of the systemic inflammatory
Furthermore, study reports have demonstrated cytokine-induced lung complications to occur early in the course of the disease, followed by kidney deterioration \[37, 65, 66\]. Iterating cascade-like bursts of cytokine overproduction eventually lead to multiple organ dysfunction syndrome (MODS) \[67\].

Organ failure may occur within the first 24 hours after onset of disease or develop during or after several days. It is generally accepted that early organ failure is caused by sterile inflammation whereas organ failure developing later (>1 week) into the disease course rather is correlated to septic complications, often in association with infected pancreatic necrosis. Accordingly, two peaks in mortality have been commonly reported. Multiple studies demonstrate most important determinants of disease severity, and consequently also mortality, to be persistent organ failure (>48 h) and the development of infected necrosis. In single organ failure the mortality is less than 10 % whereas in multi-organ failure it increases to 35-50\% \[68, 69, 70, 71\].

Figure 2
Schematic overview of the inflammatory cascade and development of systemic complications in AP
Diagnosis

For the diagnosis of AP it is generally accepted that the patient should present at least two of the following symptoms: 1) Sudden onset of upper abdominal pain, 2) elevated pancreatic amylase (or lipase) at least three times the upper level of normal and 3) signs of pancreatitis on computed tomography (CT) scan \[72, 73\].

Classification

Several attempts have emerged with the aim of reaching global consensus on terms and descriptions of AP. The Marseille and Cambridge classifications were predecessors, but it was not until the international symposium in Atlanta 1992 that a worldwide adopted system was established \[74, 75, 76\]. The Atlanta classification (AC) was a clinically based framework proposed by 40 internationally recognized experts containing descriptive terms of local and systemic complications as well as a binary stratification into mild and severe disease. However, after 20 years of assessment, studies have demonstrated considerable inconsistencies regarding the application of nomenclature and interpretation of terminology of the AC. \[77, 78, 79, 80, 81, 82\]. Additionally, substantial progress has been made in the understanding of pathophysiological pathways, disease-related complications and morbidity as well as imaging and surgical interventions. Thus, in 2012 two new classifications were introduced for assessment of AP severity: the revised Atlanta classification of 2012 (RAC) and the determinant-based classification (DBC) \[73, 83\]. Main characteristics of the AC, the RAC and the DBC are outlined in table 1.

Table 1

<table>
<thead>
<tr>
<th></th>
<th>Mild</th>
<th>Moderate/moderately severe</th>
<th>Severe</th>
<th>Critical</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Atlanta Classification</strong></td>
<td>No OF and no local complications</td>
<td>N/A</td>
<td>OF and/or local complications</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Revised Atlanta Classification</strong></td>
<td>No organ failure, No local or systemic complications</td>
<td>Transient OF and/or local or systemic complications</td>
<td>Persistent OF</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Determinant Based Classification</strong></td>
<td>No (peri)pancreatic necrosis and no OF</td>
<td>Sterile (peri)pancreatic necrosis and/or transient OF</td>
<td>Infected (peri)pancreatic necrosis or persistent OF</td>
<td>Infected (peri)pancreatic necrosis and persistent OF</td>
</tr>
</tbody>
</table>

OF: organ failure, N/A not applicable
Local complications: necrosis, abscesses and/or pseudocysts
Transient OF – resolves within 48 hours, Persistent organ failure – organ failure more than 48 hours
Since their publication, all three classifications have been validated in various settings [84, 85, 86, 87, 88, 89, 90]. Main findings of these studies are the AC being inferior in severity prediction compared to the RAC and the DBC, whereas both the latters were comparable regarding investigated outcomes.

As the Atlanta classification is now considered outdated no further review on this classification will be performed here.

**The revised Atlanta classification**

The RAC was developed through an international iterative web consultative process led by a working group [73]. In this classification a number of definitions are stated; diagnosis of AP, detailed morphological features and contrast-enhanced computed tomography (CECT) criteria of local complications (see Box 1), early and late phase of the disease as well as a trisected assessment of severity. Apart from the groups outlined in Box 1 additional local complications stated are colonic necrosis, gastric outlet dysfunction as well as splenic and portal vein thrombosis. The RAC defines systemic complications as deterioration of co-morbidities and organ failure according to the modified Marshall scoring system (table 2) [73, 91].

**Box 1**

Morphological features and CECT criteria in AP according to the RAC

<table>
<thead>
<tr>
<th>Morphology groups</th>
<th>CECT criteria</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intersitial oedematous (\text{pancreatitis}) (IEP)</td>
<td>Homogenous enhancement of the pancreatic parenchyma, normal or minor inflammatory changes of the peripancreatic tissue (see below - APPC or pancreatic pseudocyst)</td>
<td>-</td>
</tr>
<tr>
<td>Necrotising pancreatitis</td>
<td>Heterogenous enhancement of the pancreatic parenchyma and/or peripancreatic tissue necrosis (see below – ANC or WON)</td>
<td>-</td>
</tr>
<tr>
<td>Acute peripancreatic fluid collection (APFC)</td>
<td>Homogenous fluid density. No complete wall. No necrosis. Associated with IEP. Solely extrapancreatic location.</td>
<td>≤ 4 weeks</td>
</tr>
<tr>
<td>Pancreatic pseudocyst</td>
<td>Homogenous fluid density. Fully encapsulated. No necrosis. Associated with IEP. Solely extrapancreatic location.</td>
<td>&gt; 4 weeks</td>
</tr>
<tr>
<td>Acute necrotic collection (ANC)</td>
<td>Heterogenous and non-liquid density. No complete wall. Associated with necrotising pancreatitis. Intra- or extrapancreatic location</td>
<td>≤ 4 weeks</td>
</tr>
<tr>
<td>Walled-off necrosis (WON)</td>
<td>Heterogenous and non-liquid density. Fully encapsulated. Associated with necrotising pancreatitis. Intra- or extrapancreatic location</td>
<td>&gt; 4 weeks</td>
</tr>
</tbody>
</table>

**CECT contrast enhanced computed tomography**
Table 2
Modified Marshall scoring system for organ failure - A score of 2 or more equals organ failure

<table>
<thead>
<tr>
<th>Organ system</th>
<th>No organ failure</th>
<th>Organ failure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Respiratory (PaO₂/FIO₂)</strong></td>
<td>&gt; 400</td>
<td>301-400</td>
</tr>
<tr>
<td><strong>Renal</strong></td>
<td>≤ 134</td>
<td>134-169</td>
</tr>
<tr>
<td>(serum creatinine µmol/l)</td>
<td>&lt; 1.4</td>
<td>1.4-1.8</td>
</tr>
<tr>
<td>(serum creatinine mg/dl)</td>
<td>&lt; 1.4</td>
<td>1.4-1.8</td>
</tr>
<tr>
<td><strong>Cardiovascular</strong> *</td>
<td>&gt; 90</td>
<td>&lt; 90</td>
</tr>
<tr>
<td>(systolic blood pressure, mm Hg)</td>
<td>RF</td>
<td>Not RF</td>
</tr>
</tbody>
</table>
| * Without inotropic support     | RF responsive to fluids, Not RF not responsive to fluids

The determinant-based classification

Whilst the RAC is built on clinical descriptions the DBC uses determinants of severity in AP [83]. The DBC was processed through three stages; a meta-analysis of published studies, a global web-based survey and an international consensus symposium. In the DBC, AP is classified into four categories of severity based on the main determinants organ failure and infected pancreatic necrosis. For the definition of OF, the DBC uses the Sepsis-related Organ Failure Assessment score (SOFA-score) [92].

Correlation between severity grades and outcome

The grades of severity are clinically crucial as they differ significantly in prognosis and outcome. The overall mortality in AP is approximately 5 % [93]. Mild disease is self-limiting with almost negligible mortality and long-term morbidity. The vast majority of the AP patients, 70-80%, pertain to this group [94]. As has been demonstrated in multiple validation studies both mortality and morbidity increases significantly with each degree of severity. In the moderately severe (RAC) and moderate (DBC) groups morbidity rises with more frequent admission to the intensive care unit (ICU), prolonged hospital stay, need for interventions and nutritional support, whereas mortality remains low (0-6%). For the groups with severe and critical AP the level of morbidity is 100 % accompanied by very high numbers in mortality, in particular for the critical group (22-80%) [88, 89, 95, 96].
Imaging

Computed tomography (CT) is the gold standard imaging modality when assessing both diagnostic uncertainty and suspected complications in AP \cite{97}. However, despite major technical advancement during the last decades, a CT obtained during the first days after onset of disease cannot evaluate the severity of the disease and is thus not recommended to be performed routinely \cite{72, 98, 99}. As AP is a dynamic disease the assessment of radiological changes are highly time-dependent. In mild AP characteristic morphological findings are peripancreatic fat stranding (increased attenuation) and diffuse enlargement, however with homogeneous enhancement, of the gland \cite{99, 100}. In moderately severe and severe disease the process of collections and necrosis takes days to evolve. Necrosis is defined as focal or diffuse non-enhancing pancreatic parenchyma and/or heterogeneous density of collections. Peripancreatic necrosis is more difficult to identify as fat perfusion is not detectable on CT and diagnosis has to based on secondary signs \cite{99}.

Figure 3
Peripancreatic necrosis
The pancreas enhances heterogeneously (asterisks) but no apparent necrosis can be identified. Peripancreatic necrotic collections are present in the retroperitoneal pancreatic compartment and transverse mesocolon (arrowheads pointing at the borders)
CT has good capacity for defining collections as incompletely or fully encapsulated which is essential for decisions on interventions and classification. Whether the collections contain fluid or necrotic debris is of importance for correct AP classification, see Box 1. Presence of infection is seen as gas bubbles in necrotic areas on CT, in general 2-4 weeks after onset of AP [60,101].

Abdominal ultrasound is indicated during the initial phase for the detection of gallstones in all patients with AP [98]. Magnetic resonance cholangiopancreatography (MRCP) is, with the identical purpose, a commonly utilized complementary instrument in suspected biliary disease [102]. Magnetic Resonance Imaging (MRI) has equal sensitivity to CT regarding severity evaluation and diagnosis of AP. For the assessment of collections (fluid and/or necrosis) MRI is superior to CT [103]. Endoscopic ultrasound (EUS) offers the possibility of both diagnosis and intervention in the same séance. As it is superior compared to MRCP in the exclusion of small (<5 mm) gallstones EUS is a complimentary method likely to become more developed and utilized in the future [72,104].

**Correlation between CT findings and disease severity**

Like the clinical degrees of severity, the morphologic types (Box 1) of AP also differ considerably in clinical outcome, treatment strategies and prognosis. Several radiological scoring systems have been developed for the prediction of severity grades [105,106,107,108]. However, none reach sufficient stratification accuracy and thereby a CT upon admission is not a general recommendation for the AP patients [109].

Mortality in interstitial disease is reported to be 0-3% with strong correlation to comorbidity [94,110,111,112]. Although morphological and clinical severity are not necessarily consistent, the level of morbidity in interstitial AP is low as clinical symptoms most frequently resolve within a week [94]. Corresponding figures in necrotising pancreatitis have been a matter of debate. Necrotising AP afflicts 20-30% of the AP patients and it is of general conclusion that presence of necrosis is related to high morbidity, however influenced by anatomical location and extent of morphological changes [80,93,110,113]. Necrosis has additionally been associated with high mortality, in some studies up to 40% [113,114]. Other studies claim death in necrotising pancreatitis to be solely associated with organ failure and not necrosis [77,93,115,116]. As both pancreatic necrosis and organ failure are established determinants of severe AP, the relationship between these entities has been repeatedly investigated. Several studies present significantly increased organ failure in patients with necrotising pancreatitis, particularly with exceeding extent of necrosis [117,118,119]. However, other works show no correlation making general
conclusions difficult [115, 116, 120, 121]. The relationship is thus considered not fully understood. It is likely that the pathophysiological processes of necrosis propel the development of organ failure when simultaneously reduced perfusion due to organ failure contributes to the evolution of necrosis [122].

Among the patients with necrotising pancreatitis 33 % have infected necrosis [93]. Mortality in this group has been reported to be very high, thereof the introduction of the critical category of the DBC, which was mainly based on a large review by Petrov et al. [61].

**Severity prediction**

Early identification of patients in risk of developing severe disease is important as it is recognised that these patients benefit from prompt management in the ICU. Despite large efforts and multiple studies no method exists to, upon admission, determine the severity of AP with sufficiently high sensitivity and specificity [9, 123]. Age, aetiology, obesity and comorbidity are established risk factors for severe disease [124, 125, 126, 127]. A number of scoring systems and predictive markers have been proposed, however none have demonstrated adequate prognostic capacity [72, 128, 129]. SIRS (table 3), in particular persistent SIRS at 48 hours, has proven to be the best predictor of severe disease with strong correlation to increased morbidity and mortality [70, 130, 131, 132]. According to recent guidelines the superior approach for severity prediction is thus to combine the information on 1) risk factors, 2) SIRS and 3) response to initial therapy [72].

<table>
<thead>
<tr>
<th>SIRS = Presence of two or more of the following criteria</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>&lt; 36° or &gt; 38°</td>
</tr>
<tr>
<td>Heartrate</td>
<td>&gt; 90 / min</td>
</tr>
<tr>
<td>Respiratory rate</td>
<td>&gt; 20 / min</td>
</tr>
<tr>
<td>White blood cells</td>
<td>&lt; 4 x10⁹ / L or &gt; 12 x 10⁹ / L</td>
</tr>
</tbody>
</table>
The Harmless Acute Pancreatitis Score (HAPS, Box 2) is developed for the identification of patients with mild AP upon admission. Although simple and validated this scoring system is not yet applied in general clinical practice\textsuperscript{[133, 134]}. 

**BOX 2**

The Harmless Acute Pancreatitis Score (HAPS)*\textsuperscript{[132]}

<table>
<thead>
<tr>
<th>Parameters of severity to be estimated upon admission:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Signs of peritonitis (rebound tenderness or guarding in abdominal examination)</td>
</tr>
<tr>
<td>• Serum creatinine &gt; 2 mg/dL</td>
</tr>
<tr>
<td>• Hematocrit level &gt; 43% for men and &gt;39.6% for women</td>
</tr>
</tbody>
</table>

If the patient is negative in all parameters the clinical course is categorised as harmless.

If the patient is positive in any of the parameters the clinical course is categorised as non-harmless.

*Predictive biomarkers*

The ideal prognostic biomarker in AP holds several features. It is clinically available, inexpensive, insensitive to inter-individual differences and accurately predicts severe cases upon admission. The search for such a biomarker ranges over decades and numerous promising candidates have been proposed. However, still no marker has repeatedly, with a generally defined and accepted cut-off level, been able to predict complicated disease with sufficient precision.

C-reactive protein (CRP) is the most commonly used marker due to its clinical availability and low analysis cost. Being produced by the hepatocytes, through initiation of IL-1 and IL-6, CRP peaks as late as 72 hours after onset of disease. There is no established cut-off level for severe disease, at 150 mg/dL (within 24 hours from admission) 100\% sensitivity and 81.4 \% specificity was found for necrotising pancreatitis\textsuperscript{[4, 135]}.

As the interleukins are established mediators of the inflammatory reaction vast investigations have been made within this group of cytokines. IL-6 peaks 24-36 hours before CRP and, in the absence of a complicated course, decreases rapidly\textsuperscript{[31]}. According to two meta-analysis elevated levels are associated with severe disease\textsuperscript{[136, 137]}.

Zhang et al reported (for admission samples) a pooled sensitivity of 91 \% and pooled specificity of 79 \% whereas corresponding figures from Aoun et al were 83.6 \% and 75.6 \%. The chronological profile of IL-8 resemble the one of IL-6 with an early rise, making it a theoretically interesting marker. However, meta-analysis have demonstrated widespread sensitivities and specificities for complicated disease making general conclusions difficult. As being a powerful
inflammatory mediator, IL-1β has repeatedly been associated with complicated AP manifesting increased levels in severe disease [31, 47, 138]. However, further studies are needed to understand the significance of this marker. The cardinally investigated anti-inflammatory marker is IL-10. IL-10 is related to both organ failure and other complications but general cut-off levels have been difficult to set [23, 31, 45, 47, 48].

Numerous other interleukins (IL-11, IL-12, IL-15, IL-17, IL-18 etc) have been investigated in the search for prognostic markers in AP [128]. However, since previous studies are generally small and results not sufficiently promising these biomarkers will not be further discussed in this thesis.

With regard to being the first cytokine to be released during the course of AP TNF-α has been repeatedly studied. Increased levels have been associated with a complicated course in AP but the opposite has also been demonstrated [23, 139, 140]. The inconsistent results might be explained by TNF-α having a short plasma half-life making it an unstable marker when measured in blood [27]. Additionally, positive results are in general inferior to findings for IL-6 and IL-8.

The acute phase reactant procalcitonin is a widely used marker for sepsis. According to a meta-analysis by Mofidi et al superior capacity was primarily demonstrated for the identification of infected pancreatic necrosis. For this complication the pooled sensitivity and specificity was 0.80 and 0.91 respectively [141]. Moreover, results for procalcitonin in predicting severe AP have been inconsistent [140].

As coagulopathy is a contributing factor to the disease course in AP haemostasis parameters related to disseminated coagulation have been investigated as possible predictive markers. D-dimer has been studied in multiple settings demonstrating promising results with a sensitivity of 81-90% and specificity of 67-89% for complicated disease [63, 64, 142]. However, referred studies present a wide range of cut-offs making general conclusions and clinical applicability difficult.

With an odds ratio of 40.8 (95% CI 8.5-195) MCP-1 has been correlated with severe disease already 24 hours into the course of the disease [39, 40].

In a large post-hoc analysis by Koutroumpakis et al the prognostic performances of hematocrit and blood urea nitrogen (BUN) in AP were analyzed [143]. Compared with the acute physiology and chronic health evaluation II (APACHE II) and creatinine at 24 hours increased BUN and a hematocrit level >44% better predicted pancreatic necrosis and organ failure.
Management of the patient

As no causal medical treatment exists for AP, management is empirical and unspecific with the primary aim to support vital functions and prevent complications. Early identification of organ deterioration and non-delayed admission to an intermediate care unit or the ICU is crucial. The initial caretaking focuses on hemodynamic stabilisation, principally through volume resuscitation and monitoring of fluid losses. To maintain adequate perfusion, and thus oxygen supply of vital organs, aggressive (250-500 ml/h) infusion of fluids (crystalloids) needs to start already in the emergency department. Early resuscitation is associated with reduced rates of SIRS, organ failure, MODS and mortality. Clinical response to fluid resuscitation is considered obtained if the following parameters are maintained: heartbeat <120/min, mean arterial pressure >65 mm Hg and a urinary output of >0.5 ml/kg/hour. The requirements of fluids should be repeatedly evaluated to confirm adequate response but also to avoid over-resuscitation and oedema.

Effective pain management is a deficiently studied field in AP and there is no consensus on preferred analgesic or administration method. Pain reduction should nevertheless be of priority immediately upon admission.

The impact of enteral nutrition on SIRS and organ failure is yet to be clarified. However, it has been demonstrated that early enteral nutrition decreases infectious complications, organ failure and mortality. Additionally, in mild AP early oral feeding is considered safe.

Meta-analysis and reviews have not found any benefits for systemic antibiotics in the prevention of infections in AP. Prophylactic antibiotics are thus not recommended as routine treatment, neither for mild, moderately severe nor severe AP. Future studies will need to expose whether antibiotic treatment should be supported in certain subgroups.

The aetiology of AP needs to be established upon admission by reason of severity prediction and patient treatment. Laboratory test and upper abdominal ultrasound should be performed for complete investigation. If the aetiology cannot be determined during the first attack of AP further examinations (CT, MRCP, EUS, genetic counseling) should be performed. Elimination of the causative factor is important to reduce the risk of recurrent AP. In biliary pancreatitis same-admission cholecystectomy have demonstrated a 72 % reduced risk for complications related to gallstones.

During the last twenty years, in accordance with international consensus, there has been a shift towards a more conservative approach and less surgical interventions.
in AP. For infected necrosis open necrosectomy has been the standard of care for decades. However, recent studies have demonstrated a reduction in both mortality and numbers of serious complications when a minimally invasive approach is applied.

ERCP is indicated when there are signs of cholangitis. There is no consensus or clear evidence for ERCP in patients with ongoing AP and persistent cholestasis without cholangitis.
Aims

The general aim of this thesis was to investigate stratification assessment of defined severity categories in AP using clinical features and biomarkers.

Specific aims were

*Paper I*
To, in a routine clinical setting, evaluate and validate the morphological CT-criteria defined by the revised Atlanta classification on AP.

*Paper II*
To select prognostic biomarkers in AP from the literature and study these using preset cut-off levels based on the results from previous studies.

*Paper III*
To investigate the capacity of biomarkers to differentiate mild from non-mild (moderately severe plus severe) AP during the initial phase of the disease.

*Paper IV*
To analyse the temporal development of biomarkers in AP with regard to the severity categories of the revised Atlanta classification.
Materials and Methods

Paper I

Study design and cohort

Paper I is based on a multicentre study originated from a research project during the Pancreas 2000 course – an international course in pancreatology. Adult patients with AP were retrospectively enrolled at 6 European centres (Skåne University Hospital, Malmö, Sweden; Hospital del Mar, Barcelona, Spain; University Hospital of Emergency Medicine “Pigorov”, Sofia, Bulgaria; Lithuanian University of Health Sciences, Kaunas, Lithuania; East Tallinn Central Hospital, Tallinn, Estonia and Helsinki University Hospital, Helsinki, Finland) going backwards from January 2013 to January 2012. All included patients had at least one CECT performed during the course of the disease, and all CECTs up until 3 months after date of admission were included. Thereby CTs from the whole spectra of the disease were obtained. Patients who had had an invasive intervention (except from ERCP) were not included, neither were patients with a history of chronic pancreatitis. All CECTs were performed in the pancreatic and/or the portal venous phase.

At each centre one local radiologist, specialised in gastroenterological radiology, reviewed the CECTs using a protocol based on the morphological criteria stated by the revised Atlanta classification (see Box 1 in the Introduction) [73]. All local radiologists received an instruction sheet as interpretation support before scoring the CECTs. All CECTs were additionally reviewed according to the same protocol by a central radiologist (Thomas L Bollen), who is an expert in pancreatic radiology. The local and central radiologists had information on time elapsed from onset of disease until CECT performance but were blinded to other clinical data.

The following information was obtained from the medical records; SIRS upon admission, highest level of CRP during hospitalisation, need for invasive intervention, organ failure (according to the RAC) and in-hospital mortality. All patients were retrospectively classified according to the RAC [73].
Statistics

Continuous data were presented as median (range) and in analysis Mann-Whitney U test was performed. Paired samples were analysed using Wilcoxon rank test. P<0.05 was considered statistically significant. In sample size calculation 360 CECTs were estimated to give sufficient statistical power based on the aspect that the RAC has several categories divided into 2, 3 or 4 parts. 50 patients were enrolled at each centre, with the assumption that numerous patients would have had more than one CECT performed during the course of the disease. From scoring results, the interobserver agreement between the local and central radiologists was calculated using Cohens kappa test. For defined kappa values and corresponding levels of agreement see table 4. All analysis was performed using IBM SPSS Statistics for Windows version 21 and 22, Armonk, NY:IBM Corp.

<table>
<thead>
<tr>
<th>Kappa values</th>
<th>Level of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00-0.20</td>
<td>Slight</td>
</tr>
<tr>
<td>0.21-0.40</td>
<td>Fair</td>
</tr>
<tr>
<td>0.41-0.60</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.61-0.80</td>
<td>Substantial</td>
</tr>
<tr>
<td>0.81-1.00</td>
<td>Good</td>
</tr>
</tbody>
</table>

Paper II-IV

Study design and cohort

Papers II-IV are based on a mutual cohort, originating from a study named the AMY-study. Patients > 18 years old with AP were prospectively and consecutively enrolled at Skåne University Hospital, Malmö, from January 2010 until September 2013. Patients where more than 72 hours had passed since onset of disease were excluded. The grade of severity was either classified according to the Atlanta classification of 1993 or the revised Atlanta classification of 2013 [73, 76]. Clinical data including exact time for onset of disease (i.e. onset of pain) and validated questions on comorbidities were collected upon inclusion and through medical notes.
**Biomarkers**

Serum samples were taken upon arrival to the emergency room and daily as long as the pancreatic amylase was increased. Samples for analysis of IFN-γ, IL-1β, IL-6, interleukin-6 receptor (IL-6R), IL-8, IL-10, interleukin-12 (IL-12), TNF-α, MCP-1, procalcitonin and d-dimer were collected in plasma separator tubes, centrifuged (2000 rounds, 25°C, 10 minutes) and stored at -80°C until analysed. IFN-γ, IL-1β, IL-6, IL-8, IL-10, IL-12 and TNF-α were analysed with a human proinflammatory 7-plex ultrasensitive kit and IL-6R with 1-plex human IL-6R ultrasensitive kit (K15008C, K151ALC, Meso Scale Diagnostics LLC, Rockville, MD, USA). MCP-1 was assessed through human CCL2 (MCP-1) Elisa Ready-set-go kit (88-7399-88, aBioscience, San Diego, CA, USA) and d-dimer with human D-Dimer Elisa kit (D2D, 20870, Bmassay, Beijing, China). All analyses were assessed according to the manufacturer’s instructions. Procalcitonin was determined with an accredited Elisa method based on monoclonal anti-procalcitonin antibodies in accordance with routine methods at the department of Clinical Chemistry, Skåne University Hospital Malmö.

**BOX 3**  
Biomarker units

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-1β, IL-6, IL-6R, IL-8, IL-10, IL-12, IFN-γ, TNF-α</td>
<td>pg/ml</td>
</tr>
<tr>
<td>MCP-1, d-dimer, procalcitonin</td>
<td>ng/ml</td>
</tr>
<tr>
<td>Albumin, hemoglobin</td>
<td>g/L</td>
</tr>
<tr>
<td>CRP</td>
<td>mg/L</td>
</tr>
<tr>
<td>Calcium, glucose, lactate</td>
<td>mmol/L</td>
</tr>
<tr>
<td>Thrombocytes, white blood cells</td>
<td>x10⁹/L</td>
</tr>
</tbody>
</table>

Albumin, calcium, CRP, glucose, hematocrit, haemoglobin, lactate, thrombocytes and white blood cells were analysed in accordance with certified standard analysis at the department of Clinical Chemistry, Skåne University Hospital Malmö (ISO 15189:2012, accreditation number 1309).

In Paper II the cut-off levels for analysed biomarkers were set through literature review. Solely admission samples were utilised in calculations.

In Paper III samples taken between 13-36 hours after onset of disease were used for analysis. The Harmless Acute Pancreatitis Score (HAPS) was set for all patients.\(^{[133]}\)

In Paper IV samples taken at 0-24 hours and 25-48 hours after onset of pain were compared.
Statistics

Data is presented as median (with range and interquartile range), numbers (and percentages) or as mean with standard deviation (SD). All statistic analysis were performed using nonparametric tests. For comparison between two groups the Mann-Whitney U test was used and $p<0.05$ was considered statistically significant. Kruskal Wallis test was performed for comparison between three groups with an adjusted level of significance; $p<0.017$. Differences in biomarker levels between paired samples were analysed through Wilcoxon's test.

Optimal cut-off levels for individual biomarkers were acquired from Receiver Operating Characteristic curves (ROC-curves). Through selected cut-off levels sensitivity, specificity, negative and positive likelihood ratios as well as negative and positive predictive values were calculated. For evaluation of biomarker performances areas under the curves (AUCs) were obtained from the ROC-curves.

All analysis were performed using IBM SPSS Statistics for Windows, version 22 and 23, Armonk, NY:IBM corp.
Results

Paper I

Patients and CECTs

In total 301 patients with 405 CECTs were included at the six centres. Due to insufficient CECT quality 16 patients were excluded leaving 388 CECTs from 285 patients for analysis. Several of the patient characteristics like gender, age and aetiology of AP differed substantially between the centres. As all patients had had a CECT performed the grades of severity were slightly higher compared to an unselected AP cohort. Median time, for all centres, from onset of disease to CECT was seven days (range 0-90).

Interobserver agreement

As seen in table 5 the centre dependent kappa values varied distinctly between the centres. In seven categories the level of agreement was substantial (0.61-0.80); Necrosis-neck, Necrosis-body, Necrosis-tail, Presence of collections, Location of collections, Presence of wall and Presence of intraluminal gas/fluid level. Moderate agreement (0.41-0.60) was found for Parenchymal necrosis as well as Necrosis-head. Finally fair agreement (0.21-0.40) was reached on the categories Type of pancreatitis, Extrapancreatic necrosis, Characteristics of collection and Collection – most appropriate term.

Further investigation of the categories with fair level of agreement demonstrated large discrepancies in the identification of extrapancreatic necrosis and necrotic contents of collections. The central expert identified significantly more cases with necrosis compared to the local radiologists. The clinical outcome of the patients had superior correspondence with morphological findings according to the central expert.
### Table 5
Centre independent (All) and dependent (Centre B-G) kappa values for all categories scored

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of pancreatitis</td>
<td>0.370</td>
<td>0.317</td>
<td>0.342</td>
<td>0.309</td>
<td>0.098†</td>
<td>0.838*</td>
<td>0.360</td>
</tr>
<tr>
<td>Parenchymal Necrosis</td>
<td>0.539</td>
<td>0.380</td>
<td>0.319†</td>
<td>0.609</td>
<td>0.731*</td>
<td>0.663</td>
<td>0.465</td>
</tr>
<tr>
<td>Necrosis - Head</td>
<td>0.516</td>
<td>0.669</td>
<td>0.345</td>
<td>0.323†</td>
<td>1.00 *</td>
<td>0.660</td>
<td>0.646</td>
</tr>
<tr>
<td>Necrosis - Neck</td>
<td>0.618</td>
<td>0.922*</td>
<td>0.577</td>
<td>0.822</td>
<td>0.660</td>
<td>0.364</td>
<td>0.236†</td>
</tr>
<tr>
<td>Necrosis – Body</td>
<td>0.628</td>
<td>0.766</td>
<td>0.611</td>
<td>0.687</td>
<td>0.873</td>
<td>0.392†</td>
<td>0.570</td>
</tr>
<tr>
<td>Necrosis – Tail</td>
<td>0.617</td>
<td>0.451</td>
<td>0.626</td>
<td>0.687</td>
<td>0.409†</td>
<td>0.806*</td>
<td>0.532</td>
</tr>
<tr>
<td>Extrapancreatic Necrosis</td>
<td>0.326</td>
<td>0.321</td>
<td>0.504</td>
<td>0.293</td>
<td>0.120</td>
<td>0.877*</td>
<td>0.024†</td>
</tr>
<tr>
<td>Collections</td>
<td>0.756</td>
<td>0.780</td>
<td>0.750</td>
<td>0.624†</td>
<td>0.864*</td>
<td>0.627</td>
<td>0.625</td>
</tr>
<tr>
<td>Location of Collections</td>
<td>0.633</td>
<td>0.728</td>
<td>0.761*</td>
<td>0.508</td>
<td>0.694</td>
<td>0.604</td>
<td>0.439†</td>
</tr>
<tr>
<td>Characteristics of Collections</td>
<td>0.408</td>
<td>0.397</td>
<td>0.485</td>
<td>0.293</td>
<td>0.305</td>
<td>0.744*</td>
<td>0.251†</td>
</tr>
<tr>
<td>Wall</td>
<td>0.675</td>
<td>0.638</td>
<td>0.638</td>
<td>0.588†</td>
<td>0.777*</td>
<td>0.726</td>
<td>0.632</td>
</tr>
<tr>
<td>Intraluminal gas/fluid level</td>
<td>0.764</td>
<td>0.774</td>
<td>0.671†</td>
<td>0.764</td>
<td>0.887*</td>
<td>0.837</td>
<td>0.675</td>
</tr>
<tr>
<td>Collection – most appropriate term</td>
<td>0.356</td>
<td>0.385</td>
<td>0.480</td>
<td>0.136†</td>
<td>0.218</td>
<td>0.673*</td>
<td>0.261</td>
</tr>
</tbody>
</table>

* Highest kappa value for center B to G for each category
† Lowest kappa value for center B to G for each category

### Paper II

#### Patients

To enable comparison with previous studies the patients were retrospectively classified as having mild or severe disease according to the Atlanta classification of 1993 [76]. Among the 232 patients enrolled 193 (83.2%) had mild and 39 (16.8%) severe AP. The only significant difference in baseline characteristics found between the groups was level of CRP within the first 72 h which was higher in the severe group (p<0.0001).
Biomarkers and cut-off levels

The cut-off levels presented in Box 4 were assessed from literature review.

**BOX 4**

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Preset cut-off level</th>
<th>Biomarker</th>
<th>Preset cut-off level</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-1β</td>
<td>1 pg/ml</td>
<td>TNF-α</td>
<td>10 pg/ml</td>
</tr>
<tr>
<td>IL-6</td>
<td>100 pg/ml</td>
<td>MCP-1</td>
<td>500 ng/ml</td>
</tr>
<tr>
<td>IL-8</td>
<td>40 pg/ml</td>
<td>Procalcitonin</td>
<td>0.5 ng/ml</td>
</tr>
<tr>
<td>IL-10</td>
<td>7.5 pg/ml</td>
<td>d-dimer</td>
<td>0.5 ng/ml</td>
</tr>
</tbody>
</table>

IL-1β (p=0.008), IL-6 (p=0.013) and IL-10 (p=0.009) differed significantly between the patients with mild and severe disease upon admission.

Application of the preset cut-off levels on our cohort did not result in any acceptable sensitivity, specificity and positive or negative predictive value. A considerable number of the patients with severe AP were found below cut-off levels.

To acquire optimal cut-off levels for our cohort ROC curves were performed with results as presented in table 6.

**Table 6**

Areas under curves and optimal cut-off levels for the individual biomarkers

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Optimal cut-off</th>
<th>AUC</th>
<th>Biomarker</th>
<th>Optimal cut-off</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-1β</td>
<td>0.9 pg/ml</td>
<td>0.640</td>
<td>TNF-α</td>
<td>n.a</td>
<td>0.459</td>
</tr>
<tr>
<td>IL-6</td>
<td>72 pg/ml</td>
<td>0.623</td>
<td>MCP-1</td>
<td>399 ng/ml</td>
<td>0.604</td>
</tr>
<tr>
<td>IL-8</td>
<td>49 pg/ml</td>
<td>0.601</td>
<td>Procalcitonin</td>
<td>0.35 ng/ml</td>
<td>0.552</td>
</tr>
<tr>
<td>IL-10</td>
<td>12 pg/ml</td>
<td>0.638</td>
<td>d-dimer</td>
<td>n.a</td>
<td>0.516</td>
</tr>
</tbody>
</table>

*n.a. not applicable*

The biomarkers were additionally analysed using the revised Atlanta classification without reaching superior results.
Paper III

Patients

In the established cohort (Paper I) 175 patients had their admission sample taken between 13-36 hours after onset of pain. The patients were divided into mild (70.9%) and non-mild (moderately severe + severe) (29.1%) AP in accordance with the revised Atlanta classification [73].

Biomarkers

Routine (Albumin, calcium, CRP, glucose, hematocrit, haemoglobin, lactate, thrombocytes, white blood cells) and non-routine biomarkers (IFN-γ, IL-1β, IL-6, IL-6R, IL-8, IL-10, IL-12, TNF-α, MCP-1, procalcitonin and d-dimer) were analysed. Significant difference (p<0.05) between the mild and the non-mild group was found for IL-1β, IL-6, IL-6R, IL-10, MCP-1, calcium and CRP. In ROC curve analysis, CRP and IL-6 presented superior ability to differentiate the non-mild patients with AUCs of 0.808 and 0.742 respectively. With the high morbidity and mortality of the disease taken into consideration optimal cut-off levels for CRP (57 mg/L) and IL-6 (pg/ml) were selected, resulting in predictive capacities for non-mild AP as presented in table 7. Additionally, the result of combining CRP and IL-6 is demonstrated.

Table 7
Predictive capacity of CRP, IL-6 and the combination of both together.

<table>
<thead>
<tr>
<th></th>
<th>Cut-off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>LR+</th>
<th>LR-</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRP</td>
<td>57</td>
<td>0.88</td>
<td>0.5</td>
<td>1.76</td>
<td>0.24</td>
<td>0.42</td>
<td>0.87</td>
</tr>
<tr>
<td>IL-6</td>
<td>23.6</td>
<td>0.89</td>
<td>0.54</td>
<td>1.87</td>
<td>0.20</td>
<td>0.39</td>
<td>0.93</td>
</tr>
<tr>
<td>CRP + IL-6</td>
<td>57 + 23.6</td>
<td>0.98</td>
<td>0.53</td>
<td>2.1</td>
<td>0.06</td>
<td>0.49</td>
<td>0.98</td>
</tr>
</tbody>
</table>

LR+ positive likelihood ratio, LR negative likelihood ratio, PPV positive predictive value, NPV negative predictive value

CRP in mg/L, IL-6 in pg/ml

Application of the HAPS and the combination of CRP and IL-6 with regard to mild AP are presented in table 8. One patient classified as harmless according to the HAPS developed moderately severe AP with transient multiorgan failure. The patient with non-mild disease and CRP and IL-6 values found below assessed cut-off levels had a deterioration of comorbidity (short period of atrial fibrillation).
**Table 8**
HAPS score and CRP + IL-6 for the prediction of mild disease

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Sens</th>
<th>Spec</th>
<th>LR+</th>
<th>LR-</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAPS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harmless</td>
<td>10</td>
<td>1</td>
<td>0.8</td>
<td>0.98</td>
<td>4</td>
<td>0.94</td>
</tr>
<tr>
<td>Not harmless</td>
<td>114</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRP+IL-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>71</td>
<td>1</td>
<td>0.53</td>
<td>0.98</td>
<td>26.5</td>
<td>0.48</td>
</tr>
<tr>
<td>Non-mild</td>
<td>53</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

HAPS harmless acute pancreatitis score, Non-mild moderately severe plus severe AP, Sens sensitivity, Spec specificity, LR+ positive likelihood ratio, LR- negative likelihood ratio, PPV positive predictive value, NPV negative predictive value

**Paper IV**

**Patients**

Among the 232 patients enrolled in the AMY-study, 115 had samples taken at both 0-24 hours and 25-48 hours after onset of pain. This smaller cohort of 115 patients was used for paired comparisons.

In the whole cohort 70.3% developed mild, 22.8% moderately severe and 6.9% severe AP. The group with paired samples (n=115) did not differ significantly from the entire cohort regarding basal characteristics. There were more severe cases in the paired group (9.6% vs 6.9%), admission to the ICU was slightly lower (72.2% vs 81.35%), however mortality in the severe group was higher with 45.5% compared to 31.3% in the large original cohort.

**Biomarkers**

Using the human proinflammatory 7-plex ultrasensitive kit IFN-γ, IL-1β, IL-6, IL-8, IL-10, IL-12, TNF-α were measured. In the whole cohort significant differences between severity grades were found for IL-1β, IL-6, IL-8 and IL-10, thus these biomarkers were further investigated. Among admission samples (of the large group) the variations were most expressed for IL-1β, IL-6 and IL-10.

Paired samples (n=115) of IL-1β, IL-6, IL-8 and IL-10 were analysed separately. The mean value of each severity group was calculated for the individual biomarkers at 0-24 and 25-48 hours after onset of disease. Significant changes in
mean values were found for IL-8 and IL-10 in the mild group and for IL-1β and IL-6 in the moderately severe and severe groups, see table 9.

Table 9
Mean values of paired samples of biomarkers at 0-24 h and 25-48 h in the different severity groups

<table>
<thead>
<tr>
<th></th>
<th>MAP (n= 71)</th>
<th></th>
<th>MSAP (n= 33)</th>
<th></th>
<th>SAP (n= 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-1β</td>
<td>2 ±0.6</td>
<td>2 ±0.3</td>
<td>ns</td>
<td>2.4 ±0.7</td>
<td>5.9 ±1.5</td>
</tr>
<tr>
<td>IL-6</td>
<td>194 ±84</td>
<td>175 ±88</td>
<td>ns</td>
<td>394 ±254</td>
<td>221 ±44</td>
</tr>
<tr>
<td>IL-8</td>
<td>188 ±69</td>
<td>76 ±21</td>
<td>0.0001</td>
<td>188 ±64</td>
<td>102. ±15</td>
</tr>
<tr>
<td>IL-10</td>
<td>108 ±64</td>
<td>40 ±14</td>
<td>0.0001</td>
<td>202 ±118</td>
<td>27 ±6</td>
</tr>
</tbody>
</table>

MAP mild AP, MSAP moderately severe AP, SAP severe AP, ns not significant, h hours
Values are in mean (± SD)
All interleukins are in pg/ml

Differences in mean values (of 0-24 and 25-48 hours) of the biomarkers were named delta-values. We investigated whether the delta-values for the individual markers differed between severity groups, see table 10. Although all values appeared to differ significantly, statistical significance was only found for IL-1β, IL-6 and IL-10 when comparing the mild and severe groups.

Table 10
Delta-values and differences between severity groups

<table>
<thead>
<tr>
<th></th>
<th>Delta-values</th>
<th>p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAP n=71</td>
<td>MSAP n=33</td>
</tr>
<tr>
<td>IL-1β</td>
<td>0.24 ±4.4</td>
<td>3.5 ±8</td>
</tr>
<tr>
<td>IL-6</td>
<td>16 ±679</td>
<td>173 ±1326</td>
</tr>
<tr>
<td>IL-8</td>
<td>1112 ±411</td>
<td>85 ±331</td>
</tr>
<tr>
<td>IL-10</td>
<td>68 ±399</td>
<td>175 ±826</td>
</tr>
</tbody>
</table>

MAP mild AP, MSAP moderately severe AP, SAP severe AP, ns not significant
Values are in mean (± SD)
All interleukins are in pg/ml

42
Discussion

Acute pancreatitis

Clinical expressions of acute pancreatitis varies from the mildly ill patient with pain in the upper abdomen alone to extremely severe cases with multiorgan failure who dies in the ICU after weeks, or sometimes just hours, of struggling against the disorder. How the disease profile and critical pathophysiological processes differ between these patients is still unclear. In the present thesis these problems are addressed from a routine clinical perspective.

Radiological assessment

The elaboration of the revised Atlanta classification started in 2007 through a web-based consultation. A wide range of pancreatology experts from national and international associations were invited to join the project. The initial draft was revised in response to comments, and the statements of the final version were based on published evidence [73]. Regardless, the content of the classification was established through experts opinions. This severity stratification system has been multiply validated with good results [86, 87, 88, 95]. However, the clinical applicability of morphological features and radiological assessment of the classification have been questioned [161, 162].

Paper I Significant inter-observer variation in the diagnosis of extrapancreatic necrosis and type of pancreatic collections in acute pancreatitis – An international multicenter evaluation of the revised Atlanta classification

The revised Atlanta classification acknowledge both peripancreatic necrosis and necrotic debris of collections to be difficult to identify in the initial course of the disease [73]. The process of enhancement impairment normally evolves over days
making early CECTs difficult to interpret and thus unreliable for morphologic categorisation.

With this in mind we specifically addressed this particular problem in the instruction sheet provided to all local radiologists prior to CECT interpretation. Additionally, for all parameters in the scoring sheet the category “indeterminate” was available. In cases of doubt regarding the existence of necrosis the local radiologists were encouraged to select “indeterminate”.

When analysing the kappa values we found in general sufficient agreement between the central expert and the local radiologists concerning morphological assessment of AP. For clinically important findings such as presence of parenchymal necrosis, signs of infection and presence of collections the level of agreement was good (kappa values 0.54, 0.76 and 0.76).

However, large differences were seen in the interpretation of four categories; Type of pancreatitis, Extrapancreatic necrosis, Characteristics of collections and Collections - most appropriate term. The central expert diagnosed peripancreatic necrosis, acute necrotic collections and consequently necrotising AP to a much larger extent. He also selected the term indeterminate more often, in particular concerning presence of extrapancreatic necrosis (12% vs 4%). We suggested that the local radiologists were unfamiliar with the new morphology, in particular when interpreting CECTs from the early phase of the disease. Additionally, specific interest in pancreatic radiology as well as yearly number of cases investigated are likewise important. The radiologist from centre F has since previously a special focus on pancreatic radiology, mirrored by the excellent interobserver agreement with the central expert (table 5). Centre F also has (among the participating centres) annually the largest amount of patients with AP admitted including a substantial number of severe cases.

Correct classification is important for several reasons. The clinical outcome differs between the morphological types of AP and although radiological findings not always changes the management strategies it is still an essential part for the understanding of the clinical picture of patient and the disease. In our study the scoring by the central expert to a larger extent correlated with the clinical outcome of the patients.

For research purposes, development of medical therapies and comparison of inter-institutional data the interpretation of findings need to be identical. One of the aims with the revised Atlanta classification was to clarify the definitions on AP. Nevertheless our study indicates that the morphological assessment pertaining to the identification of extrapancreatic necrosis and necrotic debris remains inadequate with a risk of misclassifying the AP patients.
Biomarkers

Countless numbers of attempts have been made to perceive predictive biomarkers in AP. Despite promising results the conduction into clinical applicability has been scarce. Whether additional studies are of value is thereby a legitimate question. Considering the large amount of efforts and the minimal impact on routine management we need to address the primary issues of the disease. Is the pathogenesis in AP too complex or the inter-individual variation overly significative for the possibility to make general conclusions? Or is it rather about shortcomings in study methodologies?

Although there have been advances, knowledge on the decisive steps of the pathophysiological process and profound insight into the role of inflammatory markers are still lacking. Additionally, when studying biomarkers in AP an evident fact is that all severity categories contain both high and low responders, making the range for each marker widely scattered also within severity groups. The results from the AMY-study analogically demonstrated explicitly heterogenetic circumstances. For all biomarkers we found what seemed to be a distinct continuous rise in mean value when comparing mild against moderately severe and severe AP. However, for several of the biomarkers statistical analysis could not demonstrate any significant differences, giving the impression that the individual disparities override the discrepancies of severity groups.

**Paper II Predictive capacity of biomarkers for severe acute pancreatitis**

The setup of this study was to compile cut-off levels from previous studies on biomarkers in AP and apply them on our unselected cohort which also has a normal distribution of severity grades. By this approach we aimed to resemble the routine clinical situation.

The predetermination of cut-off levels was more complicated than expected as we found earlier works to be expressly disparate in setups and outcomes making comparisons difficult. Finally only 14 studies met our inclusion criteria, still the cut-off levels of these studies were substantially diversified and approximations had to be made in value assessment. When applied on our cohort the preset cut-off levels demonstrated suboptimal capacity for severity prediction with low AUCs and insufficient sensitivity and specificity. In recalculation, using ROC curves, optimal cut-offs for selected biomarkers were within acceptable range from those presented by previous works. The prognostic ability, estimated in AUCs, as well as sensitivities and specificities, were however still on a low level.
The reasons for biomarkers demonstrating various results pertaining to predictive capacity are likely multiple. Numbers of patients with severe AP as well as age, gender and aetiology are factors that are likely to have impact. The temporal development of the disease and the aspect of time concerning sample collection might additionally influence. In our study no difference was found between the mild and severe group regarding these basic characteristics. These findings strengthen the assumption that the AMY-cohort consists of unselected AP patients.

**Paper III**  
**IL-6 and CRP are superior in early differentiation between mild and non-mild acute pancreatitis**

With the aim of investigating the initial phase of AP samples from the AMY-study taken between 13-36 hours into the course of the disease were used. The time frame was set based on the results from a large review by Staubli et al where initial results were presented \[140\]. As the group with mild AP is large but fairly uninvestigated and concurrently identification of the non-mild (moderately severe and severe) patients is prioritised we aimed to explore early differences between these two groups. To the best of our knowledge such a study had not been previously executed. The 175 patients included were thus separated into mild and non-mild AP. Biomarkers demonstrating positive performance in earlier studies in AP were analysed \[40, 45, 137, 138, 141\]. We found that a combination of IL-6 and CRP identified the non-mild group with high AUC, good sensitivity and acceptable specificity (table 7). Due to the high morbidity and possible mortality of AP cut-off levels for IL-6 and CRP were chosen to obtain high sensitivity.

Comparison was additionally made with the HAPS which is the only score compiled for mild AP \[133\]. By applying IL-6 plus CRP (using defined cut-off levels) the specificity was equal to what was found for the HAPS. However in all other analysis the combination of IL-6 and CRP was superior in the identification of mild cases (table 8). The HAPS has previously been validated with good results \[134\]. Nevertheless this scoring system has not reached clinical practice. In our study the HAPS certainly demonstrated high specificity for the mild cases but the sensitivity of 0.08 is far too low to be clinical usable.

Since there is a lack of predictive scoring systems the vast majority of the AP patients are currently hospitalised with supportive care and various interventions \[163\]. As the incidence of AP is increasing future identification of both mild and non-mild cases will be utterly essential, both due to the high morbidity of the disease but also for health care costs \[6\]. As CRP and IL-6 are fairly inexpensive and generally clinically available we find the results of this study valuable and worth further investigation, preferably in an international multicenter study.
The temporal dynamics of biomarkers in AP has not been extensively investigated in clinical studies. In general, exact amount of hours from onset of pain to collection of blood sample is rarely reported and all admission samples up until 72 hours after onset of disease are analysed together. However, in a disease with such rapid alterations a difference in hours between sample collections is likely to matter. By obtaining information on exact time lapse from disease onset to sample we aimed to give a detailed description of biomarker development in the initial phase of AP.

In this study we compared mean values of serum samples taken at 0-24 hours and 25-48 hours after onset of disease to examine the development of inflammatory reactions between but also within severity grades. According to discrepancies in mean values all biomarkers seemed to differ significantly from 0-24 to 25-48 hours into the disease. However, in statistical analysis this was not the case due to the widespread individual variation in biomarker levels (table 9). We found a significant raise in mean level between day one and two for IL-1β and IL-6 regarding the moderately severe and severe AP patients and likewise a significant decrease for IL-8 and IL-10 in the mild group. In comparison of severity groups, the differences in mean values (named delta-values) of IL-1β, IL-6 and IL-10 varied significantly between the mild and the severe group (table 10).

Herein we have elucidated the development and temporal changes of interleukins during the initial phase of AP. Our results indicate that interleukins, in particular IL-1β and IL-6, could be valuable tools in the early prediction of AP. However, further investigations, preferably large international multicenter studies, are needed to clinically establish this statement.

Methodological considerations

This study contained retrospectively enrolled patients, whereas the CECTs were scored prospectively. All CECTs were performed between January 2012 and 2013, thus imaging techniques and patient management remained unchanged during the short inclusion period. Likewise collection of clinical data was uncomplicated. Due to the retrospective design, correlation between observed inconsistencies and
actual clinical decision-making was not investigated. An overall prospective setting would have been preferable but most likely not executable with regard to the time frame set for the project.

It is recognised that necrotic material, in particular within collections, is often disregarded on CT \cite{73}. Both ultrasound and MRI are superior diagnostic methods for the depiction of necrotic material \cite{98, 103}. Still, CT is globally the dominant imaging modality emphasising the importance of knowledge regarding possible areas of indeterminate or ambiguous interpretation.

As described previously in this thesis, the process of the revised Atlanta classification started as an iterative web process. The base of the classification is built on experts opinions and not validated data, although all statements were derived from accepted publications. It is generally considered to be a clear and simple classification, developed with the clinician in mind. However, the moderately severe category has been criticised for being diffuse as it contains a wide spectra of disease severity - from patients with deterioration of a comorbidity to those with multiorgan failure.

In this study only one central expert radiologist scored the CECTs potentially introducing bias. However high level of agreement was obtained with the local radiologist from centre F, who has a similar level of expertise. Additionally the central expert had an important role in the development of the RAC and his interpretation demonstrated superior correlation with clinical outcome.

\textit{Paper II-IV}

The AMY-study was a prospective cohort study with consecutively enrolled, and thus unselected, AP patients included from January 2010 to September 2013. By this approach we aimed to assemble a cohort representative of the disease regarding aetiologies, gender, age and number of severe cases. The information on exact time for onset of pain was specifically requested upon inclusion.

All serum samples were analysed within a short time frame by one single laboratory engineer (Anne-Marie Rohrstock) using established commercial methods, thereby diminishing the amount of variability sources.

To assess the predictive capacity of biomarkers several analytical approaches exist. In our works we have used ROC-curves, AUCs, sensitivity, specificity, positive and negative predictive values and positive and negative likelihood ratios to describe the individual performances of the biomarkers. By applying various methods we aimed to present the results from different angles for extended understanding.
The setting for Paper II was designed in 2009, at a time when the framing of the RAC was yet in its infancy and the valid classification was still the AC. Thus Paper II was designed as a validation study of cut-off levels from previous studies on biomarkers in AP. All earlier works classified the patients according to the 1993 Atlanta classification and for this reason our patients were separated accordingly, although the AC was out of date at the time for our final analysis.

All studies based on the AMY-study are of an exploratory and descriptive nature, implying some limitations. In Paper III the capacity of biomarkers was investigated and cut-off levels for the separation of mild versus non-mild disease was suggested. However the aim of the study was not to generate a prognostic score, then a validation set would have been required. Additionally, the most interesting group of patients, from a clinical perspective, are those who develop severe AP. Nevertheless, this is also the smallest group consequently increasing the risk for type II errors with false negative results due to low numbers of patients. In general our findings, in particular those of Paper III and IV, thus need to be confirmed in larger studies.

In paper IV only a subgroup (n=115) and not the whole cohort had samples taken at both 0-24 and 25-48 hours after onset of disease. However this subgroup did not differ significantly in any basic characteristic compared to the large group. In this study the group with severe AP (in both the whole cohort and the subgroup) was significantly older compared to those with mild and moderately severe disease. As age is an established risk factor for severe disease supplementary multivariate analysis have been performed, demonstrating no association in our material between age and grade of severity.
Conclusions

*Paper I*
There are areas of controversy regarding the morphologic CT-criteria established by the revised Atlanta classification, especially pertaining to the identification of extrapancreatic necrosis and necrotic debris of collections.

*Paper II*
The variations in approaches and outcomes of earlier studies on predictive biomarkers in AP are substantial making comparisons complicated. In a cohort with consecutively enrolled patients application of previously defined cut-off levels demonstrates in general insufficient prognostic capacity of biomarkers.

*Paper III*
In the differentiation of mild and non-mild (moderately severe plus severe) AP CRP and IL-6 demonstrate good predictive capacity during the initial course of the disease.

*Paper IV*
During the first 48 hours of AP mean levels of IL-1β and IL-6 increase significantly in patients with moderately severe and severe disease. Differences in mean values between day one and two were explicit for IL-1β, IL-6 and IL-10 when comparing severity groups.
Future perspectives

The incidence of AP is increasing globally, accompanied by patient suffering and rising costs [6]. Thus there is a substantial need for urgent development in severity stratification methods and therapy approaches. A recent review of the quality of medical and surgical care of AP patients in the United Kingdom stated four areas of improvement; disease prevention, management of antibiotics, utilisation of existing prognostic scores and finally development of national networks and treatment strategies [164].

Prevention
The rise in number of AP patients can be explained by a wider accessibility of assessment tools, i.e sampling and imaging. Another important cause is likely the global epidemic in obesity. Obesity (body mass index ≥ 30) has been correlated to an increased risk ratio of 2.5 for gallstones but also an independently augmented risk ratio for AP (2.2) [165]. Additionally, obesity has demonstrated to correlate with level of disease severity [125, 160]. Preventive interventions to reduce the numbers of overweight and obese humans would likely have an impact on AP incidence. Additionally it has been stated that patients with gallstones clearly benefit from same-admission cholecystectomy, however these recommendations are to a large extent not complied with [72, 156, 164].

Pathophysiology and biomarkers
Although considerable advances have been reached much is still left regarding our understanding of the pathogenesis in AP. Profound knowledge on molecular events is important not only for early severity stratification but also for the development of targeted therapies. Currently numerous experimental efforts are being made to proceed in this direction, however to reach clinical impact results need to be confirmed in large prospective clinical multicenter studies. For the future, shortening of the distance between the laboratory and patients, i.e implementation of basic science into clinical practice, is crucial [167]. Additionally, new biomarkers have to be explored. Recently the interest in free fatty acids have
increased due to their cytotoxicity and capacity to up-regulate pro-inflammatory mediators. Also, the distribution of visceral fat has gained attention [168, 169, 170]. Genetic mechanisms in pancreatic disease have mainly been investigated regarding chronic pancreatitis and pancreatic adenocarcinoma. However, in recent years the interest in genetic alterations have increased also for AP, a development which might expand our understanding of the disease [167, 171].

Management
An essential challenge for the future is the improvement of clinical management of the AP patients. Currently all patients are hospitalised and treated identically. Overall, interventions are symptom-related as there is no specific therapy towards the disease itself.

In a recent Cochrane review of pharmacological therapy in AP the conclusion was that none of the investigated randomised controlled trials could demonstrate any benefits in the interventions groups regarding mortality [172]. Additionally, due to substantial variations in study approaches comparisons of morbidity outcomes were inapplicable. In general the quality of evidence was low and numerous systematic errors and biases were found.

Development of targeted therapies together with well-planned and strict clinical studies are necessary to reach success regarding novel treatments. Patients need to be correctly classified and the results of experimental studies should be confirmed clinically [167].
Populärvetenskaplig sammanfattning

Akut pankreatit, inflammation i bukspottskörteln, är en relativt vanlig sjukdom som ökar världen över. Sjukdomen orsakas i de flesta fall av gallsten eller alkoholöverkonsumtion. Symtomen är hastigt insändande smärta i övre delen av buken. Diagnos sätts med hjälp av blodprov, där man ser förhöjt bukspottskörtelenzyme, i kombination med att patienten har ont i magen. Majoriteten av patienterna drabbas av en mild, själveliminerande form av sjukdomen. Cirka 30 % får dock en medelsvår eller svår akut pankreatit vilket leder till en lång sjukhusvistelse med organsvikt och ofta vård på intensivvårdsavdelning. Mellan 20-40% av patienterna i den svåra gruppen avlider i sjukdomen. Vid den medelsvåra och svåra formen är risken också stor att drabba av livslånga besvär, såsom diabetes och kroniska buksmärtor.

Ännu idag vet man inte varför endast vissa patienter blir svårt sjuka. Trots mycket forskning har man inte heller funnit någon metod (t.ex blodprov eller undersökning) som tidigt i förloppet kan avslöja vilken svårighetsgrad patienten kommer att utveckla. Därför behandlas alla som kommer till sjukhus med akut pankreatit som att de skulle kunna bli svårt sjuka vilket för en stor del av patienterna innebär överbehandling. Det finns inte heller någon specifik medicin mot sjukdomen utan endast understödande behandling när olika organ sviktar.

Tidigare genomgick många patienter stora kirurgiska ingrepp, t.ex för att ta bort död vävnad runt bukspottskörteln, medan man nu försöker hålla en mer konservativ linje. Hos de patienter där den akuta pankreatiten orsakats av gallsten rekommenderas dock att man tar bort gallblåsan i anslutning till aktuellt sjukdomstillfälle.

dock fortfarande okända vilket medför att man inte heller har kunnat utveckla målinriktade mediciner och andra behandlingar.

Patienterna delas oftast in i svårighetsgraderna mild, medelsvår och svår akut pankreatit enligt en klassifikation (den reviderade Atlanta klassifikationen) som publicerades 2013. Denna klassifikation är en reviderad variant av en tidigare version (Atlanta 1993) som ansetts utdaterad. När man inom forskning och studier vill jämföra olika patientgrupper och deras resultat är det av yttersta vikt att patienterna klassificeras likadant oavsett vilket forskningscentra man befinner sig vid.

Syftet med denna avhandling var dels att validera några av de nya röntgenbaserade kriterier som införts i den reviderade Atlanta klassifikationen. Vi ville också undersöka olika biomarkörer vid akut pankreatit och på olika sätt kunna visa på deras förmåga att avgöra vilka patienter som kommer att får mild, medelsvår eller svår sjukdom.

I delarbete I evaluerade sex stycken röntgenläkare (lokalradiologer) vid olika europeiskt universitetssjukhus 388 skiktröntgenbilder från patienter med olika stadier och olika svårighetsgrad av akut pankreatit. Bilderna bedömdes utifrån de morfologiska kriterier som anges i den reviderade Atlanta klassifikationen. Samma röntgenbilder utvärderades sedan på samma sätt av en röntgenläkare (expertradiologen) som är speciellt inriktad på just bukspottskörteln. Hur väl de lokala radiologernas och expertradiologens bedömningar stämde överens beräknades statistiskt. Vi kunde se att man var eniga gällande flera viktiga fynd, t.ex tecken på infektion och vävnadsdöd i körteln. De lokala radiologerna hade emellertid svårt att identifiera död vävnad i det fett som omger bukspottskörteln samt i de vätskeansamlingar som bildas vid medelsvår och svår sjukdom. Kännedom om dessa skillnader i tolkning av röntgenbilder är viktigt eftersom variation i bedömning gör att patienterna klassificeras olika. Samma patient kan alltså värderas som mer eller mindre sjuk beroende på vem som granskar hens röntgenbilder. Att bilder tolkas likadant är av stor vikt för korrekt medicinsk behandling men även för jämförelse av patientgrupper med samma svårighetsgrad av sjukdomen.


I delarbete III delades patienterna upp i mild och icke-mild (medelsvår plus svår) sjukdom enligt den reviderade Atlanta klassifikationen. Vi tittade sedan på om nivåerna hos biomarkörerna skilde sig mellan dessa två grupper och fann att genom att använda två vanliga markörer, CRP och IL-6, kunde patienter med mild och icke-mild sjukdom identifieras tidigt i sjuksomsförloppet.

I delarbete IV analyserade vi olika biomarkörers utveckling mellan första och andra dygnet vid akut pankreatit. Vi tittade också på om nivåerna hos biomarkörerna skilde sig mellan mild, medelsvår och svår sjukdom. Resultaten visade att medelvärdena hos markörerna IL-1β och IL-6 ökade tydligt från dag ett till dag två i den medelsvåra och svåra gruppen. Vi fann även att skillnaderna i medelvärde (så kallade deltavärden) hos IL-1β, IL-6 och IL-10 var statistiskt olika i den milda jämfört den svåra gruppen. Våra resultat tyder på att dessa biomarkörer kan vara intressanta att undersöka vidare för att tidigt kunna skilja ut de patienter som riskerar att bli svårt sjuka i sin akuta pankreatit.
Acknowledgements

Sara Regnér – my super supervisor. Clever, wise, cool, patient and always positive. I hope we have many years of collaboration ahead of us!

Henrik Thorlacius - my co-supervisor, thanks for your intellectual support and for reminding me that there are a lot of unenlightened reviewers out there…

Lucas Ilzarbe, Robert C Verdonk, Anne Penttilä, Povilas Ignatavicius, Alexandra Dimova, Peeter Koiva and Jonas Rosendahl - my co-authors, my brothers and sisters from my second family, the best Pancreas 2000 group ever! Thank you for your everlasting positive thinking, friendship and never-ending nights of fun. May we enjoy many future years together, love you!

The Pancreas 2000 squad - Bengt Holmberg, Matthias Löhr and Thomas Gasslander. Thank you so much for giving me the opportunity to join this course that have changed my life in so many perspectives!

All the mentors at the Pancreas 2000 course 6 - thanks for all your wise comments and good friendship. A special thank you to Marco del Chiaro for your honesty and wisdom, for pushing me out into world, making me face my fears and for so generously sharing your enormous global network.

Hannes Hartman - my coauthor and friend – thank you for your intellectual contributions, encouragement and good laughs!

Dorthe Johansen, Anne-Marie Rohrstock and Ingrid Palmquist – coauthor, data collectors, laboratory maestros and ladies of diligence and control. Without you nothing in this book would have been possible. Thank you!
Sophia Zackrisson, Eila Lantto, Guadalupe Aguilar, Tonis Loigom, Vanya Strahinova, Kristina Zviniene and especially Thomas L Bollen – radiologists and coauthors from all over Europe. Thank you for offering and sharing your wide knowledge and expertise in the Atlanta project.

Current and past colleagues in Malmö and Lund – thank you all for so liberally sharing your surgical competence daily.

My friends – thanks for all the energy, the wise words and the fun!

Berit, Jan and all the other relatives from the Sternby/Eilard/Jönsson/Lundsgård clan - thanks for valuable discussions and your never-ending positive thinking.

Cicki, Lennart, Per o Mats – my nearest and dearest, thank you for your continuous support.

Viktor, Oskar and Alfred – my kings, my happiness, my earth and my breath. You will always be stars - never cut down on your fantasy, your energy and your sunshine.

Jakob – you are the true definition of patience, love and support. Thank you.
References


34 Tran DD. Interleukin-6 in the early assessment of acute pancreatitis. *Gut* 1993;34:1467.


93 Banks PA, Freeman ML, Practice Parameters Committee of the American College of G. Practice guidelines in acute pancreatitis. Am J Gastroenterol 2006;101:2379-400.


96 Guo Q, Li M, Chen Y, Hu W. Determinant-based classification and revision of the Atlanta classification, which one should we choose to categorize acute pancreatitis? Pancreatology 2015;15:331-6.


Shen Y, Cui N, Miao B, Zhao E. Immune dysregulation in patients with severe acute pancreatitis. *Inflammation* 2011;34:36-42.


Tse F, Yuan Y. Early routine endoscopic retrograde cholangiopancreatography strategy versus early conservative management strategy in acute gallstone pancreatitis. *Cochrane Database Syst Rev* 2012;CD009779.


Significant inter-observer variation in the diagnosis of extrapancreatic necrosis and type of pancreatic collections in acute pancreatitis – An international multicenter evaluation of the revised Atlanta classification

Hanna Sternby a, *, Robert C. Verdonkb, Guadalupe Aguilar c, Alexandra Dimovad, Povilas Ignataviciue, Lucas Ilzarbef, Peeter Koivag, Eila Lanttoh, Tonis Loigomi, Anne Penttiläi, Sara Regnér a, Jonas Rosendahl k, Vanya Strahinoval, Sophia Zackrisson m, Kristina Zviienie, Thomas L. Bollen o

a Department of Surgery, Institution of Clinical Sciences Malmö, Lund University, Malmö, Sweden
b Department of Gastroenterology, St. Antonius Hospital, Nieuwegein, The Netherlands
c Department of Radiology, Hospital del Mar, Barcelona, Spain
d Department of Surgery, University Hospital for Emergency Medicine “Pirogov”, Sofia, Bulgaria
e Department of Surgery, Lithuanian University of Health Sciences, Kaunas, Lithuania
f Department of Gastroenterology, Hospital del Mar, Barcelona, Spain
g Department of Gastroenterology, East Tallinn Central Hospital, Tallinn, Estonia
h Department of Radiology, Helsinki University Central Hospital, Helsinki, Finland
i Department of Radiology, East Tallinn Central Hospital, Tallinn, Estonia
j Department of Surgery, Helsinki University Central Hospital, Helsinki, Finland
k Department of Internal Medicine, Neurology and Dermatology, Division of Gastroenterology and Rheumatology, Leipzig, Germany
l Department of Emergency Radiology, University Hospital for Emergency Medicine “Pigorov”, Sofia, Bulgaria
m Department of Radiology, Institution of Clinical Sciences Malmö, Lund University, Malmö, Sweden
n Department of Radiology, Lithuanian University of Health Sciences, Kaunas, Lithuania
o Department of Radiology, St. Antonius Hospital, Nieuwegein, The Netherlands

ARTICLE INFO

Article history:
Received 10 February 2016
Received in revised form 14 June 2016
Accepted 11 August 2016
Available online 13 August 2016

Keywords:
Acute pancreatitis
Classification
Computed tomography
Extrapancreatic necrosis
Necrotising pancreatitis

ABSTRACT

Background: For consistent reporting and better comparison of data in research the revised Atlanta classification (RAC) proposes new computed tomography (CT) criteria to describe the morphology of acute pancreatitis (AP). The aim of this study was to analyse the interobserver agreement among radiologists in evaluating CT morphology by using the new RAC criteria in patients with AP.

Methods: Patients with a first episode of AP who obtained a CT were identified and consecutively enrolled at six European centres backwards from January 2013 to January 2012. A local radiologist at each center and a central expert radiologist scored the CTs separately using the RAC criteria. Center dependent and independent interobserver agreement was determined using Kappa statistics.

Results: In total, 285 patients with 388 CTs were included. For most CT criteria, interobserver agreement was moderate to substantial. In four categories, the center independent kappa values were fair: extrapancreatic necrosis (EXPN) (0.326), type of pancreatitis (0.370), characteristics of collections (0.408), and appropriate term of collections (0.356). The fair kappa values relate to discrepancies in the identification of extrapancreatic necrotic material. The local radiologists diagnosed EXPN (33% versus 59%, \( P < 0.0001 \)) and non-homogeneous collections (35% versus 66%, \( P < 0.0001 \)) significantly less frequent than the central expert. Cases read by the central expert showed superior correlation with clinical outcome.

**Abbreviations:** AP, acute pancreatitis; CECT, contrast-enhanced computed tomography; Central exp, Central expert; CRP, C-reactive protein; CT, computed tomography; EXPN, extrapancreatic necrosis; IEP, Interstitial Oedematous Pancreatitis; IQR, interquartile range; Local rad, local radiologists; No, number; RAC, the revised Atlanta classification; SIRS, systemic inflammatory response syndrome.

* Corresponding author. Department of Surgery, Skåne University Hospital, Lund University, Jan Waldenströms gata 44, 205 02 Malmö, Sweden. Tel.: +46 736 297981.
E-mail address: Hanna.Sternby@med.lu.se (H. Sternby).

http://dx.doi.org/10.1016/j.pan.2016.08.007
1424-3903/© 2016 IAP and EPC. Published by Elsevier B.V. All rights reserved.
1. Introduction

Acute pancreatitis (AP) is a complex disease with potentially severe and fatal outcome [1,2]. Simple but clear definitions of the disease are crucial in interdisciplinary consultation, communication, and in reporting of clinical research. Such were the incentives to update the 1992 Atlanta Classification on AP [1]. Besides redefining the disease into three levels of clinical severity, the 2012 revised Atlanta Classification (RAC) has put substantial efforts into clarifying the terminology on the morphologic subtypes of AP and associated peripancreatic collections based on computed tomography (CT)-based criteria [1]. Two morphologic types of AP are discriminated: acute interstitial oedematous pancreatitis and acute necrotising pancreatitis. Acute necrotising pancreatitis is subdivided into three forms: pancreatic parenchymal necrosis, extrapancreatic necrosis (EXPN), and combined necrosis. Peripancreatic collections are classified into four types depending on content and maturation. Acute peripancreatic fluid collections and pancreatic pseudocysts are composed of fluid only and occur in interstitial oedematous pancreatitis. On CT, these collections show a homogeneous fluid density with no or incomplete well-defined wall (acute peripancreatic fluid collection) or a complete wall (pseudocyst). Acute necrotic collections and walled-off necrosis are associated with acute necrotising pancreatitis and contain varying amounts of fluid and necrotic material. On CT, these collections have various densities (fat, fluid, solid material) with no or incomplete well-defined wall (acute necrotic collection) or a complete wall (walled-off necrosis) [1,3-5]. The RAC provides approximate time frames for these peripancreatic collections. Acute peripancreatic fluid collection and acute necrotic collection pertain to the first four weeks of disease after which they usually turn into a completely encapsulated pseudocyst and walled-off necrosis, respectively.

It is well established that the morphologic types of AP differ in outcomes, therapies, and prognosis. For prognostication, stratification, and comparing of interinstitutional data, accurate assessment of AP morphology in the different stages of disease is imperative [1]. The extent of variation in interpretation of the new CT criteria is, however, unknown [6-8]. The aim of this study was to assess the interobserver agreement among radiologists in the evaluation of CT morphology using the RAC criteria.

2. Methods

2.1. Patients and study design

Patients >18 years with a first episode of AP were consecutively identified at six European study centres, going backwards from January 2013 to January 2012. Each center included 50 patients in whom at least one contrast-enhanced CT (CECT) was performed. The cases were anonymously enrolled and each patient obtained a code blinded for all investigators except for the referring center. CECTs performed within 3 months from date of admission were recorded and subsequently reviewed and scored by a local radiologist at each center. The time frame of 3 months was chosen because most CTs are performed within this period and controversies in nomenclature and management of pancreatic collections are most evident during this phase. Exclusion criteria were insufficient quality of the CECT, signs of chronic pancreatitis (i.e. pancreatic calcifications) or patients with prior pancreatitis-related invasive intervention, except from endoscopic retrograde cholangiography. Each CECT was performed in the pancreatic and/or in the portal venous phase (see Supplementary file 1 for CT specifications). Severity and CT morphology of AP were defined according to the RAC (see Box 1 for definitions) [1].

The following clinical data was collected from review of medical notes: systemic inflammatory response syndrome (SIRS) upon admission, highest level of C-reactive protein (CRP) during hospitalisation, need for invasive intervention, organ failure (persistent and transient, in line with the RAC), and in-hospital mortality. The six participating local radiologists had expertise in the field of abdominal radiology, each with more than five years’ experience. A short instruction sheet was provided to local radiologists to assist in interpretation (Supplementary file 2). All individual CECTs were scored according to a protocol based on the parameters stated in the RAC (Supplementary file 3). Subsequently, all CECTs were reviewed and scored (using the same scoring sheet) by a central expert radiologist (T.L.B) using open source DICOM viewer software (32-bit OsiriX version 3.3, Geneva, Switzerland). Local and central reviewers were blinded to any clinical data except for the timing (number of days after onset of symptoms) of each CECT. Formal approval of the local medical ethical committee was requested and obtained at each study center.

Box 1

<table>
<thead>
<tr>
<th>Morphology groups</th>
<th>CECT criteria</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interstitial oedematous pancreatitis (IEP)</td>
<td>Homogenous enhancement of the pancreatic parenchyma, normal or minor inflammatory changes of the peripancreatic tissue (see below – APFC or pancreatic pseudocyst)</td>
<td>–</td>
</tr>
<tr>
<td>Necrotising pancreatitis</td>
<td>Heterogenous enhancement of the pancreatic parenchyma and/or peripancreatic tissue necrosis (see below – ANC or WON)</td>
<td>–</td>
</tr>
<tr>
<td>Acute peripancreatic fluid collection (APFC)</td>
<td>Homogeneous fluid density. No complete wall. No necrosis. Associated weeks with IEP. Solely extrapancreatic location.</td>
<td>≤ 4</td>
</tr>
<tr>
<td>Pancreatic pseudocyst</td>
<td>Homogenous fluid density. Fully encapsulated. No necrosis. Associated weeks with IEP. Solely extrapancreatic location.</td>
<td>&gt; 4</td>
</tr>
<tr>
<td>Acute necrotic collection (ANC)</td>
<td>Heterogenous and non-liquid density. No complete wall. Associated weeks with necrotising pancreatitis. Intra- or extrapancreatic location</td>
<td>≤ 4</td>
</tr>
<tr>
<td>Walled-off necrosis (WON)</td>
<td>Heterogenous and non-liquid density. Fully encapsulated. Associated weeks with necrotising pancreatitis. Intra- or extrapancreatic location</td>
<td>&gt; 4</td>
</tr>
<tr>
<td>CECT – contrast enhanced computed tomography.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.2. Statistical analysis

Sample size calculations for variables with 2, 3, and 4 categories were performed because most of the variables of the revised Atlanta criteria are based on 2, 3, or 4 categories. Kappa values of >0.40 and > 0.60 were used since they represent at least moderate and substantial agreement [9]. Based on such a calculation, in a test for agreement between two raters, a sample size of 360 CECTs would provide a 95% confidence interval for the \( \kappa \) statistic with a width not greater than 0.20. Assuming that several patients would have more than one CECT performed, a total of 50 patients were included per center. Interobserver agreement was calculated between the local and the central radiologist, using Cohen’s kappa test, for each of the categories scored on the radiology sheet. Agreement levels were defined as: \( \kappa \) level 0.00–0.20 slight; 0.21–0.40 fair; 0.41–0.60 moderate; 0.61–0.80 substantial; and 0.81–1.00 almost perfect. Continuous data analysis was conducted using Mann-Whitney U test. Wilcoxon Signed Rank test was used for paired sample analysis. \( P < 0.05 \) was considered statistically significant. All analysis was performed using IBM SPSS Statistics for Windows, version 21 and 22, Armonk, NY:IBM corp.

3. Results

3.1. Patients

In total, 301 patients were included at six European centres of whom 159 (56%) were male with a median age of 58 years (range 18–92). Sixteen patients were excluded due to reasons stated in the CECT section below. Baseline characteristics of the remaining 285 patients are summarised in Table 1. Etiology of AP differed substantially between the centres. According to the RAC, 37.5% of the patients had mild AP, 51.5% moderately severe AP and 11.0% severe AP [1]. Overall fourteen patients died (4.9%, range 0–14%), whereas mortality within the severe group was 32.3% (range 0–78%). Each center admits 175–250 patients with AP annually, except for center F where the figure is approximately 470.

3.2. Contrast-enhanced computed tomography

A total of 405 CECTs derived from 301 patients were collected. Seventeen CECT studies were excluded due to insufficient quality of the CECT or signs of chronic pancreatitis, leaving a study cohort of 285 patients with 388 CECTs. Data on CECTs for the separate centres are presented in Table 2. Median time from onset of disease to CECT for all centres was 7 days (range 0–90, interquartile range 3–13).

3.3. Interobserver agreement

For specific information concerning all categories evaluated we refer to Supplementary file 2. The kappa values representing the interobserver agreement are shown in Table 3. There was substantial agreement in seven categories: “Necrosis – Neck” (0.618); “Necrosis – Body” (0.628); “Necrosis – Tail” (0.617); presence of “Collections” (0.756); “Location of Collections” (0.633); presence of “Wall” (0.675); and presence of “Intraluminal Gas and/or Fluid level” (0.764). Moderate agreement was reached on “Parenchymal Necrosis” (0.539) and “Necrosis – Head” (0.516). Finally, there was fair agreement on the categories “Type of Pancreatitis” (0.370), “Extrapancreatic Necrosis” (0.326), “Characteristics of Collection” (0.408), and “Collection – most appropriate term” (0.356). The center dependent kappa values differed considerably between centres.

Discrepancies in the identification of EXPN are shown in Table 4. For image samples see Fig. 1a,b and 2a–c. The expert radiologist diagnosed EXPN significantly more often than the local radiologists (59% vs. 33%, \( P < 0.0001 \)). Table 4 shows that this difference in total number of EXPN stems from the subgroup of isolated EXPN. Since the RAC acknowledges that EXPN might be difficult to diagnose within the first week, interobserver agreement was recalculated for the categories with low kappa values excluding CECTs performed within 72 h, seven days and two weeks after onset of disease (see Supplementary file 4). In this subanalysis, kappa values did improve only for CECTs performed after two weeks.

 Morphological findings scored by the central and local radiologists were correlated with clinical outcome parameters (see Table 5). Cases read as interstitial oedematous pancreatitis and isolated EXPN by the central expert correlate significantly better with clinical outcome than scoring by local radiologists. Given the good interobserver agreement for pancreatic parenchymal necrosis, results did not differ significantly between central and local radiologists for this subgroup (Supplementary file 5).

4. Discussion

The RAC proposed a new set of morphologic CT-based criteria to account for alleged shortcomings of the 1992 Atlanta classification [1,2]. One of the major aims of the RAC was to ease and ensure
consistency in the investigation and reporting of data in clinical research [1]. However, the degree of interobserver agreement in the interpretation of CT findings using these new RAC criteria has been questioned [6–8].

Main findings of our study are twofold: on the one hand, the morphologic assessment of the RAC generates overall moderate to good interobserver agreement (range 0.516–0.764) among European radiologists in 9 out of 13 items evaluated. Importantly, agreement among raters was good in evaluating clinically important CT findings in patients with AP, such as the presence of parenchymal necrosis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.

Several explanations exist for the fair agreement in diagnosing EXPN and for characterisation of pancreatic collections on CT. The CT findings in patients with AP, such as the presence of necrotising pancreatitis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.

Several explanations exist for the fair agreement in diagnosing EXPN and for characterisation of pancreatic collections on CT. The CT findings in patients with AP, such as the presence of necrotising pancreatitis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.

Several explanations exist for the fair agreement in diagnosing EXPN and for characterisation of pancreatic collections on CT. The CT findings in patients with AP, such as the presence of necrotising pancreatitis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.

Several explanations exist for the fair agreement in diagnosing EXPN and for characterisation of pancreatic collections on CT. The CT findings in patients with AP, such as the presence of necrotising pancreatitis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.

Several explanations exist for the fair agreement in diagnosing EXPN and for characterisation of pancreatic collections on CT. The CT findings in patients with AP, such as the presence of necrotising pancreatitis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.

Several explanations exist for the fair agreement in diagnosing EXPN and for characterisation of pancreatic collections on CT. The CT findings in patients with AP, such as the presence of necrotising pancreatitis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.

Several explanations exist for the fair agreement in diagnosing EXPN and for characterisation of pancreatic collections on CT. The CT findings in patients with AP, such as the presence of necrotising pancreatitis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.

Several explanations exist for the fair agreement in diagnosing EXPN and for characterisation of pancreatic collections on CT. The CT findings in patients with AP, such as the presence of necrotising pancreatitis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.

Several explanations exist for the fair agreement in diagnosing EXPN and for characterisation of pancreatic collections on CT. The CT findings in patients with AP, such as the presence of necrotising pancreatitis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.

Several explanations exist for the fair agreement in diagnosing EXPN and for characterisation of pancreatic collections on CT. The CT findings in patients with AP, such as the presence of necrotising pancreatitis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.

Several explanations exist for the fair agreement in diagnosing EXPN and for characterisation of pancreatic collections on CT. The CT findings in patients with AP, such as the presence of necrotising pancreatitis and gas bubbles. On the other hand, only fair agreement (range 0.326–0.408) was obtained for items pertaining to necrosis of extrapancreatic tissues. The central expert diagnosed EXPN significantly more frequent than the local radiologists (59% versus 33%, P < 0.0001) with better correlation with patient outcome. Our findings suggest that radiologists are largely unfamiliar with the newly defined entity of EXPN.
primarily on subjective secondary findings, such as ‘heterogeneity’ or the detection of various densities (liquid and non-liquid) within collections, rather than using the more objective and reproducible criteria of perfusion characteristics, which is used for detecting pancreatic parenchymal necrosis. Perfusion characteristics, however, cannot be used to diagnose EXPN because normal extrapancreatic fat does not enhance. In addition, the RAC acknowledges that EXPN is difficult to diagnose initially but becomes easier when the disease process evolves over time [1]. This is in line with our results with improved kappa values for EXPN diagnosis two weeks after symptom onset. Finally, reader expertise and familiarity seem equally important for diagnosing EXPN exemplified by the excellent interobserver agreement between the central expert and the local radiologist affiliated with center F that admits the highest number of patients with AP annually.

In a previous study on CT assessment of morphologic features of AP, good to excellent interobserver agreement was found in 55 cases of AP [12]. However, this study used selected cases biased towards severe disease (likely associated with more established peripancreatic collections) reviewed by tertiary experts. Although imaging is rarely required for mild AP, most patients who undergo CT for evaluation of AP turn out to have interstitial pancreatitis. Our study more closely resembles clinical practice in different European countries by enrolling patients with AP consecutively and evaluating unselected CECTs encompassing the full spectrum of morphologic abnormalities in AP, including mild and equivocal cases.

Previous studies show considerable differences in clinical outcome, treatment strategies, and prognoses between the various morphological types of AP [13–17]. Clinical outcome of EXPN is worse compared with acute interstitial pancreatitis, but better than pancreatic necrosis [13,15,17]. Patients with EXPN stay in hospital considerably longer, develop more organ failure, and undergo interventional therapy significantly more frequent than those with interstitial pancreatitis. Moreover, when infection ensues of necrotic collections in EXPN patients, outcome, therapy, and prognosis are similar to those with infected pancreatic necrosis [15]. In our study, the interpretation by the central expert more closely corroborated with actual clinical outcome. As such, accurate differentiation between the types of AP is important both from a clinical perspective as for consistent reporting of research and reliable comparison of inter-institutional data.

This study has some limitations. First, a single central expert radiologist served as standard of reference, potentially introducing bias. We considered this a limited risk because of his extensive experience in pancreatic imaging, his involvement in the development of the RAC, the superior correlation with patient outcome, and the excellent agreement with a local radiologist with

Table 5
Correlation of morphologic findings with clinical outcome by central and local radiologists.

<table>
<thead>
<tr>
<th></th>
<th>Total (N)</th>
<th>IEP</th>
<th>Isolated EXPN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Local</td>
<td>Central</td>
<td>P-value</td>
</tr>
<tr>
<td>All cases</td>
<td>199</td>
<td>107</td>
<td>0.0001</td>
</tr>
<tr>
<td>Persistent</td>
<td>20 (10.0%)</td>
<td>5 (4.7%)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Transient</td>
<td>15 (7.5%)</td>
<td>5 (4.7%)</td>
<td>0.002</td>
</tr>
<tr>
<td>Mortality</td>
<td>9 (4.5%)</td>
<td>4 (3.7%)</td>
<td>0.025</td>
</tr>
<tr>
<td>Intervention</td>
<td>11 (5.5%)</td>
<td>6 (5.6%)</td>
<td>ns</td>
</tr>
<tr>
<td>CRP (mg/l)</td>
<td>271</td>
<td>165</td>
<td>ns</td>
</tr>
<tr>
<td>SIRS</td>
<td>41 (20.6%)</td>
<td>12 (11.2%)</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Cases scored as ‘indeterminate’ are not accounted for.

IEP = Interstitial Edematous Pancreatitis.
Isolated EXPN = Extrapancreatic Necrosis without Parenchymal Necrosis.
Local = Local radiologists.
Central = Central Expert.
CRP = C-reactive protein, highest value.
SIRS = Systemic Inflammatory Response Syndrome.

Fig. 1. 70-year-old male with acute pancreatitis (a, b). The pancreas enhances heterogeneously (asterisks) but no apparent necrosis was observed by both reviewers. Peripancreatic collections are present in the retroperitoneal pancreatic compartment and transverse mesocolon (arrowheads pointing at the borders). The local reviewer scored this as interstitial pancreatitis with acute peripancreatic fluid collections; the central reviewer as necrotising pancreatitis, subtype EXPN, with acute necrotic collections.
similar expertise. Second, given the retrospective design of this study, we did not investigate to what extent the inconsistencies observed eventually affected clinical decision-making. Future studies should focus on this interesting topic. Finally, we merely studied the interobserver agreement of morphological abnormalities in AP and did not correlate imaging findings with histopathology.

Results of this study have revealed areas of controversy when using the RAC criteria for CT assessment, especially pertaining to distinguishing interstitial pancreatitis from EXPN only. There are several options for improving consistent reporting in AP. Both radiologists and clinicians need to become better familiar with imaging features of EXPN (i.e. by education or training). Second, the definition of EXPN should preferably be redefined such that stronger interrater agreement will be achieved, even among readers with varying expertise. For example, by adding a time interval of 2 weeks before its diagnosis or by using an alternative imaging (MRI or US) modality as these are better capable of detecting necrotic material within collections [10]. Third, a greater role should be attributed to MRI for overall evaluation of AP. Finally, as has been alluded to in previous reports, a three-degree morphologic classification system (‘interstitial pancreatitis’ refers to normal enhancing parenchyma without collections; ‘EXPN’ refers to normal perfused pancreatic parenchyma with pancreatic collections, and ‘necrotising pancreatitis’ refers to parenchymal necrosis with or without associated collections) could potentially lead to less interobserver variability as the differentiation between the various types of pancreatic collections becomes less of an issue [18,19]. Additionally, Such a system would likely be more in concordance with clinical grades of severity [13–17,20].

In conclusion, this study found only moderate interrater agreement for identification of EXPN. For correctly identifying EXPN and necrotic collections on CT, a diligent search for heterogeneity within pancreatic collections is crucial for accurate and consistent reporting of imaging findings (see Figs. 1 and 2).

Disclosure

The authors declare no conflict of interests.

Funding resources

No funding was involved in any part of the study or manuscript.

Acknowledgements

The authors thank Luigi Santoro, Division of Epidemiology and Biostatistics, European Institute of Oncology, for assistance with statistical analysis. The group was part of Pancreas 2000 – a research program for future pancreatologists, who in part also supported this study (www.pancreas2000.org).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.pan.2016.08.007.

References

Predictive Capacity of Biomarkers for Severe Acute Pancreatitis

Hanna Sternby\textsuperscript{a} Hannes Hartman\textsuperscript{b} Dorthe Johansen\textsuperscript{a}
Henrik Thorlacius\textsuperscript{a} Sara Regnér\textsuperscript{a}

Departments of \textsuperscript{a}Surgery and \textsuperscript{b}Gastroenterology, Institution of Clinical Sciences Malmö, Lund University, Malmö, Sweden

Key Words
Acute pancreatitis · Biomarkers · Cut-off · Prediction

Abstract

\textbf{Background:} Early prediction of severe acute pancreatitis (SAP) substantially improves treatment of patients. A large amount of biomarkers have been studied with this objective. The aim of this work was to study predictive biomarkers using preset cut-off levels in an unselected population of patients with acute pancreatitis (AP). \textbf{Methods:} 232 patients (52.2\% males, median age 66 years) with AP admitted to Skåne University Hospital, Malmö, were consecutively enrolled. Blood samples were collected upon admission and clinical data were gathered both prospectively at inclusion and through review of medical notes. Cut-off levels were defined based on the reports of prior studies, and through their results eight biomarkers (IL-1\textbeta, IL-6, IL-8, IL-10, TNF-\alpha, MCP-1, procalcitonin and D-dimer) were selected for analysis. \textbf{Results:} Of the patients, 83.2\% had mild AP and 16.8\% had SAP. Levels of IL-1\textbeta, IL-6 and IL-10 were significantly (p < 0.05) higher upon admission in the group with SAP. When applying the preset cut-off levels on our material, sensitivity and specificity for prediction of severity were low. Receiver operating characteristic curves showed that selected cut-off levels were acceptable, but areas under the curves were inferior compared to other studies. The results did not improve when using the revised Atlanta 2012 classification. \textbf{Conclusions:} Previous studies on severity prediction of AP are difficult to compare due to large variations in setups and outcomes. Calculated cut-offs in our cohort were in acceptable range from preset levels, however areas under the curves were low, indicating suboptimal biomarkers for the unselected population investigated. For comparable results and possible clinical implementations, future studies need large consecutive series with a reasonable percentage of severe cases. Additionally, novel biomarkers need to be explored.
**Introduction**

Acute pancreatitis (AP) is one of the most common gastrointestinal disorders requiring hospitalization worldwide. Data indicate a rising incidence, at present approximately 50 cases/100,000 inhabitants [1–3]. From 1992 to 2012, patients with AP were subdivided into mild and severe disease according to the Atlanta 1992 classification [4]. Most cases of AP are classified as mild, but approximately 15–20% of the patients will develop severe AP (SAP), characterized by rapidly progressing organ failure and local complications [5, 6]. Mortality within this group is high, 20–30%, and the need for supportive treatment and resuscitation already in the emergency room is of significance [7, 8]. Early detection of SAP is thus crucial and numerous clinical scoring systems have been developed for this purpose [9–13]. However, these methods have proven either unpractical, too costly or insufficient in sensitivity and specificity for prediction of severe disease [14–16]. Additionally, it has been shown that in the early stage of AP, physicians are poor at identifying the patients that will develop SAP [8, 17].

With the aim of early prediction of SAP, multiorgan dysfunction syndrome and risk of death, numerous biomarkers have been iteratively studied [18]. The majority of these are inflammatory mediators reflecting the complex pathophysiological pathway of AP. Several promising results have been presented and a number of studies additionally report cut-off levels for prediction of severity for individual biomarkers in their specific cohort.

However, despite great efforts and numerous attempts, the biomarker that predicts SAP upon admission with high reliability is yet to be found. Additionally, for clinical applicability, a cut-off level with good predictive power needs to be identified. Previous studies used post hoc cut-offs, which implies a decline in sensitivity and specificity of severity prediction when applied in a prospective setting. However, such an approach does not mirror the ordinary situation with the AP patient in the emergency room. Consequently, results have been difficult to implement in clinical practice.

Among earlier studies none have, to our knowledge, used preset cut-off values applied on multiple biomarkers in a consecutive series of patients with AP. The aim of this study was to prospectively, in an unselected population, evaluate the most promising non-routine predictive biomarkers by using preset cut-off levels based on the results of earlier works.

**Materials and Methods**

**Patients and Study Design**

Patients >18 years with AP admitted to the Department of Surgery, Skåne University Hospital, Malmö, Sweden from January 2010 to September 2013 were consecutively enrolled. For the diagnose of AP two out of three criteria needed to be fulfilled: (1) acute characteristic upper abdominal pain, (2) serum amylase ≥3 times the upper limit or (3) characteristic findings of AP on CT scan, abdominal ultrasound or MRI. Only patients admitted within 72 h from disease onset were included. To enable comparison with previous studies, AP was classified as mild or severe according to the Atlanta 1992 classification [4]. The patients were additionally classified according to the revised Atlanta 2012 classification [19]. Clinical data were obtained from the patients at inclusion and retrospectively from review of medical notes.

**Cut-Off Levels**

PubMed was searched for human studies on predictive biomarkers in AP with SAP as outcome where cut-off levels for predicted severe disease were reported. Additionally, levels of sensitivity, specificity, positive and negative predicted values for each cut-off level had to be presented. Solely studies utilizing the Atlanta 1992 classification were considered [4]. In total, 16 studies fulfilled these criteria and were selected for the predetermination of cut-off levels [15, 20–34]. The biomarkers analyzed in these publications were interleukin-1beta (IL-1β), interleukin-6 (IL-6), interleukin-8 (IL-8), interleukin-10 (IL-10), tumor necrosis...
factor-alpha (TNF-α), monocyte chemoattractant protein-1 (MCP-1), procalcitonin and D-dimer. A majority of studies presented various cut-off levels for identical biomarkers, most frequently regarding IL-6. In those cases cohort characteristics (i.e. gender, etiology, incidence of severity) were considered when predetermining the cut-off for an unselected AP population.

**Blood Samples and Measurement**

Serum samples were collected upon admission, i.e. in the emergency room, centrifuged (2,000 rounds, 25°C, 10 min) and stored at −80°C until analyzed. For analysis of IL-1β, IL-6, IL-8, IL-10 and TNF-α the human proinflammatory 7-plex ultrasensitive kit K15008C (Meso Scale Diagnostics LLC, Rockville, Md., USA) was used. MCP-1 was assessed through human CCL2 (MCP-1) ELISA Ready-set-go kit 88-7399-88 (aBioscience, San Diego, Calif., USA) and D-dimer by using Human D-Dimer (D2D) ELISA kit 20870 (Bmassay, Beijing, China). All analyses were assessed according to the manufacturer’s instructions. Procalcitonin was determined with an accredited ELISA method based on monoclonal anti-procalcitonin antibodies in accordance with routine methods at the Department of Clinical Chemistry, Skåne University Hospital, Malmö.

**Statistical Analysis**

Sample size calculations were based on results from an earlier study on AP within the same population, i.e. patients with AP admitted to the University Hospital in Malmö [29]. In that study 11.5% of the cohort developed SAP. To reach a power of 80% with an α value of 0.05 and a sensitivity and specificity >85% for MCP-1, 37 patients (4 with SAP and 33 with mild AP [MAP]) were required. With 150 enrolled patients significant results would be reached for the most investigated markers. For more unusual biomarkers at least 200 patients were considered necessary for an adequate level of sensitivity and specificity.

For comparison of continuous data between the two groups, the Mann-Whitney U test was used and p < 0.05 was considered statistically significant. For determination of the optimal cut-off value for each specific biomarker in our cohort, receiver operating characteristic (ROC) curves were performed and areas under the curves (AUGs) were calculated. All statistical analyses were performed using IBM SPSS Statistics for Windows, version 21 and 22 (IBM Corp., Armonk, N.Y., USA).

| Table 1. Patient characteristics of all patients and groups with mild and severe disease |
|---------------------------------|---------------------------------|-----------------|
| All patients 232 | MAP 193 (83.2%) | SAP 39 (16.8%) |
| Gender | | |
| Male | 121 (52.2%) | 101 (52.3%) | 20 (51.3%) |
| Female | 111 (47.8%) | 92 (47.7%) | 19 (48.3%) |
| Age, years | 66 (19–97) [51–77] | 66 (19–97) [51–77] | 63 (29–90) [52–80] |
| Body mass index | 26.3 (13.6–47.0) | 26.0 (15.2–45.4) | 26.4 (13.6–47.0) |
| Etiology | | |
| Biliary | 131 (56.5%) | 114 (59.1%) | 17 (43.6%) |
| Alcohol | 39 (16.8%) | 30 (15.5%) | 9 (23.1%) |
| Idiopathic | 26 (11.2%) | 18 (9.3%) | 8 (20.5%) |
| Other | 36 (15.5%) | 31 (16.1%) | 5 (12.8%) |
| Onset to admission, h | 9 (0–72) [3–23] | 10 (0–72) [3–24] | 5 (0–72) [2–18] |
| Highest CRP, mg/l | 150 (1–553) | 111 (1–397) | 330 (111–553) |
| Mortality | 5 (2.2%) | 0 | 5 (12.8%) |
| Intensive care unit | 19 (8.2%) | 0 | 19 (48.7%) |
| Organ failure | 33 (14.2%) | 0 | 33 (84.6%) |
| MODS | 15 (6.5%) | 0 | 15 (38.5%) |

Values are n (%) or median (range); figures in brackets are interquartile range.

Highest CRP = Highest level of CRP within the first 72 h; MODS = multiorgan dysfunction syndrome.
Results

Patients

During a 3.5-year period, 233 patients were included. On retrospective review one patient was excluded for not meeting the criteria of AP. The baseline characteristics of the 232 patients are as shown in table 1. Gender distribution was almost equal with 52.2% men and 47.8% women. Etiology was divided into biliary (56.5%), alcohol (16.8%), other (i.e., post-endoscopic retrograde cholangiopancreatography pancreatitis, tumors, other strictures) (15.5%) and idiopathic (11.2%). According to the Atlanta 1992 classification, 83.2% had MAP and 16.8% had SAP. Mortality was 2.2% overall and 12.8% within the severe group. Of the patients with SAP, 84.6% were considered to have organ failure and 48.7% were admitted to the intensive care unit. A significant difference between the parameters of the mild and severe groups in table 1 was exclusively found for the highest level of C-reactive protein (CRP) within the first 72 h (p < 0.001). For baseline characteristics according to the revised Atlanta classification see the online supplementary file (www.karger.com/doi/10.1159/000444141).

Biomarkers and Cut-Off Levels

Upon admission significant differences between median values of the groups with MAP and SAP were found for IL-1β (p = 0.008), IL-6 (p = 0.013) and IL-10 (p = 0.009). IL-8 (p = 0.056) and MCP-1 (p = 0.062) were close to significant. The results of the application of the preselected cut-off levels on the cohort are shown in table 2. ROC curves were performed to obtain the optimal cut-off level for each biomarker of this population; the results are presented in table 3 and figure 1. The ROC curve of CRP (upon admission) was added to the results since it has been repeatedly proven to be the best predictive marker for severe disease. For most parameters the preset cut-off values were within an acceptable range of those inferred from the ROC curves of the present study cohort. For IL-6, sensitivity and specificity of the preset cut-off value of 100 pg/ml were 31 and 81%, respectively. The optimal cut-off in the study population, 72 pg/ml, would have given a sensitivity of 40% and a specificity of 78%. Similarly, for procalcitonin there would have been a change in sensitivity from 18 to 52% and in specificity from 82 to 66% if the optimal cut-off, 0.35 ng/ml, had been used instead of the preset value of 0.5 ng/ml. Additionally, the AUCs of the biomarkers were inferior to those of previous studies, which might indicate that their prognostic ability is not sufficient in such a consecutive series of patients. None of the biomarkers were able to predict mortality in this cohort (see positive predictive value, table 2).

Table 2. Result of the application of preset cut-off levels on severe cases

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Preset cut-off</th>
<th>SAP</th>
<th>SAP below cut-off</th>
<th>SAP above cut-off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>PPV Mors</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-1β, pg/ml</td>
<td>1</td>
<td>36</td>
<td>(16.0%)</td>
<td>14 (38.9%)</td>
<td>61%</td>
<td>49%</td>
<td>0.19</td>
<td>0.87</td>
<td>0.04</td>
</tr>
<tr>
<td>IL-6, pg/ml</td>
<td>100</td>
<td>35</td>
<td>(15.6%)</td>
<td>24 (68.6%)</td>
<td>31%</td>
<td>82%</td>
<td>0.24</td>
<td>0.87</td>
<td>0.04</td>
</tr>
<tr>
<td>IL-8, pg/ml</td>
<td>40</td>
<td>34</td>
<td>(15.1%)</td>
<td>8 (23.6%)</td>
<td>78%</td>
<td>35%</td>
<td>0.19</td>
<td>0.89</td>
<td>0.03</td>
</tr>
<tr>
<td>IL-10, pg/ml</td>
<td>7.5</td>
<td>36</td>
<td>(16.0%)</td>
<td>5 (13.9%)</td>
<td>86%</td>
<td>29%</td>
<td>0.19</td>
<td>0.92</td>
<td>0.02</td>
</tr>
<tr>
<td>TNF-α, pg/ml</td>
<td>10</td>
<td>35</td>
<td>(15.6%)</td>
<td>29 (82.9%)</td>
<td>17%</td>
<td>78%</td>
<td>0.125</td>
<td>0.84</td>
<td>0.04</td>
</tr>
<tr>
<td>MCP-1, μg/l</td>
<td>500</td>
<td>32</td>
<td>(14.2%)</td>
<td>23 (71.9%)</td>
<td>28%</td>
<td>86%</td>
<td>0.26</td>
<td>0.87</td>
<td>0.03</td>
</tr>
<tr>
<td>PCT, ng/ml</td>
<td>0.5</td>
<td>33</td>
<td>(14.7%)</td>
<td>27 (81.8%)</td>
<td>18%</td>
<td>82%</td>
<td>0.15</td>
<td>0.85</td>
<td>0.05</td>
</tr>
<tr>
<td>D-dimer, μg/l</td>
<td>0.5</td>
<td>28</td>
<td>(12.4%)</td>
<td>7 (25.0%)</td>
<td>75%</td>
<td>12%</td>
<td>0.12</td>
<td>0.74</td>
<td>0.02</td>
</tr>
</tbody>
</table>

NPV = Negative predictive value; PCT = procalcitonin; PPV = positive predictive value; PPV Mors = positive predictive value mortality.
Since 2013 the revised Atlanta 2012 classification is considered the international standard for severity classification in AP [19]. Thus we performed a post hoc analysis using this trisected classification. Significant differences were found for IL-1β (p = 0.022) and IL-10 (p = 0.006) when comparing the group with SAP versus those with mild and moderately severe disease together. Optimal cut-off levels and AUCs of the investigated biomarkers are presented in table 4 and pertaining ROC curves in figure 2. The results did, in general, not improve when applying the new classification.
Discussion

The hunt for a prognostic biomarker in AP has been ongoing for decades. A multitude of studies have emerged for this reason and several have presented promising results. Nevertheless, the conclusions are still far from attaining clinical practice.

The aim of this study was to use cut-off levels for severity prediction reported in earlier works and to apply them in a prospective study on an unselected population with AP. The vast majority of earlier studies have been exploratory with retrospective analysis of cut-off levels for each specific cohort. However, this is not applicable on the routine situation for the physician. Investigation of predictive biomarkers with preset cut-off levels would more accurately resemble clinical reality. Given the large amount of studies conducted on promising biomarkers, a consensus on cut-off levels for these markers ought to be conceivable. Our study does, however, indicate that considerable efforts still remain to be made in order to reach adequate results in this respect.

Most studies on biomarkers for prediction of SAP only report median values of the cytokines examined. Comparatively few, for some biomarkers limited to one single study, present ROC curves, AUCs, cut-off levels and concordant levels of sensitivity, specificity, positive predictive value, negative predictive value and accuracy. For this reason the selection of biomarkers remaining to be investigated with our aim is limited. Additionally, in cases where cut-off levels are reported, sample sizes are small, study settings are dissimilar or not accurately presented and various outcomes are used. Together this makes comparison and general

---

**Table 3. Optimal cut-off and AUC for each biomarker using the Atlanta 1992 classification**

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Optimal cut-off</th>
<th>ROC AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-1β, pg/ml</td>
<td>0.9</td>
<td>0.640</td>
</tr>
<tr>
<td>IL-6, pg/ml</td>
<td>72</td>
<td>0.632</td>
</tr>
<tr>
<td>IL-8, pg/ml</td>
<td>49</td>
<td>0.601</td>
</tr>
<tr>
<td>IL-10, pg/ml</td>
<td>12</td>
<td>0.638</td>
</tr>
<tr>
<td>TNF-α, pg/ml</td>
<td>n.a.</td>
<td>0.459</td>
</tr>
<tr>
<td>MCP-1, μg/l</td>
<td>399</td>
<td>0.604</td>
</tr>
<tr>
<td>PCT, ng/ml</td>
<td>0.35</td>
<td>0.552</td>
</tr>
<tr>
<td>D-dimer, μg/l</td>
<td>n.a.</td>
<td>0.516</td>
</tr>
<tr>
<td>CRP, mg/l</td>
<td>59.5</td>
<td>0.720</td>
</tr>
</tbody>
</table>

n.a. = Not applicable.

**Table 4. Optimal cut-off and AUC for each biomarker using the revised Atlanta classification**

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Optimal cut-off</th>
<th>ROC AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-1β, pg/ml</td>
<td>0.99</td>
<td>0.689</td>
</tr>
<tr>
<td>IL-6, pg/ml</td>
<td>90.8</td>
<td>0.620</td>
</tr>
<tr>
<td>IL-8, pg/ml</td>
<td>135.1</td>
<td>0.612</td>
</tr>
<tr>
<td>IL-10, pg/ml</td>
<td>15.5</td>
<td>0.725</td>
</tr>
<tr>
<td>TNF-α, pg/ml</td>
<td>n.a.</td>
<td>0.497</td>
</tr>
<tr>
<td>MCP-1, μg/l</td>
<td>425</td>
<td>0.519</td>
</tr>
<tr>
<td>PCT, ng/ml</td>
<td>0.35</td>
<td>0.517</td>
</tr>
<tr>
<td>D-dimer, μg/l</td>
<td>n.a.</td>
<td>0.422</td>
</tr>
<tr>
<td>CRP, mg/l</td>
<td>113</td>
<td>0.802</td>
</tr>
</tbody>
</table>

n.a. = Not applicable.
conclusions complicated and imprecise. The weaknesses of this study are thus the lack of flexibility in the selection of prognostic biomarkers, in particular those of recent interest, and the risk of predetermined cut-off values being indistinct.

However, our analysis showed that cut-off levels of earlier studies were reasonably well in concordance with those found in our cohort (table 2). Dissimilarities are expected since the cut-off levels are specific for a particular cohort, while applied on a different population their predictive ability will be impaired. The total predictive capacity of the biomarkers, measured as the AUC, was generally inferior to those of previous studies. There are probably numerous reasons for these results. SAP is reported to develop in 15–20% of patients, which is comparable to the 16.8% found in our cohort [5, 6]. However, in the studies used for predetermi-
nation of cut-off levels these figures were in general much higher (median 39.5%, range 11.4–53.0%). Such a difference naturally influences the results – the greater the numbers of SAP, the more likely the finding of significant differences between the mild and severe group. A high amount of cases with SAP indicates cohort selection biases, for example referral centers or non-consecutive enrollment of patients. Another reason might be differing interpretations of the Atlanta 1992 classification. Bollen et al. [35, 36] showed extensive inappropriate use of this classification and frequent application of alternative definitions of severe disease. A clear shortcoming of our study is the utilization of this now outdated classification. We did, however, consider it necessary in this work to enable comparison with previous studies for the determination of cut-off levels. Future application of the revised Atlanta 2012 classification will hopefully make the division of patients more stringent and uniform, thus facilitating comparison of reports regarding cases of SAP [19]. Post hoc analyses of investigated biomarkers using current international standards are presented as ROC curves, AUCs and optimal cut-off levels in figure 2 and table 4. The results indicate that CRP remains the best prognostic marker and additionally demonstrate a stronger predictive ability of IL-1β and IL-10.

In our material only 45.2% of the patients with SAP were admitted to the intensive care unit, indicating less severe disease of this group than in comparative studies. However, these low numbers are due to the existence of an intermediate ward at the University Hospital in Malmö with the capacity to handle transient organ failures and local complications as well as elderly severely ill patients where intervention or admittance to the intensive care unit is not considered. Nevertheless these patients all have SAP according to the Atlanta 1992 classification.

An important, but previously not particularly investigated, factor is the time aspect. In this material we present a median time from disease onset to hospital admission of only 9 h (table 1). All blood samples were taken immediately upon arrival at the emergency room. Earlier studies have not specified this exact time interval and the time lapse from disease onset is in most cases unknown. The importance of exact information in this matter is not sufficiently studied and thus not yet clarified. Still the rather short time found in our material might be an explanation for some biomarkers not yet showing significant difference between the mild and severe group. Similarly the time frame could play a role for the outcome of cut-off levels. Blood samples in our material were probably taken rather early compared to the studies selected for predetermination of cut-off levels.

Previous studies have presented significant differences between MAP and SAP concerning clinical characteristics such as age, body mass index and etiology. No such difference was found in our cohort. The main explanation for these discrepancies is the application of an unselected population of patients with AP containing a low proportion of SAP compared to other works.

The strengths of this prospective study are its large scale of consecutively enrolled patients with AP from an unselected population, showing the whole spectrum of the disease without referral cases or selection bias. We also report detailed information on exact numbers of hours from disease onset to collection of blood samples. Analysis of biomarkers with preset cut-off levels and no post hoc assumptions make this work to a large extent resemble the difficulties of routine clinical work.

We have investigated the predictive capacity of biomarkers in AP. Despite many years of research physicians still lack instruments to anticipate the significant share of patients who will develop severe disease. Several biomarkers have been investigated and evaluated for more than two decades but are yet to be adopted into routine clinical use. To reach general conclusions the possibility of comparing study results is a necessity. Our findings indicate, as has been alluded in previous reports, a need for a higher degree of standardization and
uniformity in the design and execution of future studies [18, 37]. Implementation of research findings into routine clinical application requires results from consecutive studies on unselected populations of patients. Additionally, this work illuminates the need for continuous exploration of novel prognostic biomarkers in AP.

Acknowledgements

The authors wish to thank Anne-Marie Rohrstock for valuable support in sample analysis and Ingrid Palmquist for assistance in the handling of samples. This research was supported by Einar och Inga Nilsson, Magnus Bergvall, Lundgren, Schyberg and Mossfelt Foundation, Malmö Hospital Foundation, Skåne University Hospital and Lund University.

Statement of Ethics

This study was approved by the regional ethics committee at Lund University (2009/415). Oral and written consent was provided by all included patients.

Disclosure Statement

The authors declare that they have no conflicts of interest.

References

15 Khanna AK, Mehra S, Prakash S, Tiwary SK, Singh U, Srivastava A, Dixit VK: Comparison of Ranson, Glasgow, MOSS, SIRS, BISAP, Apache-II, CTSI scores, IL-6, CRP, and procalcitonin in predicting severity, organ failure, pancreatic necrosis, and mortality in acute pancreatitis. HPB Surg 2013;2013:367581.


IL-6 and CRP are superior in early differentiation between mild and non-mild acute pancreatitis

Hanna Sternby a, Hannes Hartman b, Dorthe Johansen a, Henrik Thorlacius a, Sara Regnér a,b, *

a Department of Surgery, Institution of Clinical Sciences, Malmö, Skåne University Hospital, Lund University, Sweden
b Department of Gastroenterology, Institution of Clinical Sciences, Malmö, Skåne University Hospital, Lund University, Sweden

A R T I C L E   I N F O

Article history:
Received 17 February 2017
Received in revised form 3 May 2017
Accepted 29 May 2017
Available online xxx

Keywords:
Acute pancreatitis
Biomarkers
Classification
Stratification

A B S T R A C T

Background: The revised Atlanta classification on acute pancreatitis (AP) presents distinct criteria for severity categorization. Due to the lack of reliable prognostic markers, a majority of patients with AP are currently hospitalized and initially managed identically. As incidence and financial costs are rising the need for early severity differentiation will increase. This study aimed to investigate the capacity of biomarkers to stratify AP patients during the initial course of the disease.

Methods: Patients with AP were prospectively enrolled and dichotomized into mild or non-mild (moderately severe and severe AP) according to the revised Atlanta classification. Serum samples taken within 13–36 h after onset of disease were analyzed for 20 biomarkers. Through receiver operating curves cut-off levels were set for 5 biomarkers whose stratifying ability was further analyzed. Additionally, the patients were classified according to the harmless acute pancreatitis score (HAPS).

Results: Among the 175 patients, 70.9% had mild and 29.1% non-mild AP. CRP and IL-6 combined, with cut-off levels 57.0 and 23.6 respectively, demonstrated superior discriminative capacity with an area under the curve of 0.803, sensitivity 98%, specificity 54% and a positive and negative likelihood ratio of 2.1 and 0.06 for the non-mild group. Regarding the mild group likelihood ratios were positive 26.5 and negative 0.48. The identification potential of the HAPS was generally inferior when compared to CRP plus IL-6.

Conclusions: In this study CRP and IL-6 demonstrate a clinically relevant capacity to differentiate mild from non-mild AP early in the course of AP.

© 2017 IAP and EPC. Published by Elsevier B.V. All rights reserved.

Introduction

At a cost of several billion dollars per year, acute pancreatitis (AP) is reported to be the leading cause of hospital admission among gastrointestinal diseases in the United States [1,2]. Globally, the incidence of the disease is rising with the most frequent etiologies being gallstones and alcohol [1,3].

The revised Atlanta classification of 2012 categorizes AP into mild, moderately severe and severe disease [4]. The main focus in research has been on the severely ill group due to the associated organ failure and high mortality rate. The majority of patients will, however, develop a self-limiting mild form of the disease. Mild AP is acknowledged to have no complications and mortality is extremely rare [4–6]. Nevertheless all patients with AP are hospitalized with supportive care and various interventions. Early severity stratification would reduce overtreatment of the mild group but also give legitimate attention to those in risk of developing non-mild (moderately severe and severe) AP. When the incidence rises, identification of both groups will be necessary for accurate medical surveillance and treatment as well as to reduce health care costs.

Clinical judgment in early severity assessment has been demonstrated to be poor [7,8]. Hence multiple studies have investigated the role of biomarkers in risk stratification of AP [9,10]. Previous works are mainly based on the Atlanta classification of 1992, and no biomarker has demonstrated sufficient prognostic...
capacity early in the course of the disease [9–11]. In 2009 Lankisch et al. introduced the Harmless Acute Pancreatitis Score (HAPS) to enable prompt identification of patients with mild AP [12]. Although simple and validated, this scoring system has not been embraced by general clinical practice.

To our knowledge, no study has yet examined the prognostic capacity of biomarkers based on the revised Atlanta classification. In this work we aimed to explore the potential in using biomarkers to separate mild AP from non-mild during the initial phase of the disease.

Methods

Patients and study design

Patients >18 years with AP admitted to the Department of Surgery, Skåne University Hospital, Malmö, Sweden, from January 2010 to September 2013 were prospectively and consecutively included. For the diagnose of AP two out of three criteria needed to be fulfilled: 1) acute characteristic upper abdominal pain, 2) serum amylase ≥3 times the upper limit or 3) characteristic findings of AP on CT scan, abdominal ultrasound or MRI. The patients were retrospectively classified as having mild, moderately severe or severe AP according to the revised Atlanta classification of 2012 [4]. Clinical data, including exact time for onset of disease (equivalent to onset of pain) and validated questions on etiology, were obtained from the patients upon inclusion and retrospectively through review of medical notes as described previously [13]. HAPS was analyzed for all patients [12].

The study was approved by the regional ethics committee at Lund University (2009/415). Informed consent (oral and written) was obtained from all participants included in the study.

Blood samples and biomarkers

Serum samples were collected upon admission to the emergency room. As the aim was to analyze the initial course of the disease, only specimens obtained between 13 and 36 h after onset of disease were included in analysis. The time interval was chosen based on a large review by Staubli et al. on laboratory biomarkers in AP [10]. Albumin (g/L), calcium (mmol/L), c-reactive protein (CRP, mg/L), glucose (mmol/L), hematocrit, hemoglobin (g/L), lactate (mmol/L), thrombocytes (x10⁹/L) and white blood cells (x10⁹/L) were analyzed in accordance with certified standard analysis at the department of Clinical Chemistry, Skåne University Hospital Malmö (ISO 15189:2012, accreditation number 1309). Blood samples for analysis of interferon gamma (IFN-γ), interleukin-1β (IL-1β), interleukin-6 (IL-6), interleukin-6 receptor (IL-6R), interleukin-8 (IL-8), interleukin-10 (IL-10), interleukin-12 (IL-12), tumor necrosis factor-alpha (TNF-α), monocyte chemoattractant protein-1 (MCP-1), procalcitonin and D-dimer were collected in plasma separator tubes, centrifuged (2000 rounds, 25 °C, 10 min) and stored at −80 °C until analyzed. IFN-γ, IL-1β, IL-6, IL-8, IL-10, IL-12 and TNF-α were analyzed with human proinflammatory 7-plex ultrasensitive kit and IL-6R with 1-plex human IL-6R ultrasensitive kit (K1508C8, K151ALC, Meso Scale Diagnostics LLC, Rockville, MD, USA). MCP-1 was assessed through human CCL2 (MCP-1) Elisa Ready-set-go kit (88-7399-88, AbiScience, San Diego, CA, USA) and D-dimer with human D-Dimer Elisa kit (D2D, 20870, Bmassay, Beijing, China). All analyses were assessed according to the manufacturer’s instructions. Procalcitonin was determined with an accredited Elisa method based on monoclonal anti-procalcitonin antibodies in accordance with routine methods at the department of Clinical Chemistry, Skåne University Hospital Malmö. Unit of all interleukins, IFN-γ and TNF-α is pg/ml whereas MCP-1, d-dimer and procalcitonin are in ng/ml.

Statistical analysis

All patients were dichotomized into mild and non-mild (moderately severe and severe) AP according to the revised Atlanta classification [4]. For continuous data, comparison between two groups, Mann-Whitney U test was used and p < 0.05 was considered statistically significant. The predictive performance of individual biomarkers were assessed through Receiver Operating Characteristic curves (ROC-curves), Areas Under Curves (AUCs) and cut-off levels with corresponding sensitivity, specificity, positive and negative likelihood ratio as well as positive and negative predictive value.

All statistical analysis was performed using IBM SPSS Statistics for Windows, version 21 and 22, Armonk, NY:IBM corp.

Results

During the inclusion period in total 245 patients with AP were admitted to the department of surgery. Among the 70 patients not enrolled 57 had samples taken before 13 h or after 36 h into the course of the disease, 5 patients refused to participate and 8 patients had insufficient level of language comprehension. Thus, 175 patients with samples taken within 13–36 h after onset of disease were included. Basal cohort characteristics are presented in Table 1. Median age was 66 years (range 19–97) and just over half were women. Most frequent etiology (54.8%) was biliary whereas a fifth was alcohol-induced. Remaining etiologies were idiopathic and other (post-ERCP, tumors, strictures etc). According to the revised Atlanta classification 70.9% had mild, 23.4% (41 patients) moderately severe and 5.7% (10 patients) severe AP. The non-mild group thus amounted to 29.1% of the patients. One patient with mild AP was admitted to the intensive care unit (ICU) due to delirium tremens. Among the 25 patients with organ failure, 12 were admitted to the ICU and 13 were managed at an intermitteance ward. Significant difference (p = 0.0001) between the mild and non-mild groups was solely found for length of hospital stay.

Of the 20 biomarkers investigated, significant difference between the mild and non-mild groups was found for seven biomarkers: IL-1β, IL-6, IL-6R, IL-10, MCP-1, calcium and CRP, see Table 2. ROC-curve analysis demonstrated AUCs below 0.5 for calcium and IL-6R, results for the remaining biomarkers are presented in Table 3. For graphical description of ROC-curves see Figs. 1–5 in supplementary material. CRP and IL-6 demonstrated superior results with AUCs of 0.808 and 0.749 respectively. Optimal cut-off levels, considering the risk of morbidity and mortality in AP, were identified from ROC-curves (Table 4). For each biomarker corresponding positive and negative likelihood ratios as well as positive and negative predictive values were calculated.

As CRP and IL-6 demonstrated superior predictive capacity in ROC-analysis a combination of these biomarkers was further investigated and the results are presented in Table 4. Only one non-mild patient with deterioration of co-morbidity (without need for admission to the ICU) was identified below the cut-off levels of both CRP (57.0) and IL-6 (23.6). Overall, severity stratification was improved when the combination of CRP and IL-6 was applied compared with individual biomarkers.

Analysis of the cohort according to the HAPS and corresponding figures using CRP plus IL-6 are presented in Table 5. The single non-mild patient classified as harmless developed moderately severe AP with multiple transient organ failure and needed management in the ICU. The HAPS demonstrated high specificity but low sensitivity for the mild AP patients. Application of selected cut-off levels for CRP and IL-6 combined resulted in equally high specificity as for the
Table 1: Baseline characteristics of all patients and mild and non-mild groups of AP.

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Mild (70.9)</th>
<th>Non-mild (29.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n (%)</td>
<td>175</td>
<td>124</td>
<td>51</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>85 (48.6)</td>
<td>60 (48.3)</td>
<td>25 (49.0)</td>
</tr>
<tr>
<td>Female</td>
<td>90 (51.4)</td>
<td>64 (51.6)</td>
<td>26 (51.0)</td>
</tr>
<tr>
<td>Age, years</td>
<td>*66 (19–97)</td>
<td>*65 (19–97)</td>
<td>*66 (29–92)</td>
</tr>
<tr>
<td>BMI</td>
<td>*25.7 (13.6–47)</td>
<td>*25.1 (15.2–45.4)</td>
<td>*26.6 (13.6–47)</td>
</tr>
<tr>
<td>Hours from onset to admission</td>
<td>*24 (13–36)</td>
<td>*24 (13–36)</td>
<td>*23 (13–36)</td>
</tr>
<tr>
<td>Etiology, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biliary</td>
<td>96 (54.8)</td>
<td>70 (56.5)</td>
<td>26 (51.0)</td>
</tr>
<tr>
<td>Alcohol</td>
<td>36 (20.6)</td>
<td>24 (19.4)</td>
<td>12 (23.5)</td>
</tr>
<tr>
<td>Other</td>
<td>23 (13.1)</td>
<td>18 (14.5)</td>
<td>5 (9.8)</td>
</tr>
<tr>
<td>Idiopathic</td>
<td>20 (11.4)</td>
<td>12 (9.7)</td>
<td>8 (15.7)</td>
</tr>
<tr>
<td>Organ failure, n (%)</td>
<td>25 (14.3)</td>
<td>0</td>
<td>25 (49.0)</td>
</tr>
<tr>
<td>ICU, n (%)</td>
<td>13 (7.4)</td>
<td>1 (0.8)</td>
<td>12 (23.5)</td>
</tr>
<tr>
<td>Mortality, n (%)</td>
<td>7 (4.0)</td>
<td>0</td>
<td>7 (13.7)</td>
</tr>
</tbody>
</table>

Values are in median (range).

†Significant difference (p = 0.0001) was found for length of hospital stay.

Table 2: Comparison of biomarker levels between mild and non-mild groups.

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>P value</th>
<th>Median range</th>
<th>Median range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mild</td>
<td>Non-mild</td>
</tr>
<tr>
<td>IFN-γ</td>
<td></td>
<td>2.6 (0.4–282.8)</td>
<td>2.6 (0.1–94.5)</td>
</tr>
<tr>
<td>IL-1β</td>
<td>0.0001</td>
<td>1.2 (0.0–32.7)</td>
<td>2.5 (0.0–61.0)</td>
</tr>
<tr>
<td>IL-6</td>
<td>0.0001</td>
<td>21.9 (6–3776.0)</td>
<td>120.5 (12.0–3662.0)</td>
</tr>
<tr>
<td>IL-8</td>
<td>0.003</td>
<td>37994 (13125–84080)</td>
<td>32169 (9300–61028)</td>
</tr>
<tr>
<td>IL-10</td>
<td>0.002</td>
<td>11.9 (0.9–4429.7)</td>
<td>23.4 (3.3–3579.0)</td>
</tr>
<tr>
<td>IL-12</td>
<td>ns</td>
<td>1.3 (0.0–4761.5)</td>
<td>0.8 (0.0–411.5)</td>
</tr>
<tr>
<td>TNF-α</td>
<td>ns</td>
<td>6.0 (2.3–916.7)</td>
<td>5.2 (1.9–103.0)</td>
</tr>
<tr>
<td>MCP-1</td>
<td>0.044</td>
<td>188.0 (0.0–7299.0)</td>
<td>236.5 (54.0–7080.0)</td>
</tr>
<tr>
<td>D-dimer</td>
<td>ns</td>
<td>787.5 (333.0–3448.0)</td>
<td>1013 (329–2270)</td>
</tr>
<tr>
<td>Procalcitonin</td>
<td>ns</td>
<td>1.6 (0.0–48.0)</td>
<td>0.8 (0.0–60.0)</td>
</tr>
<tr>
<td>Albumin</td>
<td>ns</td>
<td>31.4 (20–43)</td>
<td>31.4 (21–41)</td>
</tr>
<tr>
<td>Calcium</td>
<td>0.013</td>
<td>2.1 (2.0–2.4)</td>
<td>2.4 (2.0–2.6)</td>
</tr>
<tr>
<td>CRP</td>
<td>0.0001</td>
<td>58 (1–385)</td>
<td>211 (7–489)</td>
</tr>
<tr>
<td>p-glucose</td>
<td>ns</td>
<td>7.8 (4.2–16.4)</td>
<td>8.1 (4.7–17.3)</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>ns</td>
<td>140 (102–18)</td>
<td>145 (115–171)</td>
</tr>
<tr>
<td>Hematocrit</td>
<td>ns</td>
<td>0.44 (0.32–0.56)</td>
<td>0.44 (0.36–0.54)</td>
</tr>
<tr>
<td>Lactate</td>
<td>ns</td>
<td>1.9 (0.5–4.0)</td>
<td>1.7 (0.8–7.0)</td>
</tr>
<tr>
<td>Thrombocytes</td>
<td>ns</td>
<td>233 (107–416)</td>
<td>233 (88–550)</td>
</tr>
<tr>
<td>White blood cells</td>
<td>ns</td>
<td>11.4 (3.8–20.1)</td>
<td>12.7 (1.0–24.8)</td>
</tr>
</tbody>
</table>

Table 3: Results of ROC-curve analysis.

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>AUC</th>
<th>p-value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-1β</td>
<td>0.577</td>
<td>0.0001</td>
<td>0.583–0.770</td>
</tr>
<tr>
<td>IL-6</td>
<td>0.740</td>
<td>0.0001</td>
<td>0.674–0.825</td>
</tr>
<tr>
<td>IL-10</td>
<td>0.654</td>
<td>0.002</td>
<td>0.568–0.739</td>
</tr>
<tr>
<td>MCP-1</td>
<td>0.609</td>
<td>0.044</td>
<td>0.509–0.710</td>
</tr>
<tr>
<td>CRP</td>
<td>0.808</td>
<td>0.0001</td>
<td>0.730–0.886</td>
</tr>
</tbody>
</table>

AUC, area under curve; CI, confidence interval.

Discussion

With the introduction of the revised Atlanta classification in 2012, the severity groups of AP are more explicitly defined than previously [4,11]. Historically, focus has been on the severely ill patients. The search for prognostic biomarkers in AP is since decades an ongoing process. Several studies have presented promising results, however still none have reached clinical practice [14].

The mild group is not extensively investigated, despite containing the vast majority of cases. By providing a triaged severity distribution, the revised Atlanta classification enables more detailed comparison between AP patients. The objective of this study was to examine the capacity of biomarkers in differentiating mild from non-mild AP during the initial phase of the disease.

Within the interval 13–36 h after onset of disease, levels of 7 out of 20 investigated biomarkers differed significantly between the mild and non-mild groups. Predictive accuracy of non-mild disease was superior for CRP (AUC 0.808) and IL-6 (AUC 0.749) in ROC-curve analysis. Due to the potentially devastating consequences of severe AP, cut-off levels were selected to obtain high sensitivity. With 57 mg/L for CRP and 23.6 pg/mL for IL-6 sensitivities of almost 98% and 53% respectively. Additionally improved likelihood ratios as well as predictive values were obtained.

Currently all patients with AP are hospitalized with supportive care and various interventions [14]. Data show a 100% rise in the overall admissions for AP in the US during the last 2 decades with a subsequent augmentation in hospital expenses [2]. The increasing incidence of AP emphasizes the need for early identification of not only severe but also mild cases. Moreover, studies have demonstrated significant increases in hospital-acquired infections, length of hospital stay and even mortality also for cases of mild AP [15,16].
If identified upon admission, early discharge or adequate treatment as outpatients with standardized follow-up plans is a possibility for this group of patients [17]. Additionally major medical attention would be directed towards those at risk of developing more severe disease. Such a strategy is, however, currently generally inapplicable as most centers do not have sufficient logistic structure to care for the mild AP patients as outpatients.

Herein is demonstrated good capacity of CRP and IL-6 in identifying mainly mild but also non-mild cases in AP. Our findings are in line with previous efforts to differentiate patients with AP, although the aim of most prior studies have been prediction of severe disease according to the Atlanta classification of 1993 [10,11,18]. Biomarkers analyzed in this study have earlier demonstrated acceptable performance in early assessment of AP [19–23].

The single scoring system developed for prediction of mild AP is the HAPS. Several studies have reported high specificity for the identification of mild cases when validating this algorithm [12,24]. When applied on our cohort high specificity but very low sensitivity was found. 92% of the patients with mild AP were classified as not harmless (Table 5). As the HAPS is developed for the Atlanta classification of 1993 comparison with the mild group in this work is not fully feasible. Investigation of mild cases using CRP plus IL-6 resulted in an equally high specificity as when applying the HAPS. However, clinically much more relevant figures were obtained regarding sensitivity, positive predictive value and positive likelihood ratio.

The major limitation of this study is its exploratory nature, as the aim was not to establish a new score for the identification of non-mild or mild AP. We have examined the capacity of biomarkers in separating mild from moderately severe and severe AP early in the course of the disease. Our results, with high discriminative ability for CRP and IL-6, indicate clinical applicability of these biomarkers in determining accurate level of care for the patients during the initial phase of the disease. Additionally, our findings regarding mild AP were superior to what was obtained from the HAPS.

Although being a common disease the research on AP is declining [25]. As the incidence is rising new efforts are needed for medically correct but also economically efficient managing of the patients. This study is, to our knowledge, the first study investigating biomarkers in severity stratification of mild and non-mild AP according to the revised Atlanta classification. According to our results two generally available and inexpensive markers obtained from regular blood samples can be of relevance when differing mild from non-mild disease. Further international multicenter studies, preferably investigating not only the combination of biomarkers but also the clinical outcome of an early discharge for mild AP patients, are needed to confirm our findings.

### Table 5
Analysis of mild AP according to HAPS and CRP + IL-6.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>HAPS</th>
<th>CRP + IL-6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Harmless</td>
<td>Not Harmless</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.08</td>
<td>0.53</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>PPV</td>
<td>0.30</td>
<td>0.48</td>
</tr>
<tr>
<td>NPV</td>
<td>4.00</td>
<td>26.5</td>
</tr>
</tbody>
</table>

LR+ positive likelihood ratio; LR-, negative likelihood ratio; PPV, positive predictive value; NPV, negative predictive value.

### Funding
This work was supported by grants from the Swedish Medical Research Council (2012-3685), Bergwalls-, Ihre-, Lundgrens-, Einar och Inga Nilssons-, Stewens-, Schyberg- and SUS foundations.

### Disclosure statement
The authors declare to have no conflicts of interest.

### Acknowledgements
The authors wish to thank Anne-Marie Rohrstock for valuable support in sample analysis and Ingrid Palmquist for assistance in the handling of samples.

### Appendix A. Supplementary data
Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.pan.2017.05.392.

### References


