Acute Abdominal Pain

HELENA LAURELL
Dissertation presented at Uppsala University to be publicly examined in Enghoffsalen, Thoraxcentrum, ingång 50, Akademiska sjukhuset, Uppsala, Friday, October 13, 2006 at 13:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in Swedish.

Abstract

The aim was to identify diagnostic difficulties for acute abdominal pain at the emergency department and during hospital stay. A total of 3349 patients admitted to Mora Hospital with acute abdominal pain of up to seven days duration, were registered prospectively for history and clinical signs according to a structured schedule. The preliminary diagnosis from the attending physician at the emergency department, any investigations or surgery and final diagnosis were registered at a follow-up after at least one year.

There were no differences in diagnostic performance between physicians with 0.5 to 5 years of medical experience. The information collected and a careful examination of the patient was more important than formal competence. The main differential diagnostic problem was non-specific abdominal pain; this was the same for diagnoses requiring surgery. Patients originally diagnosed as not needing surgery had a median delay before operation of 22 hours (mean 40 hours, with 95% confidence interval of 30-50 hours), compared to 8 hours (mean 15 hours, 95% confidence interval of 12-28 hours) for patients with the same final follow-up diagnosis as the preliminary diagnosis. Constipation was a diagnostic pitfall, as 9% of the patients considered constipated required surgery for potentially life threatening reasons and 8% were later found to have an abdominal malignancy. Both the preliminary diagnosis and the discharge diagnosis were less reliable for elderly patients than for younger patients.

Elderly patients often had specific organ disease and arrived at the emergency department after a longer history of abdominal pain.

This study confirms that assessment of suspected appendicitis can still be based on clinical judgements combined with laboratory tests. Classical clinical findings indicating localised inflammation, such as isolated pain in the right iliac fossa, rebound tenderness, right-sided rectal tenderness, pain migration to the right iliac fossa, local guarding and aggravation of pain when moving, were reliable for predicting acute appendicitis. A CT scan can be saved for the more equivocal cases of acute abdominal pain. A generous strategy regarding CT scan among elderly patients with acute abdominal pain, even in the absence of pronounced signs of an inflammatory intra-abdominal process, is recommended.

Keywords: acute abdominal pain, diagnostic pitfalls, decision-making, competence, sensitivity, specificity, high age, acute appendicitis

Helena Laurell, Department of Surgical Sciences, Akademiska sjukhuset, Uppsala University, SE-75185 Uppsala, Sweden

© Helena Laurell 2006

ISSN 1651-6206
ISBN 91-554-6664-8
urn:nbn:se:uu:diva-7161 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-7161)
The diagnostic problem of today
Has greatly changed –
    the change has come to stay;
We all have to confess, though with a sigh,
On complicated tests we much rely
And use too little hand and ear and eye.

Sir Zachary Cope (1881-1974)
Abdomen in Rhyme, 1947
List of papers

This thesis is based on the following papers, which are referred to in the text by the Roman numerals given below (I-IV):

I. Acute abdominal pain in a defined population – the impact of the formal competence of the physician and the use of a structured schedule for investigation.
Laurell H, Hansson L-E & Gunnarsson U, submitted

II. Diagnostic pitfalls and accuracy of diagnosis in acute abdominal pain.
Laurell H, Hansson L-E & Gunnarsson U, Scan J of Gastroenterology 2006;41:1-6

III. Acute abdominal pain among elderly patients.
Laurell H, Hansson L-E & Gunnarsson U, Gerontology 2006;52(6):339-344

IV. Manifestations of acute appendicitis – a prospective study on acute abdominal pain.
Laurell H, Hansson L-E & Gunnarsson U, manuscript

Reprints were made with the permission of the publishers.
Contents

Introduction ................................................................................................... 11
Definitions .................................................................................................... 11
Historical remarks .................................................................................... 11
Modern research on AAP ....................................................................... 16
The physiology of abdominal pain ......................................................... 18
Epidemiology ............................................................................................. 19
The process of decision-making ............................................................ 20
Diagnostic aids ......................................................................................... 21

Aims of the study ...................................................................................... 23

Patients and methods ............................................................................. 24
Physicians .................................................................................................. 24
Inclusion criteria ....................................................................................... 24
Computer registration and validation of data ....................................... 24
Baseline study .......................................................................................... 25
Diagnoses .................................................................................................... 26
Follow-up .................................................................................................... 26
Patients included ....................................................................................... 26
Flow charts ................................................................................................ 28
Statistics ..................................................................................................... 30
Ethics .......................................................................................................... 30
The structured schedule for investigation ............................................ 31

Results ........................................................................................................... 35
Paper I ....................................................................................................... 35
Paper II ...................................................................................................... 38
Paper III ..................................................................................................... 42
Paper IV ..................................................................................................... 46

General discussion .................................................................................... 51
Measures of diagnostic performance of patients with AAP ............... 52
Duration of hospital stay ......................................................................... 54
Use of resources ....................................................................................... 54
Computer-aided diagnosis and scoring systems .................................. 55
Gender perspectives .................................................................................. 56
Abbreviations

AAP  Acute abdominal pain
CI   Confidence interval
CRP  C-reactive protein
CT   Computed tomography
ED   Emergency department
LR+  Positive likelihood ratio
LR-  Negative likelihood ratio
MRI  Magnetic resonance imaging
NPV  Negative predictive value
NSAP Non-specific abdominal pain
OMGE Organisation Mondiale de Gastroenterologie
OR   Odds ratio
PPV  Positive predictive value
UK   United Kingdom of England
VAS  Visual Analogue Scale
Introduction

Acute abdominal pain (AAP) accounts for a substantial proportion of patients arriving at a surgical emergency department (de Dombal 1979, Irvin 1989, Powers 1995). As AAP may be caused by both life threatening diseases and conditions that are spontaneously resolved, a correct diagnosis is of importance for the prognosis of the patient. AAP also affects health economy as many patients are hospitalised for what eventually turns out to be self-resolving pains, i.e. non-specific abdominal pain (NSAP).

Definitions

Acute abdominal pain (AAP) is in this context defined as abdominal pain of up to seven days duration and not caused by trauma. In the literature the duration of pain-criterion is sometimes defined to be ten days (de Dombal 1972). According to the OMGE survey AAP is stated as pain with less than one week’s duration (de Dombal 1979). The latter definition was used for the present study.

Non-specific abdominal pain (NSAP) includes all causes of self-limiting pain for which no surgical, gynaecological or other medical diagnosis is made. Usually no further investigation and no treatment is required. In the literature sometimes constipation, mesenteric lymphadenitis, dysmenorrhoea, gastroenteritis or dyspepsia, as well as patients undergoing negative laparotomy (i.e. no other surgical or other medical disease is determined and no other surgical procedure is carried out), are included in this category (de Dombal 1979). In the present papers constipation, gastroenteritis, dyspepsia and NSAP are considered to be separate entities.

Historical remarks

It is likely that AAP has affected man ever since pre-historical times. However, it is only within the last two hundred years that we have had more or less accurate knowledge of the intra-abdominal diseases that causes the AAP. The main reason for this comparatively late development of medical knowledge was that the only method of obtaining accurate information, by post-mortem examination of the intra-abdominal organs, was either forbidden or disliked by the medical authorities. Moreover, surgical operations
upon the abdomen were not performed until the beginning of the 19th century.

Hippocrates (460-370 BC), Galen (129-199 AD) and their successors observed the symptoms of their patients very carefully, but they were seriously handicapped by their ignorance of the exact state, within the abdomen, corresponding to those symptoms. They also lacked the ability to ascertain if their treatment affected the disease. We now know that many cases of attacks of abdominal pain are self-limiting and resolves spontaneously, and thus also ineffective treatments could seem successful during these ancient times. They recognised that acute colicky might get well or might pass on to more serious obstruction of the bowel (ileus); they also knew that deep inflammation could result in the formation of an abscess that might burst spontaneously or could be opened by the knife. However, every other serious abdominal condition, which was fatal, was thought to be due to an obstruction of the bowels, and was called ileus. The symptoms of ileus are well described by Hippocrates: “In ileus the belly becomes hard, there are no motions; the whole abdomen is painful, there are fever and thirst sometimes the patient is so tormented that he vomits bile …. It is an acute and dangerous disease.” (Cope 1965).

It is in the writings of Celsus (42 BC – 37 AD) we find the first description of a notable clinical sign in acute abdominal disease; resistance on pressure indicating underlying inflammation or peritonitis.

For more than a thousand years after the time of Galen no advances in the diagnosis or treatment of acute abdominal conditions took place. However, many monasteries included also education and hospital care. Notes from monks and nuns reveal that women mostly were seeking medical care for mental disorders or infertility, men for wounds and accidents and children for fever diseases and diarrhoea. During the 13th century the first universities were established. To achieve the medical authorization from Salerno also one year of practise with an experienced physician was demanded. The first official medical authorization was granted in 1140 in Sicily.

Thomas Sydenham (1624-1689) found little practical value of the scientific knowledge of medicine for the treatment of the sick, he focussed on clinical observations of patients and his greatest contribution was the many descriptions of the natural history of diseases (Nilsson 1998). It was first after the publication of Bonetus’ Sepulchretum in 1679 and the remarkable wealth of pathological material described in Giovanni Morgagni’s (1682-1771) monumental work De sedibus et causis morborum (published in 1761) that a decisive change in medical views on the causation of intra-abdominal disorders took place. William Cullen (1710-1790), in his work First Lines of the Practice of Physic (1776), used the name “peritonitis” for a condition involving the lining membrane of the abdominal cavity and its extensions to the viscera.
In 1806, Christopher Pemberton (1765-1822) published the book *A Practical Treatise on Various Diseases of the Abdominal Viscera*, which is considered a milestone in the clinical diagnosis of acute abdominal diseases. Pemberton emphasised the importance of meticulous clinical observations and the importance of careful first-hand observations, recorded at the time. He also described the differential diagnostics between peritonitis and colic pain (Cope 1965). Leopold Auenbrügger (1722-1809) in Vienna described the percussion technique: he might have got the idea from a brewery where they knocked on the barrels to judge how much beer was inside (Nilsson 1998). Shaking the patient (succussion), to prove fluid in the chest, was an ancient technique that was revived in the early 1800s, but was a quite violent method and unpleasant for the patient. In the 19th century, "topographical diagnostics" was developed for determining the size of heart and liver by palpation and percussion (Johannisson 1997). When René Laënnec (1781-1826) invented the stethoscope in 1816, the technique of auscultation was immediately accepted. During this time post-mortem examinations began to be made more frequently, and there was a rapid improvement in diagnosis. From the middle of the 19th to the middle of the 20th century, most of the pathological conditions, which give rise to acute abdominal symptoms, were recognised and their symptoms gradually differentiated.

Josef Lister (1827-1912) in London realized the importance of Louis Pasteur’s (1822-1895) work on bacteria and introduced the “antiseptic surgery” in 1867, by spraying carbolic acid in the surgery room. The reduction of postoperative infections and sepsis was remarkable, and the aseptic method was soon practised in all the countries in Europe and the USA. On 16 October 1846 ("Ether Day") at Massachusetts General Hospital, the dentist William Morton (1819-1868) administrated ether to a patient of the surgeon, John Collins Warren (1778-1856) and the operation was painless. The possibility of performing painless surgery and the increased postoperative survival, due to the aseptic method, markedly increased the possibilities of surgical treatment of acute abdominal conditions during the last decades of the 19th century (Rhodes 1985).

Wilhelm Conrad Röntgen’s (1845-1923) detection of a “new kind of rays” in 1895 provided the possibility of imaging bones inside the body. The technique spread immediately and in 1896 the first radiographs were performed in Stockholm and Uppsala. Bismuth and barium were soon shown to be impermeable to the x-rays and were used as contrast in examinations of the stomach, small and large intestines: gastric ulcers were diagnosed by this method. In 1929 the professor of radiology in Uppsala, Hugo Laurell (1884-1959) performed the first reposition of invaginated intestine by barium enema, which is probably the first radiological therapeutic method. His scientific work became the basis of radiological diagnostic methods in acute abdominal pain (Laurell 1930). The Penicillium was discovered in 1928 by Sir Alexander Fleming (1881-1955), but it was not in clinical use until 1941.
One clinical consequence of the introduction of antibiotics was the possibility to wait and observe the patient and not perform an appendectomy as soon appendicitis was suspected (Ekelund 2005). In 1970, computed tomography (CT) were introduced into clinical practice, along with magnetic resonance imaging (MRI) in 1980, both made important contributions to the diagnostic performance of abdominal diseases.

The mean hospitalisation time on surgical wards in Sweden was 14.8 days in 1950 compared to 5.9 days in 1990. The number of beds in surgical care was 8774 at 63 surgical departments in 2000 (The Swedish Federation of County Councils 2003). The total number of beds for institutional care (except psychiatry and geriatrics) in Sweden was 35,726 in 1950, 49,600 in 1970, 35,503 in 1990 (SBU 1995) and 26,076 in 2000 (The Swedish Federation of County Council 2003). That means 5.1, 6.1 and 3.4 beds per 100,000 inhabitants, 1950-1990 respectively (SBU 1995).

The growth of knowledge of appendicitis

Lorenz Heister (1683-1758) was the first to recognise that the vermiform appendix might be the site of acute primary inflammation (Ekelund 2005). Claudius Amyand (1681-1740), physician at the English court, is said to have performed the first recorded successful appendectomy in 1736, on an 11-year old boy with a perforated appendix within an inguinal hernia sac. (Amyand 1736). During the 19th century pain in the right iliac fossa was recognized as a clinical entity. However, it was first in 1886 that Reginald Fitz (1843-1913) proved that the great majority of acute inflammatory conditions around the caecum had their origin in the vermiform appendix. As a physician, he gave a description of the symptoms whereby the correct diagnosis could be made, and as a pathologist he described the varying degrees of inflammation in the appendix, the degrees of local and general peritonitis, and of abscess formation. In 1889, Charles Mac Burney (1845-1913) in New York described how an inflamed appendix could be reached and removed with small risk via a short incision. The appendectomies were lifesaving and were soon widely spread (Rhodes 1985). The mortality in appendicitis has been estimated to over 40% before the introduction of appendectomy; then decreased to around 10% in the 1910s; a few per cent in the 1940s (after the introduction of antibiotics) and a few per thousand today (Blomqvist 2001, Lally 2004). In Sweden, the first appendectomy was performed in 1889 by Karl Gustaf Lennander, who also introduced “aseptic surgery” to the country (Petrén 1932). Lennander’s successor as Head of the Surgical Department in Uppsala, Gunnar Nyström (1877-1964), wrote his thesis in 1907 on 312 patients surgically treated for non-perforated appendicitis in Uppsala during 1891-1905. Only one patient died postoperatively (of intestinal obstruction; Nyström 1907). There were 681 appendectomies performed in Sweden in 1901, increasing to 14,000 performed in 1921 (Nyström 1932).
Lorenz Heister (1683-1758) was the first to recognise that the vermiform appendix might be the site of acute primary inflammation. (Wikipedia: GNU Free Document.)
Modern research on AAP

AAP can involve different medical specialities, including surgery, gynaecology, medicine and oncology. Although AAP is a common problem, early diagnosis is often difficult and this may lead to unnecessary investigation and hospitalisations. In the late 1950s, systems appeared that aided the physician in making medical decisions and were based on methods that used decision trees or truth tables. Later, systems based on statistical methods appeared for different domains of clinical care, for example interpretations of electrocardiograms or diagnoses in internal medicine.

In 1970, Professor F Tim de Dombal introduced a computer-aided data system in Leeds, UK, which dealt with 13 different diagnoses for AAP. Based on patient data and the results of laboratory tests, this system proposed a diagnosis. In 1972, de Dombal reported a real-time comparison of physician and computer-aided diagnosis in a series of 304 prospectively collected patients suffering from AAP during 1971. The overall diagnostic accuracy of the computer system was 91.8%, and was significantly higher than the most senior clinicians, who presented an accuracy of 79.6% (de Dombal 1972). Such high diagnostic accuracy of AAP have, however, never been reproduced.

In 1986, a study was published (Adams 1986) describing computer-aided diagnosis for patients with AAP in eight UK hospitals with over 250 participating doctors and 12,662 patients. Diagnostic accuracy rose from 45.6% during the baseline period to 65.3% during the period when the computer system to aid diagnosis based on a Bayesian analysis was tested. Other centres outside England, that have followed the idea of prospective investigation of patients admitted with AAP include; Rispebjerg Hospital in Copenhagen, Denmark (Bjerregaard 1976); Rogaland Central Hospital, Stavanger, Norway (Körner 1998); Akershus Central Hospital, University of Oslo, Norway (Staniland 1980); University Central Hospital, Tampere, Finland (Ikonen 1983); Nacka Hospital, Sweden (Fenyő 2000); University Hospital, Zürich, Switzerland (Simmen 1991); Heinrich-Heine-University, Düsseldorf, Germany (Ohmann 1995); Tucson Veteran Administration Medical Centre, Arizona, USA (Orient 1986); University of Virginia Hospital, USA (Powers 1995); and Mora Hospital, Sweden (present study).

The availability of an increased number of more specified blood tests and the possibility of detailed imaging by CT and ultrasonography have had an important impact on clinical work in surgical departments. Controlled clinical trials have revealed that this technology improves diagnostic accuracy (John 1993, Rosen 2003). However, other studies (Flum 2005, Lee 2006) have not shown a decrease in misdiagnosis among women of reproductive age, patients older than 65 years or negative appendectomies during the last decades, despite the dramatically increased use of advanced diagnostic technology. Andersson et al. (1999) found that the surgeon’s attitude to explora-
tion can influence the incidence of non-perforated appendicitis and conserva-
tive management decreases the number of negative appendectomies
(Andersson 1999). In the light of these contradictory findings, it is important
to clarify the diagnostic accuracy of AAP in a population-based study, in-
cluding a careful clinical examination at the emergency department. To
achieve congruence in those clinical assessments, a detailed schedule of
patient history and clinical symptoms and signs is required for each patient.
Determining ways to define patient-categories that are more difficult to di-
agnose safely, as well as identifying those diagnoses that are difficult to de-
termine at an early stage, can be helpful to select patients for high-
technology diagnostic tools. This is important not only for economical rea-
sons, but also to reduce unnecessary exposure of the population to radiation.

Competence of the physician

In the early 1990s, there were influences in Sweden claiming that the health
care providers should not be allowed to use trainee physicians as on-calls at
the emergency departments. The Swedish National Board of Health and
Welfare (Socialstyrelsen) initiated a mapping of formal competence at the
emergency departments in Sweden (Spri 1990). They determined that al-
though the number of specialists in Sweden had increased by 30%, the por-
tion of specialists attending at the emergency departments was the same
from 1980 until 1990. The conclusion was that there are differences in the
quality of medical assessment among the emergency departments in Sweden
(Andrén-Sandberg 1991). A lively debate on medical competence took place
in Läkartidningen (Journal for Swedish Physicians) and in 1991 a quality
registration at the Department of General Surgery at Karlskrona Hospital
was performed on all patients admitted to the emergency department during
3 months. No quality differences were revealed in the assessment, whether
the attending physician was a specialist or a trainee physician of 1-3 years of
experience (Sevonius 1991 and 1994). With the use of a detailed schedule,
each physician would have the same amount of information when making
the diagnosis after history taking and examining the patient. Thus, one ques-
tion that can be posed is whether improved diagnostic accuracy is due to
medical experience or to the amount of information achieved by a careful
examination.

Another important factor, for which there is insufficient scientific know-
ledge, is the precision of the preliminary diagnosis in relation to the patient’s
time of arrival at the emergency department. The gender perspective is of
special relevance, as females of reproductive ages are considered more diffi-
cult to diagnose safely. These are all important issues in the planning of
emergency care. All data on symptoms and clinical findings in patients with
AAP, collected in a local database such as in the present study, is also a
source for epidemiological considerations.
The physiology of abdominal pain

Pain has been defined by The International Association for the Study of Pain as “an unpleasant sensory and emotional experience associated with occurred or potential tissue damage or described in terms of such a damage”. This definition emphasises the multiple aspects of pain. Pain stimulus may imply a physiological damage but is also a question of perception, sensibility and a psychological reaction (Mersky 1994).

Visceral pain

Visceral pain differs from surface pain in several aspects. In general the viscera have sensory receptors for no other modalities of sensation except pain. Highly localised types of damage to the viscera, such as a sharp cut, rarely cause severe pain. On the other hand, diffuse stimulation of pain nerve-endings in a viscus, such as distension or spasm of a hollow viscus, stretching of the ligaments or ischemia of visceral tissue, causes pain that can be severe. In cases of ischemia in visceral tissue, the formation of acidic metabolic products or tissue degenerative products, such as bradykinin and proteolytic enzymes, causes pain by stimulating the pain nerve-endings. The leakage of proteolytic acidic gastric juice through a ruptured gastric or duodenal ulcer causes widespread chemical stimulus to the visceral peritoneum and usually results in a severe pain.

Visceral pain originating in the thoracic and abdominal cavity is transmitted through type C-fibres in sensory nerves running through the sympathetic nervous system. These fibres only transmit a chronic-burning-aching type of pain. Interval pain is the result of rhythmic contraction of smooth muscle each time a peristaltic wave travels along the spastic gut, the obstructed ureter or the gallbladder. Mechanical stimulation of nerve endings or ischemia due to diminished blood flow during cramp might cause this type of pain. Over-distension of a hollow viscus can collapse the blood vessels encircling the viscus or those in the intestinal wall, thus promoting an ischemic pain. The liver parenchyma is almost insensitive to pain, however the liver capsule is extremely sensitive to both trauma and stretch (such as in hepatitis). The bile ducts are sensitive to distension, which causes severe pain (Guyton 1986). Secondary effects of visceral pain include perspiration, nausea, vomiting, anxiety and paleness.

Parietal pain

The parietal peritoneum is supplied with extensive innervation through fast conducting “delta” fibres from the spinal nerves. The brain has an ability to localise pain that is transmitted by spinal nerves; consequently, pain from the parietal peritoneum is often sharp and localised. When a disease affects a viscus, the pain is diffuse, but when the inflammation or infection spreads from the visceral peritoneum to the parietal peritoneum the sensation of pain
is sharper and easier to locate. Impulses from an inflamed appendix are referred to an area around the umbilicus and are of the aching, cramping type. When the inflamed appendix touches the abdominal wall, impulses from the parietal peritoneum make it possible for the patient to localise the pain directly over the irritated peritoneum in the right lower quadrant of the abdomen and this pain is of the sharp type. This is why the pain migrates to the right iliac fossa in the progression of appendicitis. Retro-peritoneal organs (i.e. kidneys and ureters) and portions of the mesentery are supplied by both visceral and parietal fibres (Guyton 1986). Parietal pain is often aggravated when the patient coughs or moves.

**Referred pain**

Pain from the viscera may be frequently localized to two areas of the body surface at the same time. Impulses from the viscera are transmitted via sympathetic afferent fibres entering the spinal cord at a certain segment. The same dermatome from which the visceral organ was originally derived in the embryo, also innervates certain areas of the skin. These dual pathways of pain transmission explains the referred pain of the gallbladder. Biliary duct and gallbladder pain, in addition to causing pain on the abdominal surface, frequently refers pain to a small area at the tip of the right scapula on the back (Guyton 1986).

**Epidemiology**

In the diagnosis of AAP it is important to have a rough knowledge of how common a certain diagnosis is among patients at the emergency department. Different population-based studies on AAP have reported a distribution of diseases that are remarkably similar (Ikonen 1983, Staniland 1980).

Distribution of diagnoses reported from different studies.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Tampere</th>
<th>Akershus</th>
<th>Copenhagen</th>
<th>Leeds</th>
<th>Worldwide</th>
<th>Stockholm</th>
<th>Mora</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSAP</td>
<td>50.3</td>
<td>41.5</td>
<td>37.3</td>
<td>50.5</td>
<td>44.4</td>
<td>23.8</td>
<td>37.0</td>
</tr>
<tr>
<td>Appendicitis</td>
<td>23.3</td>
<td>25.6</td>
<td>21.5</td>
<td>26.3</td>
<td>24.2</td>
<td>10.5</td>
<td>9.8</td>
</tr>
<tr>
<td>Cholecystitis</td>
<td>8.6</td>
<td>8.0</td>
<td>6.7</td>
<td>7.6</td>
<td>8.9</td>
<td>9.2</td>
<td>3.5</td>
</tr>
<tr>
<td>Pancreatitis</td>
<td>2.6</td>
<td>2.5</td>
<td>4.2</td>
<td>2.9</td>
<td>2.3</td>
<td>4.9</td>
<td>3.2</td>
</tr>
<tr>
<td>Renal colic</td>
<td>3.3</td>
<td>9.5</td>
<td>3.1</td>
<td>-</td>
<td>3.4</td>
<td>-</td>
<td>4.0</td>
</tr>
<tr>
<td>Small intest. obstruction</td>
<td>6.7</td>
<td>2.3</td>
<td>4.0</td>
<td>3.1</td>
<td>4.0</td>
<td>6.6</td>
<td>2.7</td>
</tr>
<tr>
<td>Perforated gastric ulcer</td>
<td>0</td>
<td>1.0</td>
<td>2.0</td>
<td>3.1</td>
<td>2.8</td>
<td>-</td>
<td>0.3</td>
</tr>
<tr>
<td>Diverticulitis</td>
<td>0.9</td>
<td>4.7</td>
<td>0.8</td>
<td>2.0</td>
<td>2.1</td>
<td>4.9</td>
<td>4.7</td>
</tr>
</tbody>
</table>
The numbers in the table are percentages of the total number of participating patients in each study as given above. The world wide results are from a multicenter study on the database in Leeds (Ikonen 1983). The results reported from Stockholm are a retrospective analysis on patients admitted to several hospitals in the Stockholm area (Fenyö 2000). The results from Mora represents the present study.

The mean length of life in Sweden has almost doubled since the Middle Ages. In the 19th century the mean length of life was 50 years. In 2001, expected lifetime for a newborn male in Sweden was 77.6 years and for a female 82.1 years (The Swedish Federation of County Councils 2003). Today’s population of “younger elderly”, aged 65-76 years, are more vital than the corresponding age-group living 40 years ago and their biological functions are relatively uninfluenced until the age of 75 years; however, there are important differences on an individual level. The mortality of cardiovascular diseases in Sweden today is 51%, cancer diseases is 22% and infectious diseases only 0.7%. The total population in Sweden has increased from 7 million (1950) to 9 million inhabitants (2000) and the proportion of those aged over 65 years has increased from 10.2% to 17.2%, and the proportion aged over 80 years has increased from 1.5% to 5.1% during the same period (The Swedish Federation of County Councils 2003). Thus, a growing number of patients admitted for AAP at emergency departments are of older ages. Elderly patients are considered having less specific symptoms in cases of abdominal pain and if the diagnosis is incorrect or delayed, the consequences are usually more severe than for younger patients. Thus, there is a need to outline secure pathways for the process of diagnosis and treatment of such patients.

The process of decision-making

The physician at a surgical emergency department is confronted with a patient who has a collection of symptoms and signs. Not only knowledge of the characteristics for each particular disease, but also knowledge of the prevalence of the diseases is important. Making a correct diagnosis at an early stage is favourable, but often not possible and a tentative diagnosis is made that guides the physician into the proper ordering of investigations and decisions about hospitalisation. As a delay in the assessment of the patient may be life threatening or cause severe morbidity, the first judgement based on the clinical examination is crucial. The preliminary diagnosis has to be reevaluated after further investigations are completed. The result from an investigation can be either false negative (sensitivity) or false positive (specificity). Making an early diagnosis is a matter of weighing the information collected against the probability of a suggested diagnosis; this process can be aided by decision-trees, scoring systems or computer-aided systems (de Dombal 1997). The physician will produce several diagnostic hypotheses of possible diagnoses, which he/she rejects or confirms during the interview
with the patient about the pain and earlier disease history, and during the physical examination. Often an erroneous diagnosis is due to lack of careful examination and not lack of medical skills.

Diagnostic aids

Scoring systems and computer-aided diagnosis
Improvement of diagnostic accuracy, is achieved with scoring systems or computer-aided diagnostic systems (Adams 1986, Körner 1998, Lamparelli 2000). Scoring systems, such as the Alvarado score, (Alvarado 1986) Eskelinen score (Eskelinen 1992) and Ohmann score (Ohmann 1995), are mainly used on patients with suspected appendicitis. Large population-based studies on AAP have been made using a database in Leeds, UK (de Dombal 1997). Results from a multi-centre (eight hospitals) study on 12,662 patients in 1986 (Adams 1986), reported an increase of diagnostic accuracy from 45.6% to 65.3%, comparing the baseline period to the test period with computer-aided diagnosis. When data collection forms were used (4075 patients) the initial diagnostic accuracy improved from a baseline of 45.7% to 56.7%. That the data collection form alone increased diagnostic accuracy by 10% indicates the importance of careful history taking and clinical examination.

Radiological investigations
During the last decade, the use of advanced diagnostic tools, such as CT or ultrasonography, has increased (Rosen 2003, Terasawa 2004). Radiological imaging as a screening instrument, in the differential diagnostics of AAP, is not only a waste of resources, but also carries the risk of producing irrelevant findings that may be misleading. Improved diagnostic tools do not diminish the need for high quality clinical assessment to select suitable patients for radiological imaging and to be able to ask the adequate questions to the radiologist.

Laboratory tests and clinical examination
The most commonly discussed laboratory tests in the literature about AAP are C-reactive protein (CRP) and white blood cell count (Andersson 1999, 2000 and 2004, Göransson 1991, Nordback 1988, Tepel 2003). CRP is an acute phase protein, which increases in the presence of inflammation under various clinical conditions. The combination of an elevated CRP and leucocytosis, i.e. higher levels of white blood cell count, can detect a serious inflammatory condition with a sensitivity of 90%, specificity of 89% and a positive predictive value (PPV) of 88% (Chi 1996).
Acute appendicitis is one of the most common differential diagnoses in a surgical emergency department. Nevertheless, this diagnosis is still a challenge because of the variation in its manifestations. Many symptoms traditionally associated with appendicitis, such as nausea, loss of appetite and right-sided rectal tenderness, have only been ascribed low or no diagnostic value (Andersson 1999). When Andersson et al. (1999) evaluated the diagnostic value of 21 elements of disease history, clinical findings and laboratory examinations in patients with suspected appendicitis, six of these variables were found to be especially important discriminators for appendicitis. The inflammatory variables (body temperature, white blood cell counts and differential white blood cell counts, CRP concentration) had predicting powers similar to those of the clinical findings (direct and rebound abdominal tenderness and abdominal guarding; Andersson 1999). This study was based on 502 patients admitted for suspected appendicitis. The present study, however, was based on patient history and clinical findings from all patients admitted to the emergency department with AAP.
Aims of the study

The specific aims of the study were:

To evaluate the importance of the formal competence of the emergency department physician, the patient’s time of arrival at the emergency department and the use of a structured schedule for investigation of patients with acute abdominal pain. (Paper I)

To identify differential diagnostic difficulties in patients with acute abdominal pain at the emergency department and during hospital stay. (Paper II)

To characterise differences in clinical presentation and diagnostic accuracy between younger and more elderly patients with acute abdominal pain. (Paper III)

To identify the most important clinical symptoms and signs for predicting acute appendicitis among patients with acute abdominal pain. (Paper IV)
Patients and methods

Mora Hospital, Sweden, is a rural district hospital with a catchment population of about 87,000. This population is demographically stable and almost 100% of the emergency care is provided by Mora Hospital. This hospital also provides care for a substantial number of tourists as the population temporarily increases to about 160,000 during the summer and wintertime. The hospital provides full 24-hour emergency service with surgery, X-ray, intensive care and on-call consultants also including gynaecology, anaesthesiology and internal medicine.

Physicians
At the time of the study, there was one physician on call at the department of surgery, most often undergoing continuing education but with an experienced consultant available at a few minutes’ notice. The most junior physicians were the locums with 0 to 2 years of medical experience, pre-registrar house officers with 0.5 to 1 year experience of clinical practice after university medical qualification, followed by the senior house officers with 1-5 years experience of surgery. During the nighttime, some general physicians from the primary health care services also acted as on-call physicians at the emergency department. The specialists and consultants at the hospital and the general physicians at the primary health care centres concerned in this study were generally well experienced with many years in the profession.

Inclusion criteria
Patients aged one year and over with AAP (not caused by trauma) of up to seven days duration.

Computer registration and validation of data
On arrival at the emergency department, all patients fulfilling the inclusion criteria had the study formulary included in their medical record. The attending physician registered data for history, symptoms, clinical signs and preliminary diagnosis before the patient left for admission to a ward or was discharged. Localisation of pain at onset and at the time of examination was marked on a figure on the study protocol. The physician was then asked to
provide a primary and a secondary most likely diagnosis; the first diagnosis was then compared with the diagnosis at discharge and at follow-up. In addition, results of laboratory investigations, surgery if any, duration of hospitalisation and final diagnosis at discharge were registered by the physician responsible for that decision. All schedules were then checked by a specially trained secretary and entered into a Microsoft Access® database. At this time, any obviously erroneous information detected was corrected and on computer registration logical filters detecting impossible or inconsistent combinations of data were applied. As a check of the validity of registration, 300 cases (10%) were randomised for validation. Data from the computerised register for those cases were checked against the medical records and if necessary against the hospital computer system for time of arrival, time of surgery etc. All erroneous data detected in these two steps were corrected in the register. The completeness of the registration was first checked by the secretaries, who indicated cases discharged with a history of AAP and without register formularies included in their medical records. In such cases, the physicians responsible for the present study checked to see if the case should have been included in the study, and if so, the diagnosis at discharge, any surgery and laboratory parameters were registered. During this procedure another 523 patients who should have been included in the database register were found.

At the emergency department, all patients admitted are routinely registered by the nurses according to type of symptoms. Those registrations were scanned on randomized days of admission for symptoms possibly related to AAP. Patients with a symptom related to abdominal pain were checked against the study register and if not present, the medical records were checked for possible non-participation. With this method, the overall completeness was calculated to be 79%. Before any calculations were performed, all stochastic or continuous variables were cleared of evidently erroneous information, that is, all data outside of the 75th percentile were checked against the record data.

Baseline study

A baseline registration of logistic data such as time of admission, the level of formal competence of the attending physician and the preliminary diagnosis was created during the period 1st February 1996 to 31st January 1997. During the baseline registration period, only the most likely diagnosis was registered for each patient. During the subsequent study period, 1st February 1997 to 1st June 2000, the initial formulary was supplemented with a detailed schedule for investigation and the physician provided a primary and a secondary diagnosis.
Diagnoses

There were three diagnoses used for comparison. The preliminary diagnosis was made by the attending physician at the emergency department or the primary health care centre after the clinical examination. The discharge diagnosis was made by the physician at the surgical ward when the patient was discharged after hospitalisation. The follow-up diagnosis was defined at the review of the case records one to three years after the initial visit to the emergency department or the primary health care centre. Thus, outpatients at the emergency department and primary health care patients had only preliminary diagnoses and follow-up diagnoses.

Follow-up

Records of all patients residing within the hospital catchment area were checked at least one year (mean 2.7 years) after admission. Follow-up was through checking the patients’ record at the surgical department and the primary health care centre, and if necessary records from other departments at the hospital. Further investigations were registered and the discharge diagnosis was re-evaluated according to the criteria of the World Organisation of Gastroenterology (OMGE; de Dombal 1979 and 1982, University of Leeds). This re-evaluated, final diagnosis then served as the basis for further statistical calculations of the reliability of the preliminary diagnosis registered on admission to the emergency department and the diagnosis at discharge or on admission to the primary health care centres.

Patients included

Patients older than one year of age and with abdominal pain of up to seven days duration who were admitted to the Mora Hospital or three primary health care centres in the hospital catchment area were prospectively registered in the local database during the period February 1st 1997 to June 1st 2000.

A total number of 3349 patients were registered (see the flow charts below). Of these, the primary health care centres contributed with 238 patients. At follow-up 12 did not fulfil the inclusion criteria and were excluded (abdominal pain caused by trauma, chronic pain caused by an known diagnosis, and insufficiently filled-in investigation schedule or missing medical record), thus 3337 patients were included in the study. Of these 3337, 2979 (89%) were living within the hospital catchment area. Tourists who received surgical treatment or had a diagnosis verified by radiology or endoscopy were included in the analyses, as they were considered as having an accurate diagnosis (n=94), whereas other tourists were excluded (n=264) due to difficulties in acquiring medical records from a large number of hospitals. Ac-
According to these criteria, 3073 records out of 3349 (92%) were included in the analyses and among them 222 out of 238 (93%) were primary health care patients. The corresponding number for the baseline registration year was 790 out of 881 (90%).

**Patients included in paper I**

The calculations were based on 3073 patients divided in two groups: 2851 patients admitted to the hospitals emergency department and 222 provided by the primary health care centres.

**Patients included in paper II**

A number of 2851 patients who were admitted to the hospitals emergency department were included in the study. Of these patients, 2052 were referred to a surgical ward and 799 were treated as outpatients. Calculations of sensitivity and specificity for the preliminary and discharge diagnoses, are based on the 2052 hospitalised patients.

**Patients included in paper III**

Patients living outside the hospital catchment area (n=358) and patients who were admitted to the primary health care centres (n=220) were excluded, leaving 2759 patients who were admitted to the hospitals emergency department in this study. The three study groups comprised 2289 patients above 20 years of age: 557 patients 65 to 79 years old, 274 patients aged 80 years or older and 1458 patients of ages 20-64 years who served as a reference group.

**Patients included in paper IV**

Patients admitted to the hospitals emergency department were included in this study (n=3099). Tourists who received surgical treatment were included, as they were considered as having a certain diagnosis (n=78), whereas other tourists were excluded (n=248) as they were not eligible for follow-up. For patients admitted more than once to the emergency department, and therefore having two or more study protocols (n=65), only the protocol if they had surgery or the protocol registered at the first visit was analysed. Therefore, 373 protocols were excluded, leaving a total number of 2478 patients eligible for statistical analyses in this study. Another 28 patients registered as non-participating were included in the calculation of incidence rate.
Flow charts

Paper I:

3349 registered \rightarrow 3337 included \rightarrow 3073 eligible for follow-up \rightarrow 222 from Primary Health Care Centres

\rightarrow 12 excluded \rightarrow 358 tourists \rightarrow 94 tourists with a safe diagnosis \rightarrow 2851 admitted to ED

\rightarrow 264 tourists excluded

Paper II:

3349 registered \rightarrow 3337 included \rightarrow 3073 eligible for follow-up \rightarrow 222 from Primary Health Care Centres

\rightarrow 12 excluded \rightarrow 358 tourists \rightarrow 94 tourists with a safe diagnosis \rightarrow 2851 admitted to ED

\rightarrow 264 tourists excluded \rightarrow 799 outpatients \rightarrow 2052 hospitalised
Paper III:

3349 registered

12 excluded

3337 included

358 tourists excluded

2979 from hospital catchment area

2289 above 20 years of age

220 from Primary Health Care Centres excluded

1458 aged 20-64

557 aged 65-79

274 aged 80 or more

3349 registered

12 excluded

3337 included

238 from Primary Health Care Centres excluded

3099 eligible for study

2851 admitted to ED

2478 study-patients

274 appendicites

2204 with AAP

248 tourists excluded

373 duplicates excluded

78 tourists with a safe diagnosis
Statistics

Statistica® software (Statsoft, Tulsa, USA) was used for the statistical calculations. Distribution fit of the data was checked initially and most parameters appeared to be normally distributed when there were many patients in each group, as judged by the Kolmogorow-Smirnow test; however, groups of variable size did not always fit into that distribution model. The non-parametric Mann-Whitney U-test was generally used to calculate the significance of differences in continuous variables, and the Chi-squared test was used in cases of dichotomous response parameters and to test differences in proportions between groups (Paper I). Correlation was calculated by the Spearman test (Paper III). Calculations of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+) and negative likelihood ratio (LR-) were used as measures of the diagnostic value of the preliminary and discharge diagnoses, clinical findings and laboratory tests (Paper II and IV).

The sensitivity of a symptom represents the probability that a patient who is suffering from for example appendicitis will have a certain symptom, and PPV is the probability that the patient is suffering from appendicitis, if a symptom or sign is present. Calculations of positive likelihood ratio for the preliminary diagnosis were used as an alternative measurement of diagnostic accuracy, independent of the prevalence of that diagnosis in this population. The LR+ indicates how many times greater the probability of a clinical finding is among patients with a specific diagnosis than in the baseline study population. An excellent diagnostic test has an LR+ of about 10, a fair LR+ is at least 2 and if the LR+ is less than 1 it indicates that the diagnostic test is invalid. The kappa value was calculated to measure the congruence between the preliminary and discharge diagnoses (Paper II). The kappa statistic provides a measure that varies from +1, indicating perfect congruence, to 0 (zero), an indication of no greater congruence than expected by chance (Boyd 1979). To ascertain the most important diagnostic markers for appendicitis (Paper IV), each diagnostic variable was tested first in a univariate logistic regression analysis, expressed as odds ratio (OR) with 95% confidence interval (CI), and then in a multivariate analysis with adjustments for age and gender.

Ethics

The study project was approved, in advance, by the Committee of Ethics at Uppsala University, Sweden. The database also obtained approval from the Swedish authority concerning registry data. All patients were provided information about the study at the emergency department and a separate information sheet for the patients was included in each investigation protocol.
The structured schedule for investigation

AKUT BUK STUDIEN
DALARNA

Ansvariga: L-E Hassel, U-Gunnarsson & H. Larelli
Kiropraktiker, Mera Lasootti

(Plex för patientbricks)

**Definition:** Baksmärtor av kontinuerlig eller intervallartad karaktär som debuterat under de senaste 7 dagarna. Barn under ett års exkluderas, likaså exkluderas patienter med baksmärta orsakad av buktrauma.

1. **Vårdförhållanden:**
   - Mora las., kir. klin.
   - VC Vansbro
   - VC Åsveden
   - VC Rättvik
   - VC Särna
   - VC Leksand
   - Annan

2. **Patientens ansion till skutkommittén:**
   - Datum
   - Klubbklag

3. **Patientens kön:**
   - Man
   - Kvinna

5. **Undersökare:**
   - TT Undersöker
   - AT-undersöker
   -brick
   - ST-undersöker i kirurgi
   - ST-undersöker i ortopedi

**ANAMNEX:**

6. **Smärtslokalisering:**

7. **Faktorer som förvärrar smärtan:**
   - Kolpkyssås
   - Husta
   - Djupandring
   - Födonsdag
   - Annat

8. **Faktorer som hindrar smärtan:**
   - Läget
   - Kräkning
   - Antacida
   - Födonsdag
   - Annat

9. **Har patienten använt:**
   - 19. Medicin mot baksmärta
   - 20. Antiflogistic
   - 21. Cortisonpreparat
   - 22. Drickat alkohol

10. **Smärtans svårighetsgrad:**
    - Ovetydiga smärtor
    - Värsta tankbara smärtor

31
UNDERSÖKNINGAR:

77. Temp [ ] [ ] grader          □ 90. Grav. test pos. □ Pos □ Neg
78. Pulse [ ] [ ] slag/min     □ 91. Urintick. □ 92. minit. pos
79. Blodtryck [ ] [ ] mm Hg (syst.) □ 93. rytt. pos

Markerade ordinerade undersökningar:

<table>
<thead>
<tr>
<th></th>
<th>Röntgen</th>
<th>95. Lungröntgen</th>
</tr>
</thead>
<tbody>
<tr>
<td>80. Hb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81. SIR</td>
<td></td>
<td>96. Blodöversikt</td>
</tr>
<tr>
<td>82. LpK</td>
<td></td>
<td>97. Tunntarmspassage</td>
</tr>
<tr>
<td>83. CRP</td>
<td></td>
<td>98. Ultraljud</td>
</tr>
<tr>
<td>84. B-Glukos</td>
<td></td>
<td>99. Urtgoss</td>
</tr>
<tr>
<td>85. S-Amilaas</td>
<td></td>
<td>100. CT-buk</td>
</tr>
<tr>
<td>86. S-ALAT</td>
<td></td>
<td>101. Laparoscopi</td>
</tr>
<tr>
<td>87. S-ASAT</td>
<td></td>
<td>102. Endoscopi</td>
</tr>
<tr>
<td>88. S-ALP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89. S-Bili</td>
<td></td>
<td>103. Gyn. konsul</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

104. Åtgärd:

□ Patienten sänds hem  □ Inomfattas till kirurgklinik
□ Patienten sänds hem men skall återkomma för kontrollpalpation  □ Inläggas på kirurgavdelning
□ Inläggas på observationssats  □ Rensas till annan klinik
□ Väntande rensas till kirurgklinik

105. OPERATION:

□ Nej
□ Ja 106. Op. nr. [ ] [ ] [ ] [ ] [ ] [ ] [ ] (ÅÅ MD DD)
107. Datum: [ ] [ ] [ ] [ ] [ ] [ ] [ ]
108. Klockslag: [ ] [ ] [ ] (TT MM) (ep non)

UTSKRIVNING:

109. Utskrivningdatum: [ ] [ ] [ ] [ ] [ ] [ ] [ ] (ÅÅ MD DD)
110. Patienten avlidna under vårdtiden. □ Ja □ Nej

111. DEFINITIV DIAGNOS: (endast ett alternativ)

□ Ospecific bakamära (1)
□ Gastroenteritis (2)
□ Dystipation (3)
□ Appendicitis, -grad av inflamm. ej spektererad (4)
□ Appendicitis, gangränös (5)
□ Appendicitis, perforerad (7)
□ Cholezystitis utan perforation (8)
□ Cholezystitis, perforerad/9
□ Galästenskolik (10)
□ Kolon/abd. ventriculitis utan perforation (11)
□ Kolon/abd. ventriculitis, perforerad (12)
□ Tunntarmöverdet utan strangulation (13)
□ Tunntarmöverdet med strangulation (14)
□ Kolonläsion (15)
□ Dyspepsia (16)

□ Ulcus (17)
□ Perforerat ulcer (18)
□ Akut pannenteritis (19)
□ Urinarydysinfektion (20)
□ Ureterenkolik (21)
□ Urininfektion (22)
□ Inflamed kjuren (23)
□ Inflamed nåderbråck (24)
□ Inflamed avbråck (25)
□ Bukmalningskatarrak (26)
□ Invagination (27)
□ Aortaaneursymnupte/dissection (28)
□ Mesenteriskarlsöck (29)
□ Gynäkologisk ökonom (30)
□ Övrigt: [ ] [ ] [ ] [ ] [ ] (31)

Diagnos nr: [ ] [ ] [ ] [ ] [ ] [ ] [ ] (Fyllas endast vid alt. "Övrigt")
Results

Paper I

During the study period of 40 months, 3073 patients with AAP were registered in the local database. Of 2851 patients admitted to the emergency department, 2062 patients (72%) were admitted to a surgical ward, whereas 789 (28%) were treated as outpatients. The mean age was 46 years. The male/female ratio was 0.82 (n=1382/1691). For age groups 15-45 years and over 90 years, the majority of patients were females. The ten most common final diagnoses at the emergency department were NSAP (37%), gallbladder disease (10.5%), appendicitis (9.8%), diverticulitis (4.7%), constipation (4.6%), ureteric stone (4%), gynaecological complaints (3.5%), acute pancreatitis (3.2%), acute intestinal obstruction (3.2%) and urinary tract infection (2.6%, Table 1). Detected abdominal malignancies constituted 2.8% (n=86). The sensitivities for the preliminary diagnoses at the emergency department were as follows: appendicitis 0.80, cholecystitis 0.51, gallstones 0.68, diverticulitis 0.64 and ureteric stone 0.78.

A majority of the patients attending the emergency department (88%, n=2699) were examined by a non-specialist physician with 0.5 to 5 years of experience. Most patients, 46%, were seen by pre-registrar house officers (Table 2). There was no significant difference in diagnostic performance according to category of physician (pre-registrar house officers, locums, senior house officers, general physicians and consultants). During the baseline period the pre-registrar house officers were in contact with the consultant in 34% of the cases, whereas the senior house officers had such contact in only 12% of the cases (p<0.001). Overall the rate of correct diagnoses was 54% and the corresponding value for baseline registration period was 58% (p=0.07).
Table 1 Diagnoses on admission (preliminary diagnoses) and after at least one year of follow-up (final diagnoses)

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Hospital n=2851 (preliminary)</th>
<th>Hospital n=2851 (final)</th>
<th>Primary health care, n=222 (preliminary)</th>
<th>Primary health care, n=222 (final)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>NSAP</td>
<td>641</td>
<td>22</td>
<td>1058</td>
<td>37</td>
</tr>
<tr>
<td>Gastroenteritis</td>
<td>94</td>
<td>3</td>
<td>64</td>
<td>2.2</td>
</tr>
<tr>
<td>Constipation</td>
<td>208</td>
<td>7</td>
<td>130</td>
<td>4.6</td>
</tr>
<tr>
<td>Appendicitis - unspecified</td>
<td>446</td>
<td>16</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Appendicitis - phlegmonous</td>
<td>-</td>
<td>-</td>
<td>110</td>
<td>4</td>
</tr>
<tr>
<td>Appendicitis - gangrenous</td>
<td>-</td>
<td>-</td>
<td>98</td>
<td>3.4</td>
</tr>
<tr>
<td>Appendicitis - perforated</td>
<td>-</td>
<td>-</td>
<td>69</td>
<td>2.4</td>
</tr>
<tr>
<td>Cholecystitis without perforation</td>
<td>123</td>
<td>4</td>
<td>97</td>
<td>3.4</td>
</tr>
<tr>
<td>Cholecystitis - perforated</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>0.1</td>
</tr>
<tr>
<td>Biliary stone pains</td>
<td>287</td>
<td>10</td>
<td>208</td>
<td>7</td>
</tr>
<tr>
<td>Colon-diverticulitis without perforation</td>
<td>161</td>
<td>6</td>
<td>123</td>
<td>4.3</td>
</tr>
<tr>
<td>Colon-diverticulitis - perforated</td>
<td>-</td>
<td>-</td>
<td>11</td>
<td>0.4</td>
</tr>
<tr>
<td>Obstruction of small intestine without strangulation</td>
<td>80</td>
<td>3</td>
<td>69</td>
<td>2.4</td>
</tr>
<tr>
<td>Obstruction of small intestine with strangulation</td>
<td>2</td>
<td>0.1</td>
<td>9</td>
<td>0.3</td>
</tr>
<tr>
<td>Obstruction of colon</td>
<td>27</td>
<td>1</td>
<td>14</td>
<td>0.5</td>
</tr>
<tr>
<td>Dyspepsia</td>
<td>84</td>
<td>3</td>
<td>60</td>
<td>2.2</td>
</tr>
<tr>
<td>Gastric/duodenal ulcer</td>
<td>53</td>
<td>2</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>Gastric/duodenal ulcer - perforated</td>
<td>14</td>
<td>0.5</td>
<td>8</td>
<td>0.3</td>
</tr>
<tr>
<td>Acute pancreatitis</td>
<td>70</td>
<td>2.5</td>
<td>92</td>
<td>3.2</td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>106</td>
<td>4</td>
<td>74</td>
<td>2.6</td>
</tr>
<tr>
<td>Urinary tract stone</td>
<td>181</td>
<td>6</td>
<td>107</td>
<td>4</td>
</tr>
<tr>
<td>Urinary tract obstruction</td>
<td>8</td>
<td>0.3</td>
<td>10</td>
<td>0.3</td>
</tr>
<tr>
<td>Incarcerated groin hernia</td>
<td>16</td>
<td>0.6</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>Incarcerated umbilical hernia</td>
<td>4</td>
<td>0.1</td>
<td>4</td>
<td>0.1</td>
</tr>
<tr>
<td>Incarcerated incisional hernia</td>
<td>4</td>
<td>0.1</td>
<td>5</td>
<td>0.2</td>
</tr>
<tr>
<td>Abdominal malignancy</td>
<td>22</td>
<td>0.8</td>
<td>63</td>
<td>2.2</td>
</tr>
<tr>
<td>Invagination</td>
<td>2</td>
<td>0.1</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>Aortic aneurysm</td>
<td>14</td>
<td>0.5</td>
<td>12</td>
<td>0.4</td>
</tr>
<tr>
<td>Occlusion of mesenteric artery</td>
<td>6</td>
<td>0.2</td>
<td>4</td>
<td>0.1</td>
</tr>
<tr>
<td>Gynaecological complaint</td>
<td>70</td>
<td>2.5</td>
<td>101</td>
<td>3.5</td>
</tr>
<tr>
<td>Other</td>
<td>128</td>
<td>4.5</td>
<td>199</td>
<td>7</td>
</tr>
</tbody>
</table>
The diagnostic performance was higher for outpatients both during the baseline (65%) and during the study period (63%). Accuracy rate for the preliminary diagnosis at the emergency department was lower (P<0.001) for women (52%, n=1691) than for men (58%, n=1382). The diagnostic performance at the emergency department was independent of the patient’s time of arrival. The rate of correct diagnoses from midnight to 6 a.m. was 52% (n=434), from 3 a.m. to 6 a.m. 52% (n=190), and during the daytime from 6 a.m. to midnight 55% (n=2639). During the study period, the precision of the preliminary diagnoses increased over time, 52% (n=968) in 1997, 55% (n=1090) in 1998, 56% (n=764) in 1999 and 59% (n=251) during 2000. When the second preliminary diagnoses was also included in the calculation of correct diagnoses, the overall accuracy rate increased to 59% 1997, 64% 1998, 63% 1999 and 62% during 2000. The admission frequency was lower for locums (65%, p<0.001), and for general physicians at hospital, (66%, p=0.03), than for pre-registrar house officers (70%, p=0.03), senior house officers (73%, p=0.68) and consultants (70%, p=0.74).

Table 2 Category of attending physician at the emergency department and the proportion of correct diagnoses

<table>
<thead>
<tr>
<th>Diagnoses</th>
<th>Proportion diagnoses</th>
<th>Correct diagnoses</th>
<th>Proportion correct diagnoses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Locum</td>
<td>479</td>
<td>16</td>
<td>277</td>
</tr>
<tr>
<td>Pre-registrar</td>
<td>1409</td>
<td>46</td>
<td>759</td>
</tr>
<tr>
<td>Senior house officer</td>
<td>811</td>
<td>26</td>
<td>443</td>
</tr>
<tr>
<td>Specialist/Consultant</td>
<td>40</td>
<td>1.3</td>
<td>18</td>
</tr>
<tr>
<td>General physician (hospital)</td>
<td>195</td>
<td>6.3</td>
<td>103</td>
</tr>
<tr>
<td>General physician (primary health care)</td>
<td>139</td>
<td>4.5</td>
<td>71</td>
</tr>
</tbody>
</table>

Proportion correct diagnoses between preliminary diagnosis as decided by the physician responsible on admission and final diagnosis after at least one year of follow-up.
To evaluate the diagnostic process for possible pitfalls in the management of AAP, the preliminary diagnosis from the emergency department was compared to the discharge diagnosis. Out of 2851 patients arriving at the emergency department 2052 were admitted to a surgical ward and 799 were treated as outpatients. The frequency of X-ray investigations decided at the emergency department was relatively low, and CT was ordered for only 89 (3%) of the 2851 patients. The corresponding figures were 440 (15%) for plain abdominal X-ray, 526 (18%) for ultrasonography and 139 (5%) for intravenous urography.

Table 3 Accuracy of diagnoses.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>n</th>
<th>Preliminary diagnosis</th>
<th>Discharge diagnosis</th>
<th>kappa</th>
<th>LR+</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSAP</td>
<td>1058</td>
<td>0.43</td>
<td>0.90</td>
<td>0.79</td>
<td>0.92</td>
</tr>
<tr>
<td>Appendicitis</td>
<td>277</td>
<td>0.80</td>
<td>0.91</td>
<td>0.99</td>
<td>0.996</td>
</tr>
<tr>
<td>Gallstones</td>
<td>208</td>
<td>0.68</td>
<td>0.94</td>
<td>0.78</td>
<td>0.98</td>
</tr>
<tr>
<td>Diverticulitis</td>
<td>134</td>
<td>0.64</td>
<td>0.97</td>
<td>0.83</td>
<td>0.98</td>
</tr>
<tr>
<td>Constipation</td>
<td>130</td>
<td>0.74</td>
<td>0.96</td>
<td>0.95</td>
<td>0.99</td>
</tr>
<tr>
<td>Urteric stone</td>
<td>107</td>
<td>0.78</td>
<td>0.96</td>
<td>0.96</td>
<td>0.99</td>
</tr>
<tr>
<td>Cholecystitis</td>
<td>100</td>
<td>0.51</td>
<td>0.97</td>
<td>0.79</td>
<td>0.99</td>
</tr>
<tr>
<td>Gyn. diagnoses</td>
<td>101</td>
<td>0.33</td>
<td>0.99</td>
<td>0.75</td>
<td>0.999</td>
</tr>
<tr>
<td>Pancreatitis</td>
<td>92</td>
<td>0.44</td>
<td>0.99</td>
<td>0.84</td>
<td>0.997</td>
</tr>
<tr>
<td>Intestin. obstr.</td>
<td>78</td>
<td>0.53</td>
<td>0.98</td>
<td>0.83</td>
<td>0.998</td>
</tr>
<tr>
<td>Gastroenteritis</td>
<td>64</td>
<td>0.69</td>
<td>0.98</td>
<td>0.88</td>
<td>0.99</td>
</tr>
<tr>
<td>Urinary infect.</td>
<td>74</td>
<td>0.55</td>
<td>0.98</td>
<td>0.77</td>
<td>0.99</td>
</tr>
<tr>
<td>Dyspepsia</td>
<td>60</td>
<td>0.77</td>
<td>0.99</td>
<td>0.73</td>
<td>0.998</td>
</tr>
<tr>
<td>Malignancy</td>
<td>63</td>
<td>0.14</td>
<td>0.99</td>
<td>0.40</td>
<td>0.997</td>
</tr>
<tr>
<td>Peptic ulcer</td>
<td>34</td>
<td>0.26</td>
<td>0.98</td>
<td>0.87</td>
<td>1.0</td>
</tr>
<tr>
<td>Incarcer. hernia</td>
<td>22</td>
<td>0.64</td>
<td>0.999</td>
<td>0.85</td>
<td>1.0</td>
</tr>
<tr>
<td>Aortic aneurysm</td>
<td>12</td>
<td>0.58</td>
<td>0.997</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Colonic obstr.</td>
<td>14</td>
<td>0.28</td>
<td>0.99</td>
<td>0.71</td>
<td>0.999</td>
</tr>
<tr>
<td>Mesent. occlus.</td>
<td>4</td>
<td>0.5</td>
<td>0.999</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Other diag.</td>
<td>199</td>
<td>0.30</td>
<td>0.97</td>
<td>0.81</td>
<td>0.96</td>
</tr>
</tbody>
</table>

The calculations of sensitivity and specificity for the preliminary diagnoses are based on the 2851 patients who attended the emergency department. The kappavalue is calculated between the preliminary and discharge diagnoses and is based only on the 2052 hospitalised patients. Positive likelihood ratio (LR+) is calculated on the preliminary diagnosis. Other diagnoses includes for example mesenteric lymphadenitis, inflammatory bowel disease, intra abdominal abscess, septicemia, pneumonia, pleuritis, myocardial or renal infarction, debut of diabetes mellitus, torsion and necrosis of the omentum majus, muscular back pain and addiction to analgesics.

On the basis of table 3, the five most characteristic diagnoses were selected for each of the groups described below.
**Group 1 – low predictive value of preliminary diagnosis**

Belonging to this group were diagnoses with low predictive value of the primary diagnosis at the emergency department, but with improved predictive value at discharge; these included NSAP, appendicitis, gallstones and constipation. Among the 208 patients initially diagnosed with constipation, 17 (8%) were later found to have an abdominal malignancy and 20 (10%) a diagnosis probably necessitating surgery (one patient with aortic aneurysm, one with incarcerated inguinal hernia, five patients with colonic obstruction, four with obstruction of the small intestine and nine with appendicitis). Furthermore, peptic ulcer, in spite of having a high *kappa* value and a sensitivity of 0.87 at discharge, indicated low sensitivity of the primary diagnosis. This low sensitivity partly emanated from cases treated as outpatients among whom no true positive cases were determined (none of those cases were confirmed at endoscopy). Those diagnoses together with the frequency of incorrect primary diagnoses are listed in Table 4. Calculation of LR+ revealed a low value for NSAP and moderate value appendicitis; all other diagnoses had high LR+.

Patients considered as suffering from NSAP at follow-up (*n* = 1058) had a median hospital stay of 1 day, and a total number of in-hospital days during the observation period (40 months) of 1338. Calculated on 87,000 inhabitants and a completeness of the register of 90%, this meant 43 in-hospital days per month between 100,000 inhabitants.

<table>
<thead>
<tr>
<th>Preliminary diagnosis</th>
<th>NSAP <em>n</em>=1058</th>
<th>Appendicitis <em>n</em>=277</th>
<th>Constipation <em>n</em>=130</th>
<th>Gallstones <em>n</em>=208</th>
<th>Colonic obstr. <em>n</em>=14</th>
<th>Peptic ulcer <em>n</em>=34</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSAP</td>
<td>458 (43%)</td>
<td>24 (9%)</td>
<td>14 (11%)</td>
<td>19 (9%)</td>
<td>1 (7%)</td>
<td>4 (12%)</td>
</tr>
<tr>
<td>Appendicitis</td>
<td>132 (12%)</td>
<td>222 (80%)</td>
<td>3 (2%)</td>
<td>3 (1%)</td>
<td>1 (7%)</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Gallstones</td>
<td>75 (7%)</td>
<td>2 (0.5%)</td>
<td>1 (1%)</td>
<td>141 (68%)</td>
<td>0 (0%)</td>
<td>6 (18%)</td>
</tr>
<tr>
<td>Ureteric stone</td>
<td>69 (7%)</td>
<td>1 (0.5%)</td>
<td>0 (0%)</td>
<td>3 (1%)</td>
<td>0 (0%)</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Constipation</td>
<td>56 (5%)</td>
<td>9 (3%)</td>
<td>96 (74%)</td>
<td>0 (0%)</td>
<td>3 (21%)</td>
<td>3 (9%)</td>
</tr>
<tr>
<td>Diverticulitis</td>
<td>43 (4%)</td>
<td>4 (1.5%)</td>
<td>1 (1%)</td>
<td>1 (1%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Gastroenteritis</td>
<td>36 (3%)</td>
<td>5 (2%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>33 (3%)</td>
<td>2 (0.5%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Gynaecological complaint</td>
<td>27 (3%)</td>
<td>2 (0.5%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Dyspepsia</td>
<td>20 (2%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>6 (3%)</td>
<td>0 (0%)</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Peptic ulcer</td>
<td>20 (2%)</td>
<td>1 (0.5%)</td>
<td>0 (0%)</td>
<td>7 (3%)</td>
<td>0 (0%)</td>
<td>9 (26%)</td>
</tr>
<tr>
<td>Colonic obstruction</td>
<td>4 (1%)</td>
<td>1 (0.5%)</td>
<td>3 (2%)</td>
<td>0 (0%)</td>
<td>4 (29%)</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Intestinal obstruction</td>
<td>13 (1%)</td>
<td>0 (0%)</td>
<td>5 (4%)</td>
<td>1 (1%)</td>
<td>4 (29%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Other diagnoses</td>
<td>72 (7%)</td>
<td>4 (1.5%)</td>
<td>7 (5%)</td>
<td>27 (13%)</td>
<td>1 (7%)</td>
<td>6 (17%)</td>
</tr>
</tbody>
</table>

Numbers within parentheses are percentages of cases with each final diagnosis.
Group 2 – predictive value still low at discharge

The final diagnoses of malignancy, gynaecological complaints, dyspepsia, urinary tract infection and diverticulitis had good concordance between the preliminary and discharge diagnosis, but the predictive value at discharge was low compared to the final diagnosis at follow-up. In addition to malignancies identified during the hospital stay, abdominal malignancies were detected in another 27 patients at follow-up. Furthermore, despite considerable improvement in validity during hospital-stay, the judgement that NSAP was still the reason for the patients’ complaints at discharge had low sensitivity.

Surgery

On evaluation of patients judged as needing surgery during the hospital stay, 104 patients had an initial diagnosis usually not necessitating surgery (Table 5): these patients constituted 22% of the total 479 (23% of patients admitted to a ward) undergoing operations. In six of these operations no cause was determined for the abdominal pain and the complaint was thus classified as NSAP. As a comparison, 12 (3%) of the 375 patients operated on who had a preliminary diagnosis necessitating surgery were finally diagnosed with NSAP; whereas, a gynaecological reason for the pain was determined during 16 (4%) of these operations. Patients given a preliminary diagnosis of NSAP (n=42), constipation (n=19), other diagnoses (n=13), pancreatitis (n=6) or ureteric stone (n=6) had significantly longer time before surgery, compared with the surgically-treated patients in whom the final diagnosis was the same as the preliminary diagnosis. Patients with an initial diagnosis not necessitating surgery (n=104) had a median delay of 22 hours before operation (mean was 40 hours, 95% confidence interval 30-50 hours), compared to 8 hours (mean was 15 hours, 95% confidence interval 12-18 hours) for patients with the same final diagnosis at follow-up as for the preliminary diagnosis (n=266).
Table 5 Diagnosis at surgery and diagnosis at discharge for 104 surgically-treated patients, who were given a preliminary diagnosis usually not necessitating surgery.

<table>
<thead>
<tr>
<th>Preliminary diagnosis</th>
<th>Appendicitis</th>
<th>Intestinal obstruct.</th>
<th>Diverticulitis</th>
<th>Incarc. hernia</th>
<th>Gyn. Diagn.</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSAP (n=42)</td>
<td>24</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>8</td>
<td>42</td>
</tr>
<tr>
<td>Constipation</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>19</td>
</tr>
<tr>
<td>Gastroenteritis</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Pancreatitis</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Ureteric stone</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Gyn. diagnoses</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Dyspepsia/peptic ulcer</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Other diagnoses</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>13</td>
</tr>
</tbody>
</table>
Paper III

Of the patients admitted to the emergency department, with AAP, 2289 were above 20 years of age. Among these, 831 were 65 years of age or older and constituted the total study group, which was divided into one group of 65-79 years old (n=557) and one group aged 80 years or older (n=274). Patients aged 20 to 64 years (n=1458) served as a reference group. There was a higher proportion of women (Table 6) among patients 20-64 years and ≥ 80 years of age than in the group aged 65-79 years. The frequency of surgery among all patients was independent of age group. NSAP was the most common preliminary diagnosis with no significant difference between the reference group and the two study groups (p=0.3). The proportion of NSAP in each group was similar at discharge (Table 6), whereas at follow-up (Table 7) 64% (n=934) of the patients in the reference group were assigned a specific diagnosis, compared with 76% (n=421) of the 65-79 years group and 78% (n=214) of the ≥ 80 years group (p<0.001).

Duration of pain before admission was related to age (p<0.003; reference versus the total study group; Table 6) and there was no difference in pain duration between the two study groups (p=0.9). The frequency of hospitalisation was higher in the both study groups (≥65 years); the group aged 65-79 years (459 of 557 patients, 82%; p<0.0001) and the group of oldest patients (241 of 274 patients, 88%; p<0.0001), than in the reference group (996 of 1458 patients, 68%; Table 6). The calculated duration of hospital stay was 170 days per 100 emergency admissions for the reference group; whereas in the two study groups the duration was 320 days for the 65-79 years group and 458 days for the ≥80 years group. The postoperative stay was significantly longer for the study groups (7.6 days for 65-79 years and 10.7 days for ≥80 years) than for the reference group (3.7 days, p<0.0001; Table 8).

Older patients (the total study group) were more often misdiagnosed (429 of 831 patients, 52%) at the emergency department than reference patients were (656 of 1458 patients, 45%; p=0.002), but there was no difference between the two study groups (p=0.06; Table 6). At discharge, the diagnosis was more accurate (853 of 996 hospitalised patients, 86%) in the reference group (20-64 years) than in the older patients (539 of 700 hospitalised patients, 77%; p<0.0001). There was no difference between the two study groups aged ≥65 years (78% for 65-79 years and 75% for ≥ 80 years, p=0.3; Table 6). Hospital mortality increased with age (p<0.0001; Tables 6) when the reference group (2 deaths) was compared with the two study groups (23 deaths), and there was a difference in mortality between the two older study
groups (10 patients aged 65-75 years and 13 patients aged ≥ 80 years; p=0.02). However, among patients older than 65 there was no difference in mortality between those diagnosed correctly at the emergency department and those misdiagnosed in that department (p=0.9). Furthermore, there was no relation between pre-admission duration of pain and mortality (p=0.9).

Table 6 Basic demographic data and characteristics

<table>
<thead>
<tr>
<th>Age</th>
<th>20-64 years</th>
<th>65-79 years</th>
<th>≥80 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>1458</td>
<td>557</td>
<td>274</td>
</tr>
<tr>
<td>Male/Female ratio</td>
<td>631/827 (0.8)</td>
<td>296/261 (1.1)</td>
<td>114/160 (0.7)</td>
</tr>
<tr>
<td>NSAP – preliminary diagnosis</td>
<td>287 (20%)</td>
<td>92 (16%)</td>
<td>56 (20%)</td>
</tr>
<tr>
<td>NSAP – discharge diagnosis</td>
<td>298 (20%)</td>
<td>107 (19%)</td>
<td>49 (18%)</td>
</tr>
<tr>
<td>Surgery</td>
<td>243 (17%)</td>
<td>77 (14%)</td>
<td>38 (14%)</td>
</tr>
<tr>
<td>Hospital mortality</td>
<td>2 (0.1%)</td>
<td>10 (1.8%)</td>
<td>13 (4.7%)</td>
</tr>
<tr>
<td>Duration of pain before admission (days)</td>
<td>1.4 (1.3-1.5)</td>
<td>1.7 (1.5-1.9)</td>
<td>1.7 (1.4-1.9)</td>
</tr>
<tr>
<td>Hospitalised patients</td>
<td>996 (68%)</td>
<td>459 (82%)</td>
<td>241 (88%)</td>
</tr>
<tr>
<td>Hospital stay (days)</td>
<td>2.5(2.3-2.7)</td>
<td>3.9 (3.5-4.2)</td>
<td>5.2 (4.6-5.9)</td>
</tr>
<tr>
<td>Correct preliminary diagnosis</td>
<td>802 (55%)</td>
<td>282 (51%)</td>
<td>120 (44%)</td>
</tr>
<tr>
<td>Correct discharge diagnosis</td>
<td>853 (86%)</td>
<td>359 (78%)</td>
<td>180 (75%)</td>
</tr>
</tbody>
</table>

NSAP – Non-specific abdominal pain. Preliminary diagnosis is the diagnosis at the emergency department. Numbers within parentheses are 95% confidence intervals unless otherwise stated.

There was no difference in duration of abdominal pain before admission between the two study groups or between patients being operated and those not being operated. There was no difference between the two separate study groups in the interval between admission and surgery (1.7 for 65-79 years and 2.1 days for ≥ 80 years; p=0.5). However, the interval was considerably longer for the total study group (1.8 days) than for the reference group (0.9 days, p=0.0001; Table 8).

A CT scan was ordered for 34 patients (2%) in the reference group and for 33 (6%) in the 65-79 years group and 20 (7%) in the ≥80 years group: the corresponding figures for ultrasonography were 263 (18%) for the reference group, 162 (29%) for the 65-79 years group and 77 (28%) for the ≥80 years group.
Symptoms, signs and laboratory examinations

Patients who were under 65 years of age and submitted to surgery generally had higher levels of CRP on admission than those who were not submitted to surgery did (44 for operated and 27 for non-operated; p<0.0001). In contrast, among patients older than 65, there was no significant difference in CRP between patients undergoing surgery (CRP 60) and those who did not (CRP 45; p=0.12). However, there was a difference in leucocyte counts, irrespective of age group, between patients treated surgically and those not treated surgically. Patients aged 20-64 years who were surgically treated had a leucocyte count of 13 versus 10 for patients who were not surgically treated and for patients aged over 65 the leucocyte count was 13 versus 11 if surgically-treated or not (p<0.001). Body temperature among surgically treated patients was independent of age, although it was higher among surgically treated patients in the reference group (37.6 vs 37.2°C; p<0.0001) than in the study groups (37.4 vs 37.2°C; p=0.06).

In a logistic regression analysis on diagnoses usually known to induce an intra-abdominal inflammatory response, clinical signs such as rebound tenderness (p<0.0001), local rigidity (p=0.003) and rectal tenderness (p=0.004) were less common with increasing age. Nevertheless, there was no age-related difference in the distribution between general and localised abdominal tenderness on clinical investigation in the emergency department. Age in itself did not correlate to symptoms such as vomiting (p=0.4), constipation (p=0.7) or diarrhoea (p=0.7, Table 8).

---

Table 7 Distribution of common final diagnoses at follow-up in the two study groups (≥65 years) and in the reference group (20-64 years)

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>20-64 years</th>
<th>65-79 years</th>
<th>≥80 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>1458</td>
<td>557</td>
<td>274</td>
</tr>
<tr>
<td>NSAP</td>
<td>524 (36%)</td>
<td>136 (24%)</td>
<td>60 (22%)</td>
</tr>
<tr>
<td>Appendicitis</td>
<td>159 (11%)</td>
<td>15 (3%)</td>
<td>7 (3%)</td>
</tr>
<tr>
<td>Biliary colic</td>
<td>138 (9%)</td>
<td>42 (8%)</td>
<td>22 (8%)</td>
</tr>
<tr>
<td>Cholecystitis</td>
<td>34 (2%)</td>
<td>49 (9%)</td>
<td>17 (6%)</td>
</tr>
<tr>
<td>Gynaecological complaints</td>
<td>82 (6%)</td>
<td>3 (1%)</td>
<td>3 (1%)</td>
</tr>
<tr>
<td>Constipation</td>
<td>31 (2%)</td>
<td>36 (5%)</td>
<td>26 (9%)</td>
</tr>
<tr>
<td>Ureteric stone</td>
<td>82 (6%)</td>
<td>22 (4%)</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Diverticulitis</td>
<td>75 (5%)</td>
<td>48 (9%)</td>
<td>11 (4%)</td>
</tr>
<tr>
<td>Pancreatitis</td>
<td>56 (4%)</td>
<td>24 (4%)</td>
<td>12 (4%)</td>
</tr>
<tr>
<td>Other diagnoses</td>
<td>277 (19%)</td>
<td>182 (33%)</td>
<td>115 (42%)</td>
</tr>
</tbody>
</table>

NSAP – Non-specific abdominal pain.
Table 8 Characteristics of patients requiring surgery

<table>
<thead>
<tr>
<th></th>
<th>20-64 years</th>
<th>65-79 years</th>
<th>≥80 years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n</strong></td>
<td>243</td>
<td>77</td>
<td>38</td>
</tr>
<tr>
<td><strong>Pre-admission pain duration (days)</strong></td>
<td>1.2 (1.0-1.4)</td>
<td>1.7 (1.2-2.1)</td>
<td>1.5 (1.0-1.9)</td>
</tr>
<tr>
<td><strong>Time from admission to surgery (days)</strong></td>
<td>0.9 (0.7-1.0)</td>
<td>1.7 (1.2-2.1)</td>
<td>2.1 (1.2-3.1)</td>
</tr>
<tr>
<td><strong>Body temperature (°C)</strong></td>
<td>37.6 (37.5-37.7)</td>
<td>37.4 (37.2-37.6)</td>
<td>37.6 (37.3-38.0)</td>
</tr>
<tr>
<td><strong>Leucocytes</strong></td>
<td>13.5 (12.9-14.1)</td>
<td>13.1 (12.0-14.3)</td>
<td>12.4 (9.6-15.3)</td>
</tr>
<tr>
<td><strong>CRP</strong></td>
<td>44 (37-50)</td>
<td>62 (40-83)</td>
<td>57 (34-81)</td>
</tr>
<tr>
<td><strong>General abdominal tenderness</strong></td>
<td>36 (15%)</td>
<td>19 (25%)</td>
<td>15 (39%)</td>
</tr>
<tr>
<td><strong>Rebound tenderness</strong></td>
<td>150 (62%)</td>
<td>27 (35%)</td>
<td>11 (29%)</td>
</tr>
<tr>
<td><strong>Rectal tenderness</strong></td>
<td>78 (32%)</td>
<td>13 (17%)</td>
<td>7 (18%)</td>
</tr>
<tr>
<td><strong>Abdominal rigidity</strong></td>
<td>104 (43%)</td>
<td>27 (35%)</td>
<td>13 (34%)</td>
</tr>
<tr>
<td><strong>Vomiting</strong></td>
<td>97 (40%)</td>
<td>37 (48%)</td>
<td>22 (58%)</td>
</tr>
<tr>
<td><strong>Constipation</strong></td>
<td>27 (11%)</td>
<td>20 (26%)</td>
<td>10 (26%)</td>
</tr>
<tr>
<td><strong>Diarrhoea</strong></td>
<td>40 (16%)</td>
<td>10 (13%)</td>
<td>8 (21%)</td>
</tr>
<tr>
<td><strong>VAS</strong></td>
<td>6.4 (6.1-6.8)</td>
<td>6 (5.2-6.9)</td>
<td>5 (3.4-6.6)</td>
</tr>
<tr>
<td><strong>Hospital stay (days)</strong></td>
<td>3.7 (3.3-4.1)</td>
<td>7.6 (6.4-8.8)</td>
<td>10.7 (8.4-13.0)</td>
</tr>
<tr>
<td><strong>Correct preliminary diagnosis</strong></td>
<td>149 (61%)</td>
<td>36 (47%)</td>
<td>19 (50%)</td>
</tr>
</tbody>
</table>

CRP – C-reactive protein. VAS – Visual Analogue Scale, used here as an instrument for the patient to estimate the severity of pain. Numbers within parentheses are 95% confidence intervals unless otherwise stated.
Paper IV

Differential diagnostics and diagnostic outcome

Acute appendicitis was suspected in 432 patients out of the 2478 patients admitted with AAP. Of these, 221 patients were submitted to surgery and appendicitis was confirmed. In 53 patients eventually diagnosed as having appendicitis, a different preliminary diagnosis was suggested at the emergency department. Consequently, 274 patients had confirmed appendicitis; 271 had their appendix surgically removed and three had an appendiceal abscess diagnosed by CT or ultrasonography and were treated with only antibiotics. None of the patients with an abscess had an appendectomy within the follow-up period (at least one year). Advanced appendicitis ( gangrenous or perforated) was identified in 165 patients (60%) and the perforation rate was 25% (n=68).

Of 432 patients with suspected appendicitis, 423 (98%) were admitted to the surgical ward, two were referred to the gynaecological department (one had diverticulitis and one had torsion of the right ovary) and seven patients were sent home but agreed to return to the emergency department the next morning (none of these patients were eventually diagnosed as having appendicitis). One hundred and sixty patients with a preliminary diagnosis of appendicitis were not surgically treated, as the pain had disappeared (111 patients) or another diagnosis was confirmed (49 patients). At follow-up, 211 of the 432 patients with appendicitis as a preliminary diagnosis had a different final diagnosis, these were: 122 patients suffering from NSAP; 24 with gynaecological complaints; 13 with diverticulitis (2 with perforation); 10 with mesenteric lymphadenitis; 8 with urinary infection; 7 with biliary stone disease; 5 with gastroenteritis; 4 with ureteric stone; 3 with constipation; 3 with upper respiratory tract infection; 2 with intestinal obstruction; 2 abdominal malignancies; 2 torsion of colonic epiploicae; 1 with aortic dissection; 1 with perforated gastric ulcer; 1 dyspepsia; 1 with Meckel’s diverticulitis; 1 with Crohn’s disease, and 1 with pneumonia. Of these 211 patients, 53 underwent surgery and among 316 appendectomies the result was negative in 45 cases (14%). Of the 432 patients with suspected appendicitis, preoperative CT scan was performed in only six cases and ultrasound in 36 cases.

The preoperative clinical diagnosis of appendicitis had a sensitivity of 0.81, specificity of 0.90, PPV of 0.51, LR+ of 8.1, diagnostic accuracy of 0.89 and a kappa value of 0.78. The incidence rate of patients treated surgically for appendicitis was 102 per year and 100,000 inhabitants.
Gender and age differences

Males predominantly suffered from appendicitis (57%), whereas, females (57%; p<0.001) were predominant among patients with AAP but without appendicitis. Patients with appendicitis were generally younger (mean age of 32) than patients without appendicitis (mean ages of 47; p<0.001).

Clinical symptoms

Upper gastrointestinal complaints, such as loss of appetite (anorexia), nausea and vomiting, were more common in the appendicitis group (p<0.001), compared to other patients with AAP. There was no difference in duration of pain between the two groups (p=0.95). In appendicitis patients, the pain often had a gradual onset and improvement of the pain was rare. Furthermore, continuous pain and dull pain were common, and pain-free intervals, fluctuating pain and colic pain were not. Migration of pain to the right iliac fossa was recorded in 16% of appendicitis patients and previous episodes of similar pain in 15%. Pain migration to the right iliac fossa was uncommon among patients not diagnosed with appendicitis (3%; p<0.001), whereas previous episodes of pain were more common in this group (37%; p<0.001, Table 9).

Table 9 Differences in characteristics of abdominal pain between appendicitis patients and other patients with AAP.

<table>
<thead>
<tr>
<th>Symptoms</th>
<th>n (%)</th>
<th>Sens</th>
<th>Spec</th>
<th>Acc</th>
<th>PPV</th>
<th>NPV</th>
<th>LR+</th>
<th>LR-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain duration (mean)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>274</td>
<td>2204</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of pain &gt;48 h</td>
<td>27 (24-30)</td>
<td>36 (34-38)</td>
<td>0.22</td>
<td>0.68</td>
<td>0.63</td>
<td>0.08</td>
<td>0.87</td>
<td>0.69</td>
</tr>
<tr>
<td>Rapid onset of pain</td>
<td>60 (22)</td>
<td>711 (33)</td>
<td>0.27</td>
<td>0.67</td>
<td>0.63</td>
<td>0.09</td>
<td>0.88</td>
<td>0.82</td>
</tr>
<tr>
<td>Gradual onset of pain</td>
<td>75 (27)</td>
<td>729 (33)</td>
<td>0.67</td>
<td>0.42</td>
<td>0.44</td>
<td>0.12</td>
<td>0.91</td>
<td>1.16</td>
</tr>
<tr>
<td>Improvement of pain</td>
<td>26 (9)</td>
<td>497 (23)</td>
<td>0.09</td>
<td>0.77</td>
<td>0.70</td>
<td>0.05</td>
<td>0.87</td>
<td>0.39</td>
</tr>
<tr>
<td>No change in pain intensity</td>
<td>61 (22)</td>
<td>727 (33)</td>
<td>0.22</td>
<td>0.67</td>
<td>0.62</td>
<td>0.08</td>
<td>0.87</td>
<td>0.67</td>
</tr>
<tr>
<td>Aggravation of pain</td>
<td>176 (64)</td>
<td>845 (38)</td>
<td>0.64</td>
<td>0.62</td>
<td>0.62</td>
<td>0.17</td>
<td>0.93</td>
<td>1.68</td>
</tr>
<tr>
<td>Pain-free intervals</td>
<td>32 (12)</td>
<td>648 (29)</td>
<td>0.12</td>
<td>0.71</td>
<td>0.64</td>
<td>0.05</td>
<td>0.87</td>
<td>0.41</td>
</tr>
<tr>
<td>Fluctuating pain</td>
<td>81 (30)</td>
<td>1056 (48)</td>
<td>0.30</td>
<td>0.52</td>
<td>0.50</td>
<td>0.07</td>
<td>0.86</td>
<td>0.62</td>
</tr>
<tr>
<td>Continuous pain</td>
<td>184 (67)</td>
<td>1016 (46)</td>
<td>0.67</td>
<td>0.54</td>
<td>0.55</td>
<td>0.15</td>
<td>0.93</td>
<td>1.46</td>
</tr>
<tr>
<td>Burning pain</td>
<td>27 (10)</td>
<td>245 (11)</td>
<td>0.10</td>
<td>0.89</td>
<td>0.80</td>
<td>0.10</td>
<td>0.89</td>
<td>0.91</td>
</tr>
<tr>
<td>Colic pain</td>
<td>48 (18)</td>
<td>699 (32)</td>
<td>0.17</td>
<td>0.68</td>
<td>0.63</td>
<td>0.06</td>
<td>0.87</td>
<td>0.53</td>
</tr>
<tr>
<td>Dull pain</td>
<td>161 (59)</td>
<td>931 (42)</td>
<td>0.59</td>
<td>0.58</td>
<td>0.58</td>
<td>0.15</td>
<td>0.92</td>
<td>1.40</td>
</tr>
<tr>
<td>Aggr. of pain; moving</td>
<td>170 (62)</td>
<td>646 (29)</td>
<td>0.62</td>
<td>0.71</td>
<td>0.70</td>
<td>0.21</td>
<td>0.94</td>
<td>2.14</td>
</tr>
<tr>
<td>Aggr. of pain; coughing</td>
<td>120 (44)</td>
<td>337 (15)</td>
<td>0.44</td>
<td>0.85</td>
<td>0.80</td>
<td>0.26</td>
<td>0.92</td>
<td>2.93</td>
</tr>
<tr>
<td>Allev. of pain; not mov.</td>
<td>164 (60)</td>
<td>652 (30)</td>
<td>0.60</td>
<td>0.70</td>
<td>0.69</td>
<td>0.20</td>
<td>0.93</td>
<td>2.00</td>
</tr>
<tr>
<td>Migr. of pain to right IF</td>
<td>43 (16)</td>
<td>62 (3)</td>
<td>0.16</td>
<td>0.97</td>
<td>0.88</td>
<td>0.41</td>
<td>0.90</td>
<td>5.33</td>
</tr>
<tr>
<td>VAS (mean)</td>
<td>2.9 (2.5-3)</td>
<td>2.9 (2.8-3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prev. ep. of similar pain</td>
<td>42 (15)</td>
<td>819 (37)</td>
<td>0.15</td>
<td>0.63</td>
<td>0.58</td>
<td>0.05</td>
<td>0.86</td>
<td>0.41</td>
</tr>
</tbody>
</table>

1 Before seeking medical advice.
Clinical findings

Clinical signs of peritonitis such as aggravation of pain when moving or coughing and alleviation when not moving were more common (p<0.001) in appendicitis patients. There was no difference in the severity of the pain, between the appendicitis patients and other patients with AAP (Table 9). Isolated tenderness in the right iliac fossa and rebound tenderness were more frequent (p<0.001, Table 10) among appendicitis patients, as was right-sided rectal tenderness (p<0.001, Table 10). No patient with appendicitis had tenderness located only in the left iliac fossa. On abdominal examination, patients younger than 20 years had lower LR+ values for signs of appendicitis on physical examination, with the exception of the muscle response parameters (local guarding and general rigidity), than patients over 50 years of age had. When testing the same symptoms and signs according to gender, there were no differences in LR+, except for right-sided rectal tenderness and pain migration: males had a higher LR+ (6.5 and 5.3) than in females did (4.7 and 3.7). There were no differences in the degree of inflammation, except a higher LR+ for general rigidity in perforated (7.5) compared to phlegmonous (1.5) and gangrenous appendicitis (1.0).

Table 10 Differences in clinical findings from the physical examination on admission between patients with appendicitis and other patients with acute abdominal pain.

<table>
<thead>
<tr>
<th>Symptoms</th>
<th>n (%)</th>
<th>n (%)</th>
<th>Sens App</th>
<th>Spec App</th>
<th>Acc App</th>
<th>PPV App</th>
<th>NPV App</th>
<th>LR+ App</th>
<th>LR- App</th>
</tr>
</thead>
<tbody>
<tr>
<td>App</td>
<td>274</td>
<td>2204</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tend. in right IF only</td>
<td>194 (71)</td>
<td>403 (18)</td>
<td>0.70</td>
<td>0.82</td>
<td>0.80</td>
<td>0.32</td>
<td>0.96</td>
<td>3.89</td>
<td>0.37</td>
</tr>
<tr>
<td>Tend. in left IF only</td>
<td>0 (0)</td>
<td>220 (10)</td>
<td>0.13</td>
<td>0.81</td>
<td>0.73</td>
<td>0.08</td>
<td>0.88</td>
<td>0.06</td>
<td>1.07</td>
</tr>
<tr>
<td>General abdom. tend.</td>
<td>36 (13)</td>
<td>422 (19)</td>
<td>0.13</td>
<td>0.81</td>
<td>0.73</td>
<td>0.08</td>
<td>0.88</td>
<td>0.06</td>
<td>1.07</td>
</tr>
<tr>
<td>Rebound tenderness</td>
<td>201 (73)</td>
<td>516 (23)</td>
<td>0.73</td>
<td>0.77</td>
<td>0.76</td>
<td>0.28</td>
<td>0.96</td>
<td>3.17</td>
<td>0.35</td>
</tr>
<tr>
<td>Local guarding</td>
<td>107 (39)</td>
<td>277 (13)</td>
<td>0.39</td>
<td>0.87</td>
<td>0.82</td>
<td>0.28</td>
<td>0.92</td>
<td>3.00</td>
<td>0.70</td>
</tr>
<tr>
<td>General rigidity</td>
<td>15 (5)</td>
<td>39 (2)</td>
<td>0.05</td>
<td>0.98</td>
<td>0.88</td>
<td>0.28</td>
<td>0.89</td>
<td>2.50</td>
<td>0.97</td>
</tr>
<tr>
<td>Right-sided rectal tend.</td>
<td>74 (27)</td>
<td>120 (5)</td>
<td>0.27</td>
<td>0.95</td>
<td>0.87</td>
<td>0.38</td>
<td>0.91</td>
<td>5.40</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Diagnostic markers for appendicitis

The highest OR (univariate) was found for isolated tenderness in the right iliac fossa (3.29), followed by rebound tenderness (3.00), right-sided rectal tenderness (2.53), migration of pain to the right iliac fossa (2.18), local guarding (2.11), aggravation of pain on movement (1.99), and general rigidity (1.79). In the multivariate analysis these variables still had the highest ORs (Table 11).

Table 11 Multivariate analysis of the diagnostic value of symptoms and signs for diagnosis of appendicitis.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Univariate analyses Odds Ratio ± 95% CI</th>
<th>Multivariate analyses Odds Ratio ± 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upper gastro-intestinal symptoms:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anorexia</td>
<td>1.51</td>
<td>1.37</td>
</tr>
<tr>
<td>Nausea</td>
<td>1.56</td>
<td>1.41</td>
</tr>
<tr>
<td><strong>Pain characteristics:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gradual onset of pain</td>
<td>1.20</td>
<td>1.02</td>
</tr>
<tr>
<td>Impairment of pain</td>
<td>1.70</td>
<td>1.59</td>
</tr>
<tr>
<td>Continuous pain</td>
<td>1.54</td>
<td>1.45</td>
</tr>
<tr>
<td><strong>Signs of localised inflammation:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggravation of pain on movement</td>
<td>1.99</td>
<td>1.70</td>
</tr>
<tr>
<td>Migration of pain to right iliac fossa</td>
<td>2.18</td>
<td>2.02</td>
</tr>
<tr>
<td><strong>Findings at abdominal examination:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isolated tenderness in right iliac fossa</td>
<td>3.29</td>
<td>2.96</td>
</tr>
<tr>
<td>Rebound tenderness</td>
<td>3.00</td>
<td>2.45</td>
</tr>
<tr>
<td>General rigidity</td>
<td>1.79</td>
<td>1.92</td>
</tr>
<tr>
<td>Local guarding</td>
<td>2.11</td>
<td>1.81</td>
</tr>
<tr>
<td>Right-sided rectal tenderness</td>
<td>2.53</td>
<td>2.19</td>
</tr>
</tbody>
</table>

In the multivariate analyses the factors were adjusted for age and gender. Each group was processed in separate models, except for the last group, which was split into three groups: isolated tenderness in right iliac fossa and rebound tenderness; general rigidity; local guarding and right-sided tenderness.
A combination of the three most frequently recorded clinical findings with high LR+, namely isolated tenderness in the right iliac fossa, rebound tenderness and aggravation of pain when moving, gave an LR+ of 9.5 (Table 12).

Table 12 Clinical symptoms and signs diagnostically important for appendicitis, in combination with the two most common clinical findings: isolated tenderness in the right iliac fossa (=1) and rebound tenderness (=2).

<table>
<thead>
<tr>
<th>Symptoms</th>
<th>n (%)</th>
<th>n (%)</th>
<th>p value</th>
<th>Sens App</th>
<th>Spec App</th>
<th>Acc App</th>
<th>PPV App</th>
<th>LR+ App</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>274</td>
<td>2204</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,2</td>
<td>149 (54)</td>
<td>155 (7)</td>
<td>&lt;0.001</td>
<td>0.54</td>
<td>0.93</td>
<td>0.89</td>
<td>0.49</td>
<td>7.7</td>
</tr>
<tr>
<td>1,2, aggr. of pain; moving</td>
<td>103 (38)</td>
<td>92 (4)</td>
<td>&lt;0.001</td>
<td>0.38</td>
<td>0.96</td>
<td>0.89</td>
<td>0.53</td>
<td>9.5</td>
</tr>
<tr>
<td>1,2, rectal tenderness</td>
<td>48 (18)</td>
<td>41 (2)</td>
<td>&lt;0.001</td>
<td>0.18</td>
<td>0.98</td>
<td>0.89</td>
<td>0.54</td>
<td>9.0</td>
</tr>
<tr>
<td>1,2, migration of pain</td>
<td>34 (12)</td>
<td>28 (1)</td>
<td>&lt;0.001</td>
<td>0.12</td>
<td>0.99</td>
<td>0.89</td>
<td>0.55</td>
<td>12.0</td>
</tr>
<tr>
<td>1,2, local guarding</td>
<td>68 (25)</td>
<td>36 (2)</td>
<td>&lt;0.001</td>
<td>0.25</td>
<td>0.98</td>
<td>0.90</td>
<td>0.65</td>
<td>12.5</td>
</tr>
<tr>
<td>1,2, anorexia</td>
<td>116 (42)</td>
<td>86 (4)</td>
<td>&lt;0.001</td>
<td>0.42</td>
<td>0.96</td>
<td>0.91</td>
<td>0.57</td>
<td>10.5</td>
</tr>
<tr>
<td>1,2, nausea</td>
<td>109 (40)</td>
<td>79 (4)</td>
<td>&lt;0.001</td>
<td>0.40</td>
<td>0.96</td>
<td>0.90</td>
<td>0.58</td>
<td>10.0</td>
</tr>
<tr>
<td>1,2, gradual onset of pain</td>
<td>110 (40)</td>
<td>28 (4)</td>
<td>&lt;0.001</td>
<td>0.40</td>
<td>0.96</td>
<td>0.90</td>
<td>0.54</td>
<td>10.0</td>
</tr>
<tr>
<td>1,2, aggravation of pain</td>
<td>104 (38)</td>
<td>72 (3)</td>
<td>&lt;0.001</td>
<td>0.38</td>
<td>0.97</td>
<td>0.90</td>
<td>0.59</td>
<td>12.7</td>
</tr>
<tr>
<td>1,2, continuous pain</td>
<td>99 (36)</td>
<td>87 (4)</td>
<td>&lt;0.001</td>
<td>0.36</td>
<td>0.96</td>
<td>0.89</td>
<td>0.53</td>
<td>9.0</td>
</tr>
</tbody>
</table>

General discussion

Studies on patients with AAP need to include many patients, as there are a large number of diagnoses to be considered. Due to the many different etiologies of abdominal pain and varying symptoms and signs for the same diagnosis in different patients, clinical diagnosis is often difficult. Making an early and correct diagnosis is crucial because a good prognosis is dependent on urgent treatment in some cases (Silen 2005). The finding that the level of formal competence did not make any difference in diagnostic performance was unexpected. The hypothesis was that the more experienced a physician was, the better the diagnostic results would be. One interpretation is that the even inexperienced physicians have good medical knowledge and can see the patterns of symptoms and signs characteristic for different diseases. Hence, the information collected and the careful examination of the patient are more important than formal competence (Hancock 1987). The introduction of the structured schedule encouraged this behaviour in all physician categories.

Many other studies (Eskelinen 1992, Körner 1998, Lamprelli 2000) have focused on a certain diagnosis, most commonly appendicitis and patients with suspected appendicitis. Making a comparison between studies is difficult due to different manners of follow-up (or no follow-up), study design and diagnostic criteria. In some studies (Bjerregaard 1976, de Dombal 1972, Ikonen 1983, Irvin 1989, Simmen 1991, Staniland 1980) many diagnoses, in this study defined as specific entities, are grouped together with NSAP, such as gastroenteritis, constipation, mesenterial lymphadenitis, dyspepsia; thus, diagnostic accuracy can reach higher levels due to fewer possible choices of diagnoses. In the present study, the physicians had 30 defined diagnoses to choose from, according to the schedule. Furthermore, patients who did not fit in any of these categories were classified as “other diagnosis”. In several reports, fewer different diagnoses, ranging from 8 to 13, were used (Bjerregaard 1976, de Dombal 1972, Ikonen 1983, Simmen 1991, Staniland 1980). One of these reports, (Bjerregaard 1976) used 10 diagnoses, and the diagnostic performance of the attending physicians was 55.1%. With the same set of diagnoses on this group of patients, the overall diagnostic performance would have been 54%. This comparison indicates that the predictive value of the preliminary diagnosis obtained in the present study was in parity with that of other studies.
Measures of diagnostic performance of patients with AAP

An obviously poor overall diagnostic performance of about 50%, observed in several studies (Adams 1986, Bjerregaard 1976, Gunn 1976), should not necessarily be considered as bad management, as in many cases the pain is self-limiting and there is no need for surgical intervention. Instead specific outcome-measures should be introduced in this field of clinical research. With the exception of NSAP, appendicitis and gallstones, the specificity of most diagnostic entities in the present study was fairly high. Although these are only three diagnoses with low specificity, this was a quantitative problem, as those patients constituted more than half of the cases in the present study. NASP was considered as the main diagnostic problem both quantitatively and qualitatively.

With the exception of colonic obstruction and peptic ulcer, NSAP was the most frequent misjudgement: 9% of the patients finally receiving surgery were initially diagnosed as suffering from NSAP, resulting in a significant delay of surgery. However, those patients constituted only 6% of the 641 patients considered as having NSAP at the emergency department.

Overall, the diagnostic performance was independent of the time of the patient’s arrival at the emergency department. The diagnostic performance for older patients was lower than for younger patients, as has previously been reported (Kizer 1998, Geloven 2000). The consequences of this are outlined below, and the effect on handling was more pronounced for older patients as a larger proportion suffered from specific conditions often necessitating treatment.

*Mortality rate as a measure of diagnostic outcome*

Mortality rate is of general interest and not difficult to measure. In this material the in-hospital mortality was 1.2% (25 of 2052 hospitalised patients) of whom only two patients were under 65 years of age. Delay in surgery is one crucial factor for morbidity and mortality in older age groups. High age as an independent factor is not necessarily associated with an increased operative risk (Arenal 2003, Hosking 1989), but here are possibilities for improvement.

Mortality is uncommon and risk factors for mortality may therefore be difficult to evaluate (Gunnarsson 2003). Grading by a system such as the American Society of Anaesthesiologists (ASA), which was not done in the present study, is probably the best predictor of mortality (Arenal 2003). However, making such grading is difficult for an on-call physician and is not appropriate for all patients arriving at the emergency department. Hence, the ASA-grading would have low validity in a study with the present design.
Bad surgical management as a measure

Another parameter is bad surgical management (de Dombal 1997). One way of measuring this is to compare patients requiring surgical treatment and the proportion of those with an initial diagnosis of not requiring surgery. If there is a delay of more than 24 hours from the patient’s arrival at the emergency department to the operating theatre, it could be considered as a serious management error. In the present study, 104 out of 274 surgically treated patients initially received such a diagnosis, 46 of which waited more than 24 hours before surgery. However, in some cases such as intestinal obstruction or cholecystitis, the true diagnosis may be revealed during the first 24 hours, but for medical reasons it is appropriate to choose conservative management and observation instead of surgery. Hence, these cases are not to be considered as management errors.

Patients with an initial diagnosis not necessitating surgery had a median delay before operation of 22 hours (mean was 40 hours, with 98% CI of 30-50 hours), compared to 8 hours (mean 15 hours, 95% CI of 12-28 hours) for patients with the same final diagnosis at follow-up as the preliminary diagnosis. However, the specificity of diagnoses requiring surgery was high, as in the study by Rozycki et al. (Rozycki 2002) indicating that the risk for surgery in false positive cases is low.

A diagnosis of constipation at the emergency department was associated with a considerable risk for mistakes, where 9% of the patients in fact required surgery for potentially life threatening reasons and 8% were later found to have an abdominal malignancy. An active search for the cause of constipation is mandatory, especially in older patients.

Elderly patients surgically treated for AAP had a mean time of 2.1 days from admission at the emergency department to surgery, patients aged 65-79 years waited 1.7 days and patients under 65 years of age waited 0.9 days. More severe disease among older patients with AAP, as also reported by Watters et al. (Watters 1996), may be an effect of increased occurrence of concurrent diseases at higher ages and because of a delay before admission to hospital and a longer interval between admission and the decision to operate due to less specific symptoms.

Morbidity and negative appendectomies as a measure

One example of morbidity could be if the patient had a stoma, which could have been avoided, or was unnecessarily operated. The rate of negative appendectomies is one such measure. A negative appendectomy is an operation for suspected appendicitis, where a normal appendix is found and removed but no other surgical disease is detected and no other surgical procedure is performed. The negative appendectomy rate is expressed by the number of normal appendices removed in relation to all appendices removed for suspected appendicitis (de Dombal 1997). The perforation rate among patients
with appendicitis might not be a good measurement, as data from this study indicate that the appendix in many cases was already perforated on arrival at the emergency department (Andersson 2002, Hansson 2006).

Duration of hospital stay

Patients above 80 years of age stayed in hospital almost three times longer postoperatively than patients younger than 80 years. Using duration of hospital stay as a measure of diagnostic performance can be misleading, as many elderly patients remain on the surgical ward for days, waiting to receive a place in the Community Care.

Considering a short hospital stay as an indicator of good management is not desirable, as patients are then often referred for further investigations on an outpatient basis: this can lead to diagnostic delays or even missed diagnoses. In the present study, the following diagnoses were identified as having low diagnostic accuracy even at discharge: malignancy, gynaecological complaints, dyspepsia, urinary tract infection and diverticulitis. Changes in discharge diagnoses at follow-up often resulted from information obtained from outpatient investigations that were already ordered at discharge.

Use of resources

In AAP, CT is often done for patients whose diagnosis is strongly suspected on clinical grounds. For determining the additive value of CT (Gwynn 2001), information on the sensitivity and specificity of the diagnosis made by clinical examination is important, and whether the result of such a test will change the suspected diagnosis and surgical management needs to be considered. In some cases, however, a disease is strongly suspected but there is a need for performing a radiological examination that can increase the diagnostic certainty to a point where a decision can be taken. The use of CT in undetermined cases of AAP has increased (Rosen 2003, Terasawa 2004), and it is becoming commonly used as a diagnostic tool for patients with suspected appendicitis (Flum 2005). The most common differential diagnostic entities for appendicitis in this study were NSAP, gynaecological complaints, diverticulitis, mesenteric lymphadenitis, urinary infection and gastroenteritis. These diagnoses, except some cases of gynaecological complaints, did not require surgery, but in some cases required medical treatment, observation and repeated clinical examination of the abdomen to confirm the diagnosis.

This study on patients with appendicitis, confirmed that the assessment in cases of suspected appendicitis could still be based on clinical judgements in combination with laboratory tests: the CT scan can thus be saved for the more equivocal cases of AAP (Tobati 2003). One group of patients that should have extra attention at the emergency department are the elderly, as
the proportion of patients considered as suffering from NSAP at follow-up was lower in the older age groups. A possible explanation for this may be a larger proportion of patients with specific organ pathology and this justifies a more liberal use of CT among older patients presenting with AAP. Another explanation of the lower proportion of NSAP among elderly is that older patients arrive after a longer history of AAP than younger patients do, and some self-limiting pain has already been resolved.

Although in the present study the majority of aged patients were not investigated radiologically, an increase in the proportion of patients submitted for radiological examinations by the emergency department during the early period after the onset of AAP may considerably enhance the diagnostic accuracy at that stage (Geloven 2000). An increase in cost for radiology may also be warranted by the possibility of diminishing the consumption of in-hospital resources through an early correct diagnosis. A more precise identification of patients at the emergency department who would benefit from CT could lead to a shorter average hospital stay. With a growing proportion of patients reaching high ages, there may be increasing demands on clinical resources for this group of patients (Arenal 2003, Reiss 1992); although, the need for surgery after admission to the emergency department was not increased among elderly patients in this study, compared to patients of younger ages.

Computer-aided diagnosis and scoring systems

The computer-aided system in Leeds, UK, is mathematically based on probability analyses and uses Bayes’ theorem, where each patient’s symptoms are compared with cases collected in the database (Adams 1986, Hancock 1987, Horrocks 1972). As the probability of a single symptom associated with a certain disease varies with the prevalence of that disease, the database must comprise a sufficient number of patients collected from the same study population (Staniland 1980, Fenyö 1983). This might explain why the results of the following studies did not reach the initial diagnostic performance of 92%, as seen in Leeds in 1972 (de Dombal 1972), and why it has not become a widely used tool for diagnostic performance. The present database in Mora, Sweden, could serve as a base for a diagnostic computer-aided system for the Swedish population.

Scoring systems are manual recording of symptoms and signs, each giving a certain point, which is added up to produce a score. Scoring systems are mainly used to improve diagnostic accuracy in appendicitis, although the outcome of these scoring systems has not always been convincing (Al-Hashemy 2004, Denizbasi 2003, Franke 1998, Körner 1998, Lintula 2005). The best results reported (Enochsson 2004, Körner 1998) are among women of childbearing age, which is a group of patients known to have low diagnostic accuracy and where the use of CT is limited because of radiation. A retro-
spective comparison of the present results with a scoring system for appendicitis might prove beneficial.

The problem with scoring systems for diagnosing appendicitis is the high incidence of false positive diagnoses resulting in negative appendectomies: the Modified Alvarado Score appears the most reliable score, compared to Ohmann and Eskelinen (Franke 1998, Horzic 2005, Lamparelli 2000). However, the scoring systems are an aid to diagnosis and do not substitute for the clinicians’ judgement of further diagnostic investigations or treatment. Scores are not commonly used in clinical routine (Enochsson 2004, Eskelinen 1992).

Gender perspectives

Women constituted the largest proportion of patients in this study, with a male/female ratio of 0.82 (n=1382:1691). Gynaecological complaints increase the number of diagnoses for females, and may be a plausible explanation why the diagnostic performance at the emergency department was lower for women (52%) than for men (58%; p<0.001). A lower specificity of the preliminary diagnosis for women suffering from appendicitis has also been presented in other studies (Al-Hashemy 2004, Andersson 1992 and 2000, Denizbasi 2003, Flum 2002).

There were a higher proportion of women among patients aged 20-64 years and patients older than 80 years. Gynaecological complaints and self-resolving pain, judged as NSAP, constituted the major part of admittance to the emergency department in the younger age group: this is also reported in a previous study (Sheridan 1992). That women have longer life expectancy than men could also explain the dominance of women among patients aged over 80 years presenting with AAP.

Several gynaecological diseases cause local inflammation that can mimic appendicitis; however, there was no difference according to gender in the presentation of appendicitis symptoms. When testing symptoms that were common for appendicitis, there were no gender differences in LR+ for symptoms typical for appendicitis, except for right-sided rectal tenderness and pain migration, which had a higher LR+ in males (6.5 and 5.3) than in females (4.7 and 3.7).

Gynaecological diseases in this study had low sensitivity for both the preliminary diagnosis and the discharge diagnosis. The change in diagnosis at follow-up could be due to logistic assessment, as a patient referred for a gynaecological examination and presenting no symptoms of a gynaecological complaint was sent back to the emergency department. However, if the patient suffered from a gynaecological disease, she was discharged from the surgical department and the true diagnosis was later identified at the follow-up. Many of these patients were considered as suffering from NSAP as a preliminary diagnosis.
Clinical symptoms, signs and laboratory examinations

Clinical signs, such as rebound tenderness, local rigidity and rectal tenderness, in diagnoses usually known to induce an intra-abdominal inflammatory response, were less common in patients of higher ages. Patients above 65 years also generally had a longer duration of pain before arriving at the emergency department. The explanations for these findings could be differences in the habits of seeking medical advice or physiological causes of changes in the sensation of pain intensity among people of higher ages. Laboratory parameters and body temperature were of limited value among older patients, which was in concordance with earlier studies, (Bugliosi 1990, Parker 1996) and justified a more generous strategy regarding the use of CT for elderly patients.

Classic clinical symptoms for appendicitis, such as isolated pain in the right iliac fossa, rebound tenderness, aggravation of pain when moving, pain migration to the right iliac fossa, local guarding and right-sided rectal tenderness proved to be reliable in predicting acute appendicitis. Previous studies on appendicitis suggesting that right-sided rectal tenderness is invalid as a diagnostic test, with even a reverse relationship to appendicitis (Andersson 1999 and 2000, John 1993), was not supported by the present data where rectal tenderness proved to have high power in predicting appendicitis. A combination of the three most frequently recorded clinical findings with high LR+ for appendicitis, i.e. isolated tenderness in the right iliac fossa, rebound tenderness and aggravation of pain by movement, gave an LR+ of 9.5 indicating that it was a competent diagnostic test for appendicitis.

Shortcomings and strengths

A majority of the physicians participating in the present study were non-specialist physicians with a short time of experience of general surgery, the specialists were on the contrary very experienced with many years in the profession. This is the actual situation in many hospitals in Scandinavia and unfortunately, the proportion of patients assessed at the emergency department by a surgical specialist was very small (only 1.3%). In a study focusing on the level of formal competence of the physician as a determining factor, this could be considered as a disadvantage. However, the greatest improvement in diagnostic performance may be expected during the first years of medical practice.

Previous studies (Paterson-Brown 1990, Gunn 1976) have stated that structured data-sheets for history taking and physical examination improve diagnostic accuracy and should become routine. However, the introduction of a schedule for investigation in this study did not improve diagnostic accuracy compared to that in the baseline registration year. The awareness of being studied can improve the results (the Hawthorne effect; Adams 1986);
thus, even the introduction of the baseline registration schedule may have increased the diagnostic performance: the design of the study did not allow a study of this effect. The proportion of correct diagnoses increased each year during the study period, although no individual feedback on the results of the diagnostic accuracy was given to the physicians.

The strength of the present study was the follow-up, with a pertinent reviewing of each medical record. In most other studies on patients with AAP (Adams 1986, Bowrey 1997, Powers 1995, Wilson 1977), the final diagnosis that serves as the correct diagnosis for comparison is the discharge diagnosis. Another study (Staniland 1980) simply entered each patient’s symptoms and signs into the computer-aided diagnostic system in Leeds and compared the proportion of diagnoses with the Leeds’ patients.

The preliminary diagnosis was not blinded to the reviewer in this study, as the entire medical file needed to be available for making a final diagnosis.

The prospective study design, with a high inclusion rate of 79% of all patients admitted to the surgical emergency department and including patients with gynaecological complaints, makes it very likely that there is no systematic bias in the selection of the patients to this study, and might also explain why the frequency of appendicitis was lower in this population compared to other studies in Europe (Bjerregaard 1976, Staniland 1980, Simmen 1991).

Future perspectives

The emergency departments in many countries are currently staffed by emergency physicians. Trainee physicians are not allowed to be on-call without on-site supervision. This will probably also be the scenario within 10 years in many Swedish hospitals, and if so, surgeons will not be attending as the first-line physician at the emergency department (Karlsson 2004). Therefore, smaller surgery departments may have difficulties in staffing the emergency departments with emergency physicians. The data presented here supports the possibility of organising emergency departments in different ways, depending on local circumstances.

Patients with AAP will continue to be a challenge, and radiological imaging will be needed, albeit more selectively used than presently. The Guidelines for referral to Clinical Radiology, stated by the European Community (Guidelines 2001), claims that exposure to ionising radiation in the total population has to be diminished, thus, the use of CT will probably decrease in favour of MRI. One advantage of MRI is the potential of visualising inflammatory processes in the abdominal cavity, such as diverticulitis and appendicitis (Nitta 2005).

Analyses of health economy, including case-costing data will probably direct the decisions made by the clinical physicians at the emergency departments in a much more concrete way and awareness of the actual costs of different investigative methods will become common knowledge to the phy-
sicians. Determining an appropriate way of identifying patients suffering from NSAP will be the most cost effective course of action concerning patients admitted to the emergency department with AAP.
Conclusions

A structured schedule for investigation did not improve diagnostic precision in the emergency department for patients with acute abdominal pain.

The diagnostic performance was independent of the formal competence of the emergency department physician and of the patients’ time of arrival at the emergency department. The diagnostic performance was significantly lower for female patients.

Non-specific abdominal pain is the main differential diagnostic problem for the emergency department and for diagnoses requiring surgery.

Constipation is a diagnostic pitfall and when making this diagnosis a careful re-evaluation is necessary. Further research is needed to identify patients with malignancies.

Both the preliminary diagnosis at the emergency department and the discharge diagnosis were less reliable among elderly than among younger patients. Elderly patients more often had specific organic disease and arrived at the emergency department after a longer history of abdominal pain than younger patients did. A more generous strategy regarding CT scans among elderly patients with acute abdominal pain, even in the absence of pronounced signs of an inflammatory intra-abdominal process, is recommended.

Classical clinical findings indicating localised inflammation, such as isolated pain in the right iliac fossa, rebound tenderness, right-sided rectal tenderness, pain migration to the right iliac fossa, local guarding and aggravation of pain when moving, were reliable for predicting acute appendicitis. This confirms the opinion that the patient’s history of pain combined with careful clinical examination still has an important role in detecting appendicitis among patients with acute abdominal pain.


I det första delarbetet har vi tittat på panoramat av sjukdomar som är orsak till patienternas akuta buksmärta. De vanligaste diagnoserna var: ospecifik buksmärta, dvs smärta som släpper spontant och där man ej funnit någon

I delarbete 2 har vi studerat fallgropar vid diagnostik av akut buksmärta. Vi har tittat på tillförlitligheten av den preliminärdiagnos som sattes i samband med undersökningen på akutmottagningen och jämfört den med den diagnos som patienten fick vid utskrivning. Slutligen gjordes en jämförelse av utskrivningsdiagnosen med en slutlig diagnos som bestämdes vid en
genomgång av samtliga journaler tidigast ett år efter besöket på akutmottagningen. Vår bedömning var att en uppföljningstid på 1-3 år, skulle räcka för att kunna bedöma om utskrivningsdiagnosen var korrekt. De vanligaste diagnoserna delades in i grupper beroende på diagnosens tillförlitlighet. Det visade sig att diagnoser med låg tillförlitlighet på akutmottagningen, men med betydligt ökad tillförlitlighet av utskrivningsdiagnosen, dvs tillstånd där den inneliggande utredningen i hög grad hade bidragit till att klargöra orsaken till buksmärtan, var ospecific buksmärta, blindtarmsinflammation, gallstensmärta, förstoppning och magsår. Diagnoser med låg tillförlitlighet både på akutmottagningen och vid utskrivningen var bukmaligniteter, gynnekologiska sjukdomar, dyspeptiska besvär, urinvägsinfektioner och inflammation av fickor på tjocktarmen. Överensstämmelsen mellan preliminärdiagnos och utskrivningsdiagnos var hög i denna grupp, det var bara det att de inte korrelerade med den slutliga diagnos som bestämdes vid journalgenomgången. Exempel på orsaker till detta var bland annat svar på biopsier eller senare operation som ej bekräftade diagnosen, uppföljande undersökningar som gjordes polikliniskt, tex gynnekologisk undersökning, gastroskopier, tjocktarmsröntgen eller datortomografi, som visade annan genes till smärta eller normalt fynd samt urinodlingssvar som inte visade bakterieväxt. Av 479 patienter som genomgick operation under vårdtiden, fick 104 patienter en preliminärdiagnos som vanligtvis inte behandlas med operation och som under vårtdiden visade sig vara felaktig. Detta bidrog sannolikt till att genomsnittstiden från undersökningen på akutmottagningen till operation var betydligt längre i denna grupp, medeltid 22 timmar, med 95% konfidensintervall på 30 till 50 timmar, jämfört med patienter som fick en korrekt diagnos på akutmottagningen, där medeltiden till operation var 8 timmar, med 95% konfidensintervall på 12-18 timmar. En diagnos med speciella diagnosiska svårigheter är ospecific buksmärta, där smärtan viker spontant och ingen specifik orsaksdiagnos kan fastställas. Smärtan kan dock väcka oro hos patienten som vill att utredningar ska göras. Ospecific buksmärta är vanligt (i denna studie 37%) och ska egentligen inte föranleda någon extensiv utredning, men det kan i vissa fall vara svårt för läkaren att avstå från fortsatt utredning. Den vanligaste alternativa diagnosen för patienter som lades in på misstanke om blindtarmsinflammation, var ospecific buksmärta. Här förekommer ytterligare gällande behandling, där inflammation av blindtarmen ska opereras och den ospecifica buksmärta inte behandlas alls. Detta brukar dock visa sig under observationstiden på sjukhuset. En annan diagnos som uppmärksammades var förstoppning, där 9% av patienter som opererades för ett potentiellt livshotande tillstånd (bland annat tarmvred och blindtarmsinflammation) hade fått förstoppning som första diagnos. Bland patienter som fått preliminärdiagnos förstoppning visade sig 8% ha en tumör. Vår rekommendation är att beställa en utredning av tjocktarmen hos äldre patienter som fått diagnosen förstoppning.
Det tredje delarbetet handlar om diagnostik av akut buksmärta hos äldre. Vi jämförde patienter som var 65-79 år och patienter över 80 år med patienter i åldern 20-64 år. Patienter över 65 år hade generellt längre smärtduration innan de sökte vård. Orsaken till detta kan vara andra sökmönster i de äldre generationerna, långa avstånd där vissa patienter i sjukhusets nordvästra upptagningsområde har upptäckt 25 mil till sjukhuset samt fysiologiska orsaker som kan bero på en förändrad smärtsgrad i högre åldrar. Andelen felaktig preliminärdiagnos var högre bland äldre patienter, 49% för 65-79 åringar, 56% för patienter över 80 år jämfört med 45% för 20-64 åringar. Även utskrivningsdiagnosen var oftare inkorrekt bland patienter över 65 år; 22% respektive 25% jämfört med 16%. Statistiskt signifikant skillnad föreligger för dessa siffror. Diagnosen ospecific buksmärta var efter journalgenomgång signifikant ovanligare bland äldre patienter, tydande på att äldre oftare har en organspecifik orsak till akut buksmärta. Patienter yngre än 65 år som opererades hade förhöjda inflammationsprover jämfört med jämnåriga som ej opererades. Detta mönster sags ej bland patienter över 65 år. Bland patienter som hade en diagnos som ger upphov till retning av bukhinnan, var det ovanligare hos de äldre att vid palpation av buken finna de för detta tillstånd typiska fynden: släppömhet, lokalt muskelförsvar och ömhet vid undersökning av ändtarmen.

Avhandlingens fjärde och sista delarbete behandlar en av de vanligaste orsakerna till akut buksmärta; akut blindtarmsinflammation. I denna studie hade 274 (11%) av 2478 patienter som sökte på akutmottagningen blindtarmsinflammation. Av dessa hade 165 patienter (60%) avancerad inflammation, dvs gangränös eller brusten blindtarm. Av de 432 patienter som på akutmottagningen hade misstänkt blindtarmsinflammation, var de 160 som aldrig blev opererade beroende på att smärtan försvann spontant (111 patienter; 26%) eller för att en annan orsak till smärta upptäcktes (49 patienter; 11%). Totalt 211 patienter av 432 med misstänkt blindtarmsinflammation fick en annan slutdiagnos. De vanligaste diagnoserna bland dessa var: 122 patienter hade ospecific buksmärta, 24 hade gynekologisk orsak, 13 hade infiammerade fickor på tjocktarmen (varav i 2 fall hade en tarmficka perforerat), 10 hade virusorsakad körtelbuk, 8 hade urinvägsinfektion och 7 hade gallstens smärta. Av dessa 211 patienter opererades 53. Det utfördes 316 blindtarmoperationer och i 45 fall (14%) fann man ingen orsak till buksmätan. Bedömningarna av dessa 432 patienter baseras huvudsakligen på kliniska undersökningar: datortomografi gjordes endast på 6 patienter och ultraljuddundersökning på 36 patienter. Den diagnostiska tillförlitligheten för blindtarmsinflammation var trots det hög: 0,89 diagnostisk ”accuracy”. Incidensen av blindtarmsinflammation för befolkningen i Mora lasarettets upptagningsområde var 102 per år och 100 000 invånare. Blindtarmsinflammation var vanligare bland män (57%). Medelåldern var 32 år. Symtom som var vanligare bland patienter med blindtarmsinflammation jämfört med de pati-
enter som hade en annan diagnos var: nedsatt aptit, illamående, kräkning, gradvis försämring av smärtan, kontinuerlig, molande smärta, försämring av smärta vid hosta eller rörelse, isolerad smärta i bukens högre nedre del samt släppömhet vid bukpalpation och ömhet uppåt höger vid undersökning av ändtarmen. Det var ingen av patienterna med blindtarmsinflammation som hade isolerad ömhet i bukens vänstra nedre del. Smärtvandring, dvs smärta flyttar sig från centrala delen av buken och lokaliseras sig efter hand nertill på höger sida av buken, förekom i 16% av blindtarmspatienterna medan detta beskrevs endast hos 3% bland övriga patienter med buksmärta. I en multivariat regressions analys visade sig följande fynd ha störst värde vid diagnostik av akut blindtarmsinflammation; isolerad ömhet i bukens högra, nedre del (Odds Ratio 3,29), släppömhet (3,00), ömhet uppåt höger vid undersökning av ändtarmen (2,53), smärtvandring (2,18) och lokalt muskelförsvar (2,11). Kombinationen av isolerad ömhet i bukens högra, nedre del, släppömhet och försämring av smärta vid rörelse sågs hos 38% av blindtarmspatienterna och gav ett positivt likelihood ratio på 9,5.

Sammanfattning

Akut buksmärta kan orsakas av en mängd olika tillstånd. Den första bedömningsen på akutmottagningen är viktig, inte minst för patientsäkerheten. I en del fall krävs omedelbara åtgärder och i denna studie blev 17% av patienterna opererade under vårdtiden. Det vanligaste är dock att smärta är ofarlig och släpper spontant. Detta ställer stora krav på diagnostiken. Ur detta populationsbaserade material med 3349 registrerade patienter som sökt på akut buksmärta, har vi identifierat dels de diagnoser som är svåra att ställa på akutmottagningen, och dels de diagnoser som är osäkra även vid utskrivning och som kan behöva följas upp. Vi har även identifierat två patientgrupper som är svårare att bedöma; kvinnor och patienter över 65 år. Våra resultat visade att efter en noggrann klinisk undersökning på akutmottagningen stämmer diagnosen i 54% av fallen totalt. Patienter med blindtarmsinflammation är en stor grupp, där behandlingen är kirurgi. Vi har identifierat de viktigaste symptomen i en jämförelse med alla andra patienter som söker på akut buksmärta, dvs inte bara med dem som har misstänkt blindtarmsinflammation. Såväl sensitiviteten (att peka ut den som är sjuk) som specifiteten (att inte felaktigt peka ut någon) för blindtarmsinflammation var hög, 0,81 respektive 0,90. Vi menar att den kliniska undersökningen är viktig och räcker oftast som underlag för att selektera de patienter som behöver blindtarmoperation, utan onödig förröjning i väntan på radiologisk undersökning som kan bekräfta diagnosen. Vi får inte glömma den kunskap som finns om klinisk undersökning och bara förlita oss på röntgenfynd. Utifrån denna kartläggning har vi förhoppning om att kunna bidra med kunskap om de fall gropar som förekommer i den akuta handläggningen av en i akutsjukvård stor och mycket viktig patientgrupp.
Acknowledgements

I wish to express my sincere gratitude to all those who have made it possible for me to accomplish this work. My special thanks are due to:

**Ulf Gunnarsson**, my tutor, for inspiring enthusiasm, incredible working-capacity and unflinching support, for sharing your great knowledge of statistics and scientific work, for skilful database design, and for leading our research group in a very wise way of both scientific atmosphere and friendship.

**Lars-Erik Hansson**, my co-tutor, for encouragement and for valuable help in designing the studies and skilful choice of statistical analyses, and for valuable comments.

**Lars Påhlman**, my co-tutor, for giving me the opportunity to work in the colorectal team, and for sharing your deep knowledge in the scientific field of Surgery.

**Ewa Lundgren** and **Ulf Haglund**, present and former Heads of the Department of Surgery, and **Lars Wiklund**, Head of the Department of Surgical Sciences, for providing conditions that made this work possible.

**Friends and colleagues in the research group**, for sharing your knowledge of statistical analyses at our seminars, reading my papers and giving constructive criticism, and afterwards sharing time over a little meal and a beer, always at the same restaurant, and especially **Pia Jestin** for being supportive in both science and life wisdom.

**Maud Marsden** and **Sue Pajulouma**, for skilful linguistic revision and invaluable comments.

**Hans Garmo**, for supporting me in understanding the statistical analyses whenever I needed help.

**Rolf Heuman**, my tutor in surgical skills, former Head of the Department of Surgery in Mora, for encouraging me in both clinical practice and research.
Friends and colleagues at the Department of Surgery in Mora, especially those who helped to fill in the investigation schedules, and also the secretaries, for their competent work in completing the database. My special thanks are also due to Viviann Pers, who entered the information into the database.

Friends and colleagues at the Department of Surgery in Uppsala, especially my dear colleagues at the colorectal team and my colorectal roommates Ulla-Maria Gustafsson and Lars Österlund, and former roommates Erik Lundin and Johan Hansson.

The personnel at the wards 70 B1 and 70 B2, at the outpatient clinic, at the operating unit and at the emergency department, Ulla and Lena at the endoscopy unit, and personnel at Samariterhemmets Hospital, for their support and professionalism in daily (and nighttime) work.

Agneta Gustafson and Elisabeth Bergqvist, for valuable help in practical issues.

Malin Degerman-Gunnarsson, Thyra and August, family of Ulf Gunnars-son, for letting us share his time during those busy years.

My mother, Marianne and my late father, Åke, for always believing in me and encouraging me to do the things I wanted to do, and for taking good care of our children during times of work. Ingela Collin, my sister, and Mats Laurell, my brother, and their families, for encouragement.

Fredrik Clason, my father-in-law and Stina Söderlundh, my mother-in-law, and my sisters-in-law, Gertrud Tinnis and Hannah Carlshamre, for encouragement and for taking good care of our children during times of work.

Yvonne Sundin, for support, skilful training and care of my horse Gaiego.

Most of all, Isaac, my husband, for love and friendship, for your endless patience, your expert database support, for supplying me with computers and all kinds of gadgets and for sharing your great knowledge of medicine and data computing, and for being a wonderful father to our children, Martin and Lisa, who always have been cheerful and supportive, so that I could complete this work.

This work has been supported financially by the Centre for Clinical Research (CKF) in Dalarna, the Cancer Foundation, Uppsala University Hospital, the Bengt Ihre Foundation and the Department of Surgery in Mora.
References


Amyand C. Of an inguinal rupture, with a pin in the appendix coeci, incrusted with stone; and some observaions on wounds in the guts. Phil Trans Royal Soc 1736;39:329.


d de Dombal FT. Computer aided diagnosis of Acute Abdominal Pain from the Clinical Information Science Unit, University of Leeds and Media Innovations Ltd.


Karlsson H, Tjernström E. Om 5 år ska akutläkare sköta samtliga våra akutomttagningar. *Sjukhusläkaren* 2004;5.


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 174

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)