Diagnosis in acute abdominal pain and ongoing abdominal sepsis
Kiewiet, Jordy

Citation for published version (APA):
Kiewiet, J. J. S. (2016). Diagnosis in acute abdominal pain and ongoing abdominal sepsis

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
SUMMARY, GAINED INSIGHTS AND FUTURE PERSPECTIVES
Summary
The first part of this thesis focusses on diagnosis in patients presenting to the Emergency Department with acute abdominal pain (AAP). Studies were performed evaluating different steps in the diagnostic work-up. With the focus on the clinical assessment we performed a prospective multicenter cohort study (AAP study) including 294 patients presenting with acute abdominal pain at the Emergency Department. There is room for improvement of the accuracy of the clinical diagnosis. For example, if surgeons would have a higher diagnostic accuracy than surgical residents this could lower the burden of additional work-up and unnecessary resource utilization. Chapter 1 describes a study in which 126 of these patients were independently assessed by surgical resident and a surgeon. They stated their clinical diagnosis based on history taking, physical examination and laboratory results. Diagnostic accuracy of the clinical diagnosis was compared. Surgeons made the correct diagnosis in 43% of cases compared to 44% correct diagnoses by surgical residents. Surgeons and residents have comparable sensitivity, specificity, positive predictive value and negative predictive value to distinguish urgent from non-urgent diagnoses. For specific diagnoses, such as acute appendicitis and acute diverticulitis, there were slight differences between surgeons and residents. Most notable was the difference when surgeons diagnose diverticulitis they had less false positive cases (75% versus 44%) but a slightly lower sensitivity than residents (33% versus 44%). In contrast surgeons had a higher sensitivity in diagnosing acute cholecystitis (71% versus 43%) but a somewhat lower positive predictive value (45% versus 60%). Inter-observer agreement for diagnoses varied from fair to moderate (κ 0.28-0.57). Surgeons had a higher level of diagnostic certainty then residents but the accuracy of the clinical diagnosis did not improve making it unlikely that assessment by a surgeon rather than a resident would reduce the need or call for additional work-up.

Decision tools can be used to increase accuracy of the clinical diagnosis and prevent over-utilisation of imaging. In Chapter 2 the influence of the use of decision tools for appendicitis and diverticulitis on the clinical diagnosis was assessed. These tools were evaluated in the AAP study cohort whenever a surgical resident recorded the clinical diagnosis as suspected appendicitis or diverticulitis. Before and after the use of these decision tools the clinician recorded the certainty of the diagnosis. The amount of correct clinical diagnosis of appendicitis was comparable irrespective of the use of the decision tool (57% correct without versus 55% correct with decision tool). Since the amount of diverticulitis patients was limited the diagnostic accuracy before and after the use of the decision tool was not separately calculated.
The level of certainty after completion of the decision tool increased in 19% of patients with a final diagnosis appendicitis and in 14% patients without appendicitis. For diverticulitis these proportions were 36% and 38%, respectively. In none of the patients suspected of having diverticulitis but with a different final diagnosis the decision tool influenced the choice to utilize additional imaging. In a mere 18% of patients with acute diverticulitis as final diagnosis the decision tool influenced the choice of the surgical resident to order additional imaging. These results suggest that the use of the decision tool for acute appendicitis does not significantly influences the accuracy nor the certainty of the clinical diagnosis. The decision tool for acute diverticulitis has a somewhat larger influence on the diagnostic certainty but the influence on the utilization of imaging resources is limited.

Shortly after the Emergency Department (ED) decision tool for acute diverticulitis was published a second independent study presented a different Clinical Scoring (CS) tool for diverticulitis. These two tools were subjected to external validation in Chapter 3. External validation was performed in a crosswise design on the derivation datasets of the ED tool and the CS tool. In addition, they were both evaluated for their diagnostic accuracy in a third independent dataset. Since more than 90% of patients with acute diverticulitis have uncomplicated disease they do not all need imaging. A decision tool that increases the accuracy of clinical diagnosis and more specifically increases positive predictive value could assist the clinician in omitting unnecessary additional imaging. Both decision tools had a high positive predictive value in all three datasets varying from 81% to 100% for the ED tool and from 89% to 92% for the CS tool. In addition to the positive predictive value the amount of patients that can be identified with a final diagnosis of diverticulitis from the cohorts of patients with suspected diverticulitis is important. The ED tool performed better than the CS tool since it identified 24% of patients in the ED tool derivation cohort, 20% in the CS tool derivation cohort and 14% in the independent cohort compared to 19%, 6% and 9% identification by the CS tool respectively. In conclusion, both tools could be used to aid in the selection of patients suspected of acute diverticulitis in which additional imaging can be omitted based on the diagnostic accuracy. However, the ED tool identifies more cases than the CS tool.

Additional imaging as part of the diagnostic work-up in acute abdominal pain plays an increasingly important role. Diagnostic accuracy of imaging for acute appendicitis and acute diverticulitis has been studied extensively. However, in the case of acute cholecystitis, which is the third most common urgent diagnosis in acute abdominal pain, up-to-date summary estimates of diagnostic accuracy are not available. An update is warranted because new
modalities such as a computed tomography (CT) are introduced and existing modalities like ultrasound have undergone significant improvement. In the systematic literature review described in **Chapter 4** we made diagnostic summary estimates of the most frequently used imaging modalities for acute cholecystitis. For these estimates we included 57 studies evaluating cholescintigraphy, ultrasound, magnetic resonance imaging and CT. Cholescintigraphy has its logistical disadvantages but has the highest diagnostic accuracy with a sensitivity of 96% (CI: 94% to 97%) that is significantly higher than sensitivity of ultrasound (81%, CI: 75% to 87%) and MRI (85%, CI: 66% to 95%). There were no significant differences of specificity between cholescintigraphy (90%, CI: 86% to 93%), ultrasound (83%, CI: 74% to 89%) and MRI (81%, CI: 69% to 90%). Only one study could be included that evaluated CT with a sensitivity of 94% (CI: 73% to 99%) at a specificity of 59% (CI: 42% to 74%). Because ultrasound is widely used in the work-up of the diagnosis acute cholecystitis and is part of the diagnostic criteria in guidelines it is important to realize that ultrasound has a substantial margin of error.

**Chapter 5** provides an overview of the diagnostic accuracy of each step in the diagnostic work-up of intestinal ischemia. We performed a systematic review of literature for studies evaluating patients presenting with acute abdominal pain suspected of having acute intestinal ischemia. Studies including in-hospital postoperative patients and patients with strangulated bowel obstruction were excluded because these comprise an elementary different patient population. Since intestinal ischemia is an uncommon diagnosis with a population based incidence of 4.5 to 44 cases per 100,000 person-years, diagnostic accuracy could not be established for all steps in the work-up because of lack of evidence. CT had good accuracy for diagnosis of intestinal ischemia with a summarized sensitivity of 91% (CI: 80% to 98%) at a specificity of 94% (CI: 87% to 99%) in five studies. Of all laboratory tests D-dimer was most investigated and four studies were included with a sensitivity ranging from 60% to 100% at a specificity ranging from 13% to 87%. Two studies evaluating D-lactate were included with a sensitivity of 90% and 89% at a specificity of 23% and 89% respectively. One study evaluating i-FABP was included with a sensitivity of 83% at a specificity of 89%. Seen the heterogeneity of the studies evaluating laboratory tests and use of different cut-off values data was not pooled. All and all laboratory tests are not accurate enough and lack adequate specificity. No studies evaluating the diagnostic accuracy of a clinical diagnosis were found, and no studies on L-lactate, MRI or angiography that evaluated the added value of such diagnostic modality.
Part two of this thesis focusses on the detection and diagnosis of ongoing abdominal sepsis. Primary peritonitis in which the gastro-intestinal tract is intact is considered to be a fundamental different disease entity which is not studied in this thesis. Focus is on abdominal sepsis that is known as secondary peritonitis. In secondary peritonitis there is a breach in the gastrointestinal tract where surgery is the cornerstone of treatment in contrast to primary peritonitis. Chapter 6 gives an insight in the current treatment and management principles of abdominal sepsis. Details are discussed concerning terminology, incidence, impact on the healthcare system and treatment of abdominal sepsis. Main focuses are the surgical and diagnostic strategies when patients have ongoing abdominal sepsis. Detection and adequate diagnosis of patients with ongoing abdominal sepsis needing re-intervention is challenging. The RELAP trial showed that the on-demand surgical strategy should be preferred over the planned relaparotomy strategy. In the on-demand strategy patients are only submitted to relaparotomy if there is insufficient improvement or clinical deterioration. In contrast, in the planned relaparotomy strategy, a relaparotomy is performed every two to three days until there is no sign of ongoing abdominal sepsis. Difficulty of the on-demand strategy lies in the timely and accurate identification of patients with ongoing abdominal sepsis without delaying a necessary relaparotomy since this can have deleterious consequences in these severely ill patients. We performed a study, described in Chapter 7, to determine whether microbial profiles cultured from the abdomen in the initial laparotomy for abdominal sepsis were associated with ongoing infection. Patients included in the RELAP trial with available abdominal cultures were used in this study. No microbial profile showed significant relation with ongoing abdominal infection needing re-intervention. Resistance to empiric antibiotic therapy for gram-positive cocci and coliforms showed strong but no significant association with ongoing infection (OR 3.43, 95%CI: 0.95-12.38). Since Enterococcus spp. had a high prevalence and had borderline association with in-hospital death, empiric coverage should be advocated.

Several clinical scoring tools are in use to predict mortality and morbidity in severely ill patients. Some of these scoring tools have proven to have adequate accuracy to perform this task in patients with abdominal sepsis. In Chapter 8 we report a study evaluating the APACHE-II score, SAPS-II, Mannheim peritonitis index, MODS, SOFA score and the acute part of the APACHE-II score (APS) for their ability to predict ongoing abdominal infection. Only the APS on day 1 (AUC 0.61, 95%CI: 0.52-0.69) and the SOFA score on day 2 (AUC 0.60, 95% CI: 0.52-0.69) had discriminatory capacity above change. The cut-off analysis we performed showed that low cut-off values of the scoring tools are necessary to identify an acceptable
level of patients needing a relaparotomy. In term these low cut-off values would lead to an unacceptable high percentage of negative relaparotomies making the available scoring tools designed to predict mortality and morbidity unfit to aid in the selection of patients with ongoing abdominal sepsis needing a relaparotomy.

Since the available scoring tools investigated in chapter 8 are inadequate a more specific tool is desirable. **Chapter 9** describes a study in which a new tool with this specific purpose was developed. Candidate variables to be included in the tool were selected based on previous studies and common clinical sense. Data from the RELAP trial was used to construct a prediction model. This model had good discriminating capacity with an AUC of 0.80. Similar to the tools evaluated in chapter 8 this model also lead to an unacceptable proportion of negative relaparotomies when cut-off analysis was performed striving for adequate sensitivity. However, when incorporated in a clinical decision tool where outcome of the prediction model is coupled to a proposed interval of decision tool re-assessment and use of CT, the tool could uniformly stratify the underlying risk of ongoing disease and guide monitoring.

Although the RELAP trial showed that the on-demand strategy is the surgical strategy of choice in terms of outcome and cost reduction, implementation in clinical practice proved to be less straightforward. A survey has shown that although the on-demand strategy is practiced more frequent after publication of the trial results still planned relaparotomy is widely adhered leading to suboptimal care and unnecessary high costs. In **Chapter 10** a systematic analysis of factors hindering and stimulating the implementation of the on-demand strategy is reported. The problem analysis was performed in five phases of a Delphi process using expert panels. Hindering and promoting factors and accompanying interventions were formulated, quantified and ranked in a survey including 301 respondents from several hospitals and specialties involved in the care for patients with abdominal sepsis. Most important factors among the 25 factors formulated hindering implementation of the on-demand strategy were the lack of multidisciplinary discussion of patients with abdominal sepsis, lack of availability of adequate imaging and expertise of an intervention radiologist, insufficient knowledge of literature among treating clinicians and general fear for the risk of waiting too long prior to performing a relaparotomy. Providing a clinical guideline involving multidisciplinary discussion and a description of intensive monitoring was seen as the most useful interventions to promote implementation of the on-demand strategy.
Gained insights and future perspectives

Although diagnostic work-up of patients presenting to the emergency department with acute abdominal pain seems to follow a standardized pattern involving clinical assessment, laboratory test and increasing use of imaging modalities this thesis confirms that establishing an accurate diagnosis in acute abdominal remains challenging. The recent publication of a national guideline in the Netherlands guiding clinicians in the diagnostic work-up of acute abdominal pain has been an important step to further standardize the work-up. Especially the first step in the diagnostic work-up, i.e. clinical assessment, has limitations in its diagnostic accuracy. In this step it is essential to select the right patients in need of undergoing additional steps of the work-up, with a limited number of unnecessary additional imaging or missed urgent diagnoses. To improve this process first a study needs to be done in which a systematic analysis is performed that uncovers the criteria that clinicians use in current practice to select patients for additional work-up. Based on these findings a set of criteria can be compiled which can then be evaluated prospectively to provide a guideline in which patient additional steps of the work-up are needed.

We hypothesized that experienced surgeons assessing these patients would have a higher diagnostic accuracy of their clinical diagnosis leading to less unnecessary use of hospital resources without an increase in missed urgent diagnoses. However, our study has shown that accuracy of the clinical diagnosis does not improve when all patients are assessed by an experienced surgeon. Particularly disappointing was that surgeons were not more accurate than residents in distinguishing urgent and non-urgent diagnosis. In our study we performed a head-to-head comparison between surgeons and residents whereas in current practice a surgeon is predominantly involved after a resident has done the first step in the work-up. It would be of interest to perform a study evaluating what the additional value of this methodology is in terms of accuracy of diagnosis and utilization of resources compared to evaluation of patients solely by a resident.

Clinical scoring tools for diverticulitis showed reliable diagnostic accuracy at external validation. However, similar to the scoring tool for appendicitis the tool did not have a large impact on certainty of the clinical diagnosis and the subsequent course of action of doctors evaluating these patient groups. Unfortunately we did not study the factors why these tools did not influence the doctors using them. If we can uncover these factors in a future study the tools can be adjusted to improve the impact they might have.
The summarized accuracy of ultrasound for acute cholecystitis showed substantial margin of error but is included in the diagnostic criteria for this disease in the Tokyo guidelines. In our systematic review we only included one study evaluating the diagnostic accuracy of CT for acute cholecystitis. However, CT is being performed increasingly in evaluating patients with acute abdominal pain. More studies are needed that describe the diagnostic accuracy of CT to uncover if this modality is reliable enough when acute cholecystitis is suspected or that we need to focus on different imaging modalities.

For less frequent occurring diseases causing acute abdominal pain studies describing diagnostic accuracy are very limited. This is illustrated in the systematic review for diagnostic accuracy of the different steps in the work-up of patients with mesenterial ischemia. All findings in the studies we performed in this thesis underline the complexity of diagnosis in acute abdominal pain where the diagnostic process faces a twofold challenge. First distinction between urgent and non-urgent entities has to be made but at the same time adequate diagnosis of a disease is to established. Many published studies involving patients with acute abdominal pain have studied patient populations with a specific disease such as appendicitis. The majority of studies only focus on patients either suspected of having appendicitis or patients that have a final diagnosis of appendicitis. If this approach is chosen patients that were not suspected of having appendicitis or patients that have a suspicion of appendicitis but whom turn out to have a different diagnosis are left out of the evaluation. Therefore, the patient population in these studies is not representative for the population in daily practice. This leads to over-estimation of test accuracy in such an artificially narrowed study population, as in practice patients presenting with acute abdominal pain are a broader group that by enlarge are not pinpointed to one initial suspicion diagnosis. We therefore advocate to perform more studies where the diagnostic accuracy of the different steps in the clinical work-up are performed including the broad group of patients having acute abdominal pain. An example of this study design is the OPTIMA study evaluating accuracy of readily available imaging modalities.

When an urgent diagnosis of acute abdominal pain causes abdominal sepsis another diagnostic challenge arises. Source control, adequate antimicrobial treatment and resuscitation as the initial treatment are undisputed. Since the RELAP trial the on-demand relaparotomy strategy is to be preferred. This strategy however poses the challenge how to identify the patient with an ongoing abdominal infection that needs a re-intervention. Especially the
first few days after the onset of the abdominal sepsis it is hard to distinguish between systemic sepsis related symptoms such as multiple organ failure or inadequate source control needing re-intervention in these often severely ill patients. There is no documented pattern of diagnostic monitoring of patients with abdominal sepsis. For instance, CT is widely used in the initial work-up of a septic patient but is there is little evidence available regarding its diagnostic accuracy to identify patients needing a re-intervention. With a prospective study design where CT scans are made in these first days which are evaluated by the treating multidisciplinary team whether or not a re-intervention is necessary a future study could provide important information if we have can reliably use CT to select patients needing re-intervention.

Extensive evidence is available regarding the treatment of sepsis in general. A summary of this evidence is provided by the Surviving Sepsis Campaign which provides frequently updated guidelines (www.survivingsepsiscampaign.org). In these guidelines there are no recommendation on how to monitor patients with abdominal sepsis for ongoing abdominal infection. The microbial profile cultured from the contaminated abdomen at the initial operation does not predict which patient will have an ongoing intra-abdominal infection needing relaparotomy. A multivariable scoring tool we developed does have acceptable diagnostic accuracy but lacks discriminatory capacity. If the tool was to be used strictly based on the outcome it would lead to a unacceptable high number of unnecessary negative relaparotomies. Therefore, we incorporated the model scores in a decision tool that stratifies the risk of ongoing abdominal infection needing relaparotomy in actions for re-evaluation and/or additional CT.

The decision rule could provide the basis of a clinical guideline for the diagnostic work-up for patients with ongoing abdominal infection. A future study is necessary to evaluate the use of this decision rule in a prospective setting for its applicability and diagnostic accuracy. Especially since the study uncovering hindering and promoting factors for the implementation of the on-demand strategy showed that all of the involved specialists report a need for a clinical guideline.
References