Optimizing care for patients surgically treated for severe peritonitis
Boer, K.R.

Citation for published version (APA):

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.
This dissertation concentrates on patients with severe secondary peritonitis (abdominal sepsis), a condition with high mortality and disease-related morbidity. Secondary peritonitis is a clinical diagnosis requiring an emergency surgical laparotomy to confirm diagnosis and to tailor surgical treatment. With an estimated incidence in the United States of 9 cases of secondary peritonitis per 1000 emergency hospital admissions and with extensive intensive care admissions and lengthy hospital stays, these patients represent a substantial cost to the healthcare system.

Patients surviving peritonitis also continue to report substantially reduced health-related quality of life and are often readmitted to hospital with disease-related complications in the first year following their initial acute illness and surgery. Therefore in this thesis, we concentrate on both the psychological and physiological recovery of these severely ill patients.
Optimizing Care for Patients Surgically Treated for Severe Peritonitis

To my grandparents, who taught me that with a little hard work, you can accomplish anything.
Academisch proefschrift

Ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie, in het openbaar te verdedigen in de Aula der Universiteit

op

woensdag 14 november 2007, te 11.30 uur

door

Kimberly Rachel Boer

geboren te Haarlem
## Promotie Commissie

<table>
<thead>
<tr>
<th>Promotor</th>
<th>Prof. dr. P.M.M. Bossuyt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-promotor(es)</td>
<td>Dr. M.A. Boermeester</td>
</tr>
<tr>
<td></td>
<td>Dr. C.A.J.M. de Borgie</td>
</tr>
<tr>
<td></td>
<td>Dr. J.B. Reitsma</td>
</tr>
<tr>
<td>Overige leden</td>
<td>Prof. dr. H.G. Gooszen</td>
</tr>
<tr>
<td></td>
<td>Prof. dr. J.C.J.M. de Haes</td>
</tr>
<tr>
<td></td>
<td>Prof. dr. H. Obertop</td>
</tr>
<tr>
<td></td>
<td>Mw. dr. M. Olff</td>
</tr>
<tr>
<td></td>
<td>Prof. dr. E.W. Steyerberg</td>
</tr>
<tr>
<td></td>
<td>Prof. dr. M.B. Vroom</td>
</tr>
</tbody>
</table>

## Faculteit der Geneeskunde
# Table of contents

Chapter 1 General introduction
Outline of the thesis 7

**Part I** In search of an evidence-based surgical approach
Chapter 2 Comparison of on-demand versus planned relaparotomy strategy in patients with severe peritonitis. A randomized trial 19
Chapter 3 Health related-quality of life in patients with severe peritonitis. A randomized trial comparing on-demand versus planned relaparotomy following emergency laparotomy 21
Chapter 4 Costs of relaparotomy on-demand versus planned relaparotomy in patients with severe peritonitis. An economic evaluation within a randomized controlled trial 37

**Part II** Optimizing the on-demand strategy
Chapter 5 Usefulness of the APACHE II scores to predict disease outcome and to evaluate safety of surgical strategy in patients with abdominal sepsis 71
Chapter 6 Failure of available scoring systems to predict ongoing infection in patients with abdominal sepsis after their initial emergency laparotomy 73

**Part III** Health related-quality of life in patients with secondary peritonitis
Chapter 7 Health related-quality of life six months following surgical treatment for secondary peritonitis – using the EQ-5D questionnaire 109
Chapter 8 The psychometric performance of a disease-specific health related-quality of life questionnaire for patients with secondary peritonitis: The SP-QoL Questionnaire 111

**Part IV** Post-traumatic stress disorder in patients with secondary peritonitis
Chapter 9 Long-term prevalence of post-traumatic stress disorder symptoms in patients after secondary peritonitis 147
Chapter 10 Factors associated with post-traumatic stress symptoms in a prospective cohort of patients after abdominal sepsis. A nomogram 149
Chapter 11 The relationship between post-traumatic stress disorder symptoms and health related-quality of life in patients abdominal sepsis. A prospective study 165
Chapter 1
General introduction
Outline of the thesis

“Medicine is, I have found, a strange and in many ways a disturbing business. The stakes are high, the liberties taken tremendous..... What you find when you get in close, however – close enough to see the furrowed brows, the doubts, the failures as well as the successes – is how messy, uncertain and also surprising medicine turns out to be. The thing that still startles me is how fundamentally human an endeavor it is.”

Atul Gawande (Complications)
General introduction

This thesis concentrates on patients with severe secondary peritonitis (abdominal sepsis), a condition with high mortality and disease-related morbidity. Secondary peritonitis is a clinical diagnosis requiring an emergency surgical laparotomy to confirm the diagnosis and tailor the surgical treatment. With an estimated incidence for the United States of 9.3 cases of secondary peritonitis per 1000 emergency hospital admissions, these patients represent a substantial cost to the healthcare system.

Peritonitis

The term peritonitis refers to a constellation of signs and symptoms, which are the result of microbial contamination of the peritoneal cavity with subsequent inflammation. Clinical symptoms include abdominal pain and tenderness on palpation, abdominal wall muscle rigidity and systemic signs of inflammation. Symptoms often depend on the exact location, extent and source of the infection. Peritoneal infections are classified as primary (i.e., spontaneous), secondary (i.e., related to a pathologic process in a visceral organ), or tertiary (i.e., persistent or recurrent infection after a prolonged period of adequate therapy). In the studies reported in this thesis, we concentrate only on secondary peritonitis. Causes of secondary peritonitis can be nosocomial or community acquired. Nosocomial infections are usually complications following gastrointestinal surgeries, e.g., colon cancer, anastomotic leakage or other abdominal emergencies. Community-acquired secondary peritonitis patients usually suffered from gastrointestinal problems such as diverticulitis, perforation, complicated Crohn's disease or ulcerative colitis. In severe cases of secondary peritonitis complications are frequent and can lead to severe disease-related morbidity or even mortality.

Mortality

Secondary peritonitis is associated with serious morbidity and high mortality. Mortality rate from intra-abdominal infections decreased from more than 90% to approximately 40% during the period from 1890-1924, with the fundamental role of operative therapy in the treatment of peritonitis established in the early twenties. In recent decades, despite improvements in antibiotic treatments, intensive care services and improved percutaneous and surgical treatments, mortality has only improved slightly and is still halted on a troublesome 30% to 40%.

Diagnosis

Peritonitis can be recognized by symptoms of severe pain in the abdomen, which increases with body movement. Other symptoms include abdominal bloating, constipation, fever, nausea and vomiting, dizziness, shortness of breath, rapid pulse or breathing rate, dehydration and decreased urine production. Although most peritonitis diagnoses will be made by recognizing symptoms from medical
history and from performing a physical examination, diagnostic tests can assist and eventually confirm a peritonitis diagnosis. These tests include laboratory blood tests, peritoneum fluid analysis, computer tomography scans and ultrasound examination to identify signs of inflammation. Final diagnosis must be confirmed during laparotomy. A laparotomy is a surgery procedure that entails opening the abdomen and examining whether peritonitis is present and to what extent.

Treatments

If during the initial laparotomy the clinical diagnosis is confirmed, treatment starts immediately and the surgeon aims to get control of the infectious source. This becomes the main aim of the surgery. The type and extent of surgery depend on the underlying disease and the severity of infection. There is an ongoing interest to improve the survival rate by analyzing the pathogenesis and pathophysiology of secondary peritonitis. In 2001, a review called for multicenter studies to compare commonly accepted surgical strategies for source control and evaluation of secondary peritonitis.

On-demand or planned relaparotomy

In patients with severe secondary peritonitis (APACHE II score > 10), following the initial laparotomy, two clinically accepted surgical strategies are presently used in the operative management of peritonitis: planned relaparotomy and on-demand relaparotomy. Both strategies are based on the premise that some patients with peritonitis who have already undergone a primary laparotomy will need a relaparotomy due to persistent peritonitis. Presently surgeons choose between the planned relaparotomy and the on-demand strategy based on a mixture of factors involving the underlying disease process, the extent and severity of the intra-abdominal infection. The choice of surgical treatment is also highly dependent on the preference of the treating surgeon.

The planned strategy consists of relaparotomy performed every 36 to 48 hours after the initial laparotomy to inspect, drain, lavage, and perform other necessary abdominal interventions. This approach is terminated only when a macroscopically clean abdomen is found during relaparotomy, indicating negative findings (cessation of persistent/recurrent infection). In the on-demand relaparotomy strategy, a relaparotomy is performed only in patients with clinical deterioration or lack of clinical improvement with a likely intra-abdominal cause. Other infectious foci were ruled out using laboratory and imaging modalities.

Up till now, little is known about which of these surgical approaches is superior and remains a matter of heated debate in the surgical and intensive care world. Two recent retrospective studies have concluded conflicting results. The issue of timing and adequacy of relaparotomy is paramount because an improper, untimely, or incorrect operation may have a negative effect on patient outcomes.
Any delay in performing a relaparotomy in patients with an ongoing intra-abdominal infectious source may be detrimental, possibly lethal. This surgical dilemma is compounded by studies that have shown that intra-abdominal complications correlate strongly with the number of relaparotomies. Each relaparotomy is associated with a high risk of surgery-related complications. Therefore unnecessary relaparotomy should be avoided. Proponents of planned relaparotomy claim that repeated intra-operative peritoneal lavage is beneficial and that waiting for clinical symptoms of ongoing peritonitis often proves futile because the patient may reach a point of no return. Proponents of relaparotomy on-demand claim that a relaparotomy in patients in whom the inflammatory response is already switched-on may act as a ‘second hit’ precipitating multi-organ failure.

In search of an evidence-based surgical approach

Within the assessment of the effectiveness of medical treatments, including surgical procedures, a randomized controlled trial has been widely accepted. The need for randomized controlled trials to develop safe, evidence-based and cost-effective surgical practice as an essential part of the healthcare process to enable optimalization of individual patient care has become self-evident. If effectiveness is supported by evidence, then acceptance and implementation in the clinical practice can be expected to be more expeditious.

As well, due to increasing healthcare cost pressures worldwide, economic issues have become important for the introduction of new innovations and treatment approaches. Therefore, health technology assessment has become an important factor in aiding decision-making in healthcare, especially for countries with publicly funded healthcare systems. Health technology assessment is a systematic evaluation approach based on a societal perspective, including interdisciplinary views on efficacy, effectiveness, health related-quality of life, costs, safety and ethical issues of the introduced procedure.

Although numerous clinical trials comparing different antibiotic regimens in secondary peritonitis have been published, presently there are no randomized controlled trials comparing surgical strategies. This is particularly problematic as surgery is the mainstay treatment of severe secondary peritonitis. Already in 1990 the Joint Working Party of the North American and European Surgical Infection Society published a report on the management of abdominal sepsis advocating the need for a randomized study of surgical strategies. Evidence-based surgical treatment is paramount for clinical practice. Therefore, members of the Dutch Peritonitis Study Group aimed at filling this information gap about which surgical option is the best treatment for patients with severe peritonitis.

In this dissertation we describe the collaborative work of the Dutch Peritonitis Study Group and present the results from the first multicenter randomized controlled trial comparing the on-demand and planned relaparotomy surgical strategies: the RELAP trial. From a clinical perspective, determining which surgical strategy reduces mortality and or morbidity is paramount. In this thesis the assessment of both strategies in the RELAP trial is performed from a health technology assessment perspective. Therefore, we also show data on effectiveness, safety, health related-quality of life and healthcare utilization.
Outline of the thesis

Part I: In search of an evidence-based surgical approach

The Dutch Peritonitis Study Group conducted a survey amongst surgeons in the Netherlands to determine which surgical treatment strategy was preferred to treat secondary peritonitis. The survey indicated that surgeons have a slight overall preference for the relaparotomy on-demand strategy. Gastrointestinal surgeons and surgeons working in regional and smaller hospitals were significantly more in favor of the on-demand surgical strategy than their counterparts. However, there was considerable variability in treatment decisions by surgeons, with the majority of surgeons choosing a particular treatment strategy for the individual patient based on peritonitis and surgical treatment characteristics.

In chapter 1, we introduce the importance of evidence-based medicine and healthcare technology assessment in surgical studies, and present two surgical strategies; planned and on-demand surgical strategies for the treatment of severe secondary peritonitis.

In chapter 2, we present the results of the first randomized controlled trial (RCT) comparing these two surgical strategies. The primary aim of this study was to compare disease-related morbidity and mortality; secondary aims were to compare healthcare utilization and societal cost between the planned and on-demand relaparotomy surgical strategies.

In chapter 3, we present the results of the HR-QoL component of the RELAP trial. Understanding the use of HR-QoL data for clinical decision-making and differentiating between the use of generic instruments and disease-specific instruments are topics that have been touched on frequently within the last ten years. These generic and disease-specific tools are meant to complement each other and in combination to support clinical decision-making, whilst giving clinicians a better understanding of the burdens of disease from the patient's viewpoint. In trials where the clinical effects are statistically comparable, differences in HR-QoL can become a decisive factor in treatment strategies.

The results of a detailed economic evaluation of the RELAP trial are presented in chapter 4. In the design of the RELAP trial, it was decided that the appropriate type of economic evaluation would be conditional upon the results of the primary endpoints (mortality and major disease-related morbidity) and HR-QoL. If one strategy was found to be clinically superior, but associated with a substantial increase in resource utilization, a cost-effectiveness analysis (CEA) would be required to combine clinical and economic outcomes.
Part II: Optimizing the on-demand strategy

On-demand relaparotomy for all severity of disease groups

In chapter 5 we verify whether mild peritonitis (defined by an APACHE II score of 10 or lower) has indeed been well managed with the on-demand strategy. We also examine whether initial severity of disease modifies the difference in treatment effect between the on-demand and planned relaparotomy strategies for in-hospital mortality and a complicated course of disease in more severe peritonitis. Careful monitoring of patients is critical in the on-demand strategy in order to identify patients that require a relaparotomy. Currently there are no specific scoring systems developed for this purpose. Therefore, in chapter 6 we evaluate whether widely available scoring systems developed to predict mortality in critically ill patients were of clinical value in selecting patients needing a relaparotomy by predicting ongoing intra-abdominal infection in patients with secondary peritonitis.

Part III: Health related-quality of life in patients with secondary peritonitis

In Part III we continue to concentrate on aspects of HR-QoL, including determining factors that are predictive of poor HR-QoL in patients surviving secondary peritonitis. Secondly, we present the development of a disease-specific questionnaire to determine HR-QoL in secondary peritonitis patients.

If strategy does not matter in patients’ HR-QoL, what does?

It is important to determine which factors influence HR-QoL for these patients. Accurately defining the characteristics of this heterogeneous group of patients is a prerequisite for improved HR-QoL. In chapter 7, we examine which factors play an important role in patient HR-QoL 6 months following initial laparotomy.

In chapter 8 we discuss the limitations of the generic HR-QoL questionnaires such as the EQ-5D and EQ-VAS. We then present a newly compiled disease-specific quality of life questionnaire for patients with secondary peritonitis. The use of a generic questionnaire allows comparison between different patient groups, whilst the use of this disease-specific instrument in the context of a clinical trial can allow for more adequate comparisons between the different treatment arms and could illustrate different recovery rates. So far there is limited knowledge about peritonitis-specific HR-QoL.

Part IV: Post-traumatic stress disorder in patients with secondary peritonitis

In Part IV, we continue to evaluate patient outcomes in peritonitis, by determining both retrospectively and prospectively whether patients suffer from post-traumatic stress disorder (PTSD) symptoms. We determine which factors are most predictive for increased PTSD, particularly reviewing ICU stay as
PTSD symptoms have been increasingly found in critically ill patients and patients admitted to the intensive care unit (ICU)\textsuperscript{55}.

In chapter 9 we present a study of the long-term prevalence of PTSD symptoms in patients following secondary peritonitis in general and compare the prevalence of PTSD-related symptoms between secondary peritonitis patients admitted to the ICU and patients admitted only to the surgical ward\textsuperscript{56}.

In chapter 10, we prospectively determine PTSD and depression symptoms in secondary peritonitis patients. As many, but not all, ICU patients suffer from PTSD symptoms, the aim of this paper is to aid surgeons, using a reasonably simple practical model (nomogram), in determining which patients are at higher risk for the development of PTSD symptoms following their illness.

In chapter 11, we cross-sectionally review the relationship between PTSD symptoms and HR-QoL as measured with the generic questionnaire (EQ-5D, EQ-VAS) and the disease-specific questionnaire (SP-QoL) presented in chapter 8. PTSD have been shown to play an intricate role in HR-QoL, reducing both social and emotional functioning\textsuperscript{57-60}, showing that patients without PTSD symptoms report HR-QoL scores equal to that of the general population\textsuperscript{61-65}. In this study we determine whether secondary peritonitis patients with increased PTSD symptoms also report impaired HR-QoL physical functioning and physical recovery\textsuperscript{66-67}.

In chapter 12, we discuss the challenges we faced whilst designing and running the RELAP trial. As well, we discuss ideas for future studies to optimize surgical treatment for secondary peritonitis patients, which could lead to a reduction in mortality, disease-related morbidity and an increase in HR-QoL. Finally we address methods in which PTSD symptoms could be reduced following ICU stay.
References

(38) Dowie J. Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions. Health Econ. 2002;11:1-8.
(40) Feeny D. Commentary on Jack Dowie, “Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions”. Health Econ. 2002;11:13-16.
(41) Guyatt G. Commentary on Jack Dowie, “Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions”. Health Econ. 2002;11:9-12.


“Surgeons must be very careful when they take the knife! Underneath their fine incisions stirs the Culprit - Life!”

Emily Dickinson
Part I
In search of an evidence-based surgical approach
Chapter 2
Comparison of on-demand versus planned relaparotomy strategy in patients with severe peritonitis. A randomized trial

Oddeke van Ruler, Cecilia W Mahler, Kimberly R Boer, E Ascelijn Reuland, Hein G Gooszen, Brent C Opmeer, Peter W de Graaf, Bas Lamme, Michael F Gerhards, E Philip Steller, JW Olivier van Till, Corianne A de Borgie, Dirk J Gouma, Johannes B Reitsma, Marja A Boermeester

For the Dutch Peritonitis Study Group

Abstract

Context: In patients with severe secondary peritonitis, there are two surgical treatment strategies following an initial emergency laparotomy: planned relaparotomy and relaparotomy only when the patient’s condition demands it (on-demand). The on-demand strategy may reduce mortality, morbidity, healthcare utilization, and costs. However, randomized trials have not been performed.

Objective: To compare patient outcome, healthcare utilization, and costs of on-demand and planned relaparotomy.

Design, setting and patients: Randomized, non-blinded clinical trial at 2 academic and 7 regional teaching hospitals in the Netherlands from November 2001 through February 2005. Patients had severe secondary peritonitis and an Acute Physiology And Chronic Health Evaluation (APACHE) II score of 11 or greater.

Intervention: Random allocation to on-demand or planned relaparotomy strategy.

Main outcome measures: The primary endpoint was death and/or peritonitis-related morbidity within a 12-months follow-up period. Secondary endpoints included healthcare utilization and costs.

Results: A total of 232 patients (116 on-demand and 116 planned) were randomized. One patient in the on-demand group was excluded due to an operative diagnosis of pancreatitis and 3 in each group withdrew or were lost to follow-up. There was no significant difference in primary endpoint (57% on-demand [n=64] vs. 65% planned [n=73]; p=0.25), or in mortality alone (29% on-demand [n=32] vs. 36% planned [n=41]; p=0.22) or morbidity alone (40% on-demand [n=32] vs. 44% planned [n=32]; p=0.58). A total of 42% of the on-demand patients had relaparotomy vs. 94% of the planned relaparotomy group. A total of 32% of first relaparotomies were negative in the on-demand group vs. 66% in the planned group (p<0.001). Patients in the on-demand group had shorter median intensive care unit stays (7 vs. 11 days; p=0.001), and shorter median hospital stays (27 vs. 35 days; p=0.008). Direct medical costs per patient were reduced by 23% using the on-demand strategy.

Conclusion: Patients in the on-demand relaparotomy strategy group did not have a significantly lower rate of death or major peritonitis-related morbidity compared with the planned relaparotomy group but did have a substantial reduction in relaparotomies, healthcare utilization, and medical costs.

Relap registration number ISRCTN 51729393
Introduction

Secondary peritonitis is notorious for its high mortality (20-60%), long hospital stays, and high morbidity due to the development of sepsis with multiple organ failure\(^1\). Secondary peritonitis accounts for approximately 9,3 / 1000 emergency hospital admissions in the United States\(^5\). In addition, a substantial number of patients (12-16%) undergoing elective abdominal surgery develop postoperative peritonitis\(^5\). Healthcare utilization due to secondary peritonitis is extensive, with operations to eliminate the source of infection (laparotomy)\(^7\) and multidisciplinary care in the intensive care unit setting.

After the initial (emergency) laparotomy, relaparotomy may be necessary to eliminate persistent peritonitis or a new infectious focus\(^8\-\(^10\). There are two widely used relaparotomy strategies: relaparotomy when the patient's condition demands it (on-demand) and planned relaparotomy. The aim in the on-demand strategy is to perform re-operation only in those patients who are likely to benefit from this surgery, such as those with clinical deterioration and persistent lack of improvement. In the planned strategy, a relaparotomy is performed every 36 to 48 hours for inspection, drainage, and peritoneal lavage of the abdominal cavity until findings are negative for ongoing peritonitis. The planned strategy may lead to early detection of persistent peritonitis or a new infectious focus but harbors the risk of potentially unnecessary re-explorations in critically ill patients, while the on-demand strategy harbors the risk of a potentially harmful delay in the detection of ongoing infectious sources\(^11\). Monitoring of patients by combining clinical criteria, laboratory, and computed tomography (CT) results makes adequate and timely identification of patients for relaparotomy within the on-demand strategy possible. Moreover, the on-demand strategy allows for a time window to perform less invasive percutaneous, CT-guided drainage of abscesses instead of relaparotomy. There is consensus that the preferred strategy for mild peritonitis (APACHE II [Acute Physiology And Chronic Health Evaluation] score < 10) is on-demand relaparotomy\(^6\-\(^14\).

Despite lacking evidence from randomized trials, overall support for the on-demand strategy even among patients with severe peritonitis is growing\(^4\,\(^12\,\(^13\,\(^15\,\(^16\,\(^\), but both strategies are still used side-by-side in clinical practice\(^17\,\(^18\). The debate on which strategy is preferred dates back as far as the call for a randomized trial comparing both strategies by the Joint Working Party of the North American and European Surgical Infection Society in 1990\(^19\).

We performed a randomized trial comparing the on-demand strategy with the planned relaparotomy strategy following initial emergency surgery in patients with severe secondary peritonitis (APACHE II score >10). This criterion in peritonitis has been associated with a predicted mortality of >30%\(^14\). The primary endpoints for the study were 12-month mortality and major peritonitis-related morbidity. Secondary outcomes were healthcare utilization and medical costs.
Methods

Design and eligibility

In this multicenter study, patients were randomly assigned to either an on-demand or a planned relaparotomy strategy. Patients were eligible if they were diagnosed with secondary peritonitis and required an emergency laparotomy (index laparotomy). Peritonitis was defined as intra-abdominal infectious disease, verified during surgery, caused by perforation or infection of a visceral organ, or ischemia/necrosis of part of the gastrointestinal tract due to strangulation or postoperative peritoneal infection.

An APACHE II score >10 in the initial 24-hour period was required. Exclusion criteria were age younger than 18 years or older than 80 years; peritonitis due to bowel perforation after endoscopy operated within 24 hours after perforation; abdominal infection due to continuous ambulatory peritoneal dialysis catheter; peritonitis caused by pancreatitis; expected survival of less than 6 months due to malignancy; severe brain damage due to trauma or anoxia; or imperative relaparotomy (e.g., gauze packing, stapled ends without re-anastomosis).

All randomizations were completed only if the clinical diagnosis of peritonitis was confirmed during the index laparotomy. Randomization was performed centrally at the Academic Medical Center, Amsterdam, The Netherlands, using a specialized computer-generated block sequence and stratified per study site according to the APACHE II score as minimization factor (11 to 20 vs. >20). The operating surgeon was unaware of the allocated treatment strategy while performing the initial emergency laparotomy.

Informed consent was obtained from the patient or from the legal representative when patients were temporary incapacitated due to the severity of their illness. The study was approved by the medical ethics committee of all participating centers.

Surgical treatment strategies

Planned relaparotomy
Relaparotomies were performed every 36 to 48 hours after the index laparotomy to inspect, drain, lavage, and perform other necessary abdominal interventions for residual peritonitis or new infectious focus. The sequence of planned relaparotomies was terminated when a macroscopically clean abdomen was found at relaparotomy, indicating negative findings. That decision was made by the operating surgeon.

On-demand relaparotomy
Relaparotomy was only performed in patients with clinical deterioration or lack of clinical improvement with a likely intra-abdominal cause. Other (intercurrent) infectious foci (e.g., pneumonia) were ruled out.
using laboratory tests, imaging modalities, or both. The decision to perform an on-demand relaparotomy was made by the multidisciplinary medical team. To guide the decision for re-operation, the following definitions of deterioration and lack of improvement were specified in the protocol.

Deterioration after the previous operation was considered if there was an increase of more than 4 points in Multiple Organ Dysfunction Score (MODS)\textsuperscript{20} or prespecified surgical emergencies (i.e., abdominal compartment syndrome; intra-abdominal bleeding with persistent decrease in hemoglobin despite replacement and hemodynamic instability; burst abdomen; perforation of visceral organ; anastomotic leakage; intra-abdominal abscess that cannot be drained percutaneously; ischemia/necrosis of a visceral organ).

Lack of improvement of clinical signs of persistent sepsis was considered if the Multiple Organ Dysfunction Score was unchanged (+/- 2 points) for at least 48 hours following the index laparotomy or the previous relaparotomy. Abscess detected at CT imaging with positive fine-needle aspiration results (Gram stain with evidence of bacterial involvement) that could not be drained percutaneously was another reason for relaparotomy.

All participating surgeons and institutions had experience with both strategies. It was not required by protocol to keep patients in the planned relaparotomy group mechanically ventilated between operations. The decision to perform additional procedures during the relaparotomy was left to the discretion of the operating surgeon. Standardized co-interventions included direct postoperative care, intensive care unit (ICU) care, use of corticosteroids, postoperative feeding and antibiotic treatment - all directed by physicians unrelated to the study.

Outcomes and follow-up

The primary endpoint was a combination of all-cause mortality and major disease-related morbidity in surviving patients within 12-months follow-up after index laparotomy. A major morbidity endpoint in survivors was counted only if a prespecified major disease-related morbidity led to a surgical re-intervention during index admission or readmission during the 12-months follow-up (with or without the need of surgical intervention) (Appendix 1).

Additional outcomes included healthcare utilization and direct medical costs during a 12-months follow-up.

Costs

Cost-minimization analysis was used to determine economic differences\textsuperscript{21} comparing the on-demand with the planned strategy. Direct medical costs were estimated using primary data on resource utilization and included relaparotomies, percutaneous interventions, diagnostic CT scans, length of hospital stay, ICU stay with and without mechanical ventilation, days in hospital due to readmissions, administration of antibiotics, elective re-operations, length of stay in rehabilitation centers, healthcare provided by district nurses, and enterostomy care during the 12-months follow-up.
Costs per patient were calculated by multiplying volumes of resource with unit costs. Costs were assessed according to the Dutch guidelines for pharmaco-economic research. Dutch guideline unit costs were used for ICU stay, hospital stay, antibiotic medication, blood products, and visits to primary and outpatient healthcare clinicians. Unit costs for surgical procedures, enterostomy care, and diagnostic procedures were determined at the Academic Medical Center, Amsterdam, The Netherlands.

**Statistical analysis**

The sample size calculation was based on superiority of the on-demand strategy with a 10% absolute reduction of mortality and a 10% absolute reduction of morbidity in survivors, translating into a reduction in risk for the primary combined endpoint from 44% in the planned group to 28% in the on-demand group at 6-months follow-up. This expected difference of effect size in favor of the on-demand strategy was based on a systematic review and retrospective research performed in preparation of this trial. A sample size of 111 in each group would have a power of 80% to detect such a difference (one-sided alpha of 0.05). A dropout rate of 5% (6 patients per group) was expected.

All analyses were performed on the basis of the intention-to-treat principle. We extended the scope of the follow-up to 12 months to ensure capture of all relevant complications and healthcare utilization related to the initial episode of secondary peritonitis.

We compared the proportion of patients with a primary endpoint (including analyses of the separate components, mortality and morbidity in survivors) between the two strategies and tested for significance using the $\chi^2$ test. Confidence intervals for the difference between proportions were calculated with the use of a normal approximation of the binomial distribution. The number needed to treat was calculated by taking the reciprocal of the risk difference. Continuous data are presented as median with interquartile range (IQR). Survival curves were constructed with use of the Kaplan-Meier method and tested for differences using the log-rank test.

Differences in healthcare utilization were tested for significance using the $\chi^2$ test or Mann-Whitney U test, where appropriate. Confidence intervals for differences in mean costs were based on log-transformed cost data.

Prespecified subgroup analyses were performed for the APACHE II score at the time of index operation and per including hospital. We used logistic regression models to perform a formal test for interaction to determine whether treatment effects differed significantly between these subgroups.

All statistic analyses were performed using SPSS for Windows version 12.1.2 (SPSS Inc, Chicago, IL) or SAS 9.1 (SAS Institute Inc, Cary, NC). $P<0.05$ was considered statistically significant.

**Data handling and trial monitoring**

Collection and evaluation of all data regarding the initial (index) admission period and follow-up, were performed by blinded investigators not involved with patient care. All primary endpoints were
cross-checked with data from primary sources (by an independent data manager blinded for treatment allocation). An independent data and safety monitoring committee blinded for treatment assignment evaluated progress of the trial and examined safety parameters at regular intervals (every 25 patients).

Results

Patient enrollment

All patients with secondary peritonitis were assessed for eligibility between November 2001 and February 2005 and 510 were registered. Of these registered patients 228 had one or more exclusion criteria, the main reasons being an APACHE II score of 10 or lower and age younger than 18 years or older than 80 years. A total of 232 patients were randomized (116 in each group) (Figure 1).

One patient in the on-demand group was diagnosed with pancreatitis during index laparotomy and excluded after randomization. Furthermore, in both groups one patient withdrew informed consent and two patients were lost to follow-up after the initial admission period. Therefore, data were available for the initial admission period from 229 patients and from 225 patients for the follow-up period.

Figure 1  Flow diagram summarizing inclusion, allocation and follow-up.

510 patients with abdominal sepsis assessed for eligibility

278 excluded
228 met exclusion criteria
131 APACHE II score < 11
39 aged < 18 y or > 80 y
58 other
43 refused informed consent
7 decision by surgeon

232 randomized

116 randomized to on-demand relaparotomy
1 excluded (peroperative pancreatitis)
4 did not receive on-demand strategy
2 Unintentionally receiving planned strategy by surgeon
2 surgeon decided to change strategy

112 included in primary analysis

1 withdrew from study
2 lost to follow up

116 randomized to planned relaparotomy
7 did not receive first planned relaparotomy
2 died on day of index operation
1 too ill for relaparotomy
2 surgeon decided not to reoperate
2 patient declined reoperation

113 included in primary analysis

1 withdrew from study
2 lost to follow up
### Table 1  Baseline characteristics.

Demographic and index laparotomy-related data of 229 patients randomly assigned to on-demand relaparotomy or planned relaparotomy at baseline.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>On-demand (n=114)</th>
<th>Planned (n=115)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age – median (IQR)</td>
<td>65 (53 - 75)</td>
<td>70 (60 - 75)</td>
</tr>
<tr>
<td>Male sex – no. (%)</td>
<td>53 (47%)</td>
<td>56 (49%)</td>
</tr>
<tr>
<td>APACHE II score – median (IQR)</td>
<td>14.5 (12 - 17)</td>
<td>16 (13 - 19)</td>
</tr>
<tr>
<td>APACHE II score group &gt; 20 – no. (%)</td>
<td>16 (14%)</td>
<td>19 (17%)</td>
</tr>
<tr>
<td>Mannheim Peritonitis Index – median (IQR)</td>
<td>27 (23 - 32)</td>
<td>29 (24 - 33)</td>
</tr>
<tr>
<td>Major comorbidity present – no. (%)</td>
<td>64 (56%)</td>
<td>72 (63%)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>21 (18%)</td>
<td>33 (29%)</td>
</tr>
<tr>
<td>Malignancy</td>
<td>30 (27%)</td>
<td>27 (24%)</td>
</tr>
<tr>
<td>Respiratory disease (COPD)</td>
<td>14 (12%)</td>
<td>17 (15%)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>9 (8%)</td>
<td>11 (10%)</td>
</tr>
<tr>
<td>Renal disease</td>
<td>10 (9%)</td>
<td>7 (6%)</td>
</tr>
<tr>
<td>Etiology of peritonitis – no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perforation</td>
<td>64 (56%)</td>
<td>69 (60%)</td>
</tr>
<tr>
<td>Anastomotic leakage</td>
<td>36 (32%)</td>
<td>27 (24%)</td>
</tr>
<tr>
<td>Ischemia</td>
<td>6 (5%)</td>
<td>8 (7%)</td>
</tr>
<tr>
<td>Inflammation</td>
<td>4 (4%)</td>
<td>5 (4%)</td>
</tr>
<tr>
<td>Other</td>
<td>4 (4%)</td>
<td>6 (5%)</td>
</tr>
<tr>
<td>Localization of peritonitis – no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower GI tract</td>
<td>71 (62%)</td>
<td>76 (66%)</td>
</tr>
<tr>
<td>Upper GI tract</td>
<td>30 (26%)</td>
<td>28 (24%)</td>
</tr>
<tr>
<td>Biliary tract</td>
<td>2 (2%)</td>
<td>5 (4%)</td>
</tr>
<tr>
<td>Appendix</td>
<td>3 (3%)</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Pancreas (no pancreatitis)</td>
<td>5 (4%)</td>
<td>2 (2%)</td>
</tr>
<tr>
<td>Gynecological</td>
<td>2 (2%)</td>
<td>2 (2%)</td>
</tr>
<tr>
<td>Upper and lower GI tract</td>
<td>1 (1%)</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Extent of peritonitis index operation* – no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 2 quadrants</td>
<td>44 (39%)</td>
<td>43 (38%)</td>
</tr>
<tr>
<td>Diffuse</td>
<td>70 (61%)</td>
<td>71 (62%)</td>
</tr>
<tr>
<td>Nature of contamination – no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clear</td>
<td>6 (5%)</td>
<td>8 (7%)</td>
</tr>
<tr>
<td>Turbid</td>
<td>18 (16%)</td>
<td>29 (25%)</td>
</tr>
<tr>
<td>Purulent</td>
<td>43 (38%)</td>
<td>32 (28%)</td>
</tr>
<tr>
<td>Fecal</td>
<td>43 (38%)</td>
<td>43 (38%)</td>
</tr>
<tr>
<td>Bile</td>
<td>4 (4%)</td>
<td>3 (3%)</td>
</tr>
</tbody>
</table>
**Baseline characteristics**

Study groups were comparable for all patient and index laparotomy characteristics (Table 1). The median age was 69 years (IQR, 58-75 years), 48% were men, and the median APACHE II score was 15 (IQR, 13-18). The most common cause for peritonitis was gastrointestinal perforation (58%). Prevalence of major comorbidity was high (60%). Patients with failure of elimination of the infectious focus at index laparotomy were defined as ‘no infectious focus found and therefore not treated’ or ‘focus not definitively eliminated during the initial laparotomy’; such patients were equally distributed over the two study groups.

**Treatment following emergency laparotomy**

Some patients did not receive the treatment to which they were randomized (4, on-demand; 7, planned; Figure 1). The total number of relaparotomies differed between the two strategies (Table 2): 113 in the on-demand group and 233 in the planned group (p<0.001). Forty-two percent of patients in the on-demand group underwent a relaparotomy. The proportion of patients with 3 or more relaparotomies was 9% in the on-demand group compared to 24% in the planned group (p<0.001). Regarding the first relaparotomy, negative findings (no signs of persistent peritonitis nor new infectious focus) were seen in 31% of the on-demand group and in 66% of the planned group (p<0.001). The proportion of patients with positive findings (persistent peritonitis or new infectious focus) at relaparotomy was comparable in both strategies (29% of the on-demand patients vs. 32% of the planned patients; p=0.60) (Table 2). While the overall number of diagnostic CT scans was comparable between the two strategies, CT- or ultrasound-assisted percutaneous drainage was less frequent in the on-demand group (Table 2).
Mortality and major morbidity

Combined primary endpoint

The primary endpoint combining mortality from all causes and major morbidity in survivors within 12 months after index laparotomy occurred in 57% (n=64) of the patients in the on-demand group and in 65% (n=73) of the planned group (risk difference, 7.5%; 95% confidence interval [CI], -5% to 20%; p=0.25; number needed to treat, 13) (Table 2).

Mortality

Cumulative mortality during 12-months follow-up was 29% (32/112) in the on-demand group and 36% (41/113) in the planned group, corresponding with a risk difference of 7.7% (95% CI, -7.5% to 16%; p=0.23) (Table 2). The Kaplan-Meier curve for long-term survival shows that most deaths occurred within the first 60 days after the initial emergency operation, with no difference in early mortality between the on-demand and planned strategy (p=0.55 for 60 days) (Figure 2).

Major morbidity

Morbidity in survivors occurred in 40% (32/80) of the on-demand patients and in 44% (32/72) of the planned patients (risk difference, 4.4%; 95% CI, -11% to 20%; p=0.58) (Table 2). During admission, the two most frequent causes of morbidity were perforation (on-demand 11/114 [9.6%] vs. planned 10/115 [8.7%]; p=0.80) and anastomotic leakage (on-demand 7/114 [6.1%] vs. planned 11/115 [9.6%]; p=0.34), both equally distributed among study groups. Readmission during follow-up was most frequently due to incisional hernia needing surgery (on-demand 13/112 [11.6%] vs. planned 15/113 [13.2%]; p=0.71).

Healthcare utilization and direct medical costs

The proportion of patients admitted to the ICU was comparable between the two strategies (on-demand 90% vs. planned 94%). However, patients in the on-demand group had a significantly shorter ICU stay (median, 7 days vs. 11 days in planned group; p=0.001). The median number of days that patients were mechanically ventilated was shorter for the on-demand group (5 days) vs. the planned group (8 days; p=0.007). Hospital stay for the initial admission period was also shorter for patients in the on-demand group (median, 27 days) than for patients in the planned group (median, 35 days; p=0.008). Forty-six percent of the on-demand patients were readmitted to hospital compared with 40% of planned patients in the planned group (p=0.39). Patients in the on-demand group were alive and outside the hospital for a median of 302 days during the year of follow-up, while patients from the planned group were alive and outside the hospital for a median of 284 days (p=0.09; Table 2). Healthcare utilization was significantly lower in the on-demand group; in particular, patients in the on-demand group had fewer relaparotomies and fewer days in the ICU and in the hospital (Table 2). ICU stay alone accounted for 35% to 40% of the difference in costs.

The mean direct medical costs per patient after 12 months of follow-up, including index admission period, were 23% lower in the on-demand group: €62,741 (US $86,077) when treated with the on-
demand strategy and €81,532 (US $11,858) when treated with the planned strategy, an absolute difference of €18,791 (US $25,780) per patient (95% CI, €6,819 [US $9,355] to €31,166 [US $42,758]).

**Predefined subgroup analyses**

Treatment effects with respect to the primary endpoint were comparable across APACHE II score subgroups. Treatment effects were also unchanged for the primary endpoint across the nine enrolling hospitals. All tests for interactions were not significant (p > 0.50).

**Figure 2  Survival in patients with secondary peritonitis.**

Kaplan Meier curves showing survival of patients assigned to the on-demand or planned relaparotomy strategy over 12 months of follow-up. In both groups 1 patient died on the day of the index laparotomy, leaving 111 and 112 patients at risk at day 0, respectively. Log-rank test, p = 0.33.

<table>
<thead>
<tr>
<th>Number at risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planned</td>
</tr>
<tr>
<td>On-demand</td>
</tr>
</tbody>
</table>
Table 2  Primary endpoint and healthcare utilization among secondary peritonitis patients randomly assigned to on-demand or planned treatment strategy.

Combined primary outcome consisted of mortality and major disease-related morbidity.

<table>
<thead>
<tr>
<th>Primary endpoint a</th>
<th>On-demand</th>
<th>Planned</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined primary endpoint</td>
<td>64/112 (57%)</td>
<td>73/113 (65%)</td>
<td>0.25</td>
</tr>
<tr>
<td>Mortality</td>
<td>32/112 (29%)</td>
<td>41/113 (36%)</td>
<td>0.22</td>
</tr>
<tr>
<td>Major morbidity in survivors (one or more)</td>
<td>32/80 (40%)</td>
<td>32/72 (44%)</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Risk difference (RD; 95% CI) 7.5% (-0.05 - 0.20)
Number needed to treat (NNT; 95% CI) 13 (<5 to 19)
Relative risk (RR; 95% CI) 0.88 (0.72 – 1.1)

Course of disease and healthcare utilization

| Total no. of relaparotomies (range per patient) b | 113 (0 – 14) | 233 (0 – 10) | <0.001 |
| No. of relaparotomies b | | | <0.001 |
| 0 | 66/114 (58%) | 7/115 (6%) |
| 1 | 27/114 (24%) | 59/115 (51%) |
| 2 | 11/114 (10%) | 21/115 (18%) |
| 3 or more | 10/114 (9%) | 28/115 (24%) |
| Negative findings when receiving a relaparotomy b | 15/48 (31%) | 71/108 (66%) | <0.001 |
| Positive findings per strategy | 33/114 (29%) | 37/115 (32%) | 0.60 |
| Days in ICU b median (IQR) | 7 (3 – 13) | 11 (6 – 25) | 0.001 |
| Days of mechanical ventilation b median (IQR) | 5 (2 – 11) | 8 (4 – 18) | 0.007 |
| Days in hospital b median (IQR) | 27 (15 – 47) | 35 (21 – 65) | 0.008 |
| Percutaneous interventions (no. patients) | | | |
| during index admission b | 27/114 (24%) | 43/115 (37%) | 0.024 |
| study period (12 months)a | 30/112 (27%) | 44/113 (39%) | 0.052 |
| Drain placement followed by continuous drainage a | 24/112 (21%) | 32/113 (28%) | 0.23 |
| Diagnostic imaging (Computer Tomography) a no. patients (%) | 72/114 (63%) | 71/113 (63%) | 0.96 |
| no. of procedures (range) | 158/114 (0 – 11) | 169/113 (0 – 8) | 0.33 |
| median (IQR) | 1 (0 – 2) | 1 (0 – 2) | 0.98 |
| No. of patients with readmissions (overall) a | 51/112 (46%) | 45/113 (40%) | 0.39 |
| Days alive and outside hospital a median (IQR) | 302 (65 – 338) | 284 (21 – 323) | 0.09 |

(RD = Risk Difference; CI = Confidence Interval; NNT = Number Needed to Treat; IQR = Inter-Quartile Range, RR = Relative Risk)

a Data including follow-up (n=225)  b Data on the index admission period (n=229)
Comment

This randomized trial found that compared to the planned relaparotomy, the on-demand strategy did not result in statistically significant reductions in the primary outcomes of death or major peritonitis-related morbidity but did result in significant reductions in the secondary outcomes of healthcare utilization, including the number of relaparotomies, the use of percutaneous drainage, and hospital and ICU stay. These are similar to the results of a retrospective study and a systematic review on this topic\textsuperscript{4,12}. Despite a lack of statistically significant improvement in primary clinical outcome these substantial reductions in healthcare utilization and costs with the on-demand strategy suggest that it may be the preferred strategy.

Some studies have described that the planned strategy increases the risk of multiple organ failure due to amplifying the systemic inflammatory response by multiple surgical lavages, leading to increased mortality, ICU stays, and hospital stays\textsuperscript{25-26}. We also observed that patients treated with the planned strategy had longer ICU stays and had a longer overall hospital stay. The duration of mechanical ventilation was significantly longer for the planned treatment patients. However, this difference in ventilation time may in part be related to the short period between scheduled relaparotomies in the planned strategy, inherent to the planned nature of the procedures, as a result of which some patients could not be weaned off ventilation before the next operation. The number of (minimally invasive) percutaneous interventions was also significantly lower in the on-demand group. Possibly, free abdominal fluids and abscesses were more frequent after multiple surgical interventions due to the re-interventions or a modified inflammatory response. Other potential drawbacks of the planned relaparotomy strategy are the observed strong adherence of microbes residing in the peritoneum making them resistant to peritoneal lavage\textsuperscript{27}. This reduces the effectiveness of the procedure and the damaging effects of lavage to the mesothelial layer may even reduce the innate resistance to infection\textsuperscript{26,28}.

In our trial, the on-demand and planned relaparotomy strategies were equally apt to identify patients with remaining or new intraabdominal infection after the index laparotomy. This also confirms that patients in the planned group were not more frequently determined to show positive findings due to differential verification. In other words, surgeons were not more inclined to determine a planned re-operation as positive, for example, to justify this scheduled intervention. Although the on-demand strategy reduced the number of relaparotomies, there was still a 31% chance that macroscopic findings were negative in patients selected for relaparotomy. The key challenge in the on-demand strategy is to adequately select patients for relaparotomy and to prevent potentially harmful delay in re-intervention by adequate and frequent patient monitoring\textsuperscript{29-31}. A more rigorous use of CT scanning as part of the procedure of selecting patients with abdominal sepsis for relaparotomy may well reduce the proportion of patients with negative findings at relaparotomy even further.
Multiple independent variables and combination of variables have been described to predict outcome of peritonitis\textsuperscript{5,17,29,31-36}. However, results in the literature are inconclusive and the majority of studies predict disease outcome (mortality) of sepsis rather than positive findings at relaparotomy in secondary peritonitis\textsuperscript{14,20,37-42}.

In attempt to guide decision-making for relaparotomy and enhance timing of relaparotomy in the on-demand group within this trial, we prespecified the main criteria for necessity of relaparotomy as lack of clinical improvement or clinical deterioration using a quantified method (the Multiple Organ Dysfunction Score). The European Sequential Organ Failure Assessment score\textsuperscript{40} may have been an alternative prognostic score for predicting presence of persisting peritonitis. However, although scoring systems were used, the final decision to perform a re-operation on a patient in the on-demand setting was always made within a multidisciplinary team. Therefore, considerations for relaparotomy concerning clinical, laboratory, and imaging parameters were less explicit, but in line with current clinical practice. Future research should focus on optimizing adequate and timely selection of patients for relaparotomy by identification of predictive variables for positive findings at relaparotomy and evaluating the added diagnostic value of diagnostic imaging and potential biomarkers.

One of the difficulties in research on secondary peritonitis is heterogeneity of the study population regarding, e.g., severity of disease, etiology, and localization of the infectious focus\textsuperscript{12,19,43}, which often make it difficult to extrapolate study results to individual patients in clinical practice. For this reason, we have excluded disease entities with substantially different prognosis and requiring different treatment strategy, like pancreatitis, perforation due to endoscopy operated within 24 hours, and catheter-associated peritonitis. We also excluded patients with a APACHE II score of 10 or lower\textsuperscript{14}, because the on-demand strategy is already the treatment of choice in mild peritonitis\textsuperscript{4}. Within the trial, we examined whether the treatment effect with respect to the combined endpoint differed with the severity of disease at index laparotomy. We found no indication that relative treatment effects differed between patients with moderate to severe disease (APACHE II score 11-20) and those with severe disease (APACHE II score >20). Treatment effects were also comparable across all enrolling hospitals, translating to the on-demand strategy being a feasible and valid option in both moderately and severe secondary peritonitis in every hospital setting.

In conclusion, this study found that patients in the on-demand relaparotomy group did not have a significantly lower rate of adverse outcome compared with patients in the planned relaparotomy group but did have a substantial reduction in relaparotomies, healthcare utilization, and medical costs. On-demand relaparotomy may therefore be considered the preferred surgical strategy in patients with severe peritonitis.
References


Chapter 3
Health related-quality of life in patients with severe peritonitis. A randomized trial comparing on-demand versus planned relaparotomy following the emergency laparotomy

Kimberly R Boer, Oddeke van Ruler, Brent C Opmeer, Hein G Gooszen, Peter W de Graaf, Erik J Hesselink, Michael F Gerhards, E Philip Steller, JW Olivier van Till, Dirk J Gouma, Patrick M Bossuyt, Mirjam A Sprangers, Johannes B Reitsma, Marja A Boermeester, Corianne A de Borgie

For the Dutch Peritonitis Study Group

Submitted
Abstract

**Objectives:** To compare health related-quality of life (HR-QoL) in patients with severe abdominal sepsis treated by the on-demand (OD) or planned relaparotomy (PR) strategy; and to evaluate the course of HR-QoL over time during a one-year follow-up.

**Design and setting:** Randomized controlled trial in 2 academic and 7 teaching hospitals in the Netherlands.

**Participants:** 229 patients with abdominal sepsis were randomized to either on-demand relaparotomy (n=114) or planned relaparotomy (n=115) surgical strategy following their initial emergency laparotomy.

**Outcome measures:** HR-QoL was measured at 3, 6, 9 and 12 months following initial laparotomy using both generic (Euroqol-Five Dimensions and Euroqol-Visual Analogue Scale) and disease-specific health related quality of life for secondary peritonitis questionnaires (SP-QoL).

**Results:** No differences between the two surgical strategies were found in the EQ-VAS or the EQ-5D at any time points. On the SP-QoL, differences in scores between the two surgical groups were only found for the subscales nutrition and worries at 6-months follow-up, but these differences were no longer apparent by 9- and 12-months follow-up. The other SP-QoL subscales (abdomen, body image and sexuality) showed no differences between the two intervention groups, at any time point.

Over time the EQ-VAS scores improved moderately with a small increase from a mean of 64 at 3 months to 69 at 12 months in the on-demand treated patients, and from 64 to 71 in the planned group. The absolute reductions in the percentage of patient reporting problems on the various dimensions of the EQ-5D between 3 to 12 months ranged from 5% to 28%. Symptom reductions in the self-care and daily activities dimensions were most prominent, whilst changes in the pain and mood dimensions were minimal. For the SP-QoL, the subscales nutrition and worries showed substantial improvement from month 3 to 12.

**Conclusions:** Differences in HR-QoL were small and non-significant between the two surgical strategies. This means that the small differences in major clinical outcomes in favor of the on-demand strategy were not offset by opposite changes in HR-QoL. We have demonstrated in our earlier study that a substantial reduction in healthcare resource utilization and costs renders the on-demand strategy treatment of choice. Improvements in HR-QoL over time remained minimal at one-year follow-up.
Introduction

The mainstay initial treatment for patients with abdominal sepsis is an emergency laparotomy to eliminate the source of peritoneal infection. However, following this emergency laparotomy, there is no consensus on the subsequent surgical strategy. Some surgeons perform a ‘second look’ operation approximately 48 hours following laparotomy (i.e., planned relaparotomy), whereas others reserve relaparotomy for patients showing deterioration or lack of improvement (i.e., relaparotomy on-demand). At present, both of these treatment strategies are still frequently practiced worldwide and the treatment choice is usually hospital- or surgeon-dependent. The strategy choice is a worldwide debate. Already in 1990 the Surgical Infections Society Europe called for a randomized controlled trial comparing these two strategies.

The absence of randomized controlled trial (RCT) based evidence in favor of one or the other strategy was considered the main reason for variability in treatment choice among surgeons. Therefore, our study group performed the first RCT comparing the on-demand relaparotomy strategy with the planned relaparotomy strategy following initial emergency laparotomy in patients with severe abdominal sepsis. The main clinical outcomes (all-cause mortality and major disease-related morbidity) were comparable between the two surgical strategies with only a small but non-significant benefit in favor of the on-demand strategy. Patients treated with the on-demand strategy had a significant reduction in healthcare resource utilization and related costs, predominantly due to shorter intensive care unit (ICU) and hospital stays.

In studies where the clinical outcomes were comparable (no significant treatment effect for the major clinical outcomes), health related-quality of life (HR-QoL) can play a critical role in deciding upon the best treatment strategy. HR-QoL has to be considered when evaluating healthcare utilization and costs, to avoid that a reduction in costs will negatively affect patient HR-QoL. In addition, knowledge about improvement and possible deficiencies in long-term HR-QoL is important to better inform patients and their relatives about the recovery process. Prospective data on HR-QoL in patients with abdominal sepsis are scarce.

We have observed longer ICU and hospital stays and more surgical interventions in the planned relaparotomy strategy. We have also seen that a more complicated course of the disease and a more lengthy hospital stay lead to reduced HR-QoL. Based on these findings we expect HR-QoL to be significantly better in the on-demand group although a trade-off may exist that if more patients survive in the on-demand strategy this occurs at the cost of decreased HR-QoL. Also previous studies have shown that HR-QoL remains poor in patients at 6 months after initial surgery and critical illness, and survivors from severe sepsis have shown reduced HR-QoL using generic HR-QoL instruments. To date, data at one-year follow-up and disease-specific data are lacking in patients surgically treated for secondary peritonitis.

Therefore this paper presents the HR-QoL results obtained from this multicenter RCT. The aims of this study were to compare HR-QoL between patients treated with the on-demand strategy and patients treated with the planned relaparotomy strategy. As well, we have evaluated the course of HR-QoL over time during a one-year follow-up using both a generic questionnaire and a disease-specific HR-QoL questionnaire.
Methods

Study design

We measured generic and disease-specific HR-QOL within a randomized trial comparing two surgical strategies following the initial emergency laparotomy for severe abdominal sepsis. This trial was initiated by the Academic Medical Center (AMC), Amsterdam, and was conducted in 2 academic and 7 regional teaching hospitals in The Netherlands. The primary endpoints of the trial were all-cause mortality and disease related-morbidity. The medical ethics committees of all participating centers approved this study. Details of the study design are described elsewhere6.

Patients

Patients were eligible for the trial if they were between 18 and 80 years old, had an APACHE II score >10, and were suffering from secondary peritonitis (abdominal sepsis) caused by perforation or inflammation of a visceral organ or by ischemia/necrosis of part of the gastrointestinal tract or by anastomotic leakage or non-localized abscesses after gastrointestinal surgery needing an emergency laparotomy. All patients were eligible for the HR-QoL component of this RCT, if they were alive and not hospitalized at the time of planned HR-QoL measurement. Consenting patients were randomly assigned to either the planned relaparotomy strategy or the relaparotomy on-demand strategy.

Surgical treatment strategies

In the planned strategy a relaparotomy was performed every 36-48 hours after their emergency (index) laparotomy to inspect, drain, apply lavage, and perform other necessary abdominal interventions until a relaparotomy resulted in negative findings, meaning a macroscopically clean abdomen without abnormalities or abscesses. Patients in the on-demand strategy only received a relaparotomy after the index laparotomy in case of clinical deterioration or lack of clinical improvement with a likely intra-abdominal cause (details presented in van Ruler et al.)6. All other co-interventions related to the treatment of abdominal sepsis including the use of corticosteroids, enteral or parenteral feeding, cultures, antibiotics, the use of octeotride and blood products were standardized in the study protocol.

Data collection

Data were prospectively collected for all eligible patients. Clinical data included postoperative complications, ICU stay, mechanical ventilation time, hospital stay, overall survival, and disease-free survival (details presented in van Ruler et al.)6. HR-QoL was assessed at 3, 6, 9 and 12 months following index laparotomy by administering the patient self-report Euroqol 5-Dimensions (EQ-5D) and the currently

40 Part I In search of an evidenced-based surgical approach
Euroqol-Visual Analogue Scale (EQ-VAS)\(^a\). A disease-specific HR-QoL questionnaire for patients with secondary peritonitis (Appendix 2)\(^{13}\) was also administered at these time points. Data were prospectively collected for all eligible patients. In case of non-response, patients received with a reminder by phone within two weeks; after one month of non-response a new set of questionnaires with a reminder letter was sent.

**HR-QoL questionnaires**

The *EQ-5D* measures five health dimensions: mobility, self-care, daily activities, pain/discomfort, and mood consisting of both anxiety and depression. For each dimension patients can report: 0 (no problems), 1 (some problems), and 2 (many problems).

The *EQ-VAS* is a thermometer-like scale, in which patients are asked to rate their overall well-being from 0 (worst imaginable overall health) to 100 (best imaginable overall health)\(^{14,15}\). These instruments have been validated in various populations, including Dutch healthy individuals, and were recently recommended as the instrument of choice in critical care studies\(^{16-18}\) ([www.euroqol.org](http://www.euroqol.org)).

*The SP-QoL questionnaire* was compiled specifically for peritonitis patients by selecting applicable questions from the original European Organisation for Research and Treatment of Cancer (EORTC) questionnaires (Appendix 2)\(^{13}\). The SP-QoL questionnaire consists of 7 subscales, including complaints concerning the abdomen (AB, 4 questions), nutrition (NU, 5 questions), worry (WO, 6 questions), body image (BI, 3 questions), sexuality (SEX, 2 questions), defecation problems (DEF, 7 questions) and stoma-related problems (SRP, 7 questions). At each time point, patients without an enterostomy completed the defecation subscale, whilst patients with an enterostomy completed only the stoma-related subscale. These groups differed at each time point, as enterostomies were often temporary. All questions employed a four-point response scale ranging from 1 (not at all) to 4 (very much) taken from the EORTC questionnaires. Responses within each subscale are summed and linearly transformed to create a score between 0 (best imaginable) and 100 (worst imaginable).

**Statistical analysis**

In general, data contained repeated measurements on the same patient, and were therefore analyzed using linear mixed models for continuous data (EQ-VAS, SP-QoL subscales, except stoma-related problems and defecation) and generalized estimating equations models (GEE) for the EQ-5D dimensions (dichotomised into having ‘no problems’ [scores 1] and ‘some problems’ [scores 2 and 3])\(^{19}\). These models explicitly estimate and take into account the correlation that is likely to exist between measurements within the same individual\(^{20}\). No mathematical pattern was imposed on the covariance structure for measurements within the same individual (unstructured variance-covariance matrix). To study the trend over time in scores between the two treatment arms, the mixed model contained treatment (on-demand vs. planned), time (categorical with 4 levels), and the interaction between treatment and time.
Defecation and stoma-related questions were analysed in a slightly different manner, because whether or not patients have an enterostomy changed during the 12-month follow-up. Patients with an enterostomy only filled in the stoma-related problems subscale, whilst those who do not have an enterostomy filled in the defecation subscale. Patients, whose enterostomy status changed, did not respond to the same subscales at all four time points, making mixed models less appropriate for these subscales. Therefore, these subscales were analysed cross-sectionally at two time points (6 months and 12 months) using an ANOVA model.

**Differences in HR-QoL between strategies**

For the comparison of HR-QoL in the on-demand and the planned relaparotomy group we performed a single overall test jointly testing whether all differences in HR-QoL between the two strategies were zero. Only if the overall test was significant, we tested for differences in HR-QoL at specific time points to avoid (reduce) the problem of multiple testing. All analyses were performed on an intention-to-treat basis.

**Course of HR-QoL over time**

The same linear mixed and GEE models were used to examine trends over time. However, if the results of the previous analysis did not reveal an overall difference in HR-QoL between the two strategies, the models were simplified by modelling HR-QoL as a function of time only.

SAS version 9.1 was used for all data analyses. P-values < 0.05 were considered to be statistically significant.

**Results**

A total of 229 patients were included in the initial RELAP trial, of whom 114 received the relaparotomy on-demand strategy and 115 received the planned relaparotomy strategy. Cumulative mortality during 12-month follow-up was 29% (32/112) in the on-demand group and 36% (41/113) in the planned group (p=0.23). Prespecified major disease-related morbidity in survivors (Appendix 1) was observed in 40% (32/80) of the on-demand patients and in 44% (32/72) of the planned surviving patients (p=0.58). Combined mortality and morbidity occurred in 57% of the patients in the on-demand group and in 65% of the planned group (p=0.25)6.

Patient eligibility and responses to questionnaires for the HR-QoL measurement of the study by surgical strategy are presented in Table 1. Of the 229 patients who were randomly assigned to the treatment groups and still alive, 153 patients responded to one or more time points: 77 patients in the on-demand group and 76 patients in the planned group (Table 2). Patients, who were still in hospital, were not sent questionnaires. Patients that were at home but too ill to complete the questionnaires were considered non-responders, but were not sent a reminder.
### Chapter 3  Comparing surgical strategies: HR-QoL

#### Table 1

<table>
<thead>
<tr>
<th>Month</th>
<th>Number of patients randomized, eligible and responding to HR-QoL questionnaires at each time point.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Month 3</td>
</tr>
<tr>
<td>Median time of response since index surgery (P25 to P75)</td>
<td>95 days (85 to 111)</td>
</tr>
<tr>
<td>OD PR</td>
<td>OD PR</td>
</tr>
<tr>
<td>Patients randomized</td>
<td>114</td>
</tr>
<tr>
<td>Cumulative mortality n</td>
<td>25</td>
</tr>
<tr>
<td>In hospital n</td>
<td>12</td>
</tr>
<tr>
<td>Eligible patients n</td>
<td>77</td>
</tr>
<tr>
<td>Too ill † n</td>
<td>3</td>
</tr>
<tr>
<td>Refused † n</td>
<td>6</td>
</tr>
<tr>
<td>LTF* n</td>
<td>2</td>
</tr>
<tr>
<td>Other ** n</td>
<td>3</td>
</tr>
<tr>
<td>Total Response n (% of eligible)</td>
<td>63</td>
</tr>
<tr>
<td>Completed VAS n</td>
<td>59</td>
</tr>
</tbody>
</table>

† During the 12-months follow-up, if a patient did not return their questionnaire at one of the four time points, they were called to ascertain the reason, in these cases patients indicated that they were either too ill to respond to the questionnaire or they refused because they no longer wanted to continue with the HR-QoL component of the study;

* LTF = lost to follow-up or incorrect address; OD = relaparotomy on-demand; PR = planned relaparotomy;

** Other includes: Alzheimer’s patient, mentally handicapped patient, and patients unfit to fill in the questionnaire.

Baseline characteristics of the patients who responded to the HR-QoL component of this study are presented in Table 2. In general, differences between the two surgical strategies were small. Severity of disease was comparable between the two treatment strategy groups (Table 2).

As expected, in line with the treatment strategy definitions, only 35% of the on-demand treated HR-QoL responders received at least one relaparotomy, whilst almost all (96%) of the planned treated HR-QoL responders received at least one relaparotomy. For patients who responded to the HR-QoL component of the study, the median length of ICU stay was 7 days for the on-demand group and 10 days for the planned group (p=0.004). Similarly, the median index hospital stay for the on-demand group was 26 days, whilst for the planned group 41 days (p=0.022) (Table 2).
Table 2  Baseline characteristics and operative findings during emergency relaparotomy in patients responding to at least one HR-QoL measurement.

<table>
<thead>
<tr>
<th>Patient characteristics</th>
<th>OD</th>
<th>PR</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 77</td>
<td>n = 76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age; mean (SD)</td>
<td>61.6 (14.1)</td>
<td>64.8 (13.5)</td>
<td>0.157</td>
</tr>
<tr>
<td>Males; n (%)</td>
<td>41 (52)</td>
<td>38 (48)</td>
<td>0.688</td>
</tr>
<tr>
<td>APACHE II score mean (SD)</td>
<td>14.5 (3.1)</td>
<td>15.7 (4.6)</td>
<td>0.077</td>
</tr>
<tr>
<td>Mannheim Peritonitis Index (MPI) mean (SD)</td>
<td>19.2 (7.5)</td>
<td>20.5 (7.5)</td>
<td>0.271</td>
</tr>
<tr>
<td>≥ 1 major comorbidity; n (%)</td>
<td>44 (57)</td>
<td>46 (61)</td>
<td>0.671</td>
</tr>
</tbody>
</table>

**Findings at emergency relaparotomy**

<table>
<thead>
<tr>
<th>Etiology for peritonitis n (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflammation</td>
<td>3 (4)</td>
</tr>
<tr>
<td>Perforation</td>
<td>40 (52)</td>
</tr>
<tr>
<td>Ischemia / necrosis</td>
<td>2 (3)</td>
</tr>
<tr>
<td>Anastomotic leakage</td>
<td>29 (38)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (4)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hospital acquired peritonitis following surgery n (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>46 (60)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Main clinical outcomes</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with ≥ 1 relaparotomy n (%)</td>
<td>27 (35)</td>
</tr>
<tr>
<td>Number of relaparotomies; median (range)</td>
<td>0 (0-14)</td>
</tr>
<tr>
<td>Patients admitted to ICU n (%)</td>
<td>67 (87)</td>
</tr>
<tr>
<td>Length of ICU stay; median (P25 –P75)</td>
<td>7 (4 to 12)</td>
</tr>
<tr>
<td>Patients ventilated n (%)</td>
<td>67 (87)</td>
</tr>
<tr>
<td>Duration of ventilation; median, (P25 –P75)</td>
<td>5 (3 to 9)</td>
</tr>
<tr>
<td>Length of Hospital stay; median, (P25 –P75)§</td>
<td>26 (16.5 to 45.5)</td>
</tr>
<tr>
<td>Patients readmitted ≥ 1 at 12-months n (%)</td>
<td>41 (54)</td>
</tr>
<tr>
<td>≥ 1 Morbidities during 12-months follow-up** n (%)</td>
<td>31 (40)</td>
</tr>
<tr>
<td>Patients with enterostomy at 12-months n (%)</td>
<td>22 (36)</td>
</tr>
</tbody>
</table>

* Type of contamination missing one patient; ** Morbidities missing for two patients; § For two patients the exact hospital stay was unknown, due to transfer to other hospital; ‡ Non-parametric testing using the Mann-Whitney U test
Differences in HR-QoL between strategies

Overall HR-QoL, as measured by the EQ-VAS (0-100), improved slightly in the on-demand group from a mean value of 64 at 3 months to 69 at 12 months, these changes were similar in the planned group from 64 to 71. The overall test for differences was therefore not significant (p=0.86, Figure 1). The results of the EQ-5D questionnaire are shown in Figure 2. None of the EQ-5D dimensions, mobility (p=0.15), self-care (p=0.49), daily activities (p=0.59), pain (p=0.51) or mood (p=0.36) differed between patients who received the on-demand strategy and patients who received the planned strategy over the 12-months follow-up.

With the disease-specific SP-QoL, the on-demand treated group reported more problems on the nutrition subscale (p=0.031) compared to patients treated with the planned strategy. These differences were found specifically at 6-months follow-up, but were no longer present at 9- and 12-months follow-up. On the other SP-QoL subscales patients’ scores were similar between the two surgical strategies at all time points (Figure 3). Neither stoma-related problems nor the defecation subscale differed at follow-up 6 months or 12 months (Table 3).

**Figure 1**  *Model based means for scores for EQ-VAS at 3, 6, 9 and 12 months per surgical treatment.*

Between the surgical strategies at all time points p=0.86
Increase over time from 3 to 12 months for both arms p=0.023

Scores ranging from 0 to 100, with 0 indicating worse possible HR-QoL.
Course of HR-QoL over time

As there were no significant differences in HR-QoL between the two surgical strategies, we now report changes over time for the on-demand treated and planned group combined to increase precision and to avoid redundancy in results.

There was a significant time effect on the EQ-VAS in both groups. Patients reported better overall scores (71.3, SD ±16.1) at 12-months compared to 3-months follow-up (64.7, SD ±16.5) on the EQ-VAS (p < 0.001, Figure 1). Although scores increased by approximately 5 points, they remained substantially below the scores of a Dutch general population (mean EQ-VAS score of 79.7 (SD ±15.9))²¹.

On the EQ-5D, an average of 65% of the secondary peritonitis patients reported moderate to severely impaired mobility at 3 months, and although there was considerable improvement over time, still more than 50% of the patients continued to report impaired mobility at 12 months (absolute reduction of 15%, p = 0.023, Figure 2). Thirty-five percent of the secondary peritonitis patients reported problems with self-care at 3 months, with considerable improvement at 12 months (absolute mean reduction of 12%, p = 0.007, Figure 2). The majority of patients reported problems with daily activities at 3 months (80%), and although again there was considerable improvement over time with an absolute mean reduction of 30%, (p < 0.001, Figure 2). Approximately, 55% of patients reported pain at 3 months and

* Overall test for differences between on-demand strategy and planned strategy at all four-time points; results depicted are modelled means using a mixed model
** Test for overall (both groups combined) for trend over time
there was no improvement over time (absolute mean reduction of 5%, p=0.524). Forty percent of patients reported mood problems at 3 months, whilst at 12 months 30% reported mood problems with an absolute mean reduction of 10%, p=0.053; Figure 2).

Regarding the SP-QoL, there were significant reductions in complaints from 3-months to 12-months follow-up on the nutrition (mean reduction of 8%, p<0.001), worries (mean reduction of 20%, p<0.001) and sexuality (mean reduction of 6%, p=0.026) subscales. On the abdomen and body image subscales there were no statistically significant improvements over time between patient responses at 3 months compared to responses at 12 months (p>0.10, Figure 3).
Discussion

This study presents the HR-QoL results from the first randomized controlled trial comparing the on-demand versus the planned relaparotomy strategy in patients after their initial emergency operation for severe abdominal sepsis. The differences in HR-QoL between the two strategies were limited, on both the EuroQol, which measures generic HR-QoL, and the SP-QoL, a disease-specific questionnaire for patients with secondary peritonitis. On the EQ-5D and EQ-VAS there were no clinically significant differences between the surgical strategies for any of the follow-up time-points. On the SP-QoL questionnaire patients in the on-demand group reported more worries and abdominal and nutritional problems at 3 and 6 months following the initial laparotomy. However, at later follow-up measurements (9 and 12 months), these differences were no longer apparent.

Although clinical outcomes (mortality and major morbidity) were not significantly different between the on-demand and planned strategies, the number of relaparotomies and the duration of ICU and hospital stay were significantly reduced for patients treated by the on-demand strategy. In a previous paper examining HR-QoL only at 6-months follow-up for the total group, we did find that length of ICU stay and overall hospital stay were predictive factors for reduced HR-QoL. Based on these findings we expected to find an overall reduced HR-QoL in the planned relaparotomy group compared to the on-demand group during follow-up. This was not the case. When comparing the two strategies in the

---

* Overall test for differences between on-demand strategy and planned strategy at all four-time points; results depicted are modelled means using a mixed model

** Test for overall (both groups combined) trend over time

---

Figure 3  Percentage of patients with secondary peritonitis reporting problems on each SP-QoL subscale at 3, 6, 9, 12 months following emergency relaparotomy per surgical strategy group (not including the stoma-related problems and defecation subscales).

---

```
Percentage of patients reporting problems

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>6</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdomen</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nutrition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worries</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

p = 0.076*  
p = 0.099**

p = 0.031*  
p < 0.001**

p = 0.104*  
p < 0.001**
```
clinical trial the overall difference in hospital stay was 26 versus 41 days in favor of the on-demand strategy. A possible explanation is that in both groups hospital stay including an ICU period with mechanical ventilation after acute surgery already had such a large baseline impact on HR-QoL that a reduction in length of stay or number of interventions was not translated into an improved HR-QoL. Across treatment strategies differences between patients with an uncomplicated course of disease and shorter ICU and hospital stays versus those with a complicated course or disease and longer stay were apparent. The extent of the differences between individuals was probably larger than the observed differences between the treatment groups (4 days, Table 2).

In the mixed models we assumed that missing values were random and allowed the mixed model to account for potential differences. No imputation for missing values was used in this data set. For the stoma-related and defecation subscales we assumed selective missing data and therefore adjusted our analyses accordingly. Missing data in HR-QoL studies always play an important role. In the planned relaparotomy group a higher percentage of patients returned their questionnaires (7% difference) at 12-month follow-up than the on-demand patients. Although mortality was not significantly different (8% difference), it may have nevertheless been different enough to create a small bias in HR-QoL data. This could mean that patients who were more severely ill in the planned relaparotomy may have already died, and that patients in the on-demand strategy, who survived may report more impaired HR-QoL. This could account for the lack of differences between the on-demand and planned treatment strategies.
Regarding the evaluation of HR-QoL over time in all peritonitis patients, we found that the activities in daily life dimensions (mobility, self-care and daily activities) improved from 3 to 12 months. However, these activities continued to remain restricted in both groups at 12 months following initial surgery; with over 50% of patients still reporting problems. Difficulties could be due to the lengthy recovery time, numerous readmissions for disease-related morbidity (>60%) and still having an enterostomy (>35%) at the 12-month follow-up. The percentage of patients reporting activities in daily life problems is higher than those reported retrospectively by Bosscha in peritonitis patients, but more in line with recent results in patients with general sepsis.

Mood problems did not decrease over time. The number of patients complaining of mood problems remained substantially high even at 12 months following surgery (>30% at 12 months). This is consistent with number of PTSD and depression problems reported by critical ill patients in recent studies.

Over 50% of the patients continue to report pain problems at 12 months. This finding is informative for treating surgeons and general physicians. Chronic post-surgical pain is a syndrome recognized by the International Association for the Study of Pain, indicated by reporting pain for longer than 3 months. Chronic pain has been studied in patients that underwent hernia surgery and cardiovascular surgery, showing that approximately 40% of patients reported chronic pain. In a recent study in patients that underwent gastrointestinal surgery, 20% of patients continued to report chronic pain 4 years after surgery. Patients reported neuropathic pain in the abdominal region or adjacent to their old surgical wound. In our population we found slightly higher numbers of patients that continued to report pain, but we only recorded crude information on pain using just one question from the generic HR-QoL questionnaire. Future studies should be committed to assess the extent and nature of pain in severe peritonitis patients more extensively.

Most of the disease-specific subscale scores, including nutrition, worries and sexuality, of the SP-QoL improved significantly over time. However, the abdomen and body image subscales showed only minor improvement over time. This may indicate that patients continue to have problems with their body

### Table 3: Model based SP-QoL subscale, stoma-related subscale and defecation subscale, scores per treatment arm at 6 and 12 months following index operation.

<table>
<thead>
<tr>
<th>SP-QoL subscales</th>
<th>6 months</th>
<th>12 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defecation subscales</td>
<td>OD 17.4</td>
<td>16.4</td>
</tr>
<tr>
<td></td>
<td>PR 11.7</td>
<td>12.6</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>0.087*</td>
</tr>
<tr>
<td>Stoma-related problems</td>
<td>OD 38.4</td>
<td>30.4</td>
</tr>
<tr>
<td></td>
<td>PR 38.7</td>
<td>37.9</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>0.952*</td>
</tr>
</tbody>
</table>

* Individual ANOVAs for the follow-up timepoints 6 months and 12 months.
image, probably due to scarring and/or herniation. The scars can be a constant reminder of the patients’ ordeal\textsuperscript{32,33}. As well, a large percentage of patients still have an enterostomy at 12-months follow-up. Disease-specific information related to the impact of secondary peritonitis can support the treating physician in understanding the recovery process of these patients\textsuperscript{34-36}. Especially, since these are issues that the patient themselves may not offer as information, due to embarrassment, or not wanting to bother the physician with non-life threatening complaints\textsuperscript{37}. In this study we found that approximately 40\% of patients still reported problems with sexuality at 12-months follow-up. A recent self-report-based study on sexual dysfunction found similar results in survivors of intensive care treatment, showing that approximately 40-45\% of both patients and partners were unsatisfied with their sexual life\textsuperscript{38}.

This study also reiterates the necessity of long-term follow-up in HR-QoL research. Patients were followed for 12 months after their index surgery, if we had ended the follow-up at 6 months, we would have found HR-QoL differences between the two surgical strategies that had however disappeared by 9- and 12-months follow-up.

The acute and severe nature of abdominal sepsis prevented us from obtaining HR-QoL information at baseline. Although this is not a problem in determining the differences between the two surgical strategies, we were not able to assess the importance of existing problems in HR-QoL in the recovery of these patients.

In conclusion, differences in HR-QoL were small and non-significant between the two surgical strategies. This means that the small differences in major clinical outcomes in favor of the on-demand strategy were not offset by opposite changes in HR-QoL. We have demonstrated in our earlier study that a substantial reduction in healthcare resource utilization and costs renders the on-demand strategy treatment of choice, and health related-quality of life does not influence this choice. It is also of note that, although HR-QoL improves somewhat over time, peritonitis patients still report considerable reduced HR-QoL one year following surgery, regardless of the surgical strategy applied.
References


11. Granja C, Dias C, Costa-Pereira A, Sarmento A. Quality of life of survivors from severe sepsis and septic shock may be similar to that of others who survive critical illness. *Crit Care.* 2004;8:R91-R98.


Chapter 4
Costs of relaparotomy on-demand versus planned relaparotomy in patients with severe peritonitis. An economic evaluation within a randomized controlled trial

Brent C Opmeer, Kimberly R Boer, Oddeke van Ruler, Johannes B Reitsma, Hein G Gooszen, Peter W de Graaf, Bas Lamme, Michael F Gerhards, E Philip Steller, Cecilia W Mahler, Dirk J Gouma, Patrick M Bossuyt, Corianne A de Borgie, Marja A Boermeester

For the Dutch Peritonitis Study Group

Submitted
Abstract

**Background and aims:** We have recently reported the results of the first randomized trial comparing on-demand versus planned relaparotomy strategy in patients with severe peritonitis. We now provide a detailed evaluation of healthcare utilization and costs associated with both strategies.

**Methods:** The economic evaluation was based on primary resource utilization data collected within the framework of a multicenter, randomized controlled trial. Direct medical, non-medical and indirect costs associated with both strategies were estimated from a societal perspective up to one year following the index operation. The analysis was set up as a cost-minimization analysis because no clear differences in primary outcomes were observed between the on-demand and planned groups, but substantially less resources were utilized in the on-demand group. We compared the two strategies in terms of mean costs per patient, as well as in the distribution of costs across patients.

**Results:** Mean total costs per patient generated by resource utilization within one-year follow-up were substantially lower in the on-demand group (mean €65,812), as compared to the planned relaparotomy group (mean €83,476), resulting in a mean absolute difference of €17,664 per patient (95% CI €5,056 to €29,004). Relative differences (approximately 21%) were only marginally affected by sensitivity analyses using different assumptions. Planned relaparotomy was consistently generating more costs across the whole range of different courses of disease (quick recovery and few resources used on one end of the spectrum, to slow recovery and many resources used on the other end).

**Conclusion:** Considering that patients in the on-demand group had a lower (although not significant) rate of adverse clinical outcomes compared to the planned group, the substantial reduction in societal costs renders the on-demand strategy a far more efficient relaparotomy strategy in patients with severe peritonitis.
Introduction

Secondary peritonitis or abdominal sepsis is a serious condition with high mortality (estimates vary between 20-60%) and considerable major disease-related morbidity. Patients with severe peritonitis require intensive monitoring and medical treatment, often including lengthy ICU stays. With an estimated incidence for the United States of 9.3 cases of patients with secondary peritonitis per 1000 emergency hospital admissions, these patients incur substantial costs to the healthcare system.

The debate about the preferred relaparotomy strategy (on-demand vs. planned) in these patients is longstanding with both strategies having their proponents. We have recently published the results of the first randomized trial comparing these two surgical strategies and shown that patients in the on-demand group did not have a significantly lower rate of adverse clinical outcomes compared to the planned group, but that substantially less healthcare resources, such as ICU or hospital days were utilized in the on-demand group.

In this paper we present a detailed description and comparison of healthcare resource utilization and costs between the on-demand and planned relaparotomy strategy. This economic evaluation will be performed from a societal perspective, including direct medical costs, direct non-medical costs and indirect costs generated by each surgical strategy.

Patients and methods

Design and eligibility

This economic evaluation was part of the RELAP-trial, a randomized controlled trial comparing the on-demand relaparotomy strategy with the planned relaparotomy strategy in patients with severe peritonitis. Details about the design, conduct, and major clinical findings of this trial have been reported elsewhere.

The appropriate type of economic evaluation is conditional on the results of the primary endpoints (mortality, major disease-related morbidity) and health related-quality of life (HR-QoL). If one strategy was found to be clinically superior, but associated with a substantial increase in resource utilization, a cost-effectiveness analysis (CEA) would be required to combine clinical and economic outcomes. In case of comparable outcomes in clinical effectiveness, or if one strategy dominated (more effective at less resources used), a cost-minimization analysis (CMA) would suffice. As the clinical endpoints in this study were comparable or even in favor of the on-demand strategy, and the on-demand strategy generated considerably less resource utilization, this economic evaluation was set up as a cost-minimization analysis.

A total of 229 patients from nine enrolling hospitals in The Netherlands were randomly assigned to either the on-demand strategy (n=114) or the planned relaparotomy strategy (n = 115). Main criteria for eligibility were a clinical diagnosis of secondary peritonitis (excluding Continuous Ambulatory Peritoneal Dialysis (CAPD) peritonitis and pancreatitis), a severe clinical condition indicated by an
Acute Physiologic and Chronic Health Evaluation (APACHE) II score of >10 and age between 18 and 80 years.

**Economic evaluation**

Healthcare utilization and other resources were prospectively documented for individual patients, using the registration forms of the clinical study, and acquiring data from additional sources where needed. The time window for the economic evaluation was 12 months following the initial emergency laparotomy, equal to the follow-up time for the clinical outcomes.

The cost-analysis was set up from a societal perspective, which consists of three cost categories. These include direct medical costs, direct non-medical costs and indirect costs. Direct medical costs are generated by healthcare utilization and include hospital and ICU admission periods, therapeutic and diagnostic procedures, medication, and visits to primary and paramedical healthcare providers after discharge during the one-year follow-up. Direct non-medical costs reflect utilization of non-healthcare resources generated by the disease, for example travel to and from healthcare providers. Indirect costs are associated with loss of productivity due to impaired ability to work.

**Resource utilization**

Data on resource utilization during the index hospital admission included the number of surgical interventions (including relaparotomies), percutaneous drainage procedures, diagnostic procedures (abdominal computer tomography (CT), ultrasound (US), plain X-abdomen), postoperative hospital stay from index laparotomy onward and ICU stay. All data were registered using a case record form (CRF). Resource utilization after discharge was documented in specifically developed, self-administered questionnaires that were sent to surviving discharged patients at 3-, 6-, 9- and 12-months follow-up. Patients reported use of health services, including hospital readmissions, visits to physicians, medical or paramedical specialists or other healthcare providers. Working status and absence from paid work was also documented using the Health and Labour Questionnaire. Questionnaires were sent by mail, with a reminder by phone within two weeks if there was no response. After one month of non-response a new questionnaire including a reminder letter was sent.

**Costs: unit prices**

Estimates of unit costs for stay on the ward, stay in the ICU, blood transfusions, outpatient healthcare, secondary and paramedical healthcare providers, travel expenses and productivity loss were based on Dutch reference data from the handbook of the Dutch Health Council. Dutch pharmaceutical unit cost listings were used to determine the cost of antibiotic medication. Insurance reimbursement fees were used to estimate costs for microbiology diagnostics and mechanical ventilation. Top-down cost calculations were used to determine cost of index laparotomy, second-look relaparotomy, percutaneous drainage procedures and diagnostic imaging. Average top-down unit-cost were used to estimate the costs of relaparotomies including other surgical procedures (e.g. enterostomy (re)construction,
Costs: calculations

Costs were calculated for individual patients by multiplying actually used healthcare resources and unit prices (economic valuations of each volume unit). Mean costs per patient within one year of follow-up for the on-demand and planned strategy groups were calculated. Data concerning post-discharge healthcare utilization were not always complete due to non-response to the self-administered questionnaires. In the case of non-response, the average resource utilization for an out-of-hospital day was estimated within each study/treatment group, and extrapolated over the total out-of-hospital period. Costs associated with loss of productivity due to illness or recovery in patients below the age of 65 were estimated based on patient reported absences from paid labor. Productivity costs were calculated by using a friction cost approach, which assumes that each employee is replaceable in the workforce. Therefore we valued productivity loss within the one-year follow-up only for the friction period of 154 days, which is the estimated time needed to find a replacement.

Statistical analysis

All analyses were performed according to the intention-to-treat principle. Data for resource utilization were entered and prepared for analyses using SPSS 12 (SPSS Inc., Chicago, IL), unit costs were calculated in MS Excel 2003, and cost calculations, statistical analyses including sensitivity analyses were conducted using SAS 9.1 (SAS Institute, Cary, NC).

Mean volumes of resource utilization and associated costs during index admission and one-year follow-up were estimated for the two surgical strategies. Differences in volumes were tested for significance using a non-parametric Wilcoxon Mann-Whitney test.

For the assessment of the total costs, the distribution of direct medical costs generated during index admission and follow-up, direct non-medical costs and indirect costs across individual patients was estimated for the two surgical strategies. These distributions were compared as mean differences with 95% confidence intervals (95% CI) based on geometric means.

In addition to the main cost analysis, we evaluated the robustness of our results using sensitivity analyses. The following assumptions and analytical options were explored: (1) For percutaneous drainage procedures, the AMC unit cost estimates were replaced by the reimbursement fees for this procedures. (2) To avoid that the encountered cost differences where direct results from the differentiation between academic and non-academic hospitals, a weighted average unit cost per hospital ward day was used. This average was weighted by the actual ratio of academic and non-academic hospital beds in The Netherlands. (3) We differentiated between unit costs of relaparotomies with and those without other surgical procedures (such as enterostomy (re)construction, abscess drainage, colon resection, etc.) instead of using a single all-in amount for all relaparotomies. (4A-F) To
enhance generalizability of the results to other countries with publicly financed healthcare systems, Dutch reference prices for ICU days were replaced by unit costs estimated for the UK, Germany, France, Norway, Austria and Canada, respectively\(^9\)\(^{19-23}\). Additionally, (5) we compared the total costs of the two strategies when disregarding the costs of relaparotomy procedures during the index admission because these differences were intrinsic to the strategy itself, as the planned strategy involved more procedures than the on-demand strategy.

Finally, we evaluated whether differences in costs were consistent across patients with varying clinical conditions, or whether these differences were more pronounced in clinically more severely ill patients. We hypothesized that the total costs are a proxy for severity of the clinical condition, as a more severe condition and/or difficult recovery would imply a more complicated course of disease and treatment, in turn generating more resource utilization and costs. Based on this assumption, we compared the distribution of costs across patients with different clinical conditions (from those patients with a relatively quick recovery and less resource utilization to patients with a more severe course of disease with slower recovery and extensive use of healthcare resources) between the two strategies. The comparison was graphically presented by ranking patients within each study group by their total costs, and comparing total costs of patients with similar ranks.

**Results**

**Main clinical findings**

A total of 229 patients were correctly randomized and included in this study. In both surgical strategies one patient withdrew informed consent and one patient was lost to follow-up, meaning that data on the initial admission were available for 229 patients (114 on-demand and 115 planned strategy) and data on the one-year follow-up for 225 patients (112 on-demand and 113 planned strategy).

Demographic and clinical baseline characteristics of these randomized patients are presented in Table 1 together with a summary of the main clinical outcomes. The results show that morbidity and mortality were comparable. More details on the clinical outcomes can be found in the trial publication\(^6\).

**Resource utilization and costs**

Data on healthcare utilization during the index admission and readmissions during the one-year follow-up were collected using the registration system of the clinical study. As a result, data available for analysis from the index admission were almost complete. Self-reported questionnaires with additional data on resource utilization during follow-up were available from 76 patients (92% of patients alive at 6 months) in the on-demand group and 74 (94% of alive patients) in the planned group.

Results of the cost analyses are presented in Table 2, reporting mean volumes, total costs, mean costs per patient per strategy for the on-demand group and the planned group for resource utilization during the index admission, and follow-up. Mean costs per patient associated with relaparotomy procedures
during index admission were estimated as €4,617 (index laparotomy plus 113 relaparotomies in 114 patients) for the on-demand group and €6,641 (index laparotomy plus 233 relaparotomies in 115 patients) for the planned group (p<0.001). Although in the on-demand group significantly fewer relaparotomies compromised additional surgical procedures (42 in the on-demand and 54 in the planned relaparotomy group; p=0.022), the associated mean costs per patient (€1,211 for the on-demand group and €1,543 for the planned group) of these additional procedures differed only marginally (€332).

To determine whether the lower number of relaparotomies in the on-demand group was compensated for by additional US- or CT-guided percutaneous drainages (PCD) of abscesses or (infected) abdominal fluids collections, we computed the number of radiological guided percutaneous interventions (with or without continuous drainage) during the index admission. The costs for PCD were €147 (27% received PCD) for the on-demand group and €233 (39% received PCD; p=0.038)) for the planned group. Estimated mean costs per patient associated with CT scans were €346 (CT scans in 64% of patients, total of 158 CT scans) and €367 (CT scans in 61% of patients; total of 169 CT scans) in the on-demand group and the planned group, respectively.

As a large majority of patients were admitted to the ICU, substantial costs were generated by ICU stay. Mean estimated costs per patient generated by ICU stay were €21,040 (90% of the patients; mean stay 12 days) in the on-demand group and €31,248 (94% of the patients; mean stay 18 days; p=0.001) for the planned relaparotomy group (mean difference €10,208). Costs associated with mechanical ventilation showed a similar picture, with mean costs estimated as €3,080 (mean 8.3 days) for the on-demand group and €4,360 (mean 12 days; p=0.004) for the planned group (mean cost difference €1,280). Costs generated by hospital stay on the ward (excluding ICU days) were estimated as €11,609 for the on-demand group and €11,784 for the planned relaparotomy group. Although the total length

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>On-demand</th>
<th>Planned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male / Female</td>
<td>53/61</td>
<td>56/59</td>
</tr>
<tr>
<td>Mean age in years (SD)</td>
<td>63 (14)</td>
<td>66 (12)</td>
</tr>
<tr>
<td>APACHE II &gt; 20, n (%)</td>
<td>16 (14)</td>
<td>19 (17)</td>
</tr>
<tr>
<td>Mean Mannheim Peritonitis index (95% CI)</td>
<td>27 (23 to 32)</td>
<td>29 (24 to 33)</td>
</tr>
<tr>
<td>One or more comorbidity present, n (%)</td>
<td>64 (57)</td>
<td>73 (65)</td>
</tr>
<tr>
<td>No. of deaths at 1 year (%)</td>
<td>32 (29)</td>
<td>41 (36)</td>
</tr>
<tr>
<td>Severe morbidity in survivors, n (%)*</td>
<td>32 (40)</td>
<td>32 (44)</td>
</tr>
</tbody>
</table>

*Number of surviving patients: 80 on-demand, 72 planned
of hospital stay for the index admission was substantially shorter for patients in the on-demand group (mean 38 days) than for patients in the planned group (mean 45 days), length of stay on the ward was comparable (mean 26 and 27 days respectively; p=0.21).

The on-demand group used substantially less medication and material (e.g., days with enterostomy care, blood transfusions), resulting in lower costs. The mean costs per patient for medication were €474 vs. €619, for blood products €89 vs. €178, and for enterostomy care €741 and €917 comparing the on-demand group and the planned relaparotomy group. The direct medical costs during admission were significantly lower in the on-demand group, with a mean difference of €14,688 (95% CI €5,225 to €22,915).

Costs associated with readmissions during one-year follow-up were comparable (mean number of hospital days 9.5 versus 11.8 days (p=0.88), with mean costs €4,280 for the on-demand patients versus €5,083 for the planned patients). On average, enterostomy care and associated costs during follow-up were considerable. Patients had an enterostomy for a mean of 140 days (€4,449) in the on-demand group and 150 days (€4,767) in the planned relaparotomy group (p=0.44).

Distinct differences in resource utilization and associated costs of outpatient and secondary healthcare providers were found for home care supplied by district nurses and stay in rehabilitation centers, where the on-demand group received less home care (45 hours per patient vs. 71 hours in the planned relaparotomy group; p=0.16) and had shorter stays in rehabilitation centers than planned relaparotomy group (19 versus 23 days; p=0.90). Consequently the associated costs were substantially lower in the on-demand group (mean costs home care on-demand €1,836 versus planned €2,947 and rehabilitation on-demand €6,480 vs. planned €8,040). Differences between the study groups in utilization and associated costs of outpatient care and visits to secondary healthcare providers (general practitioner, company doctor) were marginal. The direct medical costs during follow-up (between discharge and 12 months following the index laparotomy) were lower in the on-demand group with a mean difference per patient of €3,782 (95% CI €-2,945 to €6,774).

For lost productivity due to absence from paid work, the data indicated that among patients reporting to have paid work; only the occasional patient reported a return to work within 12 months. For all other patients aged less than 65 years, we assumed that they were at least absent for the full friction period (154 days), as hospital (re)admission status and poor health status were the main reasons for not sending or returning the questionnaire. The mean number of days absent from paid work was estimated as 70 days in the on-demand group, versus 50 days in the planned group (p=0.038). Costs associated with lost productivity for these days were estimated as €2,854 and €2,048, respectively.

Overall, for the whole study period, mean total costs per patient associated with the on-demand strategy were €65,812 vs. €83,476 with the planned strategy (absolute difference €17,664 95% CI €5,056 to €29,004; relative difference 21%). Of these total costs, 75% was generated during the initial admission, of which 45 to 55% were ICU costs.

Subsequent sensitivity analyses showed that the results of the cost analysis were robust for changes for the various assumptions (Table 3). Absolute estimates of total costs were found to change within a limited range for each strategy (less than €8,000 for on-demand, and less than €10,000 for planned
relaparotomy), whilst the relative difference between the two strategies remained stable (21-22%). The final question remains whether this difference is consistent across patients with different courses of disease and whether cost differences were predominantly found in cases with a more severe course of disease with slow recovery and hence extensive use of healthcare resources. Figure 1 shows the distribution of total costs per patient in each group after patients have been ranked according to their total costs. Costs were found to be consistently lower in the on-demand group compared to the planned relaparotomy group across the whole range of costs. This suggests that if we employ the total costs as a proxy for the (complicated or uncomplicated) courses of disease that costs were consistently lower in the on-demand group whether looking at patients with an uncomplicated course of the disease or

Figure 1  Comparison of total costs per patient if patients were ranked to their total costs (as proxy for clinical condition and recovery) in each strategy.
patients with a more eventful and longer recovery. Only for a small number of patients at the very high end of total costs, generated by extremely high resource utilization, costs were comparable in the on-demand and in the planned relaparotomy groups.

### Table 2  
*Mean use of resources and costs in the on-demand and planned relaparotomy group during index admission and follow-up until one year after randomization*

<table>
<thead>
<tr>
<th></th>
<th>OD (n= 114)</th>
<th>PR (n = 115)</th>
<th>Difference (PR-OD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>unit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean volume</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total costs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(€)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Cost p.p. (€)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean volume</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total costs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(€)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean costs p.p. (€)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**DIRECT MEDICAL COSTS**

**Index admission**

**Admission**

<table>
<thead>
<tr>
<th></th>
<th>OD (n= 114)</th>
<th>PR (n = 115)</th>
<th>Difference (PR-OD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ward stay - index (excl ICU) day</td>
<td>26</td>
<td>1,323,448</td>
<td>11,609</td>
</tr>
<tr>
<td>ICU-stay day</td>
<td>12</td>
<td>2,356,434</td>
<td>21,040</td>
</tr>
</tbody>
</table>

**Interventions**

**(re)laparotomy**

<table>
<thead>
<tr>
<th></th>
<th>OD (n= 114)</th>
<th>PR (n = 115)</th>
<th>Difference (PR-OD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>index laparotomy</td>
<td>1.0</td>
<td>258,482</td>
<td>2,267</td>
</tr>
<tr>
<td>second-look procedure</td>
<td>0.62</td>
<td>129,855</td>
<td>1,139</td>
</tr>
<tr>
<td>with other surgical procedures procedure</td>
<td>0.37</td>
<td>138,003</td>
<td>1,211</td>
</tr>
<tr>
<td>percutaneous drainage procedure</td>
<td>0.49</td>
<td>16,755</td>
<td>147</td>
</tr>
</tbody>
</table>

**Diagnostic CT and cultures**

<table>
<thead>
<tr>
<th></th>
<th>OD (n= 114)</th>
<th>PR (n = 115)</th>
<th>Difference (PR-OD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT procedure</td>
<td>1.4</td>
<td>39,458</td>
<td>346</td>
</tr>
<tr>
<td>microbiology cultures</td>
<td>43</td>
<td>66,747</td>
<td>586</td>
</tr>
</tbody>
</table>

**Medication and other materials**

<table>
<thead>
<tr>
<th></th>
<th>OD (n= 114)</th>
<th>PR (n = 115)</th>
<th>Difference (PR-OD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>antibiotic therapy (excl ICU) day</td>
<td>6.0</td>
<td>54,023</td>
<td>474</td>
</tr>
<tr>
<td>enteroostomy care** day</td>
<td>24</td>
<td>83,009</td>
<td>741</td>
</tr>
<tr>
<td>blood products unit</td>
<td>0.61</td>
<td>10,139</td>
<td>89</td>
</tr>
<tr>
<td>mechanical ventilation*** day</td>
<td>8.3</td>
<td>351,153</td>
<td>3,080</td>
</tr>
</tbody>
</table>

**SUBTOTAL**

<table>
<thead>
<tr>
<th></th>
<th>OD (n= 114)</th>
<th>PR (n = 115)</th>
<th>Difference (PR-OD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4,880,939</td>
<td>43,580</td>
<td>6,700,851</td>
</tr>
<tr>
<td>Follow-up</td>
<td>OD (n=112)</td>
<td>PR (n=113)</td>
<td>Difference (PR-OD)</td>
</tr>
<tr>
<td>---------------------------</td>
<td>------------</td>
<td>------------</td>
<td>-------------------</td>
</tr>
<tr>
<td></td>
<td>unit</td>
<td>Mean</td>
<td>Total costs</td>
</tr>
<tr>
<td></td>
<td>volume</td>
<td>Mean</td>
<td>(€)</td>
</tr>
<tr>
<td><strong>Inpatient care</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ward stay - follow up</td>
<td>day</td>
<td>9.5</td>
<td>487,892</td>
</tr>
<tr>
<td>elective surgery</td>
<td>procedure</td>
<td>0.39</td>
<td>98,015</td>
</tr>
<tr>
<td><strong>Outpatient care</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>specialist consultation*</td>
<td>visit</td>
<td>16.3</td>
<td>145,265</td>
</tr>
<tr>
<td>CT-abdomen</td>
<td>procedure</td>
<td>0.07</td>
<td>1,998</td>
</tr>
<tr>
<td>US-abdomen</td>
<td>procedure</td>
<td>0.05</td>
<td>522</td>
</tr>
<tr>
<td>plain X-ray</td>
<td>procedure</td>
<td>0.18</td>
<td>871</td>
</tr>
<tr>
<td>enterostomy care**</td>
<td>day</td>
<td>140</td>
<td>498,291</td>
</tr>
<tr>
<td><strong>Other health care providers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>primary care physician*</td>
<td>visit</td>
<td>9.4</td>
<td>21,740</td>
</tr>
<tr>
<td>company doctor*</td>
<td>visit</td>
<td>1.8</td>
<td>4,667</td>
</tr>
<tr>
<td>paramedical specialist*</td>
<td>visit</td>
<td>29</td>
<td>75,561</td>
</tr>
<tr>
<td>district nurse*</td>
<td>hour</td>
<td>45</td>
<td>205,649</td>
</tr>
<tr>
<td>rehabilitation center*</td>
<td>day</td>
<td>19</td>
<td>725,771</td>
</tr>
<tr>
<td><strong>SUBTOTAL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2,162,354</td>
<td>19,307</td>
</tr>
<tr>
<td><strong>DIRECT NON-MEDICAL COSTS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>travel costs*</td>
<td>km</td>
<td>395</td>
<td>7,970</td>
</tr>
<tr>
<td><strong>INDIRECT COSTS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>absence from paid work*</td>
<td>day</td>
<td>70</td>
<td>319,649</td>
</tr>
<tr>
<td><strong>TOTAL COSTS (€)</strong></td>
<td></td>
<td>7,421,367</td>
<td>65,812</td>
</tr>
</tbody>
</table>

* PR n = 74 / OD n = 76: number of patients on which analyses are based, this information was then extrapolated to the entire sample per arm

** PR n = 53 patients with enterostomy; OD n = 49 patients with enterostomy

volume Average amount of resource utilized per patient
total costs Total costs of resource utilization of all patients
mean costs Average cost per patient
difference Difference between PR and OD (positive values in favor of OD)
Discussion

In this paper we presented an economic evaluation within a randomized clinical trial comparing two commonly used surgical strategies for patients with secondary peritonitis after their initial emergency laparotomy, being on-demand relaparotomy and planned relaparotomy strategy. In an earlier publication focusing on the clinical outcomes of the RELAP trial, we demonstrated that patients in the on-demand group did not have a significantly lower rate of adverse outcomes compared to the planned group, but that substantially less resources were utilized in the on-demand group\(^6\). The results of the detailed cost analyses presented here indicated that across the full range of healthcare resources, as well as across patients with different disease and recovery courses, resource utilization and associated costs generated by treatment and follow-up of severe abdominal sepsis were substantially lower for the on-demand strategy than for the planned strategy.

The observed cost differences were predominantly related to lengthier ICU stays, duration of mechanical ventilation, and hospital ward stay during the index admission period. Costs of rehabilitation centers and home care and of readmissions to the general hospital during follow-up were also substantial contributors to these costs differences. Although the planned strategy per definition involved at least one relaparotomy procedure, costs generated only by this extra procedure were only a mere fraction of the encountered cost differences. In an additional analysis in which costs associated with relaparotomy procedures were disregarded, it was clearly demonstrated that the observed cost differences between the surgical strategies could not be solely attributed to the costs associated with extra relaparotomies.

An important component of the total direct medical costs was composed of ICU stay (often involving mechanical ventilation). Consequently the estimated total costs were highly influenced by the unit costs estimate for an ICU day. We used a reference price, which is based on data from a range of general and academic hospitals in The Netherlands\(^13\). Cost estimates for ICU stay reported in the literature varied considerably, if reported at all. These differences were due to both different calculation methods, and different patient groups for which costs have been calculated. Estimated costs per ICU day may also differ due to differences between the organization and facilities (staff allocation and remuneration, equipment costs, non-clinical support services and estates)\(^9\). Differences between countries and their healthcare systems could explain part of this variation. In order to enhance the generalization of our findings to other countries, we presented the consequences of using cost estimates found for the UK\(^8\), Austria\(^21\), France\(^23\), Canada\(^20\), Germany\(^22\), and Norway\(^19\). Estimates for countries with publicly funded healthcare systems were better reported in the literature, than for countries with other types of healthcare systems (e.g., United States). Information pertaining to these costs and studies addressing the real costs of healthcare resources appeared to be lacking for non-publicly funded healthcare systems. These cross-country analyses, using international estimates of ICU stay, indicated that total costs associated with these surgical strategies could vary substantially across settings and countries, but that the relative difference between the two strategies remained stable.
The observed differences in indirect costs (productivity loss) between the surgical strategies may be a method-related effect rather than reflecting clinical differences: productivity costs only concern patients under the age of 65 (the majority of patients).

In general, resource utilization was found to be higher in the planned group than in the on-demand group. Therefore, adjustments in unit costs would result in changing total costs, rather than affecting the difference between on-demand and planned relaparotomy. Total costs indeed varied to some degree with the different assumptions regarding unit cost prices, but the relative difference between the strategies remained consistent across these analyses. On average, the on-demand strategy generated approximately 21% less costs than the planned relaparotomy. This means that per 1000 patients admitted to an emergency room with severe peritonitis, half of whom are currently operated according to the planned strategy, some 10 million euros could be saved if the on-demand strategy was employed instead.

### Table 3  Summary of sensitivity analyses.

Mean total costs and estimated absolute and relative differences between relaparotomy on-demand and planned relaparotomy across alternative assumptions and calculation methods.

<table>
<thead>
<tr>
<th>Description</th>
<th>OD</th>
<th>PR</th>
<th>difference</th>
<th>95% CI*</th>
<th>difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main</td>
<td>65,812</td>
<td>83,476</td>
<td>17,664</td>
<td>(5,056 to 29,004)</td>
<td>21.2%</td>
</tr>
<tr>
<td>1 Percutaneous drainage procedures (reimbursement fee instead of AMC estimates)</td>
<td>65,786</td>
<td>83,454</td>
<td>17,656</td>
<td>(5,050 to 28,996)</td>
<td>21.2%</td>
</tr>
<tr>
<td>2 Ward-stay unit costs (weighted average of Academic and General hospitals)**</td>
<td>62,982</td>
<td>81,042</td>
<td>18,060</td>
<td>(5,432 to 28,661)</td>
<td>22.3%</td>
</tr>
<tr>
<td>3 ICU-day unit costs (AMC Top Down calculation instead of guideline)</td>
<td>70,737</td>
<td>91,006</td>
<td>20,269</td>
<td>(5,946 to 32,151)</td>
<td>22.3%</td>
</tr>
<tr>
<td>4 With ICU-day unit costs estimated for:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a UK</td>
<td>63,279</td>
<td>79,714</td>
<td>16,435</td>
<td>(4,777 to 28,431)</td>
<td>20.6%</td>
</tr>
<tr>
<td>b Germany</td>
<td>61,585</td>
<td>77,198</td>
<td>15,613</td>
<td>(4,573 to 28,037)</td>
<td>20.2%</td>
</tr>
<tr>
<td>c France</td>
<td>69,146</td>
<td>88,427</td>
<td>19,281</td>
<td>(5,365 to 29,713)</td>
<td>21.8%</td>
</tr>
<tr>
<td>d Norway</td>
<td>77,269</td>
<td>100,491</td>
<td>23,222</td>
<td>(5,936 to 31,298)</td>
<td>23.1%</td>
</tr>
<tr>
<td>e Austria</td>
<td>63,838</td>
<td>80,544</td>
<td>16,706</td>
<td>(4,845 to 28,560)</td>
<td>20.7%</td>
</tr>
<tr>
<td>f Canada</td>
<td>59,003</td>
<td>73,364</td>
<td>14,361</td>
<td>(4,211 to 27,407)</td>
<td>19.6%</td>
</tr>
<tr>
<td>5 Exclude all costs of Relaparotomy procedures</td>
<td>62,587</td>
<td>77,939</td>
<td>15,352</td>
<td>(3,008 to 25,387)</td>
<td>19.7%</td>
</tr>
</tbody>
</table>

* based on geometric means;
** weighted by ratio of Academic and General hospital beds in the Netherlands (1:6);
At present, no other studies have reported a detailed description of costs associated with resource utilization generated by abdominal sepsis patients treated by either strategy. As no published data on costs are available for this disease, we can compare volumes only; our findings were comparable to a retrospective study comparing on-demand and planned relaparotomy, reporting similar figures for average length of hospital stay (49.5 vs. 52.0, including ICU) and average length of ICU stay (12.6 vs. 17.8 days). Duration of mechanical ventilation was a few days longer (10.3 vs. 13.9) as compared to our results (8.5 vs. 12 days), and the difference (on-demand 2.5 days shorter ventilated) was consistent with our findings.

This study has some strengths and limitations. First, the economic evaluation was performed as part of a randomized controlled trial and stratified for severity of disease, ensuring that the patients in both strategies were comparable with respect to clinical and prognostic factors. Differences in resource utilization and related costs were therefore unlikely due to systematic differences (bias) between study groups, but can truly be attributed to the surgical strategy. Furthermore, the economic evaluation was based on data on resource utilization required for the clinical trial and extended with additional relevant information acquired with self-administered questionnaires. This bottom-up strategy, based on detailed healthcare utilization data from individual patients, allowed us to obtain insight into the healthcare process and to determine the main cost driving factors. Although the majority of (direct medical) costs were generated during the index admission, the systematic documentation of the follow-up period until one-year after the index laparotomy allowed us to demonstrate that encountered differences in resource utilization remained consistent after discharge.

Currently, support of the on-demand strategy is growing, and sound empirical evidence regarding the optimal approach is now available from a prospective randomized comparison. The clinical results of the RELAP-trial, reported elsewhere, and the present economic evaluation supports further implementation of an on-demand relaparotomy strategy for treatment of patients with abdominal sepsis.

In conclusion, this economic evaluation prospectively demonstrated that resource utilization and associated costs generated during treatment and follow-up of severe peritonitis were substantially lower for an on-demand strategy compared to a planned strategy. These differences were found across the full range of healthcare resources as well as across patients with different courses of disease. Considering that patients in the on-demand relaparotomy strategy had a lower (albeit not statistically significant) rate of adverse outcomes compared to the planned relaparotomy group, the substantial reduction in costs (21%) associated with healthcare utilization renders the on-demand relaparotomy a far more efficient surgical strategy in patients with severe peritonitis. Further implementation of an on-demand relaparotomy strategy could have a major impact on the healthcare expenses for this severe and costly medical condition.
References


“Decisions compound themselves, in medicine like in anything else. No sooner have you taken one fork in the road than another and another come upon you.”

Atul Gawande (Complications)
Part II

Optimizing the on-demand strategy
Chapter 5
Usefulness of the APACHE II score to predict disease outcome and to evaluate safety of surgical strategy in patients with abdominal sepsis

Oddeke van Ruler, Manon LW Ziech, Kimberly R Boer, E Ascelijn Reuland, Johannes B Reitsma, Marja A Boermeester
Abstract

**Objective:** To evaluate the usefulness of APACHE II score to predict patient outcome (mortality and complicated course) and to evaluate the safety of surgical strategy (on-demand versus planned relaparotomy) in patients with abdominal sepsis in modern medical practice.

**Design:** Prospective, multicenter cohort study.

**Setting:** Two academic and 7 regional teaching hospitals in The Netherlands.

**Patients:** A consecutive cohort of 309 patients with abdominal sepsis, consisting of 229 patients with an APACHE II score $>10$ that were randomized between on-demand relaparotomy ($n=114$) or planned relaparotomy ($n=115$), and 80 non-randomized patients with an APACHE II score $\leq 10$ (exclusion criterion of relap trial)

**Main outcomes:** Predictive value of APACHE II score on patient outcome using 3 groups of disease-severity (mild: APACHE II $\leq 10$, moderate: APACHE II 11-20, severe: APACHE II $>20$); safety of on-demand strategy for patients with mild or severe abdominal sepsis.

**Results:** Higher APACHE II scores were associated with higher in-hospital mortality (mild 2.5%, moderate 15% and severe 57%; $p<0.001$). Mild peritonitis was almost exclusively (95%) treated by the on-demand strategy and had a significantly more fortunate course of the disease, including less complications (mild 35%, versus 59% for moderate and 47% for severe peritonitis; $p=0.002$). As well, the APACHE II score did not modify treatment effect for mortality (test of interaction $p=0.30$) or complicated course of disease (test of interaction $p=0.95$) within the group of randomized patients.

**Conclusions:** The APACHE II score is a good predictor of in-hospital mortality, and of a complicated course in patients with abdominal sepsis in the present clinical setting. Patients with mild peritonitis but also patients with severe peritonitis can be safely treated by the on-demand strategy.
Introduction

Abdominal sepsis is a serious and common medical condition, with in-hospital mortality ranging from 20% up to 60%, depending on severity of disease\(^1-^4\). Surgical elimination of the infectious source by emergency laparotomy (index operation) is the cornerstone of treatment of abdominal sepsis. Additionally, intensive care organ function support and adequate antibiotic treatment are obligatory. Also, extensive patient monitoring for adequate and timely selection of patients needing a relaparotomy, is essential to reduce mortality and morbidity additional\(^5\). An alternative approach is the planned relaparotomy strategy in which patients are reoperated every 48 hours until findings are negative.

Mortality rates are associated with the severity of peritonitis\(^2,^6,^7\). The current standard to assess severity of peritonitis is by calculating the Acute Physiology And Chronic Health Evaluation (APACHE) II at the initial emergency laparotomy\(^8\). The APACHE II score was developed by William Knaus et al. as a severity of disease classification system and is to date a well-established and properly validated instrument to predict mortality in severely ill, general ICU patients\(^3,^9-^11\).

In 1988 Bohnen et al. showed that the APACHE II score was also useful in predicting in-hospital mortality specifically for patients with abdominal sepsis\(^6\). However, due to modernized medical care and improvements in ICU within the last two decades, treatment and support of abdominal sepsis patients has significantly changed\(^12\), potentially rendering the APACHE II score, calculated early in the disease, a less accurate predictor of mortality. (Re)evaluation of the usefulness of the APACHE II score in stratifying patients with abdominal sepsis for their risk of a complicated course of disease (in-hospital mortality and/or major disease-related morbidity), regarding the current level of medical care, may be needed.

Recently, a randomized controlled clinical trial (RCT), the RELAP trial, has been performed by our study group addressing the long-standing discussion on the dilemma of surgical treatment strategy following initial emergency laparotomy for patients with moderate to severe abdominal sepsis (APACHE II >10): relaparotomy on-demand versus planned relaparotomy. The RELAP trial concluded that the on-demand relaparotomy strategy should be the preferred surgical strategy, based on a (non-significant) difference in mortality/major morbidity outcome, substantially lower healthcare utilization and significantly lower costs in favor of patients treated by the on-demand strategy during 12-months follow-up\(^13\). Patients with a low APACHE II score (≤ 10; mild peritonitis) at index operation were excluded from the RELAP trial as there is consensus in clinical practice, that the preferred surgical strategy for mild peritonitis is on-demand relaparotomy\(^2,^6,^14,^15\), but empirical data supporting this cut-off was sparse. The actual choice of treatment strategy in this group was left to the treating surgeon’s discretion.

To date the planned relaparotomy strategy is applied mostly and more premeditated in patients with severe initial disease and high APACHE II scores, although evidence to substantiate this selection is not available. To evaluate this often supported view, that more severely ill patients would be better off with a more aggressive surgical treatment strategy, we investigated whether severity of disease would affect the treatment effect of on-demand versus planned relaparotomy on in-hospital mortality and on an overall complicated course of disease.
Therefore, this study evaluates three potential functions (roles) of the APACHE II scoring in patients with abdominal sepsis: (1) the usefulness of the APACHE II score to predict in-hospital mortality and a complicated course of the disease in today's clinical setting; (2) to verify whether patients with mild peritonitis, defined by an APACHE II score of 10 or lower, are indeed safely managed by the on-demand strategy and (3) to examine whether the APACHE II score (as measure of initial severity of disease) modifies the difference in treatment effect between the on-demand and planned relaparotomy strategy for in-hospital mortality and a complicated course of disease in patients with more severe peritonitis.

Methods

Patient eligibility and study design

Patients were eligible for the RELAP trial if diagnosed with secondary peritonitis caused by perforation or infection of a visceral organ, ischemia/necrosis of (part of) the gastrointestinal tract or postoperative peritoneal infection. Exclusion criteria were: age below 18 or above 80; peritonitis due to perforation of the bowel after endoscopy operated within 24 hours; abdominal infection due to an indwelling dialysis (CAPD) catheter; appendicitis treated by minimal invasive surgery; peritonitis caused by pancreatitis; expected survival of less than 6 months due to disseminated malignancy; severe brain damage due to trauma or anoxia; imperative surgical therapy needed to eliminate the primary intra-abdominal focus (e.g. temporary end-stapled bowel loops in case of ischemia, temporary packing with gauzes in case of severe bleeding during surgery).

All eligible patients were registered at 2 academic and 7 major teaching hospitals and patients with an APACHE II score >10 were enrolled in the RELAP trial and randomly allocated to either the relaparotomy on-demand strategy (n=114) or the planned relaparotomy (n=115), following the initial emergency laparotomy13.

Patients with mild peritonitis (APACHE II score ≤10) were not included in the RELAP trial (n=80), and treatment strategy was left to the discretion of the treating medical team. However, follow-up data of these patients were collected. The entire cohort of peritonitis patients was included in the present study. This study was approved by the Medical Ethics Committees of all participating institutes.

Data collection

Data were prospectively collected. Data included patient characteristics, disease- and operation-related information and postoperative variables, including the development of major peritonitis-related complications for 12 months following index operation. The APACHE II score was assessed by adding points for the worst values of the variables of the APACHE II score in the first 24 hours of illness. In case of not having a complete 24-hour APACHE II score before initial laparotomy, but confirmed peritonitis during this first laparotomy, the APACHE II score was assessed within 24 hours after the initial laparotomy.
Chapter 5  Usefulness of the APACHE II score

Initial cohort
N= 510
Patients with secondary peritonitis undergoing an emergency laparotomy

Patients excluded
N= 201
exclusion criteria (n=151)
•pancreatitis at index operation (n=1)
•withdrawal of informed consent (n=2)
•age <18 or >80 (n=39)
•other exclusion criterion (n=109)
not included (n=50)
•no informed consent (n=43)
•decided by treating surgeon (n=7)

Eligible patients with secondary peritonitis
N = 309

Not included in RELAP trial
APACHE II ≤ 10
N=80

Included in the RELAP trial
APACHE II > 10
N=229
On-demand n=114 Planned n=115

Figure 1  Flowchart of study cohort

Treatment strategies

A relaparotomy ‘on-demand’ was performed only in case of no clinical improvement or in case of clinical deterioration, monitored by physiological, laboratory and radiology parameters. A planned relaparotomy is performed every 36 to 48 hours until the abdomen is macroscopically clean at the beginning of the final relaparotomy.
Statistical analysis

APACHE II scores were calculated for all patients at the time of study entry (24 hours before or after the index operation), using the worst outcomes of the component variables. Expected mortality accompanying APACHE II scores was calculated for each patient by Logit=-3.517+(APACHE II)*0.146; Predicted Death Rate=$e^{logit}/(1+e^{logit})$. The expected mortality for the unadjusted APACHE II scores, which was used for the (emergency) surgery for abdominal sepsis, does not include the prespecified diagnostic weight category (http://www.sfar.org/scores2/apache22.html). The observed mortality per APACHE II score category was compared to the mean expected mortality regarding the APACHE II score per patient.

The predictive value of the APACHE II score was determined using binary regression analysis for in-hospital mortality alone and for a complicated clinical course (mortality and/or morbidity) while admitted to the hospital. Complications were divided into surgery-related complications, needing surgical re-intervention, and sepsis-related complications needing additional supportive treatment. Complicated clinical course was defined as death and/or the development of one or more major complications, listed with results in Table 2 (Appendix 1).

For three clinically relevant categories of disease severity, APACHE II scores were divided in mild (0-10), moderate (11-20) and severe (>20). The three categories were compared for patient characteristics and the outcomes in-hospital mortality and the development of complications. Where appropriate the $\chi^2$ test, the Pearson's exact and the Kruskall-Wallis tests were used to detect differences between categories of severity of disease. Data are presented as medians with 25%-75% inter-quartile ranges (IQR).

Furthermore, a test of interaction was performed to determine whether the treatment strategy effect on in-hospital mortality was modified by APACHE II distribution at time of study entry. Treatment effect is given in odds ratios (OR) with 95% confidence intervals (95% CI) for the three APACHE II categories for severity of disease. Interaction was considered present if the p-value of the test for interaction was $p<0.05$.

Statistical analysis was performed using SPSS (version 12.0) for Windows (SPSS, Chicago, IL). All p-values $<0.05$ were considered statistically significant.
Results

Patient enrollment

A total of 510 patients with abdominal sepsis were registered at the nine participating hospitals of the RELAP trial between November 2001 and August 2005. These patients were evaluated for eligibility for this present study. Two hundred and one patients met one or more exclusion criteria other than the APACHE II score limit of >10 followed for the RELAP trial, leaving 309 to compile the present study cohort. Of these 309 patients, 229 patients had an APACHE II score >10 and were therefore enrolled in RELAP trial. The other 80 registered abdominal sepsis patients fit for analysis had an APACHE II score ≤10 (Figure 1).

Demographic data

The overall mean APACHE II score in the entire study population was 13 (±SD 6.2). Major differences were found between patients with mild disease defined by an APACHE II score ≤10 (n=80), patients with moderate severity of disease (APACHE II score 11 to 20; n=194) and patients with severe disease (APACHE II score >20; n=35). The study populations are described in detail in Table 1. Age was significantly higher in the APACHE II score categories above 10 (APACHE II ≤10: median age 46 (IQR 35-58) versus APACHE II 11-20: age median 69 (IQR 56-75) versus APACHE II >20: median age 72 (IQR 67-75; p<0.001) and the incidence of major comorbidity (defined as malignancy present within the past 3 years, cardiovascular disease, renal failure, respiratory disease and/or diabetes mellitus) was significantly higher in the APACHE II >20 group (80% vs. APACHE II 11-20: 56% and APACHE II ≤10: 18%; p<0.001). As both age and comorbidity are important components of the APACHE II score, this was expected. Perforation of a visceral organ as the cause of peritonitis was the most frequently encountered etiology in all groups, overall 191/309 (62%), followed by anastomotic leakage after elective abdominal surgery (24% overall). The infectious focus was significantly more often localized at the lower digestive tract in the higher APACHE II score categories (28% APACHE II ≤10 vs. 63% APACHE II 11-20 group and 71% in the APACHE II >20 group; p<0.001) and significantly less often at the appendix (38% APACHE II ≤10 versus 2% APACHE II 11-20 group and 0% in the APACHE II >20 group; p<0.001).

Predictive value of APACHE II score in abdominal sepsis

Mortality

The overall in-hospital mortality was 16.5% (52/309) with significant difference between APACHE II categories. Mortality increased from 2.5% for patients with APACHE II scores ≤10, to 16% in patients with scores ranging from 11 to 20 and 77% in patients with scores >20 (p<0.001; Table 2).
Table 1  Demographic and clinical characteristics at time of emergency laparotomy.

Categories are determined by APACHE II score: patients with mild peritonitis (APACHE II score ≤10), patients included in the RELAP trial with moderate peritonitis (APACHE II score of 11-20) and patients included in the RELAP trial with severe peritonitis (APACHE II score >20).

<table>
<thead>
<tr>
<th>Variables</th>
<th>APACHE II &lt; 10</th>
<th>APACHE II 11-20</th>
<th>APACHE II &gt;20</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean 5.1</td>
<td>mean 14.5</td>
<td>mean 23.9</td>
<td></td>
</tr>
<tr>
<td>Age (years) – median (IQR)</td>
<td>46 (35-58)</td>
<td>69 (56 – 75)</td>
<td>72 (67 – 75)</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Male</td>
<td>42 (53%)</td>
<td>94 (49%)</td>
<td>15 (43%)</td>
<td>0.62</td>
</tr>
<tr>
<td>Major comorbidity present – no. (%)</td>
<td>14 (18%)</td>
<td>108 (56%)</td>
<td>28 (80%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Malignancy</td>
<td>9 (11%)</td>
<td>49 (25%)</td>
<td>8 (23%)</td>
<td>-</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>4 (5%)</td>
<td>44 (23%)</td>
<td>10 (29%)</td>
<td>-</td>
</tr>
<tr>
<td>Respiratory disease (COPD)</td>
<td>1 (1%)</td>
<td>21 (11%)</td>
<td>10 (29%)</td>
<td>-</td>
</tr>
<tr>
<td>Renal disease</td>
<td>1 (1%)</td>
<td>13 (7%)</td>
<td>4 (11%)</td>
<td>-</td>
</tr>
<tr>
<td>Diabetes</td>
<td>0 (0%)</td>
<td>16 (8%)</td>
<td>4 (11%)</td>
<td>-</td>
</tr>
<tr>
<td>Etiology of peritonitis – no. (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.002</td>
</tr>
<tr>
<td>Perforation</td>
<td>58 (73%)</td>
<td>107 (55%)</td>
<td>26 (74%)</td>
<td>-</td>
</tr>
<tr>
<td>Anastomotic leakage</td>
<td>11 (14%)</td>
<td>60 (31%)</td>
<td>3 (9%)</td>
<td>-</td>
</tr>
<tr>
<td>Ischemia</td>
<td>4 (5%)</td>
<td>10 (5%)</td>
<td>4 (11%)</td>
<td>-</td>
</tr>
<tr>
<td>Inflammation</td>
<td>7 (9%)</td>
<td>9 (5%)</td>
<td>0 (0%)</td>
<td>-</td>
</tr>
<tr>
<td>Other</td>
<td>0 (0%)</td>
<td>8 (4%)</td>
<td>2 (6%)</td>
<td>-</td>
</tr>
<tr>
<td>Postoperative peritonitis</td>
<td>24 (30%)</td>
<td>99 (51%)</td>
<td>9 (26%)</td>
<td>0.001</td>
</tr>
<tr>
<td>Localization</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Upper GI tract</td>
<td>27 (34%)</td>
<td>49 (25%)</td>
<td>9 (26%)</td>
<td>-</td>
</tr>
<tr>
<td>Lower GI tract</td>
<td>22 (28%)</td>
<td>122 (63%)</td>
<td>25 (71%)</td>
<td>-</td>
</tr>
<tr>
<td>Biliary tract</td>
<td>0 (0%)</td>
<td>13 (7%)</td>
<td>1 (3%)</td>
<td>-</td>
</tr>
<tr>
<td>Appendix</td>
<td>30 (38%)</td>
<td>4 (2%)</td>
<td>0 (0%)</td>
<td>-</td>
</tr>
<tr>
<td>Other*</td>
<td>1 (1%)</td>
<td>6 (3%)</td>
<td>0 (0%)</td>
<td>-</td>
</tr>
</tbody>
</table>

† Kruskal – Wallis

All categorical data are tested by Chi square (χ²)

SD = Standard Deviation, IQR = Inter-Quartile Range

* ‘Other’ consisted of either an infectious focus localized at the upper as well as the lower GI tract or at a gynaecologic site.
Figure 2  Calibration of APACHE II score as predictor of in-hospital mortality.

Observed versus expected probabilities of in-hospital mortality across a range of observed APACHE II scores in the entire study population (n=309).
Predicted probabilities were calculated using the formula distributed on http://www.sfar.org/scores2/apache22.html. The associated Hosmer-Lemeshow test for model fit was p=0.154.

Observed versus unadjusted expected mortality (%) for secondary peritonitis patients (n = 309)
Figure 2 shows the observed and expected numbers of deaths across a more refined distribution of APACHE II scores. It reveals that the existing formula for predicting mortality based on APACHE II scores applies for this population as the differences in observed and expected numbers of deaths are small with a p-value of the goodness of fit test of 0.154 (Hosmer-Lemeshow). However, the APACHE II score showed to underestimate mortality for patients with mild to moderate peritonitis (higher expected mean proportion of deaths than observed for APACHE II scores ≤15) and to overestimate mortality for patients with APACHE II scores of >20 (lower mean proportion of observed deaths than expected; Figure 2).

Logistic regression analysis confirmed that the APACHE II score is a strong predictor of mortality, as a 5-point increase in APACHE II score is associated with an OR 2.5 (95% CI 2.1-2.8; p<0.001).

Complicated course of disease
Overall, 55% (169/309) of the studied patients developed one or more disease-related complications. The number of patients with a complicated course of disease (mortality and/or disease-related complications) was significantly different between the APACHE II ≤10, 11-20 and >20 categories (36% vs. 66% vs. 77%; p<0.001). Also, the number of complications per patient was significantly lower for the APACHE II ≤10 group compared to the APACHE II 11-20 and >20 groups (Table 2). In the APACHE II ≤10 group most patients developed sole surgery-related complications, whereas patients with an APACHE II score 11-20 and patients with an APACHE II score >20 developed both surgery- and sepsis-related complications (Table 2).

Higher APACHE II scores were associated with a significant increased risk of development of a complicated course of disease (per 5-point increase in APACHE II score: OR 1.6; 95% CI 1.6-2.5; p<0.001).

Evaluation of patients with mild peritonitis

Treatment strategy
In the APACHE II ≤10 group, no randomization took place and the treating surgeon was free to choose a surgical treatment strategy. As expected, the majority of patients (95%) were treated by the on-demand strategy. Only 4 patients were treated with the planned strategy and all 4 patients had a perforation of the upper gastro-intestinal tract as cause of their secondary peritonitis. The overall relaparotomy rate in the APACHE II ≤10 group was low (19%; Table 3). Among these first relaparotomies 75% was positive for (persistent) peritonitis. Regarding patients treated by the on-demand strategy, only 15% of the mild peritonitis patients receiving one or more relaparotomies, compared to 43% in the APACHE II 11-20 group and 38% in the APACHE II >20 group (p<0.001).
### Table 2  
**In-hospital mortality (observed), the development of complications and the clinical course of the disease are described for the 3 severity of disease categories: mild peritonitis (APACHE II score ≤10), moderate peritonitis (APACHE II score of 11-20) and severe peritonitis (APACHE II score >20).**

<table>
<thead>
<tr>
<th>Variables</th>
<th>APACHE II ≤ 10</th>
<th>APACHE II 11-20</th>
<th>APACHE II &gt;20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean 5.1</td>
<td>mean 14.5</td>
<td>mean 23.9</td>
</tr>
<tr>
<td></td>
<td>N = 80</td>
<td>N = 194</td>
<td>N = 35</td>
</tr>
<tr>
<td><strong>Complicated course of disease</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality and/or morbidity – no. patients (%)</td>
<td>29/80 (36%)</td>
<td>127/193 (66%)</td>
<td>27/35 (77%)</td>
</tr>
<tr>
<td>In-hospital mortality – no. patients (%)</td>
<td>2/80 (2.5%)</td>
<td>30/194 (16%)</td>
<td>20/35 (57%)</td>
</tr>
<tr>
<td><strong>Complications ≥ 1 – no. (surviving) patients (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only surgery-related complications</td>
<td>22/78 (28%)</td>
<td>44/160 (28%)</td>
<td>1/15 (7%)</td>
</tr>
<tr>
<td>Only sepsis-related complications</td>
<td>2/78 (3%)</td>
<td>18/160 (11%)</td>
<td>3/15 (20%)</td>
</tr>
<tr>
<td>Both surgery- &amp; sepsis-related complications</td>
<td>3/78 (4%)</td>
<td>34/160 (21%)</td>
<td>3/15 (20%)</td>
</tr>
<tr>
<td>No. of complications per patient (incl. death) – median (IQR)</td>
<td>0 (0-1)</td>
<td>1 (0-2)</td>
<td>3 (1-4)</td>
</tr>
<tr>
<td>Surgery-related – no. patients overall (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perforation</td>
<td>1 (1.3%)</td>
<td>15 (7.7%)</td>
<td>6 (17%)</td>
</tr>
<tr>
<td>Abscess</td>
<td>11 (14%)</td>
<td>58 (30%)</td>
<td>11 (31%)</td>
</tr>
<tr>
<td>Bleeding</td>
<td>0 (0.0%)</td>
<td>10 (5.2%)</td>
<td>1 (2.9%)</td>
</tr>
<tr>
<td>Enterostomy dysfunction</td>
<td>0 (0.0%)</td>
<td>9 (4.6%)</td>
<td>2 (5.7%)</td>
</tr>
<tr>
<td>Obstructive ileus</td>
<td>3 (4.0%)</td>
<td>0 (0.0%)</td>
<td>1 (2.9%)</td>
</tr>
<tr>
<td>Fistula</td>
<td>1 (1.3%)</td>
<td>2 (1.0%)</td>
<td>1 (2.9%)</td>
</tr>
<tr>
<td>Fascia dehiscence, burst abdomen</td>
<td>3 (4.0%)</td>
<td>12 (6.2%)</td>
<td>1 (2.9%)</td>
</tr>
<tr>
<td>Surgical Site infection</td>
<td>6 (7.5%)</td>
<td>16 (8.2%)</td>
<td>3 (8.6%)</td>
</tr>
<tr>
<td>Anastomotic leakage</td>
<td>6 (7.5%)</td>
<td>14 (7.2%)</td>
<td>4 (11%)</td>
</tr>
<tr>
<td>Ischemia/necrosis</td>
<td>0 (0.0%)</td>
<td>5 (2.6%)</td>
<td>1 (2.9%)</td>
</tr>
<tr>
<td>Evacuation hematoma</td>
<td>1 (1.3%)</td>
<td>5 (2.6%)</td>
<td>1 (2.9%)</td>
</tr>
<tr>
<td>Sepsis-related – no. patients (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pneumonia</td>
<td>3 (4.0%)</td>
<td>25 (13%)</td>
<td>6 (17%)</td>
</tr>
<tr>
<td>Acute respiratory failure</td>
<td>1 (1.3%)</td>
<td>16 (8.2%)</td>
<td>10 (29%)</td>
</tr>
<tr>
<td>ARDS</td>
<td>1 (1.3%)</td>
<td>11 (5.7%)</td>
<td>3 (8.6%)</td>
</tr>
<tr>
<td>Multi organ failure (MOF)</td>
<td>0 (0.0%)</td>
<td>12 (6.2%)</td>
<td>6 (17%)</td>
</tr>
<tr>
<td>Hepatic failure</td>
<td>0 (0.0%)</td>
<td>5 (2.6%)</td>
<td>3 (8.6%)</td>
</tr>
<tr>
<td>DIC</td>
<td>0 (0.0%)</td>
<td>6 (2.9%)</td>
<td>2 (5.7%)</td>
</tr>
<tr>
<td>Line sepsis</td>
<td>2 (2.5%)</td>
<td>5 (2.6%)</td>
<td>3 (8.6%)</td>
</tr>
</tbody>
</table>

GI = Gastro-Intestinal, DIC = Disseminated Intravascular Coagulation, ARDS = Acute Respiratory Distress Syndrome.

ª needing surgical or radiological intervention, **including death (defined as complication)
Mortality
Overall in-hospital mortality was low for the APACHE II ≤10 group (2.5%) and lower than would be expected based on APACHE II scores. In the APACHE II ≤10 in both treatment strategy arms one patient died (on-demand 1/74 (1.4%) vs. planned 1/4 (25%), p=0.003).

Clinical course
The less complicated clinical course of the APACHE II ≤10 group was reflected by significantly fewer patients being admitted to the ICU (18% vs. APACHE 11-20 91% and APACHE II >20 100%, p<0.001) and by a significantly shorter overall hospital stay (median 10, IQR 17 to 18 versus APACHE 11-20: median 28.5, IQR 17 to 57 and APACHE >20: median 27, IQR 18 to 51, p<0.001, Table 3).

On-demand treatment in more severe peritonitis
In the RELAP trial patients were stratified for severity of disease: moderate peritonitis (APACHE II score from 11 to 20) and severe peritonitis (APACHE II score >20). Although APACHE II score was a risk factor for adverse outcome, it did not interfere with the difference in treatment effect between the on-demand and the planned strategy, neither for in-hospital mortality (Figure 4a; p-value for test of interaction 0.30), nor for the outcome of an overall complicated course of disease (Figure 4b; p-value for test of interaction 0.95). These results indicate that clinical outcomes were similar across both the APACHE II score 11 to 20 and the APACHE II score >20 categories. This translates into all patients being effectively and safely treated by the on-demand strategy, including the most severe ill patients (APACHE II >20). The mild peritonitis patients (APACHE II score ≤10) were left out of these interaction analyses and are depicted merely as a validation of our clinical results as described in the previous paragraph. Comparison between the on-demand and planned strategy is not particularly meaningful in this mild peritonitis subgroup, because only a few patients was treated with the planned strategy. Also treatment allocation was not randomized in this subgroup.

Discussion
The APACHE II score was useful in stratifying patients with abdominal sepsis for severity of disease, by predicting in-hospital mortality as well as an overall complicated clinical course after the initial emergency laparotomy for peritonitis, in the present day clinical setting. Patients with mild peritonitis (APACHE II score ≤10) were indeed safely and effectively treated by the on-demand strategy, reflected by low rates of relaparotomy, lower than expected mortality and short hospital and ICU stays. Furthermore, contrary to earlier tendencies and a long-standing dogma, we produced evidence that the on-demand strategy can be safely applied also in patients with severe abdominal sepsis; those patients with APACHE II scores >20.
Results are expressed as odds ratios and 95% confidence intervals and stratifies for three APACHE II categories with a formal test for interaction for treatment effect between only moderate (APACHE II score of 11-20) and severe (APACHE II score >20) peritonitis groups as in the mild peritonitis group almost all patients (96%) received the on-demand strategy. A p-value of <0.05 was considered significant for treatment effect.

Until the late 1990’s the APACHE II score has been evaluated in secondary peritonitis by research groups and was considered the best scoring system to stratify patients for severity of disease. Bohnen et al. were the first to establish that increased APACHE II scores in abdominal sepsis were associated with higher mortality. However, actual mortality rates in the study of Bohnen were considerably higher than in our population even though the mean overall APACHE II scores were similar for both populations (13.7 Bohnen vs. 13.0 in presented study). This may suggest a positive effect of modern medical care on outcome of these critically ill patients.

To date the APACHE II score is considered to be an independent prognostic factor for in-hospital mortality in abdominal sepsis. We found an overestimation of prediction of in-hospital mortality in the mild to moderately ill patients which, again, may reflect modernized and better current standards of patient care as predicted mortality associated with APACHE II scores as presented by Knaus et al. were determined back in 1985. On the other hand, in-hospital mortality in our severely diseased population...
(APACHE II score >20) was underestimated. We argue that it may be necessary to calculate an abdominal sepsis-specific diagnostic weight to adjust the calculation of the expected mortality associated with the APACHE II score to optimize prediction of patient outcome in abdominal sepsis by APACHE II score.

**Table 3**  
Clinical course of the disease during index admission period is described for the 3 severity of disease categories: mild peritonitis (APACHE II score ≤10), moderate peritonitis (APACHE II score of 11-20) and severe peritonitis (APACHE II score >20)

<table>
<thead>
<tr>
<th>APACHE II ≤ 10</th>
<th>APACHE II 11-20</th>
<th>APACHE II &gt;20</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean 5.1</td>
<td>mean 14.5</td>
<td>mean 23.9</td>
<td></td>
</tr>
<tr>
<td>N = 80</td>
<td>N = 194</td>
<td>N = 35</td>
<td></td>
</tr>
</tbody>
</table>

**Clinical course**

<table>
<thead>
<tr>
<th></th>
<th>APACHE II ≤ 10</th>
<th>APACHE II 11-20</th>
<th>APACHE II &gt;20</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relaparotomy received – no. patients (%)</td>
<td>15 (19%)</td>
<td>133 (69%)</td>
<td>23 (66%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>1st relap positive findings – no. patients (%)</td>
<td>12 (15%)</td>
<td>60 (31%)</td>
<td>10 (29%)</td>
<td>0.101</td>
</tr>
<tr>
<td>ICU admission – no. patients (%)</td>
<td>14 (18%)</td>
<td>177 (91%)</td>
<td>35 (100%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ICU stay – median days (IQR)</td>
<td>0 (0-0)</td>
<td>10 (6-18)</td>
<td>17 (9-28)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Hospital stay – median days (IQR)</td>
<td>11 (7-18)</td>
<td>28.5 (17-57)</td>
<td>27 (18-51)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

The RELAP trial did find a benefit in clinical outcomes for the on-demand strategy, but differences were not significantly different. The question is posed whether the favorable effects of the on-demand strategy are indeed relevant for different categories of severity of abdominal sepsis. There is not much literature on the surgical treatment strategy in mild peritonitis. The relaparotomy treatment dilemma arises most prominently for the more severely ill patients or patients with generalized peritonitis. Subsequently, research was performed comparing strategies for this group of patients. Whether mild peritonitis is indeed safely and effectively treated by the on-demand strategy was never evaluated, but was preserved by clinical experience. This study confirms that the on-demand strategy is a valid approach to treat mild peritonitis patients.

As only 4 patients with mild peritonitis were treated by the planned strategy we cannot evaluate if this strategy would have given similar, worse or better patient outcomes. The 4 patients did not show any similarities to set aside a specific group of patients needing a more aggressive treatment approach. Regarding the putative negative impact of the planned strategy on patient outcome in this mildly diseased group, and feasibility of the on-demand strategy, a study addressing this issue will probably never be conducted.

But then, maybe severely diseased patients do benefit from a more aggressive planned strategy. Alternatively, however, multiple interventions may harm this already compromised group of patients. Literature addressing this issue is scarce. A previous, retrospective study by our study group showed improved survival for patients with more severe abdominal sepsis (measured by a MPI >25) when treated by the on-demand strategy. Koperna et al. found that the time until the first re-operation
positively effected survival in abdominal sepsis patients, except for severely ill patients (APACHE II score of >26)\textsuperscript{25}. These results are in line with findings of this present study: all patients with abdominal sepsis, regardless of severity of disease, benefit from the on-demand relaparotomy strategy.

The additional value of the APACHE II score in prediction aimed at selecting individual (high-risk) patients for relaparotomy is limited. To predict which patients are indeed in need of a relaparotomy for (persistent) peritonitis, future research needs to focus on prediction of individualized patient outcome and prediction of relaparotomy outcome. Other widely used scoring systems should also be evaluated for this purpose\textsuperscript{17}.

In conclusion, the APACHE II is a good indicator of disease severity and can be used to stratify patients for severity of disease not only for mortality but also for the development of complications. All patients with abdominal sepsis can be safely and effectively treated by the on-demand strategy, regardless of initial severity of disease.
Reference List

Chapter 6
Failure of available scoring systems to predict ongoing infection in patients with abdominal sepsis after their initial emergency laparotomy

Oddeke van Ruler, Kimberly R Boer, Bas Lamme, Dirk J Gouma, Marja A Boermeester, Johannes B Reitsma

Submitted
Abstract

Objective: In the treatment of abdominal sepsis timely and adequate selection of patients for relaparotomy after the initial emergency (index) laparotomy is essential to optimize patient outcome. Currently, there are no specific scoring systems aimed at selecting patients that require a relaparotomy. Therefore we examined commonly used scoring systems, designed to predict overall outcome in critically ill patients, for their ability to identify (select) those patients with ongoing infection needing relaparotomy.

Design: Area under the receiver operating characteristic (ROC) curve (AUC) analyses were employed to determine whether existing scoring systems can discriminate between patients with ongoing infection needing relaparotomy and those without.

Setting: Data from a randomized clinical trial comparing two surgical strategies – on-demand vs. planned relaparotomy – in 2 academic and 7 regional hospitals.

Patients: The trial population (n=229) consisted of patients with severe peritonitis (APACHE II score >10) who underwent an emergency relaparotomy. We excluded patients who had their first relaparotomy after more than one week (n=8). The study population consisted of 221 patients at risk for ongoing infection.

Measurements: In the planned strategy group a patient was classified as ‘ongoing infection needing a relaparotomy’ in case of positive findings at relaparotomy. In the on-demand group, patients that received a relaparotomy revealing positive findings and patients without a relaparotomy but who died within 14 days after the index operation were classified as ‘ongoing infection needing relaparotomy’. The following scoring systems were evaluated for their association with requiring a relaparotomy: Acute Physiology and Chronic Health Evaluation (APACHE) II score, simplified Acute Physiology Score (SAPS) II, Mannheim Peritonitis Index (MPI), sequential Multiple Organ Dysfunction Score (MODS), Sequential Organ Failure Assessment (SOFA) score, and the acute part of the APACHE II score (APS).

For APACHE II, SAPS II, and MPI we used a single measurement (day 1). For sequential scoring systems (MODS, SOFA and APS), we measured the score at day 1, day 2 and the absolute difference between scores on day 2 and day 1.

Main results: The overall proportion of patients requiring a relaparotomy was 32% (71/221). The SAPS II (AUC 0.61; 95% CI 0.52-0.69) and 2 sequential scores on day 1 had a discriminatory ability above chance in identifying patients with ongoing infection needing relaparotomy: SOFA (AUC 0.60; 95% CI 0.51-0.68), APS (AUC 0.62; 95% CI 0.53-0.70). However, to correctly identify 90% of all patients needing a relaparotomy (sensitivity) would require such a low cut-off value that around 80% of all patients selected by these scoring systems would have negative findings at relaparotomy. The discriminatory ability of all scoring systems did not improve when examining scores at day 2 or the difference in scores between day 2 and day 1.

Conclusions: None of the widely-used ICU scoring systems are of clinical value for the identification of patients with ongoing infection needing relaparotomy. Apparently, quantification of organ failure or changes in organ failure can not differentiate between ongoing organ failure due to abdominal sepsis despite source control and ongoing abdominal infection. There is a need to develop more specific tools that can assist surgeons in their daily monitoring of these patients after initial emergency laparotomy.
Introduction

Recently we conducted a trial among patients with abdominal sepsis comparing two surgical strategies after the initial emergency operation: the on-demand relaparotomy strategy versus the planned relaparotomy strategy (RELAP trial)\(^1\). From this trial we concluded that the on-demand strategy was preferred, based on comparable clinical outcomes (12-month mortality in on-demand group of 29% vs. 36% in planned group; \(p=0.22\)), but a substantial reduction of healthcare utilization and costs\(^1\). Planned relaparotomy yielded negative findings in 66% of patients and, thus, had no therapeutic effect in these patients. Still, there is definitely room for improvement in selection of patient for relaparotomy in the on-demand strategy as 31% of these patients had a negative relaparotomy\(^1\).

The on-demand strategy implies a vigilant observation of the postoperative peritonitis patient. Better monitoring following the initial emergency (index) laparotomy and adequate selection of patients with ongoing infection for re-intervention may further improve outcome of the on-demand strategy. Presently, the on-demand strategy includes re-operation when patients show clinical deterioration or do not improve\(^1\). As yet, these conditions are not well defined. There is no consensus or guideline on which specific parameters to monitor that can assist treating physicians in their decision which patients to select for re-operation.

Several commonly used scoring systems exist, that assess the severity of disease in critically ill patients. These scoring systems have originally been developed and proven valuable in predicting mortality. As these scores are widely incorporated in the daily treatment of ICU patients, we questioned whether these scoring systems would also be useful in early and adequate identification of patients with ongoing abdominal infection needing a relaparotomy. We have also determined whether these existing scoring systems could be beneficial in selecting patients for re-intervention, including either relaparotomy or percutaneous drainage.

Methods

Study population

The RELAP trial was a randomized controlled trial comparing the on-demand relaparotomy strategy with the planned relaparotomy strategy in patients with severe abdominal sepsis. The ‘on-demand’ strategy was defined as performing a relaparotomy only in case of clinical deterioration or lack of improvement, monitored by physiological, laboratory and radiology parameters. The planned strategy was defined as performing a relaparotomy every 36 to 48 hours until the abdomen was macroscopically clean at the beginning of the final relaparotomy\(^1\).

The main inclusion criteria were secondary peritonitis for which an emergency laparotomy was being performed and severe disease indicated by an APACHE II score \(>10\). More details have been reported in the article presenting the main results of the RELAP trial\(^1\).
The trial cohort consisted of 229 patients of which 114 were randomly allocated to the on-demand strategy and 115 to the planned strategy. For the present study we included patients from both arms of the study (n=229), but excluded those patients in whom the first relaparotomy took place more than 7 days after the initial emergency operation (n=8). We maintained this 7-day period as ongoing infection directly associated with the primary disease is presumed to occur within this time window. Problems arising beyond this 7-day period are more likely due to other causes (intra- or extra-abdominal), once elimination of the infectious focus has been reached during index laparotomy. Moreover, the essential difference between the planned and on-demand strategy lies in the first relaparotomy which in the planned strategy is being performed on the second or third day after index surgery. In contrast, a commitment for relaparotomy on the second or third day is not part of the on-demand strategy, but instead the need for a relaparotomy is determined day by day in particular this first week.

**Outcome**

**Ongoing infection needing relaparotomy**

In the planned strategy group a patient was classified as ‘ongoing infection needing a relaparotomy’ in case of macroscopic positive findings at relaparotomy. Moreover, all ‘planned patients who died within 14 days were classified as ‘ongoing infection’. A negative relaparotomy (no residual infection or new pathology) was classified as ‘no ongoing infection, not needing relaparotomy’.

In the on-demand treated patients only 48% received a first relaparotomy. For the re-operated patients applied that a patient was classified as ‘ongoing infection, needing a relaparotomy’ in case of macroscopic positive findings at relaparotomy. In the 52% of the on-demand patients who were not re-operated obviously no direct visual inspection of the abdomen has been performed. Therefore, we used a 14-day follow-up period as additional verification: if a patient without visual verification died within this 14-day period, the patient was still classified as ‘ongoing infection needing relaparotomy’. Furthermore, if patients without visual verification survived at least 2 weeks after the index surgery they were also classified as ‘no ongoing infection, not needing relaparotomy’. The assumption was that if in these surviving patients an early relaparotomy had been performed, findings would have been negative (Figure 1).

**Ongoing infection needing re-intervention**

Patients needing re-intervention (n=82) included all patients that were classified as ‘ongoing infection needing relaparotomy’ (n=71, Figure 1) supplemented by the group of patients that received a US- or CT-guided percutaneous drainage (PCD) of abdominal fluid with placement of drains for continuous drainage and lavage (n=11, Figure 1). Consequently, for this second analysis 139 patients were labeled as ‘no ongoing infection, not needing re-intervention’.
Chapter 6  Existing scoring systems to predict ongoing infection

Initial cohort
n = 229
RELAP patients with peritonitis undergoing an emergency (index) laparotomy

Excluded
Relaparotomy >7 days after index (n=8)

Study population
n = 221
Patients with peritonitis at risk for ongoing infection

Patients undergoing a relaparotomy (n=148)

Patients not undergoing a relaparotomy (n=73)
No relaparotomy but PCD with drains left in situ (n=11)
No relaparotomy and no PCD (n=62)

Patients with a positive relaparotomy (n = 63)

Patients with no relaparotomy but dead < 14 days (n=8)

Ongoing infection needing relaparotomy
n = 71 (32%)

Patients with a negative relaparotomy (n=85)

Patients with no relaparotomy with or without PCD and >14 day survival (n=65)

No ongoing infection, not needing relaparotomy
n = 150 (68%)

Figure 1  Flow chart showing patient selection and outcome definition
### Table 1  Overview of components of the various existing scoring systems

<table>
<thead>
<tr>
<th>Component</th>
<th>APACHE II</th>
<th>SAPS II</th>
<th>MPI</th>
<th>SOFA</th>
<th>MODS</th>
<th>APS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of admission</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic disease present</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malignant comorbidity</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organ failure present</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>White blood cell counts</td>
<td></td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygenation*</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Mechanical ventilation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Respiratory rate</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arterial pH</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>HCO3-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Cardiovascular state**</td>
<td></td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart rate</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Blood pressure***</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Hematocrite</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creatinine</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Urine output</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Urea</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Bilirubin</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Coagulation (thrombocytes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Potassium</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Sodium</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Glasgow coma score</td>
<td>√</td>
<td></td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease specific parameters****</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*PaO2 or PaO2/FiO2 ratio

**Includes administration and dosage of (nor)epinephrine, dopamine or dobutamine

***Systolic blood pressure, mean arterial pressure (MAP) or as part of the pressure adjusted heart rate (PAR)

****Disease-specific parameters include duration of peritonitis, origin, extent of peritonitis and type of contamination
Scoring systems

Only widely used scoring systems were evaluated for their performance in predicting the need of a relaparotomy after the index operation. The scoring systems evaluated were the Acute Physiology and Chronic Health Evaluation (APACHE) II score\(^2\), the simplified Acute Physiology Score (SAPS) II\(^3\), the Mannheim Peritonitis Index (MPI)\(^4,5\), the Multiple Organ Dysfunction Score (MODS)\(^6\), the Sepsis-related Organ Failure Assessment (SOFA) score\(^7\) and the APS, the physiological part extracted from the APACHE II score\(^2\).

The APACHE II score was assessed, following the RELAP trial protocol, using the worst values of each independent constituent in a 24-hour time frame including the index laparotomy\(^7\). The SAPS II and MPI were compiled by adding the worst values of the independent constituents in the initial 24 hours of index laparotomy.

The sequential scores, SOFA, APS and MODS were calculated on day 1 and day 2, following the index laparotomy by adding the worst values of the independent constituents measured that day (Table 1).

Single time point measurements

**Acute Physiology and Chronic Health Evaluation (APACHE) II score**

The APACHE II score has been designed by Knaus et al. in 1985 to assess initial severity of disease for patients admitted to the ICU in relation to predicting patient outcome (death)\(^2\). The score consists of a chronic part (age and comorbidity) and a physiological part evaluating the condition of various organ systems based on values of laboratory parameters. The APACHE II score is developed as single time point measurement (range of score: minimum 0 to maximum 71)\(^2\).

**Simplified Acute Physiology Score (SAPS) II**

The SAPS II is aimed at the prediction of mortality in critically ill patients admitted to the ICU\(^3\). The score consists of 17 variables: 12 physiological variables, age, type of admission and 3 variables related to underlying disease. The SAPS II is validated to predict mortality for groups of patients rather than individual patients (range of score: minimum 0 to maximum 163) as a single time point measurement\(^3,8\).

**Mannheim Peritonitis Index (MPI)**

The MPI has been developed and validated to predict mortality for patients with secondary peritonitis\(^4,5\). It consists of only non-physiological variables: general patient characteristics (age, gender, organ failure and malignancy present at time of presentation) supplemented with disease-specific parameters (duration of peritonitis, origin, extent of peritonitis and type of contamination). Assessment of risk of mortality by the MPI is also based on a single time point measurement. And again, the score is validated to be useful in risk assessment for groups of patients rather than the individual patient (range of score: minimum 0 to maximum 47)\(^5\).
Sequential scores

**Acute Physiology Score (APS) extracted from the APACHE II score**

The acute physiological part of the APACHE II score was independently evaluated as sequential scoring system (range of score: minimum 0 to maximum 58).

**Sequential Organ Failure Assessment (SOFA) score**

The SOFA score was created by a joint effort of the European Society of Intensive Care Medicine to quantify and objectify the degree of organ dysfunction in critically ill patients. The SOFA score consists of only physiological parameters associated with various clinically important organ systems, and can therefore be sequentially calculated. An increase in SOFA score is associated with an increase in failing organ systems and death (range of score: minimum 0 to maximum 24).

**Multiple Organ Dysfunction Score (MODS)**

The MODS was developed by Marshall et al. to assess the severity of the multiple organ dysfunction as outcome in critical illness. This score consists of only physiological parameters associated with various clinically important organ systems, and can therefore be sequentially calculated. An increase in MODS is associated with an increase in failing organ systems. A limitation to the MODS is that several components are only measured in the ICU, making this scoring only usable as a predictive tool for patients admitted to the ICU (range of score: minimum 0 to maximum 24).

Statistical analysis

**Baseline characteristics**

Demographic data, clinical characteristics and findings at the initial emergency operation were compared between patients who did or did not have ongoing infection needing relaparotomy. Continuous variables were expressed as mean ± standard deviation (±SD) or median (25–75% interquartile range) and compared, respectively, using Student’s t test or Mann–Whitney U test depending on the distribution of the data. Categorical variables were reported as absolute numbers (frequency with percentages) and analyzed using χ² test.

**Predicting ‘ongoing infection needing relaparotomy’**

Our aim was to assess the ability of existing scoring systems to identify patients with ongoing infection needing a relaparotomy. Therefore, we focused on the area under the receiver operating characteristic (ROC) curve (AUC) to express the ability of each scoring system to discriminate between patients who did or did not have an ongoing infection. The AUC is a single indicator of discriminative power, and its value can be interpreted as the probability that a randomly chosen patient needing relaparotomy (a ‘case’) had a higher score on an instrument than a randomly chosen patient not needing a relaparotomy (a ‘control’). If an existing scoring system has an AUC of 1 then it can perfectly discriminate between cases and controls. An area of 0.5 is synonymous with no better discrimination beyond chance.
AUC’s are presented for the APACHE II, SAPS II and MPI at index laparotomy (worst 24-hour values). AUC’s are presented for the MODS, SOFA score, and APS at day 1, day 2, and the absolute difference between day 2 and day 1. Patients who were already dead (no scores) and patients who already received a relaparotomy prior to the day of the measurements were not included in the analyses. We used logistic regression models to calculate these areas under the curves, also known as the concordance or c-statistic. These logistic regression models also provide odds ratios and 95% confidence intervals (CI) expressing the strength of the association between a risk score and the probability of ongoing infection needing relaparotomy.

Although we tried to harmonize the definition of ongoing infection needing relaparotomy across the two surgical strategies, we specifically examined whether there was a difference in predictive capability of one of the existing scoring system between patients treated by the on-demand and those treated by the planned strategy. To examine this, logistic regression models were made that contained type of surgical strategy, the scoring system, and the interaction between surgical strategy and the scoring system. A significant p-value for the interaction term would indicate that the predictive ability of such a scoring system was different between patients treated by the on-demand and those treated by planned strategy, and probably related to a difference in defining the outcome.

If a scoring system had a significantly better discriminatory ability than can be expected by chance only (AUC >0.6), we calculated a specific cut-off point that would have led to a sensitivity of 90%. In other words, applying that cut-off value would have identified 90% of all patients needing relaparotomy and consequently miss 10% of these patients. Such a cut-off analysis shows the consequences of applying a specific cut-off value, in particular it reveals the number of patients not needing a relaparotomy but who would have been operated based upon a score above such a cut-off value (1 minus specificity).

**Predicting ‘ongoing infection needing re-intervention’**

To assess the ability of existing scoring systems to identify patients with ongoing infection using the wider definition of ‘needing re-intervention’ (relaparotomy or percutaneous drainage; see section above), we determined the AUC’s as measure of predictive value for ‘ongoing infection needing reintervention’11.

**Predicting mortality**

To verify consistency of data we examined the ability of existing scoring systems to predict mortality, the very purpose for which they were originally developed and validated. We again used AUC’s to express the ability of each scoring system to discriminate between patients who died in-hospital and those who survived their hospital admission11.

**Missing values**

As the various existing scoring systems were based on a multitude of different variables that had to be measured or recorded on consecutive days, there were inevitably missing values in our dataset. Although we had a low rate of missing values in the underlying variables, missing values in a
compository score would introduce uncertainty even when values of other variables were present. Therefore, we used multiple imputation to replace missing values with a set of plausible values that represent the uncertainty about the right value to impute. The multivariate relationships between all underlying variables were used to impute the missing values. The method of imputation was Markov Chain Monte Carlo as implemented by the SAS system. Appropriate transformations were applied to individual variables to improve normality. For patients not admitted to the ICU at day 1 and/or day 2 normal values were imputed for components specifically associated with ICU care (central venous pressure, air oxygen pressure (FiO2) – in case of no oxygen supplementation – arterial oxygenation (PaO2) and Glasgow Coma Scale). The multiple imputed data sets were then analyzed (one by one) by using standard procedures (e.g. logistic regression) for complete data. In a next step the results from these analyses were combined to produce estimates and confidence intervals that properly reflect the uncertainty due to missing values. We used a total of ten rounds of imputation to estimate the final parameters with their confidence intervals.

P-values less than 0.05 were considered statistically significant. Statistical analyses were carried out using SAS 9.1 (SAS Institute, Cary, NC).

Results

Patient inclusion

The study population consisted of 221 patients at risk for ongoing infection after emergency laparotomy because of secondary peritonitis. Figure 1 shows that using the specified outcome definition 71 patients (32%) were classified as ‘ongoing infection needing relaparotomy’, whereas 150 patients (68%) were classified as ‘no ongoing infection, not needing relaparotomy’. For the sequential scoring systems, 1 patient was excluded from the analyses at day 1, because of relaparotomy within several hours of index laparotomy. For analyses at day 2 some more patients were excluded from analyses: 13 patients had already undergone a relaparotomy, and 3 patients had died, leaving 204 patients available for analyses (needing relaparotomy, n=64 and not needing relaparotomy, n=140).

Table 2 lists the demographic and baseline characteristics of both outcome groups. More than 90% of patients had been admitted to the ICU. As can be expected, patients classified as ‘ongoing infection needing relaparotomy’ had a significant longer ICU stay and high mortality rate.
Table 2  Demographic, initial peritonitis and recovery data for all 221 included patients

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Ongoing infection needing relaparotomy (n = 71)</th>
<th>No ongoing infection, not needing relaparotomy (n = 150)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age (IQR)</td>
<td>66.9 (57 to 74)</td>
<td>70.0 (57 to 76)</td>
<td>0.36</td>
</tr>
<tr>
<td>Male n (%)</td>
<td>38 (54%)</td>
<td>65 (43%)</td>
<td>0.16</td>
</tr>
<tr>
<td>Postoperative peritonitis n (%)</td>
<td>32 (45%)</td>
<td>71 (47%)</td>
<td>0.75</td>
</tr>
<tr>
<td>Diffuse extent (generalized peritonitis) n (%)</td>
<td>47 (66%)</td>
<td>88 (59%)</td>
<td>0.48</td>
</tr>
<tr>
<td>ICU admission n (%)</td>
<td>69 (97%)</td>
<td>135 (90%)</td>
<td>0.061</td>
</tr>
<tr>
<td>Length of ICU stay (median days (IQR))</td>
<td>12 (7 to 32)</td>
<td>7 (4 to 16)</td>
<td>0.001</td>
</tr>
<tr>
<td>Length of index hospital stay (median days (IQR))</td>
<td>32 (17 to 74)</td>
<td>28 (17 to 49)</td>
<td>0.28</td>
</tr>
<tr>
<td>In-hospital mortality n (%)</td>
<td>26 (37%)</td>
<td>23 (15%)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Predicting ongoing infection needing relaparotomy

None of the interactions between existing scoring systems and the applied surgical strategy (on-demand or planned relaparotomy strategy) were significant with p-values ranging from 0.194 (for the APS at day 1) to 0.986 (for the MODS at day 2). Therefore, all ROC analyses were based on the total cohort of patients from the RELAP trial (Table 3). The ROC curves showing the pairs of sensitivity and specificity for the possible cut-off values are presented in Figure 2.

The severity of disease scores, APACHE II score, SAPS II and MPI, are single time point measurements for critically ill patients. The APACHE II score showed an AUC of 0.50 (95% CI 0.42 to 0.59), indicating no predictive value for ongoing infection needing relaparotomy. The SAPS II score had an AUC of 0.61 (95% CI 0.52 to 0.69), and the MPI score revealed an AUC of 0.53 (95% CI 0.44 to 0.61) for prediction of ‘needing relaparotomy’ (Table 3, Figure 2).

The predictive value of the SOFA score was modest and similar at the 2 time points (day 1: AUC 0.60, 95% CI 0.51 to 0.68; day 2: AUC 0.60, 95% CI 0.52 to 0.69) (Table 2, Figure 2). The absolute difference (delta SOFA scores between day 2 and day 1 measurement scores) showed no discriminatory value with an AUC of 0.55, 95% CI 0.46 to 0.64. The MODS had also no discriminatory value for ongoing infection (day 1: AUC 0.58 (95% CI 0.50 to 0.66); day 2: AUC 0.56 (95% CI 0.47 to 0.64); Table 3, Figure 2).

The acute component of the APACHE II score, the APS, had a modest discriminatory ability at day 1 with an AUC of 0.62 (95% CI 0.53 to 0.70), at day 2 the AUC was 0.57 (95% CI 0.48 to 0.66; Table 3, Figure 2).
Figure 2  Receiver operating characteristic (ROC) curves associated with the area under the curves (AUC’s) depicted in Table 2.

ROC’s are depicted for the scoring systems measured at a single time point (a. APACHE-II, SAPS-II, MPI), the sequential scoring systems (MODS, SOFA, APS) measured at day 1 (b.) and at day 2 (c.).
Table 3  Ongoing infection needing relaparotomy

Mean scores with associated standard deviations (SD) compared for patients with ‘ongoing infection needing relaparotomy’ (n=71) and patients with ‘no ongoing infection, not needing relaparotomy’ (n=150). Predictive value of evaluated existing scoring systems comparing these groups is presented as area under the curve (AUC) with associated 95% confidence intervals (CI). P-values indicating whether the AUC significantly differs from 0.5 (no predictive value) are given.

<table>
<thead>
<tr>
<th>Score</th>
<th>Ongoing infection needing relaparotomy [Mean score (SD)]</th>
<th>No ongoing infection, not needing relaparotomy [Mean score (SD)]</th>
<th>AUC (95% CI)</th>
<th>P-value whether AUC different from 0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>APACHE II</td>
<td>16.1 (4.8)</td>
<td>15.7 (4.0)</td>
<td>0.50 (0.42 to 0.59)</td>
<td>0.944</td>
</tr>
<tr>
<td>SAPS II</td>
<td>40.1 (11.9)</td>
<td>35.6 (10.9)</td>
<td>0.61 (0.52 to 0.69)</td>
<td>0.012</td>
</tr>
<tr>
<td>MPI</td>
<td>28.2 (6.7)</td>
<td>27.5 (7.6)</td>
<td>0.53 (0.44 to 0.61)</td>
<td>0.532</td>
</tr>
<tr>
<td>SOFA (day 1)</td>
<td>8.2 (4.0)</td>
<td>6.8 (4.0)</td>
<td>0.60 (0.51 to 0.68)</td>
<td>0.025</td>
</tr>
<tr>
<td>SOFA (day 2)*</td>
<td>7.9 (4.4)</td>
<td>6.1 (4.0)</td>
<td>0.60 (0.52 to 0.69)</td>
<td>0.019</td>
</tr>
<tr>
<td>SOFA Delta (day 2-day 1)*</td>
<td>-0.17 (2.5)</td>
<td>-0.78 (2.7)</td>
<td>0.55 (0.46 to 0.64)</td>
<td>0.302</td>
</tr>
<tr>
<td>MODS (day 1)</td>
<td>5.0 (2.8)</td>
<td>4.2 (2.6)</td>
<td>0.58 (0.50 to 0.66)</td>
<td>0.064</td>
</tr>
<tr>
<td>MODS (day 2)*</td>
<td>4.4 (3.0)</td>
<td>3.7 (2.6)</td>
<td>0.56 (0.47 to 0.64)</td>
<td>0.208</td>
</tr>
<tr>
<td>MODS Delta (day 2-day 1)*</td>
<td>-0.47 (1.9)</td>
<td>-0.49 (1.7)</td>
<td>0.51 (0.42 to 0.60)</td>
<td>0.789</td>
</tr>
<tr>
<td>APS (day 1)</td>
<td>8.4 (4.5)</td>
<td>6.4 (3.6)</td>
<td>0.62 (0.53 to 0.70)</td>
<td>0.008</td>
</tr>
<tr>
<td>APS (day 2)*</td>
<td>6.9 (3.5)</td>
<td>5.7 (3.1)</td>
<td>0.57 (0.48 to 0.66)</td>
<td>0.107</td>
</tr>
<tr>
<td>APS Delta (day 2-day 1)*</td>
<td>-1.5 (3.9)</td>
<td>-0.73 (3.5)</td>
<td>0.55 (0.46 to 0.64)</td>
<td>0.322</td>
</tr>
</tbody>
</table>

*Only for patients admitted to ICU at 24 and 48 hours (n= 168)

Consequences of cut-off values producing 90% sensitivity

For scoring systems that performed significantly better than chance, we determined a cut-off score that would produce a sensitivity of 90% and then calculated the proportion of patients undergoing an unnecessary relaparotomy (1 minus specificity). A score of 27 determined for the SAPS II would have identified 90% of the patients requiring a relaparotomy, but the corresponding specificity was only 20%. Based on this cut-off value, the positive predictive value would have been 34%, indicating that for all 187 patients with a score above this cut-off of 27 only 64 patients would have been rightfully re-operated upon, and 123 patients would have been re-operated under suspicion of ongoing abdominal infection but resulting in negative findings at relaparotomy. The negative predictive value, based on this cut-off value, would have been 79%, indicating that 27 of the 34 patients with a score under the cut-off value of 27 would rightfully not have been re-operated. However, in 7 patients with ongoing
infection needing relaparotomy this re-operation would have been withheld (Table 4). At a 90% sensitivity cut-off the SOFA day 1 score (2.1) would have a positive predictive value of 33% and a negative predictive value of 72%. The APS day 1 score showed sensitivity and predictive values similar to those of the SAPS II (90% sensitivity cut-off value 3.1; Table 4).

Predicting ongoing infection needing re-intervention

In total 82 patients were classified as ‘ongoing infection needing re-intervention’ (Figure 1) and 139 patients were classified as ‘no ongoing infection, not needing re-intervention’. Only the APS at day 1 showed an AUC >0.6 for prediction of ‘ongoing infection needing re-intervention’ (AUC 0.61, 95% CI 0.53 to 0.68). All other scoring systems showed no discriminatory value (Table 5).

Table 4  Selection of patients for relaparotomy using 90% sensitivity cut-off values

Performance of scoring systems, moderately predictive for ‘ongoing infection needing relaparotomy’ (AUC > 6.0; SAPS II, SOFA day 1 and APS day 1), in selecting patients for relaparotomy using a cut-off value based on a 90% sensitivity.

The associated sensitivity, the positive predictive value (PPV) and negative predictive value (NPV) are presented.
Predicting mortality

All evaluated scoring systems were predictive of mortality in peritonitis patients, as AUC’s were significantly different from an AUC of 0.5. Results of predictive value for in-hospital mortality are presented in Table 6. The SAPS II was the best predictor of mortality with an AUC of 0.80 (0.73 to 0.87), whilst the disease-specific MPI was the least predictive of mortality with an AUC of 0.60 (0.51 to 0.69).

Discussion

The predictive value of available scoring systems, in particular those that can be assessed sequentially, for ongoing abdominal infection needing relaparotomy are not known. In clinical practice changes in organ functions are seen as useful triggers to employ diagnostic tools or intervene. Of the single time point scoring systems only the SAPS II had an area under the curve that was significantly above 0.5, whereas the APACHE II score and MPI had no discriminatory ability for ongoing infection needing relaparotomy. Sequential scoring systems SOFA and APS were equal but very modest performers. The performance of all sequential scoring systems did not improve when using day 2 instead of day 1 scores or using the difference between day 2 and day 1.

Even the scores with AUC’s >0.60 showed an extremely low specificity at 90% sensitivity. This low specificity is a trade-off for requesting a high sensitivity in situations where the overall performance of a score is moderate, even poor. Broadening the definition of ongoing abdominal infection by patients needing re-intervention (relaparotomy or percutaneous drainage) did not enhance identification of patients with (persistent) peritonitis.

The RELAP trial concludes that the on-demand strategy is preferred1. Stringent monitoring of patients is a vital component of the on-demand strategy. A scoring system can aid in adequate and timely identification of patients for relaparotomy. Ideally, such a prediction model should be a sequential score, rather than a score based on a single time-point measurement. Changes in organ failure may be of better value in objectifying the clinical course of the disease, in particular since postoperative (follow-up) variables are more predictive than variables that become available during index laparotomy18.

We were surprised by the low performance of these well-known scoring systems, as most of these scores quantify organ function. However, none of the evaluated scoring systems were originally developed to predict the need of a relaparotomy for ongoing peritonitis following acute peritonitis2-7. All scores, except the MPI, have been developed to predict death for ICU patients in general and for groups of patients (strata) rather than predicting death for individual patients2,3,6,7. Although the MPI is specifically developed for patients with abdominal sepsis, it is focused on prediction of death for this group of patients rather than occurrence of ongoing infection19. Also, the MPI largely consists of peritonitis-related data, determined at the initial emergency laparotomy1.
These variables are described to be less predictive than physiological post-operative variables\(^\text{18}\). All scores, indeed, did better at predicting death, as they are developed and validated to do. Apparently, quantification of organ failure or changes in organ failure can not differentiate between ongoing organ failure due to abdominal sepsis despite source control of the initial causative focus and ongoing abdominal infection.

For the patients that were not re-operated, the time frame in which the predictor status of the sequential scoring systems was assessed was less precise. Our best deduction was to evaluate the scores during the clinical phase in which the dilemma of early relaparotomy is most prominent; day 1 and day 2 after the initial emergency laparotomy, for all included patients. In view of the disappointing predictive values, it is unlikely that extension of assessment after day 2 would have revealed different results.

Patients included in this study were randomized to the on-demand or the planned strategy\(^\text{1}\). This enhances the generalizability of the results but foremost eliminates selection bias in choice of practiced treatment strategies. Differences in these treatment arms did lead to differential verification, but not necessarily to verification bias. Another option would have been to use only the planned arm of the trial, as all these patients were re-operated and had uniform outcome verification. Instead, all existing scoring systems were tested for the assumption that both the on-demand and planned strategy could be combined for the above analyses and we found no significant interaction between treatment strategies and predictive ability of the various scoring systems. This means that the predictive ability of

<table>
<thead>
<tr>
<th>Score</th>
<th>AUC (95% CI)</th>
<th>P-value whether AUC different from 0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>APACHE II</td>
<td>0.49 (0.42 to 0.57)</td>
<td>0.880</td>
</tr>
<tr>
<td>SAPS II</td>
<td>0.58 (0.50 to 0.66)</td>
<td>0.041</td>
</tr>
<tr>
<td>MPI</td>
<td>0.54 (0.47 to 0.62)</td>
<td>0.289</td>
</tr>
<tr>
<td>SOFA (day 1)</td>
<td>0.58 (0.50 to 0.65)</td>
<td>0.064</td>
</tr>
<tr>
<td>SOFA (day 2)*</td>
<td>0.58 (0.51 to 0.66)</td>
<td>0.039</td>
</tr>
<tr>
<td>MODS (day 1)</td>
<td>0.57 (0.49 to 0.65)</td>
<td>0.088</td>
</tr>
<tr>
<td>MODS (day 2)*</td>
<td>0.57 (0.49 to 0.65)</td>
<td>0.108</td>
</tr>
<tr>
<td>APS (day 1)</td>
<td>0.61 (0.53 to 0.68)</td>
<td>0.009</td>
</tr>
<tr>
<td>APS (day 2)*</td>
<td>0.55 (0.47 to 0.64)</td>
<td>0.191</td>
</tr>
</tbody>
</table>

* Only for patients admitted to ICU at 24 and 48 hours (n= 168)

These variables are described to be less predictive than physiological post-operative variables\(^\text{18}\). All scores, indeed, did better at predicting death, as they are developed and validated to do. Apparently, quantification of organ failure or changes in organ failure can not differentiate between ongoing organ failure due to abdominal sepsis despite source control of the initial causative focus and ongoing abdominal infection.

For the patients that were not re-operated, the time frame in which the predictor status of the sequential scoring systems was assessed was less precise. Our best deduction was to evaluate the scores during the clinical phase in which the dilemma of early relaparotomy is most prominent; day 1 and day 2 after the initial emergency laparotomy, for all included patients. In view of the disappointing predictive values, it is unlikely that extension of assessment after day 2 would have revealed different results.

Patients included in this study were randomized to the on-demand or the planned strategy\(^\text{1}\). This enhances the generalizability of the results but foremost eliminates selection bias in choice of practiced treatment strategies. Differences in these treatment arms did lead to differential verification, but not necessarily to verification bias. Another option would have been to use only the planned arm of the trial, as all these patients were re-operated and had uniform outcome verification. Instead, all existing scoring systems were tested for the assumption that both the on-demand and planned strategy could be combined for the above analyses and we found no significant interaction between treatment strategies and predictive ability of the various scoring systems. This means that the predictive ability of
existing scoring systems is comparable for both on-demand and planned treated patients. Importantly, also the proportion of events (positive findings at relaparotomy) was comparable for both strategies (29% for on-demand vs. 32% for planned)\(^1\).

### Table 6  In-hospital mortality

Predictive value for mortality comparing patients who died during hospital admission (\(n=49\)) and those who survived hospital admission (\(n=172\)). Predictive value is presented as area under the curve (AUC) with associated 95% confidence intervals (CI) and p-values indicating whether the AUC significantly differs from 0.5 (no predictive value).

<table>
<thead>
<tr>
<th>Score</th>
<th>AUC (95% CI)</th>
<th>P-values whether AUC different from 0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>APACHE II</td>
<td>0.74 (0.65 to 0.82)</td>
<td>(&lt; 0.001)</td>
</tr>
<tr>
<td>SAPS II</td>
<td>0.80 (0.73 to 0.87)</td>
<td>(&lt; 0.001)</td>
</tr>
<tr>
<td>MPI score</td>
<td>0.60 (0.51 to 0.69)</td>
<td>0.031</td>
</tr>
<tr>
<td>SOFA (day 1)</td>
<td>0.68 (0.59 to 0.77)</td>
<td>(&lt; 0.001)</td>
</tr>
<tr>
<td>MODS (day 1)</td>
<td>0.76 (0.68 to 0.83)</td>
<td>(&lt; 0.001)</td>
</tr>
<tr>
<td>APS (day 1)</td>
<td>0.69 (0.60 to 0.78)</td>
<td>(&lt; 0.001)</td>
</tr>
</tbody>
</table>

For clinical purposes, the discriminatory power is more important than stratification. A 90% level of sensitivity was employed, as it is considered worse to mistakenly not re-operate a patient with ongoing infection needing relaparotomy than it is to re-operate a patient on the suspicion of ongoing infection but with negative findings\(^{18,20}\). The approximate of 90% sensitivity was chosen to determine a cut-off for adequate scoring systems, reflecting the trade-off between a false positive prediction of peritonitis (negative relaparotomy) and a false negative prediction of peritonitis (no relaparotomy although one is needed). Nevertheless, performing too many negative relaparotomies should be avoided\(^21\). None of the scoring systems had a clinically important predictive value nor demonstrated a clinically useful discriminatory ability. In order not to withhold relaparotomy from too many patients who need treatment for ongoing infection, an unacceptably high proportion of inappropriate relaparotomies would be performed based on these scores.

All presented existing scoring systems lack the additional information derived from diagnostic imaging techniques which is likely valuable for selection of patients with ongoing infection needing re-intervention. For patients suspected of abdominal infection following elective abdominal surgery, CT imaging has a high diagnostic accuracy\(^22\). The exact value of diagnostic imaging in operated peritonitis patients with suspected ongoing abdominal infection is not known, as consequences from management decisions based on CT results have not been evaluated yet. Future research is needed to determine the exact accuracy of CT scanning in on-demand treated peritonitis patients who are suspected of ongoing infection.
In conclusion, none of the existing and widely used severity of disease scores, specifically developed for critically ill patients, were clinically useful in the identification of patients with ongoing infection needing a relaparotomy. Therefore, new tools need to be developed that specifically incorporate parameters indicative for ongoing abdominal infection, rather than merely ongoing organ failure, in patients with abdominal sepsis. Preferably, these specific tools combine clinical findings, laboratory measurements and results from diagnostic imaging tests to assist the multidisciplinary team in selecting patients for re-intervention to treat ongoing abdominal infection.
References


“Health is a state of complete physical, mental and social well-being, and not merely the absence of disease or infirmity.”

*World Health Organisation, 1948*
Part III

Health related-quality of life in patients with secondary peritonitis
Chapter 7
Health related-quality of life six months following surgical treatment for secondary peritonitis – using the EQ-5D questionnaire

Kimberly R Boer, Oddeke van Ruler, Johannes B Reitsma, Cecilia W Mahler, Brent C Opmeer, E Ascelijn Reuland, Hein G Gooszen, Peter W de Graaf, Eric J Hesselink, Michael F Gerhards, E Philip Steller, Mirjam A Sprangers, Marja A Boermeester, Corianne A De Borgie

For the Dutch Peritonitis Study Group

Health Qual Life Outcomes. 2007 Jul 2;5(1):35
Abstract

Objective: To compare health related-quality of life (HR-QoL) in patients surgically treated for secondary peritonitis to that of a healthy population, and to prospectively identify factors associated with poorer (lower) HR-QoL.

Design: A prospective cohort of secondary peritonitis patients was mailed the EQ-5D and EQ-VAS 6 months following initial laparotomy.

Setting: Multicenter study in 2 academic and 7 regional teaching hospitals.

Patients: 130 of the 155 eligible patients (84%) responded to the HR-QoL questionnaires.

Results: HR-QoL was significantly worse on all dimensions in peritonitis patients than in a healthy reference population. Peritonitis characteristics at initial presentation were not associated with HR-QoL at 6 months. A more complicated course of the disease leading to longer hospitalization times and patients with an enterostomy had a negative impact on the mobility (p<0.02), self-care (p<0.001) and daily activities (p=0.01). In a multivariate analysis for the EQ-VAS every doubling of hospital stay decreases the EQ-VAS by 3.8 points (p=0.015). Morbidity during the 6-months follow-up was not found to be predictive for the EQ-5D or EQ-VAS.

Conclusion: Six months following initial surgery, patients with secondary peritonitis report more problems in HR-QoL than a healthy reference population. Unfavorable disease characteristics at initial presentation were not predictive for poorer HR-QoL, but a more complicated course of the disease was most predictive of HR-QoL at 6 months.
Introduction

Secondary peritonitis has a high in-hospital mortality (24-35%), continued high post-hospital discharge mortality, as well as a considerable long-term morbidity\(^5\). Patients are hospitalized for extensive periods of time and often endure lengthy intensive care unit (ICU) stays\(^5\)\(^6\)\(^7\). Recently, improving health related-quality of life (HR-QoL) in patients with sepsis\(^1\)\(^1\)\(^1\)\(^3\)\(^4\) has become a complementary goal in patient care\(^5\). The importance of HR-QoL will continue to grow with improvement in peritonitis survival. Till now, most HR-QoL data in secondary peritonitis and abdominal sepsis have been collected retrospectively\(^4\)\(^5\)\(^6\)\(^7\)\(^8\)\(^9\). These studies have shown that peritonitis patients suffer from HR-QoL impairments both in the short term as well as the long term. Good quality data from prospective studies are necessary to identify factors related to lower HR-QoL. Insight into these factors is needed to inform patients, to develop preventive measures for high-risk patients, and to provide tailored support for individual patients.

The aims of this study were twofold. Firstly, to assess HR-QoL in patients with secondary peritonitis, and to compare this with HR-QoL reported for a general reference population\(^10\). And secondly, to determine which factors (patient, peritonitis and postoperative) are related to HR-QoL 6 months following patients with severe secondary peritonitis (APACHE II >10)\(^9\)\(^1\)\(^2\).

Methods

Study design

This study was embedded in an ongoing peritonitis trial evaluating two surgical strategies for patients with peritonitis, initiated by the Academic Medical Center (AMC), Amsterdam, The Netherlands. Patients were enrolled between December 2001 and August 2005 in 2 academic and 7 regional teaching hospitals in The Netherlands.

Patients

Patients were eligible for the RELAP trial if they had a clinical diagnosis of secondary peritonitis requiring emergency laparotomy. Peritonitis had to be caused by perforation or infection of a visceral organ, or ischemia/necrosis of part of the gastrointestinal tract or postoperative peritoneal infection. An Acute Physiology And Chronic Health Evaluation (APACHE) II score above 10 was required, as the preferred strategy for mild peritonitis (APACHE II score ≤10) is on-demand. Exclusion criteria included: age below 18 or above 80; peritonitis due to bowel perforation after endoscopy operated within 24 hours; abdominal infection due to indwelling dialysis (CAPD) catheter; acute pancreatitis; expected survival of less than 6 months due to disseminated malignancy; severe brain damage due to trauma or anoxia; imperative relaparotomy (e.g., gauze packing).

To be eligible for participation in the present HR-QoL study, patients had to be alive and out of hospital at 6 months following index laparotomy (Figure 1).
Patients included at index operation for HR-QoL study (n = 225)

Patients still alive at month 6 follow-up (n = 162)

Eligible patients for HR-QoL study & sent HR-QoL form (n = 155)

Overall patient response (n = 130)
Response rate = 84%

Patients who died (63)

Readmitted or still in hospital (6)
Address change (2)

Refused (9)
Too ill (7)
Unknown (6)
Questionnaire poorly filled in (1)
Other (2); Alzheimer (1) & Mentally handicapped (1)

Instruments

HR-QoL was assessed approximately 6 months after the index laparotomy by administering the patient self-report Euroqol Five-Dimensions (EQ-5D) questionnaire which includes five dimensions and the Euroqol-Visual Analogue Scale (EQ-VAS)\(^2\). The Euroqol instruments have been extensively validated, including Dutch healthy individuals, and were recently recommended as the instrument of choice in critical care studies\(^22\text{-}25\). EQ-5D was originally designed to complement other instruments but is now increasingly used as a ‘stand alone’ measure. The EQ-5D measures five health dimensions: mobility, self-care, daily activities, pain/discomfort, and mood consisting of both anxiety and depression. In the EQ-5D patients report: 0 (no problems), 1 (moderate problems), and 2 (extreme problems)\(^26\). The EQ-VAS is a thermometer-like scale, in which patients rate their overall well-being from 0 (worst imaginable overall health) to 100 (best imaginable overall health)\(^26\text{-}27\).
Data collection

Preoperative risk factors and postoperative morbidity data were prospectively collected for all eligible patients. HR-QoL data were collected 6 months after index operation. EQ-5D and EQ-VAS questionnaires were sent by mail to patients who survived at least 6 months, with a reminder by phone within 2 weeks if there was no response. After 1 month without response patients were phoned and then a new set of questionnaires with a reminder letter were sent.

Reference population

These same instruments (EQ-VAS and EQ-5D) were used to measure a sample of 851 healthy residents in the Netherlands\textsuperscript{18} and in this study were used as a reference population.

Data analysis

Reference population

The proportion of peritonitis patients reporting moderate or extreme problems (combined together) on each of the EQ-5D dimensions in the study group was compared to the proportion reported by the general Dutch population\textsuperscript{18} using a $\chi^2$ test. Differences in mean EQ-VAS scores were calculated between the study peritonitis patients and the general population stratified by 10-year age groups, and tested for significance using the Student's $t$ test\textsuperscript{26}.

Representatively, the sample with HR-QoL measurements for the non-responders (non-respondent analysis) was evaluated using $\chi^2$ tests to compare categorical data, and the Student's $t$ test or the Mann-Whitney U test for continuous data.

Predictive factors

An initial set of potential factors was based on two previous studies examining factors associated with increased mortality and morbidity in patients with secondary peritonitis\textsuperscript{5,13}. These candidate factors were divided into three distinct categories:

1) **General patient characteristics**: age, gender, and having one or more major comorbidities. Major comorbidities were measured by severity and included cardiovascular disease; chronic obstructive pulmonary disease (COPD); malignancy; renal disease; and diabetes mellitus (DM).

2) **Peritonitis characteristics**: severity of disease at study entry measured by the APACHE II score and severity of peritonitis measured by the Mannheim Peritonitis Index (MPI), extent (localized versus diffuse) and type of contamination (clear, turbid, purulent, fecal), etiology of peritonitis (inflammation, perforation, ischemia/necrosis, anastomotic leakage), and community-acquired versus hospital-acquired or nosocomial infection (these infections include post operative peritonitis as a complication of a previous (elective) surgical intervention or peritonitis that is the result of treatment in a hospital or hospital-like setting).
3) Postoperative characteristics: number of relaparotomies, length of stay in ICU and hospital, duration of mechanical ventilation, complications during ICU stay, i.e., acute respiratory distress syndrome (ARDS). Also factors including having an enterostomy at 6 months, the number of hospital readmissions (for peritonitis-related morbidity) and experiencing one of the predefined severe morbidities during the 6-months follow-up (including incisional hernia, bowel obstruction/herniation, burst abdomen, abdominal compartment syndrome, fistula, intra-abdominal bleeding, perforation, anastomotic leakage, ischemia/necrosis, enterostomy dysfunction, bleeding ulcer, abscess (needing drainage), renal failure, myocardial infarction/embolus/cerebral vascular accident, pneumonia or urosepsis needing readmission (see Appendix 1 for the complete list).

We used a general linear model to identify factors associated with the EQ-VAS, or with the proportion of patients reporting moderate or severe problems on any of the five dimensions of the EQ-5D. Factors associated with HR-QoL (p ≤0.1) were then entered in a multivariate model, unless predictive factors were strongly correlated with each other; then only one factor with the strongest association was chosen. The functional form of continuous predictors was graphically assessed and, in the case of pertinent non-linearity, a transformation was performed. Statistical Package for the Social Sciences (SPSS 11.01, SPSS Inc, Chicago, IL) was used for all data analysis.

Results

A total of 155 surviving patients were eligible for the HR-QoL study and questionnaires were sent to all of them. The overall response rate was 85% (130/155; see Figure 1). The average responses were provided at 6 months and 4 days after index laparotomy.

The mean age of patients at enrollment was 63 years, and 53% of the patients were male (Table 1). Patients at trial entry were generally severely ill, as reflected by a mean APACHE II score of 15.1 and mean MPI of 19.9 (Table 1).

There was no significant difference in any patient baseline characteristics; peritonitis characteristics or postoperative characteristics between patients who responded to the HR-QoL questionnaires (n=130) and patients who did not respond (n= 32, Figure 1).
### Table 1  General patient, peritonitis and postoperative characteristics (n=130).

<table>
<thead>
<tr>
<th>Patient characteristics</th>
<th>(n= 130)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age; mean (SD)</td>
<td>63 (14)</td>
<td></td>
</tr>
<tr>
<td>Males; n</td>
<td>70</td>
<td>53</td>
</tr>
<tr>
<td>&gt; 1 major comorbidity; n</td>
<td>73</td>
<td>56</td>
</tr>
</tbody>
</table>

#### Peritonitis characteristics

<table>
<thead>
<tr>
<th>Peritonitis characteristics</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>APACHE II mean (SD)</td>
<td>15.1 (4.1)</td>
<td></td>
</tr>
<tr>
<td>Mannheim peritonitis index, mean (SD)</td>
<td>19.9 (7.6)</td>
<td></td>
</tr>
</tbody>
</table>

#### Extent of contamination:

- 1 or 2 quadrants: 49 (37)
- Diffuse: 82 (64)

#### Type of contamination:

- Clear: 8 (6)
- Turbid: 29 (22)
- Purulent: 42 (32)

#### Etiology of peritonitis:

- Inflammation: 6 (5)
- Perforation: 72 (55)
- Ischemia / necrosis: 6 (5)
- Anastomotic leakage: 41 (31)
- Other: 6 (5)

#### Hospital-acquired peritonitis patients (peritonitis following earlier elective operation and/or during hospital stay)

- 69 (53)

#### Postoperative characteristics

<table>
<thead>
<tr>
<th>Postoperative characteristics</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pts with &gt; 1 relaparotomy n</td>
<td>86</td>
<td>66</td>
</tr>
<tr>
<td>Relaparotomies; median (range)</td>
<td>1.0 relaps (1 -10)</td>
<td></td>
</tr>
<tr>
<td>Pts admitted to ICU n</td>
<td>115</td>
<td>88</td>
</tr>
<tr>
<td>Length of ICU stay; median (P25 –P75)</td>
<td>9 days (6-21)</td>
<td></td>
</tr>
<tr>
<td>Patients ventilated n</td>
<td>110</td>
<td>84</td>
</tr>
<tr>
<td>Duration of ventilation; median, (P25 –P75)</td>
<td>6 days (3-12)</td>
<td></td>
</tr>
<tr>
<td>Length of hospital stay median, (P25 –P75)§</td>
<td>34 days (19-60)</td>
<td></td>
</tr>
<tr>
<td>Acute Respiratory Distress Syndrome (ARDS)</td>
<td>7</td>
<td>5.4</td>
</tr>
<tr>
<td>Patients readmitted &gt; 1 at 6 months</td>
<td>74</td>
<td>57</td>
</tr>
<tr>
<td>&gt;1 Morbidities during 6-months follow-up*</td>
<td>33</td>
<td>26</td>
</tr>
<tr>
<td>Patients with enterostomy at 6-months</td>
<td>73</td>
<td>56</td>
</tr>
</tbody>
</table>

* Information on morbidities missing for one patient (n = 129)

§ For two patients the exact hospital stay was unknown, due to transfer to other hospital
Comparison with other populations

Compared to a healthy reference population\textsuperscript{18}, the peritonitis group reported significantly more problems on all EQ-5D dimensions (\(p<0.001\) for all dimensions, see Figure 2). Patients with peritonitis showed in all age groups lower EQ-VAS scores than the reference group, indicating worse overall HR-QoL. In the RELAP group, EQ-VAS scores appeared to be low from young till old and did not particularly worsen for those who are older.

Figure 2  Percentage of HR-QoL problems reported by peritonitis study patients (\(n=130\)) compared to a general reference population from The Netherlands (Dutch ref pop) (\(n=851\))\textsuperscript{18} by EQ-5D dimensions.

Predictive factors

Results of the univariate analyses evaluating patient, peritonitis and postoperative factors as predictors for HR-QoL at 6 months are reported in Table 2.

General patient characteristics

In a univariate analysis men reported significantly fewer problems with mobility and daily activities. Increasing age decreased overall well-being and increased problems in mobility, but was protective for mood problems with younger patients scoring more mood problems. Major comorbidities at baseline were predictive for more problems related to mobility and mood at 6 months (Table 2).
Peritonitis characteristics

Peritonitis characteristics were not associated with scores on EQ-VAS or EQ-5D when looking at severity of disease or peritonitis severity, etiology or type and extent of the contamination (Table 2). There were no HR-QoL differences between patients with community-developed peritonitis and patients with hospital-acquired peritonitis.

Postoperative characteristics

Patients who stayed longer in ICU and/or surgical hospital-ward reported more problems on all functional impairment dimensions mobility, self-care and daily activities and overall well-being, but not on the pain and mood dimensions (Table 2). Although ICU stay and hospital stay are clearly associated with HR-QoL, in a univariate analysis mechanical ventilation was not.

Readmissions during the 6-months follow-up were also associated with lower HR-QoL scores. Patients who still had an enterostomy 6 months following surgery reported more problems in the functional impairment dimensions: mobility, self-care and daily activities (the combination of these dimensions is often referred to as a specific discipline within HR-QoL called Activities in Daily Life (ADL)). Overall those patients reported more well-being problems than patients without an enterostomy (Table 2).

Multivariate analysis

The following factors were entered in the multivariate analysis based on the results of the univariate association (p ≤ 0.10) with HR-QoL with at least two of the five EQ-5D dimensions or an effect on the EQ-VAS (Table 2): gender, major comorbidity, enterostomy at 6 months, length of ICU stay and length of hospital stay (a log2 transformation was done to create linearity) and severe morbidity during follow-up. From the literature, it was decided that age should always be added to the models, irrespective of the univariate analyses22,28-30. ICU stay and hospital stay were highly correlated (Spearman’s r=0.681) and therefore not both of the factors could be added to the multivariate model. Length of hospital stay was selected to best represent the accumulation of what a patient underwent following secondary peritonitis, used as an adequate proxy for poor patients recovery and potential complications. The same set of factors were included in all six models (EQ-VAS: Table 3; EQ-5D dimensions: Table 4).

In the multivariate analysis the only independent factor that was predictive for worse overall patient well-being, as measured by the EQ-VAS, was length of hospital stay (log2 transformed); every doubling of the length of hospital stay decreased the EQ-VAS (0-100) score by 3.8 points (p=0.015, Table 3).

In the logistic models for each dimension of the EQ-5D the following factors were predictive of HR-QoL (Table 4). Females reported more mobility problems (OR=2.9, p=0.013), more problems in daily activities (OR=3.7, p=0.006) and more pain and discomfort (OR=2.3, p=0.037). Increasing age was associated with fewer problems with mood (OR=0.54 per 10 years, p=0.001); whilst patients with a major comorbidity were more likely to report problems on the mood dimension (OR=3.6, p=0.007).

Length of hospital stay was associated with more problems in all ADL dimensions; a doubling of the length of hospital stay increased problems in mobility (OR=1.6, p=0.02), self-care (OR=2.5, p<0.001).
and daily activities (OR=1.9, p=0.01). Whilst severe morbidity (as experienced) during the 6-months follow-up was no longer independently associated with lower HR-QoL in the multivariate model. However, longer hospital stay is in part due to severe morbidity; so clinically it may not be possible to consider them apart.

Patients with an enterostomy at 6-months follow-up reported more problems for mobility (OR=2.8, p=0.016) and daily activities (OR=2.8, p=0.027), but not for self-care or mood.

---

**Table 2**  
*Strength of univariate association between potential predictors and reporting problems on the EQ-VAS and on the EQ-5D on which inclusion in final multivariate model is based.*

<table>
<thead>
<tr>
<th>EQ-VAS</th>
<th>EQ-5D</th>
<th>Included into Multivariate model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>VAS</td>
<td>Mobility</td>
</tr>
<tr>
<td>Patient characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Age</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Major Comorbidity</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Peritonitis characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APACHE II</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MPI</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Extent of contamination</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Etiology of peritonitis</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hospital-acquired vs. community-acquired peritonitis</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Postoperative characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pts with ≥1 Relaparotomy</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Acute Respiratory Distress Syndrome (ARDS)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Length of ICU stay</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Length of ventilation</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Length of hospital stay</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>Readmissions during follow-up</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Severe morbidity during 6-months follow-up (Appendix 2)</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Enterostomy at 6-months follow-up</td>
<td>+</td>
<td>++</td>
</tr>
</tbody>
</table>

+ Univariate significance (p < 0.10), ++ univariate significance (p<0.05)  
- No univariate significance found
Discussion

This study shows that patients treated for secondary peritonitis report considerably more complaints on all EQ-5D dimensions 6 months after initial surgery than a general reference population. Furthermore, HR-QoL at 6 months was found to be associated with several patient characteristics and particularly postoperative characteristics, whereas factors directly related initial severity of peritonitis did not affect HR-QoL\(^1\).

Comparisons with other populations

The comparison with a general reference population of healthy individuals allows for a better understanding of the extent of reduction in HR-QoL in this patient group. To give an even better perspective of the extent of the HR-QoL presented here we can compare our peritonitis patient group to a group of general sepsis patients, who were also measured at 6 months following ICU discharge using the EuroQol questionnaire\(^1\). Comparing these groups shows our peritonitis patients reported more problems with ADL, e.g. more problems with mobility and daily activities, despite having comparable APACHE II scores, hospital stay and length of ICU stay with the general sepsis patients. This difference in ADL dimensions could, at least in part, be explained by some extent of disfiguration and protracted wound healing following major surgery for patients with peritonitis in contrast to patients with sepsis (resulting from other causes). As well, the peritonitis patients often have an enterostomy for a lengthy period of time, which in this study has also been shown to reduce patients’ mobility and daily activities. In contrast, secondary peritonitis patients reported fewer mood problems

Table 3  Impact of potential predictors on EQ-VAS scores.

Results expressed as absolute changes in mean scores derived from multivariate model.

<table>
<thead>
<tr>
<th></th>
<th>Euroqol Visual Analogue Scale (n=127)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean difference in EQ-VAS score†</td>
</tr>
<tr>
<td>Gender (Male vs. females)</td>
<td>4.0</td>
</tr>
<tr>
<td>Age (per 10 years increase)</td>
<td>-2.9</td>
</tr>
<tr>
<td>Patients without major comorbidity at study entry</td>
<td>3.9</td>
</tr>
<tr>
<td>Every doubling of the length of hospital stay</td>
<td>-3.8</td>
</tr>
<tr>
<td>Patients without severe morbidity during 6-months follow-up</td>
<td>6.4</td>
</tr>
<tr>
<td>Patients with no enterostomy at 6-months follow-up</td>
<td>4.9</td>
</tr>
</tbody>
</table>

* Significant p<0.05
† Three patients were dropped due to missing VAS scores
‡ Lower EQ-VAS scores indicate poorer health status
than patients with sepsis from other causes. One possible factor that could account for this difference is the higher mean age of our peritonitis population. In this study and an earlier retrospective study we have shown that older secondary peritonitis patients report fewer mood problems.

Factors associated with lower HR-QoL

In our study, general patient characteristics played an important role in the HR-QoL at 6-months follow-up. Female patients were more likely to report problems with overall HR-QoL, mobility, daily activities, pain and discomfort and mood. Of the nine studies, involving survivors of critical illness and intensive care patients as reviewed by Dowdy et al., associations between HR-QoL and gender were found in only two studies. Peritonitis patients showed a clear association between increased age and improved emotional health, possibly related to an adjusting process. Similar findings were reported in other studies, showing that elderly patients demonstrated more positive health attitudes than younger survivors. However, in a recent review no significant associations were found between age and mental health (SF-36), anxiety/depression (EQ-5D) and/or psychosocial HR-QoL. Comorbidity, often an important determinant of health outcomes, was frequently present in this patient group. Patients with 1 or more major comorbidity reported significantly more mood problems. In these

Table 4  
Odds ratios for reporting moderate/severe problems on each of the dimensions of the EQ-5D.

<table>
<thead>
<tr>
<th>Predictive factors</th>
<th>Mobility (n=128)</th>
<th>Self-care (n=129)</th>
<th>Daily activities (n=129)</th>
<th>Pain/discomfort (n=129)</th>
<th>Mood (n=130)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR P-value</td>
<td>OR P-value</td>
<td>OR P-value</td>
<td>OR P-value</td>
<td>OR P-value</td>
</tr>
<tr>
<td>Gender (Female)</td>
<td>2.9 0.013*</td>
<td>1.5 0.296</td>
<td>3.7 0.006*</td>
<td>2.3 0.030*</td>
<td>1.7 0.176</td>
</tr>
<tr>
<td>Age (per 10 years increase)</td>
<td>1.0 0.246</td>
<td>0.99 0.534</td>
<td>0.98 0.148</td>
<td>0.74 0.078</td>
<td>0.54 &lt; 0.001*</td>
</tr>
<tr>
<td>Patients with major comorbidity at study entry</td>
<td>2.0 0.120</td>
<td>0.92 0.848</td>
<td>1.1 0.782</td>
<td>1.8 0.151</td>
<td>3.6 0.007*</td>
</tr>
<tr>
<td>Every doubling of the length of hospital stay</td>
<td>1.6 0.020*</td>
<td>2.5 &lt;0.001*</td>
<td>1.9 0.010*</td>
<td>1.1 0.537</td>
<td>0.91 0.649</td>
</tr>
<tr>
<td>Patients with severe morbidity during 6-months follow-up</td>
<td>0.71 0.484</td>
<td>0.58 0.294</td>
<td>0.83 0.719</td>
<td>2.4 0.065</td>
<td>2.0 0.130</td>
</tr>
<tr>
<td>Patients with an enterostomy at 6-months follow-up</td>
<td>2.8 0.016*</td>
<td>1.7 0.240</td>
<td>2.8 0.027*</td>
<td>1.2 0.613</td>
<td>1.5 0.320</td>
</tr>
</tbody>
</table>

Patients report moderate and/or severe problems

* Significant p<0.05
analyses we only considered major comorbidities, indicating a pre-existent more severely compromised clinical condition. Although most patients also suffered from an underlying disease (i.e., primary condition) or underwent a primary procedure prior to their secondary peritonitis, these factors were not considered in major comorbidities. Primary conditions are more likely to be the actual cause or part of the etiology of the peritonitis than an actual comorbidity; these included malignancy, diverticulitis, Crohn's disease, ulceritis and colitis ulcerosa. Disease severity measure by the APACHE II has been shown to be an adequate predictor for survival in abdominal sepsis patients\textsuperscript{5,33-35}. Studies relating disease severity with HR-QoL studies have found mixed results; in some papers preoperative severity of disease was a predictor of HR-QoL\textsuperscript{30,33,36-40}, whilst others observed no correlation\textsuperscript{28,29,41-43}. In our study higher APACHE II scores were not associated with poorer HR-QoL. This absent relation could be explained by the homogeneity of the sample with respect to disease severity; only APACHE II scores higher than 10 were included in the study, reflecting severe illness with an expected mortality around 30%\textsuperscript{35}. In this spectrum of severe illness the variability in APACHE II might be insufficient to predict future HR-QoL.

We found no relation between initial peritonitis severity (MPI), extent or type of contamination and the etiology of the peritonitis and HR-QoL at 6 months. This indicates that the HR-QoL outcome of the most severe peritonitis patients may in some cases be far better than anticipated during the initial phase. For example, if a peritonitis patient is admitted to the ICU with a high MPI score and has a diffuse fecal peritonitis then, conditional on survival, their HR-QoL at 6-months follow-up may be similar to those patients that were admitted with less severe peritonitis. This indicates that although these factors are indicators of mortality and morbidity, by themselves they are not associated with poorer HR-QoL at 6 months. As well, HR-QoL differences were not found between patients with community-developed peritonitis and patients with nosocomial peritonitis.

In this study the strongest factor associated with lower HR-QoL was length of hospitalization. This suggests that an extended and more complicated course of disease with longer ICU stay combined with severe morbidity accumulates into worse HR-QoL, most notably in problems with mobility, self-care, and daily activities. ADL problems were primarily related to an extended course of disease encountered during the hospital stay with longer ICU stay – likely related to an accumulation of factors, for example more severe organ dysfunction, such as ARDS, MOF, septic shock and critical illness neuropathy and depended on the patients' response to peritonitis – rather than the underlying etiology and extent of the peritonitis at presentation.

Contrary to expectations, experiencing disease-related morbidity during the 6-months follow-up on its own was not an independent predictor for EQ-5D or EQ-VAS outcomes. This is partially due to the multivariate nature of our analysis, where length of hospital stay probably includes ICU stay in what it measures. Findings in the literature on the relation between length of hospital or length of ICU stay and HR-QoL vary: some studies also found that length of stay was strongly related to HR-QoL\textsuperscript{22,29,30,33,40}, while other studies found no relation\textsuperscript{20-44}. Another particular sequel of the disease is having an enterostomy constructed at surgery for peritonitis, which in these patients is usually still present at 6-months follow-up. As expected, patients with an
enterostomy reported more problems with mobility and daily activities. Reduction of the length of time until restoring continuity in those with a temporary enterostomy, as well as being more liberal with primary anastomosis in some situations (i.e., diverticulitis) may improve long-term HR-QoL.

We assessed HR-QoL using a generic questionnaire, which enabled us to make comparisons to the general population and other diseased populations\textsuperscript{22}. The EQ-5D and EQ-VAS have also been recommended as the choice of generic HR-QoL patient groups, and well validated. Nonetheless, applying a disease-specific questionnaire, including peritonitis-specific symptoms and complaints, may allow for more insight into possible factors that may not be detected by a generic HR-QoL instrument\textsuperscript{24,45}.

It may be a viable option that hospitals consider investing in a tailored support network for patients with more lengthy hospitalization stays, to better prepare both the patient and the home caregivers for the period following discharge characterized by diminished HR-QoL. Younger patients of working age and patients with existing major comorbidity seem to warrant more psychosocial support when discharged from the hospital, which could in turn enable them to return to the workforce more quickly and reduce costs due to loss of productivity. Once the acute life-threatening situation has dissipated and patients are in the surgical ward or have been discharged there may be ample opportunities to consider the patients’ psychosocial network. The results also suggest that this support should be aimed at all peritonitis patients, irrespective of their severity of illness at presentation, since their 6-month HR-QoL is not different from those with a seemingly more favorable presentation, as severity of peritonitis is not an important indicator of later HR-QoL.
References


(11) Granja C, Dias C, Costa-Pereira A, Sarmento A. Quality of life of survivors from severe sepsis and septic shock may be similar to that of others who survive critical illness. *Crit Care.* 2004;8:R91-R98.


(25) EQ-5D, An instrument to describe and value health (www.euroqol.org). This website contains information about the EuroQol Group, membership and research activities, details of EQ-5D development and current status [2007.

Chapter 8
The psychometric performance of a disease-specific health related-quality of life questionnaire for patients with secondary peritonitis: The SP-QoL questionnaire

Kimberly R Boer, Brent C Opmeer, Johannes B Reitsma, Oddeke van Ruler, Cecilia W Mahler, Corianne A de Borgie, Marja A Boermeester, Mirjam A Sprangers

For the Dutch Peritonitis Study Group

Quality of Life Research, in press
Abstract

Objective: To examine the psychometric performance of a newly constructed disease-specific quality of life questionnaire (The SP-QoL) for patients with secondary peritonitis, compiled from the EORTC modules QLQ-CR38, QLQ-OES18 and QLQ-PAN26.

Design: A cohort of secondary peritonitis patients surveyed at 6 and 12 months for HR-QoL.

Setting: Patients were enrolled in 2 academic and 7 regional teaching hospitals in The Netherlands.

Results: From the 153 eligible patients, 130 patients (85%) responded to the questionnaire at 6 months and 112 patients (73%) at both 6- and 12-months follow-up. All subscales, except sexuality, showed good acceptability. The internal consistency of the SP-QoL was very good (Cronbach’s $\alpha = 0.904$). Total inter-item correlations were generally moderate, with mean inter-item correlations being stronger within each subscale than over the entire questionnaire. The sexuality, body image, abdomen, nutrition, stoma-related problems and the defecation subscales correlated poorly to moderately well with the EQ-VAS (ranging from $r = -0.29$ to $-0.48$), whilst the worries subscale correlated remarkably well ($r = -0.67$).

Conclusion: This systematic selection and compilation of existing items into a new instrument (SP-QoL) have shown to have high acceptability, to be internally consistent and to be adequately reliable and valid. Therefore, with minor adjustments, we recommend this questionnaire.

Acknowledgements

Our sincere appreciation to Prof. Dr. Jane M Blazeby, MRC Clinician Scientist and Honorary Consultant Senior Lecturer in Surgery at the department of Clinical Sciences at the South Bristol Royal Infirmary; Mr. Colin Johnson, Reader in Surgery, the University Surgical Unit, Southampton General Hospital, UK; Dr. Galina Velikova, Cancer Research UK Clinician Scientist, Senior Lecturer and Medical Oncology Consultant from the Cancer Research UK Clinical Centre at the St James’s Hospital in Leeds and finally the EORTC Quality of Life Group for allowing the use of the items of the EORTC modules (OES-18, PAN-26, QLQ-CR38) for inclusion in the SP-QoL.

Special thanks to the six surgeons from participating centers for their contribution to identify relevant questions from the three EORTC modules for this disease group.
Introduction

Secondary peritonitis or abdominal sepsis has a high in-hospital mortality (24-35%), with continued high post-hospital discharge mortality, as well as a considerable long-term morbidity. Improving health related-quality of life (HR-QoL) in patients with sepsis has recently received increased attention, recognizing that HR-QoL is an important outcome, which plays a complementary role to both mortality and long-term morbidity outcomes in the evaluation of medical (surgical) interventions.

Peritonitis patients suffer both short-term as well as long-term HR-QoL impairments; hence regain of HR-QoL in peritonitis patients has warranted further study and efforts in the treatment of these patients.

HR-QoL has become a key outcome in assessing critically ill patients, often using generic questionnaires. These generic questionnaires may, however, be insensitive to functional and disease-specific aspects of HR-QoL. Generally, we have seen from the literature that the disease-specific measures, which are based on pertinent disease-related symptoms, can play a key role when evaluating different treatment strategies for a health problem and may be able to assist clinical and management decision-making for individual patients.

There are many disease-specific questionnaires already available, but none yet for secondary peritonitis patients. Having reviewed questionnaires for their applicability for this patient group, the disease-specific QoL modules of the European Organisation for Research and Treatment of Cancer Quality of Life Group (EORTC QL group) seemed the most promising. The EORTC cancer clinical trials organization, of which the QL group is a division, has developed and validated many questionnaires for patients with similar symptoms; hence we selected items from three standardized EORTC modules for pancreatic cancer (QLQ-PAN26), colorectal cancer (QLQ-CR38) and oesophageal cancer (QLQ-OES18) patients that were also pertinent for peritonitis patients.

In this study we report the compilation of a disease-specific QoL questionnaire for secondary peritonitis patients (SP-QoL) based on these three EORTC modules and evaluate the psychometric performance of this questionnaire in a prospective cohort of secondary peritonitis patients. The aim of this study was to develop an instrument able to describe pertinent symptoms and complaints that are relevant in the recovery process of secondary peritonitis and relevant from the patients' perspective for health related-quality of life. The suitability of the instrument was assessed by determining the acceptability of the questionnaire's items; the internal consistency of the instrument as a whole and of the different subscales; the distribution of scores per subscale; the external validity by evaluating the association of scores between the generic HR-QoL questionnaire and each of the SP-QoL subscales. For this we used the Euroqol Visual Analogue Scale (EQ-VAS) and the Euroqol-Five Dimensions (EQ-5D) utility score, as these were recently recommended as the instrument of choice in critical care studies. We also reviewed the mean change over time for the EQ-VAS and EQ-5D utility score from 6- to 12-months follow-up. We assumed that these changes in time would be similar in direction and relative magnitude to the mean changes found over time in the SP-QoL subscales.
Methods

Study design

Our study was embedded in an ongoing peritonitis trial evaluating two surgical strategies for secondary peritonitis, initiated by the Academic Medical Center (AMC), Amsterdam, The Netherlands. Patients were enrolled between December 2001 and February 2004 in 2 academic and 7 regional teaching hospitals in The Netherlands.

Patients

In this multi-center study, patients were eligible for the trial if they were between 18 and 80 years old, had an APACHE II score higher than 10, and were suffering from secondary peritonitis caused by perforation or inflammation of a visceral organ or by ischemia/necrosis of (part of) the gastrointestinal tract or by anastomotic leakage or non-localized abscesses after gastrointestinal surgery needing an emergency laparotomy. Patients had to survive at least 6 months after index laparotomy and had to be discharged from the hospital to be eligible for the present study. The medical ethics committees of all the participating centers approved this study. Consent to participate in the study was obtained either from the patients themselves or from a family member or guardian of the patient.

Questionnaire construction

An initial pool of questions was formed by items from three existing disease-specific modules for colorectal, (QLQ-CR38\textsuperscript{27}), oesophageal (QLQ-OES18\textsuperscript{28}) and pancreatic cancer, (QLQ-PAN26\textsuperscript{25,26}) from the EORTC, as they covered the relevant symptoms and complaints patients are likely to report after secondary peritonitis and abdominal surgery. To obtain optimal content validity, six surgeons independently indicated which items they considered relevant for assessment during follow-up. Items selected by at least four surgeons were included, yielding 30 items. In a consensus meeting, the study group decided to remove 7 items that overlapped and to add 11 items to keep the integrity of the subscale structure of the original items. From the defecation subscale 4 items were added and the subscale ‘stoma-related problems’ was added in its entirety (QLQ-CR38\textsuperscript{27}). The resulting questionnaire (SP-QoL) comprises 7 subscales: abdomen (AB), nutrition (NU), worries (WO), body image (BI), defecation problems (DEF), stoma-related problems (SRP) and sexuality (SEX) (see Table 1 for complete details of the questionnaire compilation from the three original EORTC modules). For this study we also added an open-ended question inviting comments and possible criticisms about the questionnaire. Each item was scaled according to the originally derived scores: 1 (not at all) to 4 (very much). For recovering critically ill patients completion of the questionnaire took about 15 minutes.
Table 1  Construction of the peritonitis questionnaire

Source and original items marked by 4 or more surgeons and any items, which were subsequently excluded or added to the final version.

<table>
<thead>
<tr>
<th>Subscales+</th>
<th>Source</th>
<th>Marked &gt; 4 surgeons (appropriate items from the original EORTC questionnaires)</th>
<th>Excluded overlapping items</th>
<th>Added to keep integrity of the subscales</th>
<th>Item numbers in final SP-QoL</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>Pancreas</td>
<td>1,2,3</td>
<td></td>
<td></td>
<td>1-4</td>
</tr>
<tr>
<td></td>
<td>Oesophagus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NU</td>
<td>Pancreas</td>
<td>7</td>
<td></td>
<td></td>
<td>5-9</td>
</tr>
<tr>
<td></td>
<td>Oesophagus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WO</td>
<td>Pancreas</td>
<td>12,13,22,23</td>
<td>20,22</td>
<td></td>
<td>10-15</td>
</tr>
<tr>
<td></td>
<td>Oesophagus</td>
<td>20,21,22</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Colorectal</td>
<td>9,16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BI</td>
<td>Pancreas</td>
<td>19,20</td>
<td>19,20</td>
<td></td>
<td>16-18</td>
</tr>
<tr>
<td></td>
<td>Colorectal</td>
<td>13,14,15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEX</td>
<td>Pancreas</td>
<td>26,27</td>
<td></td>
<td></td>
<td>19-20</td>
</tr>
<tr>
<td></td>
<td>Colorectal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEF</td>
<td>Pancreas</td>
<td>17,18</td>
<td>17,18</td>
<td></td>
<td>22-28</td>
</tr>
<tr>
<td></td>
<td>Colorectal</td>
<td>24,25,26</td>
<td>27-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRP</td>
<td>Pancreas</td>
<td>.</td>
<td></td>
<td></td>
<td>29-35</td>
</tr>
<tr>
<td></td>
<td>Colorectal</td>
<td></td>
<td></td>
<td></td>
<td>31-37</td>
</tr>
<tr>
<td>Total items</td>
<td></td>
<td>30</td>
<td>7</td>
<td>11</td>
<td>34*</td>
</tr>
</tbody>
</table>

*Item numbers in final SP-QoL includes one additional item (21: patient has a stoma) not included in this overview.
+ Abdomen (AB); Nutrition (NU); Worries (WO); Body image (BI); Sexuality (SEX); Defecation (DEF); Stoma-related problems (SRP)

**Instruments**

In addition to this SP-QoL questionnaire, we administered a generic HR-QoL questionnaire, the Euroqol Visual Analogue Scale (EQ-VAS) and the Euroqol-Five Dimensions (EQ-5D). The EQ-VAS is a thermometer-like scale, in which patients rate their overall well-being from 0 (worst imaginable overall health) to 100 (best imaginable overall health). The EQ-5D measures five health dimensions: mobility, self-care, daily activities, pain/discomfort, and mood consisting of both anxiety and depression. In the EQ-5D patients report: 0 (no problems), 1 (moderate problems), and 2 (extreme problems)\textsuperscript{23-33,34}. In this study we calculated an EQ-5D utility score by using the method described by Dolan\textsuperscript{29} with a correction from Lamers specifically derived for a Dutch population\textsuperscript{30}. These health utility scores ranged from 1 (full health, no problems indicated in the mobility, self care, usual activities, pain/discomfort or the mood...
dimension) to -0.33 (poorest health state scores, reporting extreme problems on the mobility, self care, usual activities, pain/discomfort and mood dimensions)\textsuperscript{30}.

**Data collection**

Preoperative and postoperative data were prospectively collected for all eligible patients. The questionnaires were distributed by mail approximately 6 and 12 months after index operation, with a reminder by phone within two weeks and a new questionnaire with a reminder letter after 1 month.

**Data analysis**

Preoperative and postoperative data were expressed as percentages for categorical data, means (SD) for normally distributed variables and median (interquartile range [IQR]) for non-normally distributed continuous variables. Statistical significance was considered $p < 0.05$.

For the non-respondent analysis, $\chi^2$ tests were used to compare categorical data, and the Student’s $t$ test or the Mann-Whitney U test was used for continuous data.

*The following psychometric properties of the SP-QoL were evaluated:*  
*The acceptability of the items* was determined by examining the frequency in which specific items were left unanswered amongst overall responding patients. The defecation subscale and stoma-related problems subscale were filled in by different subgroups of patients, and therefore these two subscales were analyzed separately.  
*Internal consistency* of the instrument as a whole and of the different subscales was assessed by Cronbach’s alpha (in group comparisons a value of 0.7 is considered the lower bound of internal consistency)\textsuperscript{35,36}. Due to the ordinal nature of the SP-QoL subscales, the Cronbach’s alpha is the lower boundary of the reliability estimate. Further item selection was done by examination of the Cronbach’s alpha where each item was iteratively removed. If the Cronbach’s alpha increased following removal, then the item was deleted from the subscale (items reduced by 0.01 were removed). In addition, the pattern of mean inter-item correlations and item-total correlations per subscale and for the instrument as a whole were examined (0.30 was considered moderate)\textsuperscript{35}.  
*The distributions* of scores per subscale were presented by box-plots to determine the subscales skewness and potential for floor or ceiling effect. All original raw scale scores presented in these box plots were linearly converted to a 1 to 100 scale, with higher scores indicating worse levels of functioning or well-being. Each box shows the median, quartiles, and extreme values within a SP-QoL subscale. For all analyses the SP-QoL subscales were performed using the linearly converted scores.  
*Convergent validity* was evaluated by measuring the association of scores between the EQ-VAS (0 to 100), the EQ-5D utility score and each SP-QoL subscales using Spearman’s rank correlations ($r > 0.4$ is
considered a moderate correlation; \( r > 0.75 \) is considered a high correlation). Our expectation was that the worries, abdomen and nutrition subscales would correlate well with the EQ-VAS, whilst the other subscales would correlate poorly.

Convergent validity was also tested using the SP-QoL worries subscale and the mood (anxiety/depression) dimensions of the EQ-5D. We expected patients who scored higher on the EQ-5D mood (anxiety/depression) dimension to also score higher on the worries dimension of the SP-QoL. Comparing the mobility dimension to the SP-QoL worries subscale tested divergent validity, as we expected patients that scored higher on the mobility dimension to show no effect on the worries subscale. Since the worries subscale is normally distributed we have presented this using means and standard deviations and used an ANOVA for statistical testing.

We also reviewed the mean changes and effect sizes over time for the EQ-VAS and EQ-5D utility score from 6- to 12-months follow-up\(^{13,32}\). We assumed that these changes in time would be similar in direction and relative magnitude to the mean changes found over time in the SP-QoL subscales. Effect size is defined as the mean change from pretest to follow-up found in a variable, divided by the standard deviation of that variable at pretest\(^{17}\). Thereby an effect size of 0.2 is considered small, 0.5 medium, and 0.8 indicates a large effect size (ES)\(^{18}\).

Statistical Package for the Social Sciences (SPSS 11.01, SPSS Inc, Chicago, IL) was used for all data analysis.

**Results**

From the eligible 153 patients (Figure 1), 130 responded to the questionnaire at 6 months, a response rate of 85%. A total of 112 patients responded at both 6 and 12 months (Figure 1). Those patients without a stoma completed the defecation subscale \((n=54)\), whilst those patients with a stoma filled out the stoma-related subscale \((n=73)\), a total response rate of 83%.

The baseline characteristics for patients responding at month 6 are presented in Table 2. The mean age of patients at enrollment was 63 years, and 53% of the patients were male. Patients at trial entry were generally severely ill, as reflected by high scores on both the APACHE II (median of 14.0) and Mannheim Peritonitis Index (MPI, median of 19.5). Fifty-six percent of patients had pre-existing major comorbidities. Following the index laparotomy for peritonitis 88% of patients \((n=115)\) were admitted to the ICU. The median ICU stay for those patients was 9 days \((IQR 6 \text{ to } 22)\) and the median ventilation time of patients admitted to ICU was 6 days \((IQR 3 \text{ to } 12)\). The median hospital stay was 35 days \((IQR 20 \text{ to } 61)\), and 55% of all patients reported still having an enterostomy 6 months following surgery. There were no differences in baseline or postoperative characteristics between patients who responded at 6 months and patients who did not (data not shown).
Patients included at index operation for HR-QoL study (n = 229)

Patients available at month 6 follow-up (n = 159)

Patients who passed away (n = 64)
Readmitted or still in hospital (n = 6)

Address change or no address (n = 6)

Patients sent QoL questionnaires (n = 153)

Six months:
Returned & complete questionnaires (n = 130)

Unknown (n = 3)
Refused (n = 8)
Too ill (n = 7)
Other (n = 5)

Twelve months:
Returned & complete questionnaires (n = 112)

Patients passed away (n = 6)
Unknown (n = 1)
Readmitted (n = 2)
Refused (n = 7)
Too ill (n = 1)
Lost to follow-up (n = 1)

**Acceptability of items**

Questions from the sexuality subscale, questions 19 and 20, were left unanswered by 14% and 15% of the patients at month 6 and month 12, respectively. In the open question at the end of the SP-QoL questionnaire, patients added comments concerning the sexuality questions indicating that many of these patients were widowed and felt that these questions were not applicable to them. All other subscales showed greater than 95% response on all items.
Internal consistency

The internal consistency of the abdomen, nutrition, worries, self-image and sexuality subscales showed Cronbach’s $\alpha$ ranging from 0.78 for worries to 0.90 for the sexuality subscale (see Table 3A for details). The removal method did not reveal any particular items that decreased the overall per subset Cronbach’s $\alpha$ for any of the subscales (Table 3A). The mean inter-item correlation ranged from 0.37 on the worries scale to 0.83 on the sexuality scale. Total inter-item correlations were generally moderate. Only the defecation subscale showed a low internal consistency (Cronbach’s $\alpha=0.60$) and a low mean inter-item correlation (mean inter-item correlation=0.17); questions 26 and 28 showed item correlations below 0.30 (Table 3B).

---

**Table 2** Patient and post-operative characteristics of study population ($n = 130$)

<table>
<thead>
<tr>
<th>Patient characteristics</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age; mean (SD)</td>
<td>62.8 (13.9)</td>
</tr>
<tr>
<td>Males; n</td>
<td>70</td>
</tr>
<tr>
<td>APACHE II median (IQR)</td>
<td>14 (12 to 17)</td>
</tr>
<tr>
<td>MPI median (IQR)</td>
<td>19.5 (15 to 25)</td>
</tr>
<tr>
<td>≥ 1 Major comorbidity; n</td>
<td>73</td>
</tr>
<tr>
<td>Postoperative peritonitis (after elective surgery); n</td>
<td>69</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Postoperative characteristics</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relaparotomies; n</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>44</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
</tr>
<tr>
<td>2 - 3</td>
<td>24</td>
</tr>
<tr>
<td>≥ 4</td>
<td>13</td>
</tr>
<tr>
<td>Patients admitted to ICU; n</td>
<td>115</td>
</tr>
<tr>
<td>Length of ICU stay; median (IQR)</td>
<td>9.0 days (6-22)</td>
</tr>
<tr>
<td>Patients ventilated; n</td>
<td>110</td>
</tr>
<tr>
<td>Duration of ventilation; median (IQR)</td>
<td>5.0 days (3-12)</td>
</tr>
<tr>
<td>Length of hospital stay; median (IQR)</td>
<td>35 days (20-61)</td>
</tr>
<tr>
<td>Enterostomy at 6 months at 6-months follow-up; n</td>
<td>73</td>
</tr>
</tbody>
</table>
Cronbach's α, inter-item correlations and corrected item-total correlations per item to subscale and overall questionnaire for the subscales abdomen, nutrition, worries, body image, and sexuality.

<table>
<thead>
<tr>
<th></th>
<th>Cronbach's α</th>
<th>Mean Inter-item correlation</th>
<th>Corrected Item-total correlations for subscales</th>
<th>Corrected Item-total correlations for SP-QoL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total SP-QoL* (n=121):</td>
<td>0.904+</td>
<td>0.360</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdomen (n = 130)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Reduced appetite</td>
<td>0.811</td>
<td>0.518</td>
<td>0.604</td>
<td>0.622</td>
</tr>
<tr>
<td>2 Unpleasant feeling in the abdomen</td>
<td></td>
<td></td>
<td>0.592</td>
<td>0.590</td>
</tr>
<tr>
<td>3 Bloated feeling in the abdomen</td>
<td></td>
<td></td>
<td>0.671</td>
<td>0.626</td>
</tr>
<tr>
<td>4 Felt full too quickly</td>
<td></td>
<td></td>
<td>0.649</td>
<td>0.603</td>
</tr>
<tr>
<td>Nutrition (n = 129)</td>
<td>0.863</td>
<td>0.564</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Eat solid foods</td>
<td></td>
<td></td>
<td>0.765</td>
<td>0.690</td>
</tr>
<tr>
<td>6 Drink liquids</td>
<td></td>
<td></td>
<td>0.497</td>
<td>0.503</td>
</tr>
<tr>
<td>7 Limited in the type of food</td>
<td></td>
<td></td>
<td>0.749</td>
<td>0.578</td>
</tr>
<tr>
<td>8 Limited in the amount of food</td>
<td></td>
<td></td>
<td>0.767</td>
<td>0.675</td>
</tr>
<tr>
<td>9 Food and drink taste different</td>
<td></td>
<td></td>
<td>0.708</td>
<td>0.587</td>
</tr>
<tr>
<td>Worries: (n = 127)</td>
<td>0.776</td>
<td>0.367</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Lost weight</td>
<td></td>
<td></td>
<td>0.391</td>
<td>0.402</td>
</tr>
<tr>
<td>11 Concerned that weight is too low</td>
<td></td>
<td></td>
<td>0.392</td>
<td>0.472</td>
</tr>
<tr>
<td>12 Weakness in arms and legs</td>
<td></td>
<td></td>
<td>0.569</td>
<td>0.601</td>
</tr>
<tr>
<td>13 Demanding sickness</td>
<td></td>
<td></td>
<td>0.625</td>
<td>0.580</td>
</tr>
<tr>
<td>14 Future situation</td>
<td></td>
<td></td>
<td>0.570</td>
<td>0.621</td>
</tr>
<tr>
<td>15 Limited in planning activities</td>
<td></td>
<td></td>
<td>0.602</td>
<td>0.590</td>
</tr>
<tr>
<td>Body Image (n = 126)</td>
<td>0.834</td>
<td>0.626</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 Feeling less attractive</td>
<td></td>
<td></td>
<td>0.721</td>
<td>0.485</td>
</tr>
<tr>
<td>17 Feeling less masculine/ feminine</td>
<td></td>
<td></td>
<td>0.730</td>
<td>0.513</td>
</tr>
<tr>
<td>18 Dissatisfied with body</td>
<td></td>
<td></td>
<td>0.635</td>
<td>0.497</td>
</tr>
<tr>
<td>Sexuality (n = 101)</td>
<td>0.904</td>
<td>0.826</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 Interest in sex</td>
<td></td>
<td></td>
<td>0.826</td>
<td></td>
</tr>
<tr>
<td>20 Sexual enjoyment</td>
<td></td>
<td></td>
<td>0.826</td>
<td></td>
</tr>
</tbody>
</table>

*Response range is from 1-4 for all questions, all questions refer to the last week, all questions were asked originally in Dutch and have been translated for this paper (using the official EORTC translations), the Dutch version is presented in Appendix 2.

+In this analysis we removed questions 19 and 20 due to poor response (when included the Cronbach’s α was 0.884 for n = 98).
### Table 3b  Psychometric properties of the defecation subscale divided into two versions based on the presence or absence of a stoma (Question 21)

Cronbach’s a and inter-item correlations, corrected item-total correlations per item to the subscales defecation, and stoma-related problems.

<table>
<thead>
<tr>
<th>Question</th>
<th>Cronbach’s a</th>
<th>Corrected Item-total correlations to subscale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Defecation (patients without an enterostomy, n = 54)</strong>:</td>
<td>0.598</td>
<td>0.171+</td>
</tr>
<tr>
<td>22 Frequency of bowel movement/ day</td>
<td></td>
<td>0.446</td>
</tr>
<tr>
<td>23 Frequency of bowel movement/ night</td>
<td></td>
<td>0.470</td>
</tr>
<tr>
<td>24 Urge without producing stool</td>
<td></td>
<td>0.333</td>
</tr>
<tr>
<td>25 Unintentional release of stools</td>
<td></td>
<td>0.322</td>
</tr>
<tr>
<td>26 Blood with stools</td>
<td></td>
<td>0.032</td>
</tr>
<tr>
<td>27 Difficulty in moving bowels</td>
<td></td>
<td>0.343</td>
</tr>
<tr>
<td>28 Painful bowel movements</td>
<td></td>
<td>0.279</td>
</tr>
<tr>
<td><strong>Stoma-related problems (patients with a enterostomy, n= 73)</strong>:</td>
<td>0.860</td>
<td>0.468+</td>
</tr>
<tr>
<td>29 Afraid about the stoma noise</td>
<td></td>
<td>0.601</td>
</tr>
<tr>
<td>30 Afraid about the smell of stools</td>
<td></td>
<td>0.607</td>
</tr>
<tr>
<td>31 Worry about possible leakage</td>
<td></td>
<td>0.638</td>
</tr>
<tr>
<td>32 Caring for stoma</td>
<td></td>
<td>0.685</td>
</tr>
<tr>
<td>33 Irritated skin</td>
<td></td>
<td>0.523</td>
</tr>
<tr>
<td>34 Embarrassment</td>
<td></td>
<td>0.699</td>
</tr>
<tr>
<td>35 Feeling less complete</td>
<td></td>
<td>0.647</td>
</tr>
</tbody>
</table>

*Response range is from 1-4 for all questions, all questions refer to the last week, all questions were asked originally in Dutch and have been translated for this paper (using the official EORTC translations), the Dutch version is presented in Appendix 2.  
+Mean inter-item correlation for the subscale.

### Distribution of scores per subscale

Box-plots of the linearly transformed scores on each subscale are given in Figure 2. The worries and body image subscales scores are well distributed, showing use of the entire subscale. Abdomen and nutrition subscales showed floor effects, with only a few outliers in the higher ranges of the subscales. The defecation subscale showed a wide range of scores, whilst stoma patients only responded to the lower half the stoma-related subscale, again indicating a floor effect (Figure 2).
Validity

Correlations between the EQ-VAS and EQ5D utilities score and each subscale at month 6 are shown in Table 4. The SP-QoL subscales, sexuality, body image, stoma-related problems and the defecation subscales correlated moderately with the EQ-VAS (ranging from r=-0.29 to -0.39) and with the EQ-5D utility score (ranging from r=-0.27 to -0.37). The nutrition and abdomen subscales correlated well with the EQ-VAS (r=-0.41 to -0.48, respectively) and the EQ-5D utility score (r=-0.43 to -0.55, retrospectively). The worries subscale correlated very well with both the EQ-VAS (r=-0.67) and the EQ-5D utility score (r=-0.62) (details presented in Table 4). These scores are negative because the EQ-VAS and EQ-5D utility scores range from low to high, with high scores indicating better HR-QoL, whilst the SP-QoL subscales scores range from high to low, with low score indicating better HR-QoL. One hundred and twenty-nine patients filled in both the EQ-VAS and EQ-5D dimensions for which the EQ-5D utility score was calculated; these scores also correlated quite well (0.70, p<0.001).
Change over time

One hundred and twelve patients responded at both 6 and 12 month of follow-up (73% of the eligible patients). EQ-VAS and EQ-5D utility scores improved slightly between 6 and 12 months with a mean improvement for the EQ-VAS of 3.7 (1.3 to 6.0, n=110, Table 5) and a small effect size of 0.21, and a mean improvement of 0.041 (0.1 to 0.08, n=112, Table 5) for the EQ-5D utility score. In the SP-QoL, improvement over time was not found in most subscales, presented in Table 5; only the nutrition, worries and the sexuality subscales showed statistically significant improvement over time; mean nutrition improvement was minimal at -3.4 (-6.4 to 0.44, n=112); with an effect size of 0.16. The mean improvement for the worries subscale was quite large at -11.6 (-15.4 to –7.7, n=108), with a medium effect size of 0.51. The mean sexuality subscale improvement was -8.1 (-15.5 to 0.8, n=75), with a small effect size of 0.20. All the other SP-QoL subscales were low (ranging from 0.08 tot 0.20, Table 5).

Table 4 Correlation in scores between the EQ-VAS, EQ-5D utility score and each subscale of the SP-QoL at 6-months follow-up

<table>
<thead>
<tr>
<th>SP-QoL subscales</th>
<th>AB</th>
<th>NU</th>
<th>WO</th>
<th>BI</th>
<th>SEX</th>
<th>DEF</th>
<th>SRP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>127</td>
<td>126</td>
<td>125</td>
<td>125</td>
<td>100</td>
<td>53</td>
<td>70</td>
</tr>
<tr>
<td>EQ-VAS correlation coefficient*</td>
<td>-0.48</td>
<td>-0.41</td>
<td>-0.67</td>
<td>-0.37</td>
<td>-0.29</td>
<td>-0.39</td>
<td>-0.35</td>
</tr>
<tr>
<td>P-value</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.003</td>
<td>0.003</td>
<td>0.002</td>
</tr>
</tbody>
</table>

| Number of patients | 129 | 128 | 126 | 125 | 100 | 53  | 72  |
| EQ-5D correlation coefficient* | -0.55 | -0.43 | -0.62 | -0.37 | -0.27 | -0.324 | -0.29 |
| P-value | < 0.001 | < 0.001 | < 0.001 | < 0.001 | 0.007 | 0.018 | 0.013 |

*Spearman’s correlation coefficient these correlations are negative because the EQ-VAS ranges from 0 to 100 and the EQ-5D utility score ranges from -0.33 to 1 with high scores indicating better HR-QoL, whilst the SP-QoL has been converted to 1 to 100 scale with higher scores indicating poorer HR-QoL. The strengths of the associations are not altered by the direction.

+Abdomen (AB); Nutrition (NU); Worries (WO); Body image (BI); Sexuality (SEX); Defecation (DEF); Stoma-related problems (SRP).
This questionnaire is based on items systematically selected from three standardized EORTC modules for oesophageal cancer (QLQ-OES24\textsuperscript{24}), pancreatic cancer (QLQ-PAN26\textsuperscript{25,26}) and colorectal cancer (QLQ-CR38\textsuperscript{27,29}). The rationale of using questionnaires designed for cancer patients to assess our patient group was three-fold. Firstly, peritonitis is caused by infectious disease at various locations of the digestive tract and once (partly) recovered these patients suffer from (fairly) similar gastrointestinal and other symptoms and complaints as patients with cancer of the gastrointestinal tract. Secondly, although these patients suffer acutely, cancer patients also suffer an immediate threat to life when being confronted with a cancer diagnosis and related surgery; their recovery process or persistence of complaints may be very similar. And finally, there is precedent for using an EORTC cancer questionnaire for patients with benign disease; the PAN26\textsuperscript{25,26} has been also been validated for use in chronic pancreatitis patients.

The internal consistencies of the items of six of the seven subscales were high, with moderate to high mean inter-item correlations and high corrected item total correlations\textsuperscript{35,36}. These item-total correlations are often interpreted as providing an estimate of the unidimensionality of the scale; the conventions are that values lower than 0.30 suggest non-unidimensionality. Only the defecation subscale showed an unacceptable internal consistency and a low mean inter-item correlation and corrected item-total correlation. This indicates that reporting problems on one question of the defecation subscale does not automatically increase the probability that a patient would report problems on the remaining questions of the subscale. In particular, question 26, which determines whether blood is present in a patient’s stool, could be considered for removal. However, problems reported on a single question (such as 28) could still indicate relevant disease-related problems that need to be recognized and addressed by the clinician and indicate reduced HR-QoL. Another solution for this would be to consider the defecation subscale as an inventory of symptoms, but not as a subscale, because the cohesion of the questions is low. The EORTC QL has used individual items more often and can be found in many of the disease-specific HR-QoL questionnaires (e.g., EORTC QL-OES18).

The abdomen, nutrition and sexuality subscales showed a considerable floor effect, meaning that patients only used the lower end of the subscale. This indicates that these subscales may not be sensitive in detecting differences between patients with good and fair HR-QoL. It could, however, also mean that there were few problems reported in this area by our study group. The worries, body image, defecation and stoma-related problems subscales showed good use of the entire subscale.

Abdomen, nutrition and worries subscales showed moderate correlation with the EQ-VAS and EQ-5D, whilst the dimensions we expected to give more disease-specific information (body image, sexuality, defecation and stoma-related problems) showed relatively low correlations. Owing to the nature of generic and disease-specific HR-QoL questionnaires, we suspected that there would be considerable, but not perfect, association between the two questionnaires. We expected the disease-specific questionnaire to be more sensitive to the particulars of this patient group, hence its use. We expected that parallel to the perceived clinical recovery between 6- and 12-months follow-up, there would also
be significant improvement in HR-QoL. This was not actually the case; the EQ-VAS and EQ-5D utility score found little improvement over time, with surprising low effect size. Patients have a long recovery time and once discharged from the hospital, signs of further improvement appear to be minimal; it is possible that we may have overestimated both the clinical and HR-QoL improvement patients would undergo from 6 to 12 months, since all but one of the SP-QoL subscales showed a small effect size. Only the worries subscale showed a moderate \(^{38}\) median effect size from 6- to 12-months follow-up, indicating a potentially clinically relevant improvement over time\(^{38,31,32}\). Showing that with the exception of the worries sub-scale the EQ-VAS was more sensitive than the SP-QoL and the EQ-5D utility more sensitive than the SP-QoL, with exception of the worries and sexuality subscales. This may be due to the fact that the response in the worries subscale was more adequately dispersed over the entire subscale (difference within patients), whilst the abdomen and nutrition subscales showed a considerable floor effect. In the subscales body image and sexuality we did not actually expect an improvement over time; we would not have been surprised if patients had even reported more problems at 12 months than 6 months time.

This illustrates that, the added value of the disease-specific questionnaire can depend on the application of the questionnaire and the variability and heterogeneity in the patient group. For example when comparing this patient group to other severely ill patients the EQ-5D may be the preferred

### Table 5  Mean differences EQ-VAS, EQ-5D utility score and the SP-QoL over time, 6 to 12 months following surgery

<table>
<thead>
<tr>
<th>Mean difference from 6 to 12 months (95 % CI)*</th>
<th>Effect sizes**</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Euroqol</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQ-VAS</td>
<td>3.7 (1.3 to 6.0)</td>
<td>0.21</td>
</tr>
<tr>
<td>EQ-5D utility score</td>
<td>0.041 (0.01 to 0.08)</td>
<td>0.19</td>
</tr>
<tr>
<td><strong>SP-QoL</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdomen (AB)</td>
<td>-2.9 (-7.1 to 1.4)</td>
<td>0.13</td>
</tr>
<tr>
<td>Nutrition (NU)</td>
<td>-3.4 (-6.4 to 0.44)</td>
<td>0.16</td>
</tr>
<tr>
<td>Worries (WO)</td>
<td>-11.6 (-15.4 to -7.7)</td>
<td>0.51</td>
</tr>
<tr>
<td>Body image (BI)</td>
<td>-2.3 (-4.4 to 1.4)</td>
<td>0.08</td>
</tr>
<tr>
<td>Sexuality (SEX)</td>
<td>-8.1 (-15.5 to 0.8)</td>
<td>0.20</td>
</tr>
<tr>
<td>Defecation (DEF)</td>
<td>-1.5 (-5.0 to 1.4)</td>
<td>0.12</td>
</tr>
<tr>
<td>Stoma-related problems (SRP)</td>
<td>-2.2 (-4.1 to 8.4)</td>
<td>0.08</td>
</tr>
</tbody>
</table>

* Results only from patients that responded at both 6- and 12-months follow-up, linear transformation, 0 -100.

** Mean differences from 6- to 12-months follow-up divided by the standard deviation at 6-months follow-up.
instrument because of its parsimony and generalizability. However, when comparing different treatments the SP-QoL may show added value in information pertaining to symptoms and complaint that could potentially differ between the two groups. Because we have validated this questionnaire in a group of only severe peritonitis patients, we recommend that this questionnaire be further evaluated in all peritonitis patients not only the most severe group.

Understanding the use of HR-QoL data for clinical decision-making and differentiating between the use of generic instruments and disease-specific instruments is a topic that has been touched on frequently within the last 10 years\textsuperscript{10,18-20}. These generic and disease-specific tools are meant to complement each other and in combination to support clinical decision-making\textsuperscript{39,40}, whilst also giving clinicians a better understanding of the burden of disease from the patient's viewpoint\textsuperscript{41-43}. The generic questionnaire has obvious limitations, which the disease-specific questionnaire can supplement. However the generic questionnaire remains the preferred choice when comparing patient groups with different diseases. When this instrument is used in the context of a clinical trial, the questionnaire should allow a more adequate comparison between the different treatment arms and could illustrate different recovery rates. This is the first disease-specific questionnaire for patients with peritonitis, and has shown to be robust. This questionnaire may assist in better understanding of these patients and better ability to choose between treatment options for these patients; as so far there is limited knowledge about peritonitis specific HR-QoL.

Limitations

Unfortunately, due to the acuteness of the disease it is not possible to perform a full validation study for this questionnaire\textsuperscript{44}, as three pertinent pieces of information are missing\textsuperscript{45,46}. Firstly, it is not possible to collect baseline data, as these patients are already severely ill at hospital presentation. Secondly, test-retest information could not be obtained, as the population was too ill to burden with filling in the questionnaires more often. Finally, in addition to asking the surgeons, we could also have asked patients to indicate which questions should be included in the questionnaire or whether other adjustments should be made. However, we attached an open question for patients to indicate what they thought of the questionnaire as a whole; we encouraged patients to complete this open question and asked the patients if they felt the questionnaire was complete in its present form\textsuperscript{47}. This opportunity for comments, particularly on the sexuality questions was used by many of the patients. From these patient comments, the low response and the high correlation between the two sexuality questions, we would recommend the removal of question 20, patient enjoyment in sexual relations. We do not feel this justifies the removal of both questions, as sexuality, although not applicable to all patients (hence the low response), can be a very important part of HR-QoL.
Conclusion

The SP-QoL is a disease-specific questionnaire for secondary peritonitis patients, based on standardized modules from the EORTC QL group. The EORTC instruments have been demonstrated to be valuable tools in understanding patient HR-QoL and recovery. This systematic selection and compilation of existing items into a new instrument (SP-QoL) has shown to have a high acceptability, be internally consistent and be adequately reliable and valid. We recommend this questionnaire for the evaluation of HR-QoL outcomes in patients with secondary peritonitis to complement standard generic questionnaires.
References


(6) Granja C, Dias C, Costa-Pereira A, Sarmento A. Quality of life of survivors from severe sepsis and septic shock may be similar to that of others who survive critical illness. Crit Care. 2004;8:R91-R98.


(18) Dowie J. Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions. Health Econ. 2002;11:1-8.

(19) Feeny D. Commentary on Jack Dowie, “Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions”. Health Econ. 2002;11:13-16.

(20) Guyatt G. Commentary on Jack Dowie, “Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions”. Health Econ. 2002;11:9-12.


“The greatest mistake in the treatment of diseases is that there are physicians for the body and physicians for the soul, although the two cannot be separated.”

Plato
Part IV
Post-traumatic stress disorder in patients with secondary peritonitis
Chapter 9
Long-term prevalence of post-traumatic stress disorder symptoms in patients after secondary peritonitis

Kimberly R Boer, Cecilia W Mahler, Cagdas Unlu, Bas Lamme, Margreeth B Vroom, Mirjam A Sprangers, Dirk J Gouma, Johannes B Reitsma, Corianne A de Borgie, Marja A Boermeester

Abstract

**Introduction:** The aim of this study was to determine the long-term prevalence of post-traumatic stress disorder (PTSD) symptomology in patients following secondary peritonitis and to determine whether the prevalence of PTSD-related symptoms differed between patients admitted to the intensive care unit (ICU) and patients admitted only to the surgical ward.

**Method:** A retrospective cohort of consecutive patients treated for secondary peritonitis was sent a postal survey containing a self-report questionnaire: the Post-traumatic Stress Syndrome-10 Questions Inventory (PTSS-10). From a database of 278 patients operated for secondary peritonitis between 1994 and 2000, 131 patients were long-term survivors having a follow-up period of at least four years and eligible for our study in a tertiary referral hospital in Amsterdam, The Netherlands.

**Results:** The response rate was 86% leading to a cohort of 100 patients of whom 61% had been admitted to ICU. PTSD-related symptoms were found in 24% (95% CI 17% to 33%) of patients with a cut-off of 35, whilst the prevalence of PTSD symptomology including borderline patients who scored above 27 points was 38% (95% CI 29 to 48%). In a multivariate analyses controlling for age, gender and APACHE II, number of relaparotomies and length of hospital stay, the odds for ICU patients having PTSD symptomology was 4.3 higher (95% CI 1.11 to 16.5) using the PTSS-10 questionnaire cut-off above 35. Older patients and males were less likely to report PTSD symptoms.

**Conclusion:** Nearly a quarter of patients receiving surgical treatment for secondary peritonitis develop PTSD symptoms. Patients admitted to the ICU had a significantly higher risk of PTSD symptoms after adjusting for baseline differences, in particular age.
Introduction

Peritonitis or abdominal sepsis is a severe disease with a high mortality rate (approximately 30%)\(^1,2\). Intensive care unit (ICU) and hospital admissions may be lengthy and morbidity extensive. Hence, experiencing peritonitis is a major life event. Patients who survive critical illness often report poor quality of life (QoL) and the presence of post-traumatic stress disorder (PTSD) symptomology in the post-clinical period\(^3,4\). PTSD symptoms include intrusive recollections; avoidant/numbing symptoms; and hyperarousal symptoms from exposure to one or more traumatic events\(^5\). Patients with PTSD (symptoms) have reduced quality of life\(^6,8,10-12\) and frequently suffer from depression\(^8,13\). Therefore monitoring PTSD symptomology in ICU patients could complement hospital and long-term survival outcomes, guide early socio-psychological interventions and improve long-term patient care. Hence, PTSD is worthy of evaluation to help understand the complex nature of long-term outcomes\(^14\).

Many survivors of critical illness and its treatment, suffer from continuous traumatic memories and re-live adverse experiences of their illness, such as respiratory distress, anxiety, pain, or loss of control, which are all associated with an increased risk of development of PTSD\(^3,6\). Several studies have reported prevalence of 15% to 38% of PTSD related symptoms in patients who had been admitted to ICU\(^4,8\). Some authors have argued that specific circumstances and memories during ICU stay can serve as a trigger for developing PTSD symptoms rather than having a severe underlying illness itself. However, the majority of studies examining the relation between ICU stay and PTSD symptoms used cohorts in which all patients have been admitted to ICU, meaning that these studies cannot differentiate between these two possible patient experiences.

In addition, data on the prevalence of PTSD-related symptoms after secondary peritonitis are lacking. It is unknown whether the prevalence of symptoms related to PTSD or memories of traumatic experiences differ between peritonitis patients after ICU admission (operation, ICU stay and hospital ward stay) and patients without ICU admission (operation and hospital ward stay only).

The aim of this study was first to determine the long-term prevalence of PTSD symptomology in patients 4 to 10 years after secondary peritonitis based on a self-report questionnaire. We also aimed to compare the prevalence of PTSD-related symptoms between patients admitted to the ICU and patients admitted only to the surgical ward. Finally, we examined whether the prevalence of PTSD symptomology in these patients was increased due to the traumatic memories that patients had during their ICU and/ or hospital stay\(^1\).

Patients and methods

Study population

A retrospective cohort of 278 consecutive patients, who were treated surgically for secondary peritonitis between January 1994 and January 2000, was the starting point of this study\(^1\). All patients were treated at the Department of Surgery in the Academic Medical Center at the University of
Amsterdam, The Netherlands. All patients still alive were eligible for inclusion. These patients were informed about this study by telephone to improve the response rate. Due to the non-interventional nature of the study, the Institutional Review Board waived the need for informed consent.

**Data collection**

All patients still alive were eligible for the study (n=118) and received a standardized instrument for assessing symptoms related to PTSD, the Post-traumatic Stress Scale–10 (PTSS-10). In addition, they received a 4-question Adverse Experiences Questionnaire. Each questionnaire referred to the patient’s feelings over the last 14 days.

Patients who had been admitted to the ICU during their hospital stay for peritonitis were sent a questionnaire that specifically asked the patient to consider their feelings during the last 14 days whilst keeping their past ICU stay in mind. Patients not admitted to ICU were asked to fill in the questionnaire for the last 14 days keeping in mind their past stay in the general ward following their peritonitis.

A separate questionnaire was included to collect relevant clinical data since discharge from the hospital for peritonitis (including readmissions since discharge after surgical treatment for secondary peritonitis and use of medication in the last few years, and newly developed diseases and their treatment).

Patients who returned incomplete questionnaires were contacted by phone within two weeks in an attempt to complete the questionnaire by phone. Patients who did not return the questionnaires were sent the questionnaires two more times within a 6-week period. After these attempts patients who had given initial telephone consent were contacted again by phone to obtain information regarding their motivations for not responding.

Demographic and clinical data at the time of the index operation (the emergency laparotomy performed at initial presentation of peritonitis) were collected from hospital charts and a computerized registration system. The following information was collected: patient characteristics, which included age, gender, comorbidity, use of medication, Acute Physiology and Chronic Health Evaluation II score (APACHE II) before operation and Mannheim Peritonitis Index (MPI). Disease and operation characteristics contained etiology of peritonitis, origin of peritonitis, surgical treatment strategy and number of relaparotomies. Postoperative characteristics consisted of the number of days in hospital, days in ICU, days of mechanical ventilation, ‘open abdomen’ (laparostomy) during admission, number and type of complications, number of readmissions and the mean follow-up time. Patient recall was checked using the hospital information and medication system to check readmission and use of medication. Out of hospital medication such as prescribed by the family physician was obtained only by questionnaire.

**Instruments**

Post-traumatic Stress Syndrome-10 Questions Inventory (PTSS-10): The PTSS-10 was originally designed to diagnose PTSD according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) II criteria in victims of natural disasters and subsequently validated in Norwegian seamen after...
torture in Libya. The PTSS-10 has since been validated in acute respiratory distress syndrome (ARDS) patients after ICU treatment using the Structured Clinical Interview for DSM-IV (SCID) II, an interview technique according to DSM-IV. The PTSS-10 is now a widely used validated self-report questionnaire assessing symptoms related to PTSD with a reported sensitivity of 77% and a specificity of 97.5% for the diagnosis of PTSD.

The questionnaire consists of 10 items, each with a Likert scale ranging from 1 (‘never’) to 7 (‘always’). A sum score with a range between 10 and 70 is calculated, with higher scores indicating more PTSD-related symptoms. A sum-score of 35 or higher is considered an adequate cut-off for PTSD related symptomology, whilst patients with scores between 27 and 35 on the PTSS-10 were considered borderline PTSD symptomology. The validated English version was translated into Dutch according to a forward-backward translation procedure.

Adverse events/ traumatic experiences questionnaire: The 4-item adverse experiences questionnaire assesses the presence of four types of traumatic memories during their stay in ICU or hospital ward: anxiety, respiratory distress, pain and/or nightmares. Patients scored the frequency in which they experienced these traumatic events (of memories as they remembered them) during their stay in the ICU or hospital ward using a 4-point response scale of 1 (none), 2 (sometimes), 3 (regularly) or 4 (often).

**Statistical analysis**

Ninety-five percent confidence intervals around estimates of prevalence were calculated using the method of Wilson. Clinical characteristics and the prevalence of PTSD symptoms between patients that were admitted to ICU during their initial stay and patients treated solely at the surgical ward were compared. Depending on the nature of the clinical variables, we either used Pearson $\chi^2$, Student t test or Mann-Whitney U test, where appropriate.

We built multivariate logistic regression models to assess the association between ICU stay and the presence of PTSD symptomology (PTSS-10 sum score above 35) after adjusting for other factors. We adjusted for factors related to patients characteristics: age (continuous) and gender; disease characteristics: APACHE II score at baseline (continuous) and patients with ≥1 relaparotomy (yes/no dichotomous); postoperative characteristics: days in hospital (to the base 10 transformed to improve linear relationship with the outcome). These factors were chosen either because they were identified in earlier PTSD studies and literature, i.e. age, gender and comorbidity, or they were chosen due to univariate significance ($p<0.1$) with the dependent factor (PTSD symptomology) in our study; i.e. APACHE II score, patients with greater than one relaparotomy, days in hospital. If factors were highly correlated, we only selected one of the correlated factors in the multivariate model to avoid the problem of co-linearity. Odds ratios with 95% confidence intervals were used to quantify the strength of the association. To determine the fit of the final multivariate logistic model, we calculated the area under the ROC curve (AUC), also known as the concordance statistic, and performed the Hosmer-Lemeshow goodness-of-fit test.
To determine whether traumatic memories during the patient’s stay in hospital or ICU played a role in the development of PTSD symptomology in our peritonitis population, we examined the percentage of patients with PTSD symptomology within each level of response on the traumatic memories questions. Because of the ordered response on the traumatic memories question, we used the $\chi^2$ test for trend to examine this relation. P-values of less than 0.05 were considered statistically significant.

**Figure 1 Flowchart of study inclusion**

![Flowchart](image)

**Results**

From the initial cohort of 278 patients with secondary peritonitis, 118 patients were long-term survivors. These patients received the set of questionnaires, and 104 patients (88%) responded (Figure 1). Of the 14 patients who did not respond to the questionnaire, 5 patients were not willing to complete the questionnaire and 9 patients, who were initially informed about the study by phone before the mailing, could not be contacted again to find out their reason for non-response to the questionnaire.
Four patients were excluded because too many data were missing (Figure 1). No significant differences in operation, hospital or postoperative characteristics were found between the patients suitable for analysis (n=100) and the eligible patients still alive that did not respond (n=32). Although patient characteristics in the non-responding group indicated that these patients were younger (mean 51 vs. 40 years old, p<0.001), presented with fewer comorbidities (comorbidity was present in 65% vs. 30%, p<0.001) at initial operation and had both lower APACHE II scores (9.5 vs. 7.5, p=0.049) and MPI scores (22 vs. 18.6, p=0.024) than the responders. There was no difference between responding patients and non-responders for ICU admittance.

For responding patients the average time since index operation and follow-up was 5.3 years for ICU and non-ICU patient groups. Comorbidity was present in 65% of patients, and nearly 80% of patients were on some type of medication (Table 1). The mean APACHE II score at time of the index operation was 9.5 (SD±5) score and the mean MPI score was 21.9 (±SD-7). Seventy-six percent of patients were treated according to an on-demand (OD) relaparotomy strategy and 24% according to a planned relaparotomy (PR) strategy, and overall, 59% of patients underwent one or more relaparotomies.

**PTSD symptomology**

The median PTSS-10 score among all patients was 22, with 25% of the patients having a score below 13 and 25% of patients with a score above 33. Using the recommended cut-off value for PTSD symptomology of 35 points on the PTSS-10 questionnaire17,19,22,23, the overall prevalence of PTSD-related symptoms was 24% (95% CI 17% to 33%). The overall prevalence of PTSD symptomology including borderline patients who scored above 27 points was 38% (95% CI 29 to 48%).

**Comparison between ICU and non-ICU patients**

Patient, disease and operation characteristics for ICU patients (61%) and non-ICU (39%) are presented in Table 1. Patients who had stayed in ICU were on average 7.5 years older than patients not admitted to ICU (p=0.011). ICU patients also had higher APACHE II (mean difference 2.2 points, p=0.037) and MPI (mean difference 3.2 points, p=0.036). Patients with an ICU stay had received a laparostomy (open abdomen) in 36% of cases, whereas only 8% of the ward patients had been treated with a laparostomy (i.e., in 92% of patients admitted to the surgical ward primary abdominal closure was done; p=0.001). A relaparotomy was significantly more common in the ICU group than in the non-ICU group (72% vs. 40%, p<0.001).

With respect to postoperative characteristics, patients had a median stay in hospital of 37 days. ICU survivors had a longer hospital stay compared to non-ICU survivors (49 versus 27 median days p=0.001) and had more non-surgical complications (57% versus 8%, p<0.001). Fifty-four (89%) patients were mechanically ventilated during their ICU stay. These patients were ventilated for a median of 11 days. Four patients from the ICU admitted patients suffered from early ARDS (within 96 hours) following peritonitis.
Table 1  Characteristics of study population.

<table>
<thead>
<tr>
<th>Patient characteristics at index operation</th>
<th>Overall</th>
<th>Non-ICU</th>
<th>ICU</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age mean (± sd)</td>
<td>51.1 (±14)</td>
<td>46.7 (±15)</td>
<td>54.2 (±13)</td>
<td>0.011*</td>
</tr>
<tr>
<td>Gender Male % (n)</td>
<td>59% (60)</td>
<td>51% (20)</td>
<td>64% (39)</td>
<td>0.21†</td>
</tr>
<tr>
<td>Comorbidity % (n)</td>
<td>65% (66)</td>
<td>67% (26)</td>
<td>65% (40)</td>
<td>0.91†</td>
</tr>
<tr>
<td>Use of any medication % (n)</td>
<td>77% (78)</td>
<td>82% (32)</td>
<td>78% (46)</td>
<td>0.62†</td>
</tr>
<tr>
<td>Apache II score mean (ssd)</td>
<td>9.5 (± 5.2)</td>
<td>8.2 (± 5)</td>
<td>10.4 (± 5)</td>
<td>0.037</td>
</tr>
<tr>
<td>MPI mean (ssd)</td>
<td>21.9 (± 7.4)</td>
<td>20.0 (± 8)</td>
<td>23.2 (± 7)</td>
<td>0.036*</td>
</tr>
</tbody>
</table>

Disease and operation characteristics

<table>
<thead>
<tr>
<th>Etiology of peritonitis % (n)</th>
<th>0.29†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perforation</td>
<td>39% (39)</td>
</tr>
<tr>
<td>Anastomotic leakage</td>
<td>26% (26)</td>
</tr>
<tr>
<td>Ischemia</td>
<td>5% (5)</td>
</tr>
<tr>
<td>Pancreatitis</td>
<td>7% (7)</td>
</tr>
<tr>
<td>Bile leakage</td>
<td>7% (7)</td>
</tr>
<tr>
<td>Abcess</td>
<td>9% (9)</td>
</tr>
<tr>
<td>Other</td>
<td>7% (7)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Origin of peritonitis % (n)</th>
<th>0.28†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colon</td>
<td>36% (36)</td>
</tr>
<tr>
<td>Small intestine</td>
<td>25% (25)</td>
</tr>
<tr>
<td>Pancreas</td>
<td>13% (13)</td>
</tr>
<tr>
<td>Appendix</td>
<td>4% (4)</td>
</tr>
<tr>
<td>Gall bladder</td>
<td>4% (4)</td>
</tr>
<tr>
<td>Stomach / Duodenum</td>
<td>6% (6)</td>
</tr>
<tr>
<td>Other</td>
<td>12% (12)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment strategy % (n)</th>
<th>0.11†</th>
</tr>
</thead>
<tbody>
<tr>
<td>OD</td>
<td>76% (77)</td>
</tr>
<tr>
<td>PR</td>
<td>24% (24)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Laparostomy (open abdomen) during admission % (n)</th>
<th>0.001†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with ≥ 1 relaparotomy % (n)</td>
<td>59% (60)</td>
</tr>
</tbody>
</table>

Patient characteristics at time of peritonitis; Disease and operation characteristics at time of peritonitis; Postoperative characteristics after surgical treatment for peritonitis.
PTSD symptoms

In an univariate analysis, using the PTSS-10 questionnaire above 35 points as the cut-off we found a prevalence of PTSD symptomology of 18% (7/39) in the non-ICU group and 28% (17/61) in the ICU group (p=0.21). Several factors were examined whether they confounded the strength of the relationship between ICU stay and the probability of having relevant PTSD symptomology (Table 2). After controlling for age, gender, APACHE II score, relaparotomy and length of hospital stay in a multivariate analysis, patients admitted to the ICU had a 4.3 times higher odds (95% CI 1.11 to 16.5) of reporting PTSD symptomology compared to ward only patients on the PTSS-10 questionnaire. Other factors that were significantly associated with more PTSD symptoms in the multivariate model included: Females had a 3.5 times higher odds of developing PTSD symptoms (95% CI 1.2 to 10.6) compared to males; younger patients, with every year decrease in age the odds decreased by 0.93 (95% CI 0.89 to 0.98); and more severe disease at the initial surgery, with every point increase in APACHE II score the odds increased by 1.1 (95% CI 1.002 to 1.25). Therefore, the main reason for finding a stronger relation between ICU stay and PTSD symptomology in the multivariate model is that older patients are less likely to develop PTSD symptoms. Because ICU patients on average were older than
non-ICU patients, the unadjusted relationship underestimated the effect of ICU on PTSD symptoms. Males were also less likely to report PTSD symptomology (OR 0.95; 95% CI 0.91 to 0.98), but because of the comparable gender distribution in ICU and non-ICU patients it did not confound the relation between ICU stay and PTSD symptomology (Table 2). Length of hospital stay was associated with more PTSD symptomology and was therefore also a confounder for the relation between ICU stay and PTSD symptomology because hospital stay was markedly longer in ICU patients compared to non-ICU patients. The area under the ROC curve (AUC) for the final multivariate model was 0.77 (95% CI 0.66 to 0.88). This indicates that if we would randomly choose one patient above the PTSS-10 cut-off value and one patient below, the probability that the patient above the cut-off would have a higher predicted risk for PTSD symptomology based on the model is 77%. Differences in observed versus predicted probabilities were small with the Hosmer-Lemeshow test having a p-value of 0.41.

Traumatic memories and PTSD symptoms

In the total study population, traumatic memories were associated with more PTSD symptomology (Table 3). Patients reporting more traumatic memories during their ICU or hospital stay reported significantly more PTSD symptoms on the PTSS-10. Patients with nightmares, panic attacks, intense pain and difficulty breathing during ICU or hospital ward stay had higher median scores than patients reporting no traumatic memories from the ICU or hospital ward (Table 3). There were, however, no statistically significant differences between the ICU group and the non-ICU group of patients with respect to reporting traumatic memories (nightmares, \( \chi^2=5.84, p=0.12 \); fear or panic attacks, \( \chi^2=6.9, p=0.075 \); pain \( \chi^2=1.01, p=0.80 \); difficulty breathing, \( \chi^2=5.3, p=0.15 \)).

Discussion

Our cohort of patients experiencing the same acute disease includes both patients who have been admitted to ICU and those who were treated on the surgical ward only. This enables a more detailed analysis of the impact of ICU stay on long-term PTSD symptomology. We found a high overall prevalence of long-term PTSD symptomology as screened by the PTSS-10 questionnaire many years after surgical treatment for secondary peritonitis. The proportion of patients scoring above the 35-point threshold on the PTSS-10 was 24%. The PTSS-10 is an instrument specifically designed for determining PTSD symptoms in ICU patients. The prevalence of PTSD symptoms found in our patients was similar to that found in a retrospective study in ARDS patients in 1998 using the PTSS-10 questionnaire, and similar to the ARDS patients studied in 2004 with a median of 8 years follow-up in which 24% of patients suffered from full-blown PTSD diagnosed by SCID\(^{11}\). Past studies have observed a lifetime prevalence of 7.8% to 8.3% in the general population in the 1990s in the USA\(^{24}\), but more recently a study conducted in 6 European countries (the ESEMeD study\(^{25,26}\)) estimated a considerably lower prevalence of PTSD varying between 0.9% and 2.9%. Compared to these general populations, the proportions of PTSD symptomology in an ICU population after a lengthy hospital discharge time are high and costly\(^{13,25-28}\).
Table 2  Multivariable logistic regression analysis for factors associated with the presence of PTSD symptomology

<table>
<thead>
<tr>
<th>PTSS-10 Sum score &gt; 35</th>
<th>Patients with sum scores above 35 (n = 24)</th>
<th>Patients with sum scores below 35 (n = 76)</th>
<th>Adjusted OR** (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients admitted to ICU (%)</td>
<td>71%</td>
<td>57%</td>
<td>4.3 (1.11 to 16.5)</td>
</tr>
<tr>
<td>Females (%)</td>
<td>54%</td>
<td>37%</td>
<td>3.5 (1.2 to 10.6)</td>
</tr>
<tr>
<td>Age mean, (sd)</td>
<td>52.9 (14.9)</td>
<td>59.8 (14.1)</td>
<td>0.93† (0.89 to 0.98)</td>
</tr>
<tr>
<td>APACHE II mean (sd)</td>
<td>10.7 (5.8)</td>
<td>9.1 (5.0)</td>
<td>1.1‡ (1.002 to 1.25)</td>
</tr>
<tr>
<td>Relaparotomy ≥ 1 (%)</td>
<td>63%</td>
<td>59%</td>
<td>3.8 (0.86 to 16.8)</td>
</tr>
<tr>
<td>Hospital stay+ (median, P25–P75)</td>
<td>46 (28–54)</td>
<td>33 (21–59)</td>
<td>2.2 * (0.8 to 5.8)</td>
</tr>
</tbody>
</table>

*One patient was missing hospital discharge dates and had therefore missing information pertained to length of hospital stay; imputation was done using the mean duration stay for the non-ICU stay group; One patient was missing APACHE II score data. **The odds ratio (OR) has been adjusted for gender, age, APACHE II score, 1 or more relaparotomy and length of hospitals stay. †Per 1-point increase in APACHE II score, the odds for having PTSD symptoms increased by 1.1. ‡Per 1-year increase in age the odds for having PTSD symptoms decreased by 0.93.

We found that patients who responded to the PTSS-10 questionnaire showed higher APACHE II scores, MPI scores and increased comorbidity than patients who did not respond to the questionnaire. These differences may have led to a small overestimation in PTSD symptoms (n=100). Simultaneously, our patient group had an overall lower mean APACHE II score than reported in other ICU populations with similar PTSD symptom prevalence. Although the APACHE II scores of the study patients admitted to the ICU are lower compared to other studies on PTSD symptoms using the PTSS-10 questionnaire3,4, the APACHE II scores are not particularly low for a population of peritonitis patients1.

In a univariate analysis we found no significant differences in the prevalence of PTSD symptoms between ICU (28%) and non-ICU patients (18%) on the PTSS-10, but ICU stay was independently associated with PTSD symptomology after adjusting for other factors related to PTSD, in particular age.

As expected, when comparing ICU patients to non-ICU patients, differences were found in patient, disease, operation and postoperative characteristics. ICU patients were older and had a more severe disease according to the registered APACHE II score, more surgical interventions, and longer hospital stay, all of which could have affected their eventual PTSD symptomology. To control for these differences and determine whether ICU was an independent factor for PTSD we created a multivariate model. When controlling for age, gender, APACHE II score, 1 or more relaparotomy and length of hospital stay in the postoperative period, we found a significant difference in the prevalence of PTSD symptomology according to the PTSS-10 between
patients with and without ICU stay. Higher age and being male played a protective role for PTSD, whereas higher APACHE II scores led to more PTSD symptoms. These results contradict earlier results in which there were no associations found between higher APACHE II scores and a higher probability of developing PTSD symptoms4,5. It is important to note that even the non-ICU group reports a relatively high prevalence of PTSD-related symptoms. This suggests not only the ICU environment, but secondary peritonitis in itself may be enough of a life-threatening traumatic event for a patient to develop PTSD.

Due to the fact that mechanical ventilation has been associated in the past with the development of more PTSD-like symptoms after ICU treatment29, this may be the reason that our ICU patients also report more PTSD symptomology than surgery ward only patients. Because nearly all our ICU patients were mechanically ventilated, we could not determine the independent impact of these two factors. Due to the retrospective nature of this study, details concerning the severity of sepsis such as septic shock status on admission and hydrocortisone use during ICU stay could not be ascertained as risk factors in all patients12,23. These factors could be potentially important in the development of PTSD symptoms in ICU patients. The importance of hydrocortisone use in the ICU and the development of PTSD symptoms has been previously highlighted12,23. A randomized study has shown that the introduction of hydrocortisone treatment during into the ICU stay regime reduces the later development of PTSD symptoms23. In past studies ARDS has been demonstrated to be an independent predictor of

Table 3  Traumatic memories during ICU/hospital stay in relation to the PTSS-10

<table>
<thead>
<tr>
<th>Traumatic memories or adverse experiences during ICU/hospital stay</th>
<th>% of patients with sum scores above 35 (n =24)</th>
<th>P-value for trend*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nightmares</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Never (n=42)</td>
<td>9.5</td>
<td>0.001</td>
</tr>
<tr>
<td>- Sometimes (n=29)</td>
<td>24.1</td>
<td></td>
</tr>
<tr>
<td>- Regularly (n=20)</td>
<td>45.0</td>
<td></td>
</tr>
<tr>
<td>- Often (n=9)</td>
<td>44.4</td>
<td></td>
</tr>
<tr>
<td>Fear or panic attacks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Never (n=53)</td>
<td>9.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>- Sometimes (n=23)</td>
<td>26.1</td>
<td></td>
</tr>
<tr>
<td>- Regularly (n=16)</td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td>- Often (n=8)</td>
<td>62.5</td>
<td></td>
</tr>
<tr>
<td>Intense Pain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Never (n=27)</td>
<td>7.4</td>
<td>0.007</td>
</tr>
<tr>
<td>- Sometimes (n=35)</td>
<td>17.1</td>
<td></td>
</tr>
<tr>
<td>- Regularly (n=15)</td>
<td>60.0</td>
<td></td>
</tr>
<tr>
<td>- Often (n=23)</td>
<td>30.4</td>
<td></td>
</tr>
<tr>
<td>Difficulty breathing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Never (n=50)</td>
<td>12.0</td>
<td>0.014</td>
</tr>
<tr>
<td>- Sometimes (n=30)</td>
<td>33.3</td>
<td></td>
</tr>
<tr>
<td>- Regularly (n=9)</td>
<td>44.4</td>
<td></td>
</tr>
<tr>
<td>- Often (n=11)</td>
<td>36.4</td>
<td></td>
</tr>
</tbody>
</table>

* x² test for linear trend
PTSD symptoms; patients suffering from ARDS reported more PTSD symptoms\textsuperscript{6,10,32}, however in this study we only had data on development of ARDS within the first 4 days following peritonitis. Risk factors of the ICU environment such as ARDS, septic shock and mechanical ventilation (the vast majority of the study patients admitted to the ICU were ventilated) could – at least in part – account for the differences in PTSD symptoms reported by patients in the ICU versus surgical wards only.

There were no differences in the number of traumatic memories reported by the ICU patients versus the surgical ward only patients, although we found a clear positive linear association between more traumatic memories and higher sum scores on the PTSS-10. This relation between traumatic memories and the PTSS-10 score was also found in two earlier studies in ICU patients\textsuperscript{6,30}. We had anticipated that the ICU environment would create more traumatic memories, which would in turn lead to more PTSD-related symptoms. However, the proportion of patients with traumatic memories was comparable between ICU and non-ICU patients.

Limitations

Ideally PTSD is diagnosed using a Structured Clinical Interview (SCID)\textsuperscript{23} according to the DSM-IV\textsuperscript{9}. The SCID II is a semi-structured diagnostic interview designed to allow clinicians and researchers make reliable DSM-IV psychiatric diagnoses. In recent studies, it has been established that a self-report PTSS-10 questionnaire can be as useful a tool in determining which patients are suffering from PTSD symptomology\textsuperscript{5,17}. These studies found significantly higher PTSS-10 scores in patients with a SCID II PTSD diagnosis than patients without. The sensitivity in these studies varied from 77\% to 100\%, and specificity from 92\% to 98\% when using a cut-off score of 35\textsuperscript{5,17}. However, these estimates were imprecise because of the small sample size in these studies. It is unclear to what extent sensitivity and specificity of the PTSS-10 instrument for PTSD may vary according to disease and other characteristics\textsuperscript{31}.

The distinction between PTSD symptoms captured by the PTSS-10 and a PTSD diagnosis is vital in understanding that this questionnaire does not give a DSM-IV diagnosis, but only an indication of the level of symptomology. Clinically this would indicate that patients who score above the cut-off, should be referred by the attending physician to a psychologist to carry out a SCID II\textsuperscript{9}.

Our results are suggestive that the (persisting) presence of traumatic memories is likely to be relevant for the development of PTSD-related symptoms following a traumatic event and not ICU alone, as we have seen a strong linear relationship between traumatic memories and PTSS-10. We assessed these traumatic memories (or adverse experiences) in accordance to the patients’ recollections. This can limit the conclusions one can make, as it is possible that the patients’ perceived traumatic experience may contribute to their long-term PTSD symptomology, hence making a causal conclusion impossible. Information concerning other non-related traumatic experiences or life events that may have occurred after hospital admission was not collected. Therefore, the influence of super-imposed trauma cannot
be ruled out. Also, as this was a retrospective study it was also not possible to collect PTSD data on patients prior to their peritonitis. However, considering the acute nature of peritonitis, even in a prospective trial these data would be hard to collect. Given the impact of a severe life-threatening illness such as peritonitis a relationship with the development of PTSD symptoms is plausible, but causality cannot be established when no information is available on other life events.

**Conclusion**

Nearly a quarter of patients receiving surgical treatment for secondary peritonitis developed PTSD symptoms. Considering the high long-term prevalence of PTSD, patients admitted to the ICU had a higher risk of PTSD symptoms, but only after taking their higher age into account. Early detection of PTSD in peritonitis patients by questionnaires such as the PTSS-10 deserves attention.
References


Chapter 10
Factors associated with post-traumatic stress symptoms in a prospective cohort of patients after abdominal sepsis. A nomogram

Kimberly R Boer, Oddeke van Ruler, Arnold A van Emmerik, Mirjam A Sprangers, Sophia E de Rooij, Margreeth B Vroom, Corianne A de Borgie, Marja A Boermeester, Johannes B Reitsma

For the Dutch Peritonitis Study Group

Intensive Care Medicine, 2007, provisionally accepted
Abstract

Objective: To determine to what extent patients who survived abdominal sepsis suffer from symptoms of post-traumatic stress disorder (PTSD) and depression, and to identify potential risk factors for PTSD symptoms.

Design and setting: PTSD and depression symptoms were measured using the Impact of Events Scale-Revised (IES-R), the Post-Traumatic Symptom Scale–10 (PTSS-10) and the Beck Depression Inventory-II (BDI-II).

Patients and participants: A total of 135 peritonitis patients were eligible for this study of which 107 (80%) patients completed the questionnaire. The median APACHE II score was 14 (12 to 16) and 89% were admitted to the ICU.

Measurements and results: The percentage of patients with ‘moderate’ PTSD symptom scores was 28% (95% CI 20 to 37), whilst 10% (95% CI 6 to 17) of patients had ‘high’ PTSD symptom scores. Only 5% (95% CI 2 to 12) of the patients expressed severe depression symptoms. Factors associated with increased PTSD symptoms in a multivariate ordinal regression model were younger age (0.74 per 10 years older, p=0.082), length of ICU stay (OR=1.4 per doubling of duration, p=0.003) and having some (OR=4.9, p=0.06) or many (OR=55.5, p<0.001) traumatic memories of the ICU or hospital stay.

Conclusion: As many as 38% of patients after abdominal sepsis report elevated levels of PTSD symptoms on at least one of the questionnaires. Our nomogram may assist in identifying patients at increased risk for developing symptoms of PTSD.
Introduction

Post-traumatic stress disorder (PTSD) is the development of psychological and physical symptoms following exposure to one or more traumatic events. PTSD symptoms include intrusive recollections (re-experiencing the trauma in flashbacks, memories or nightmares); avoidant and numbing symptoms (including diminished emotions and avoidance of situations that are reminders of the traumatic event); and hyperarousal (including increased irritability, exaggerated startle reactions or difficulty sleeping or concentrating). PTSD symptoms have a major impact on life, illustrated by the fact that these patients have a reduced quality of life and frequently suffer from depression. In addition, PTSD can lead to substance abuse, job instability and even suicide.

Events that typically trigger the development of PTSD include exposure to violent events such as rape, domestic violence, child abuse, war, accidents, natural disasters and political torture, all of which include a threat to life. Increasingly PTSD has also been found in patients who survived a major, life-threatening disease and patients who were admitted to the intensive care unit (ICU) for an extensive period. Severe peritonitis (or abdominal sepsis) is such a disease where patients have an episode of acute and severe illness, frequently followed by a lengthy ICU stay and a long recovery period that often includes multiple surgical and non-surgical interventions. This combination of factors could make this patient group particularly vulnerable for developing PTSD symptoms. To date, little is known about the presence and severity of PTSD and possible risk factors in patients recovering from severe peritonitis.

As it is neither efficient nor feasible to have all patients with secondary peritonitis, who have recently undergone surgery, to be interviewed by a psychologist, we have aimed to make a prognostic model to aid surgeons in making a selection of high-risk patients based on PTSD symptomology (screening tool) to determine which patient should be further assessed for a PTSD diagnosis and eventually treatment. PTSD and depression often present with similar symptoms and can both be found following ICU stay. Due to their potential overlap it is important to determine the extent and development of not only PTSD symptoms but also depression symptoms following abdominal sepsis.

Therefore, our aim was to determine the presence and level of symptoms of PTSD and depression in patients surviving abdominal sepsis. In addition we attempted to identify factors associated with the development of PTSD symptomology, identifying these factors could be important in determining potential targets for interventions aimed at early detection and potential treatment of PTSD symptoms in peritonitis patients and other critically ill patients and avoid burdening all patients with in-depth and distressing psychological assessment interviews.
Methods

Study design

Our study was embedded in an ongoing randomized clinical trial (RELAP trial) evaluating two surgical treatment strategies for patients with secondary peritonitis after the initial emergency laparotomy, initiated by the Academic Medical Center, University of Amsterdam, The Netherlands. Patients were enrolled between December 2001 and February 2005 in 2 academic and 7 regional teaching hospitals. All patients were followed for 12 months after initial (index) laparotomy. The study was approved by the Medical Ethical Committee of the Academic Medical Center, Amsterdam. All patients gave informed consent to participate in this study.

Study population

Patients were eligible for the RELAP trial if they had a clinical diagnosis of secondary peritonitis requiring emergency laparotomy. Peritonitis had to be caused by perforation or infection of a visceral organ, or ischemia/necrosis of a part of the gastrointestinal tract, or postoperative peritoneal infection. An Acute Physiology And Chronic Health Evaluation (APACHE) II score above 10 was required for study inclusion. Exclusion criteria included being below 18 or above 80 years old; peritonitis due to bowel perforation after endoscopy operated within 24 hours; abdominal infection due to indwelling dialysis (CAPD) catheter; acute pancreatitis; expected survival of less than 6 months due to disseminated malignancy; severe brain damage due to trauma or anoxia; and imperative relaparotomy (e.g. for the removal of gauze packing). Symptoms of PTSD and depression were measured 12 to 18 months after their initial emergency laparotomy.

Data collection

All self-administered PTSD and depression questionnaires were distributed by mail to patients who survived at least 12 months, with a reminder by phone within two weeks in case of no response. After 1 month without response a new questionnaire including a reminder letter was sent. General patient characteristics and disease and postoperative characteristics were prospectively collected over a 12-month follow-up period for all included patients using case record forms (CRF).
Outcomes

Instruments assessing the level of PTSD and depression symptoms

We used two instruments for measuring PTSD symptoms in research settings: the Post-Traumatic Stress Scale-10 (PTSS-10) and the Impact of Event Scale-Revised (IES-R). Both the PTSS-10 and IES-R are designed to assess the presence and intensity of specific post-traumatic stress symptoms and are suitable for mapping symptoms and reactions after traumatic events. Both instruments have shown good psychometric characteristics. To determine the extent of depression symptoms, we have used the well-validated Beck Depression Inventory-II (BDI-II).

Post-traumatic Stress Syndrome Scale-10 (PTSS-10) was originally designed to diagnose PTSD according to Diagnostic and Statistical Manual of Mental Disorders-III (DSM-III) criteria in victims of natural disasters. The PTSS-10 has since been validated in Acute Respiratory Distress Syndrome (ARDS) patients.
after ICU treatment using the Structured Clinical Interview (SCID) for the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV)\textsuperscript{14,36}. The PTSS-10 is now a widely used self-report questionnaire assessing symptoms related to PTSD, particularly in critically ill and ICU patients\textsuperscript{14,15}.

The PTSS-10 consists of 10 items of which each item ranges from 1 (none) to 7 (always) points. The total score ranges from 10 to 70, with higher scores indicating more symptoms. Items include those indicating sleep disturbances, nightmares, depression, hyper-alertness, withdrawal, general irritability, frequent mood swings, guilt, avoidance of activities prompting the recall of possible traumatic events, and increased muscle tension\textsuperscript{28}. Stoll and colleagues demonstrated criterion validity by receiver operating characteristic curve analysis and showed that a cut-off of 35 points had a sensitivity of 77\% and a specificity of 97.5\% for the diagnosis of PTSD\textsuperscript{14,36}, therefore scores of 35 or above were considered indicative of PTSD.

The Impact of Events Scale–Revised (IES-R) is one of the most commonly used self-report questionnaires for determining PTSD symptomology following a trauma\textsuperscript{29}. The IES-R can discriminate between stress reactions at different time points after an event, and has convergent validity with observer-diagnosed PTSD, and is valuable for detecting individuals who require treatment\textsuperscript{37}. The IES-R consists of 22 items, each item ranging from 0 (no problems) to 4 (frequent problems) with the total score ranging from 0 to 66. Scores above 24 points are generally considered indicative of PTSD with higher scores indicating more symptoms\textsuperscript{30}. The IES-R has been developed based on DSM-IV criteria and therefore has three distinct subscales, the avoidance subscale (eight questions), the intrusion subscale (eight questions) and the hyperarousal subscale (six questions)\textsuperscript{30,38}, and it is one of the most frequently used self-report questionnaires in both the clinic and in PTSD research\textsuperscript{29}.

The Beck Depression Inventory-II (BDI-II) is a well-established questionnaire for measuring the presence and severity of depressive symptoms consisting of 21 questions\textsuperscript{39}. It employs a 4-point response scale ranging from 0 to 3; the total score ranging from 0 to 63, with higher scores indicating increased symptoms of depression. In accordance to the BDI-II manual, depression is classified as: none to minimal depression (sum score 0-13), mild depression (sum score 14-19), moderate depression (sum score 20-28) and severe depression (sum score 29-66).

**Predictors**

Potential risk factors were selected from previous studies examining factors for increased mortality and morbidity\textsuperscript{20-25,40,41} in secondary peritonitis supplemented with specific factors mentioned in the PTSD literature\textsuperscript{10,12-14,17,42,43}. This pre-selection was based on findings from earlier studies\textsuperscript{44}. We divided these factors into three distinct categories.

**General patient characteristics** include age, gender, and the presence of major comorbidity (cardiovascular disease; COPD; renal failure; diabetes; malignancy).
Disease characteristics and postoperative course include severity of disease measured at the time of initial laparotomy using the APACHE II score. As several components of the APACHE II score are already considered in a univariate analysis (namely age and comorbidity); we have chosen to replace the APACHE II score with the APS score as a potential predictor of PTSD. The APS is composed of only the acute components of the APACHE II score; so without age and comorbidity. Postoperative characteristics included administration of hydrocortisone during ICU stay, development of acute respiratory distress (ARDS), number of relaparotomies, length of ICU stay and hospital admission period, the development of a disease-related major morbidity during 6-months follow-up and an enterostomy present after 6-months follow-up.

Traumatic memories of ICU/hospital stay were assessed using the 4-item adverse experiences questionnaire, which captures four types of traumatic memories of their stay in the ICU or hospital ward: nightmares, fear and panic, pain and difficulty breathing. Patients scored the frequency of traumatic memories of their stay in the ICU or hospital ward using a 4-point response scale of 0 (never), 1 (sometimes), 2 (regularly) or 3 (often). These were subsequently summed and classified into 3 graded categories of traumatic memories, 0 (no traumatic memories), 1 to 4 (some traumatic memories), and more than 4 (many traumatic memories). The questions concerning traumatic memories of the ICU/hospital stay were administered concurrently with the PTSS-10, IES-R and BDI-II questionnaires at least 12-months follow-up.

We also collected data on whether patients had experienced other traumas or whether a close family-member or friend had experienced a trauma within the last three years. We used questions 29 and 30 from the Life Stressor Checklist-Revised which were also administered at the same time as the PTSS-10, IES-R and BDI-II questionnaires. Responses were given dichotomously as yes or no and patients were subsequently asked to specify the event type. These questions were asked to determine to what extent the PTSD symptomology found in this patient group was due to their peritonitis or due to other traumatic events.

Data analysis

General patient characteristics, disease characteristics and postoperative course were expressed as percentage for categorical data, mean (SD) for normally distributed numerical data and median (interquartiles, IQR) for non-normally distributed continuous data.

We used two instruments aimed at measuring the presence and severity of PTSD symptoms in our population, each with their own cut-off value. Combining data from two instruments measuring the same construct (PTSD symptoms) may lead to a more robust classification of patients. To preserve the natural ordering of patients who scored below the cut-off value on both questionnaires (‘low-scoring patients’), patients scoring above the cut-off on only one of these questionnaires (‘moderate-scoring patients’), and patients scoring above the cut-off on both questionnaires (‘high-scoring patients’), we applied ordinal regression modeling. The proportion of patients in each of these three categories is presented with 95% confidence intervals (95% CI) using the method of Wilson.
The prevalence and severity of depression were determined by the proportion of patients who scored within the cut-offs reported in the BDI-II manual. The prevalence of depression symptoms was presented as percentages with 95% CI. Furthermore, mean BDI-II scores per PTSD symptom group (‘low-scoring patients’, ‘moderate-scoring patients’, ‘high-scoring patients’) were presented and compared using an ANOVA.

Potential predictors for PTSD symptoms were analyzed using an ordinal logistic regression model. This ordinal regression model is an extension of the binary logistic model and is appropriate when a continuous trait is grouped into several categories by using cut-offs.

All potential predictors for PTSD symptoms were first examined in univariate ordinal regression models. Factors with a p-value of less than 0.1 were entered in a multivariate ordinal logistic regression model. If variables within a group of predictors were strongly correlated, only the factor with the strongest univariate relationship and/or most relevant clinical interpretation was added to the model. Due to clinical relevance from the literature on PTSD and ICU studies, age and gender were always included in the multivariate model regardless of the strength of their associations.

In addition, the factor comprised of non-related other traumas that the patient experienced within the last 3 years was also included in the final model to assess its potential confounding role.

The fit and validity of the model was checked in the following ways. The discriminatory properties of the model were examined by visualizing the distribution and degree of overlap in risk scores of individual patients within and between the three outcome categories. The proportional odds assumption was evaluated by performing the test for parallel lines.

Calibration was checked by comparing expected and observed number of patients in each of the three outcome categories across deciles of expected risk and tested for significance by using an extension of the Hosmer-Lemeshow goodness-of-fit statistic.

**Nomogram**

A nomogram was developed to visualize the prognostic strength of the different factors from the multivariate model in a single figure. A nomogram allows readers to calculate an expected distribution of PTSD symptomology (‘low-scoring patients’, ‘moderate-scoring patients’, and ‘high-scoring patients’) based on a specific profile of a patient. The number of points for each predictor was based on the original coefficient from the multivariate ordinal model by multiplying it by ten and rounding it to the lowest whole number. The total number of points derived by specifying values for all predictors was used to calculate the expected probabilities that a patient would be a ‘low-scoring patient’, a ‘moderate-scoring patient’ or a ‘high-scoring patient’.

Analyses were performed using SAS software version 9.1 (SAS Institute Inc., Cary, NC, USA)
Results

A total of 132 patients were eligible for this study and thus received questionnaires. A total of 108 (80%) patients responded to the questionnaire. Reasons given for not completing the PTSD questionnaire were: refusal (11), too ill (5), other reasons (4), and unknown (4) (Figure 1). On average the responses were provided approximately 12.5 months following initial emergency laparotomy. There were no significant differences in any of the patient- or disease-characteristics between respondents and non-respondents.

Patients were on average 66.8 median years old (IQR 57 to 73) and 54% were male. Patients were severely ill, with a median APACHE II score of 14 (IQR 12 to 16) and a median APS score of 6 (IQR 4 to 8) and 5% had a major comorbidity (Table 1). Ninety-six patients (89%) were admitted to ICU and their median ICU stay was 7 days (IQR 4 to 15 days) and these patients were mechanically ventilated for a median of 5 days (IQR 1 to 8 days). Patients were hospitalized for a median period of 28 days (IQR 19 to 55 days).

Fifty-one percent of patients also underwent a trauma themselves in the 3 years prior to filling-out the questionnaires (these included leg and breast amputations; rib and hip fractures; heart attacks and other heart complications; brain aneurysm; bleeding ulcer; streptococcal infection; kidney failure) or a family member or friend undergone a recent traumatic event (including sibling, child, parent or partner being severely ill, and sibling, child, parent or partner died).

Prevalence of PTSD and depression symptoms

The percentage of ‘moderate-scoring’ PTSD patients was 28% (95% CI 20 to 37%), whilst 10% (95% CI 6 to 17%) of patients were ‘high-scoring’ patients (Table 1). The proportion of patients classified as mildly depressed, according the BDI-II, was 22% (95% CI 16 to 31), whilst 7% (95% CI 3 to 13) were classified as moderately depressed and 5% (95% CI 2 to 10) as severely depressed. There was a positive association between the presence and severity of PTSD and depression symptoms. Only one of the ‘low-scoring’ PTSD patients (1.5%) reported moderate or severe depression symptoms, whilst 27% (11/41) of ‘moderate- to high-scoring’ PTSD symptom patients reported moderate to severe depression symptoms (Table 2).
Table 1 Association between severity of PTSD symptoms (3 categories) and patient, disease operative and postoperative characteristics

Results from univariate ordinal regression models

<table>
<thead>
<tr>
<th></th>
<th>PTSD symptoms ** (n=107)</th>
<th>Univariate Ordinal regression*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>None to mild</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(n=66)</td>
</tr>
<tr>
<td><strong>General patient’s characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median age (IQR)</td>
<td>67 (57 to 73)</td>
<td>70 (60 to 74)</td>
</tr>
<tr>
<td>Male gender (%)</td>
<td>54%</td>
<td>53%</td>
</tr>
<tr>
<td>Major comorbidity present (%)***</td>
<td>53%</td>
<td>55%</td>
</tr>
<tr>
<td><strong>Peritonitis and postoperative characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Initial:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median APS score (IQR)</td>
<td>6 (4 to 8)</td>
<td>6 (4 to 8)</td>
</tr>
<tr>
<td>Hydrocortisone in first 14 days in ICU (median days)</td>
<td>2 (0 to 7)</td>
<td>1.5 (0 to 8)</td>
</tr>
<tr>
<td>ARDS</td>
<td>6%</td>
<td>3%</td>
</tr>
<tr>
<td>1 or more relaparotomies</td>
<td>67%</td>
<td>70%</td>
</tr>
<tr>
<td>Admitted to the ICU</td>
<td>89%</td>
<td>85%</td>
</tr>
<tr>
<td>Median length of ICU stay (IQR)</td>
<td>7 (4 to 15)</td>
<td>7 (4 to 12)</td>
</tr>
<tr>
<td>Median ventilation time (IQR)</td>
<td>5 (1 to 8)</td>
<td>4 (1 to 7)</td>
</tr>
<tr>
<td>Median length of hospital stay (IQR)</td>
<td>28 (19 to 55)</td>
<td>26 (18 to 47)</td>
</tr>
<tr>
<td><strong>Follow-up</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease-related major morbidity at 6-months follow-up</td>
<td>15%</td>
<td>9%</td>
</tr>
<tr>
<td>Enterostomy at 6-months follow-up</td>
<td>51%</td>
<td>47%</td>
</tr>
</tbody>
</table>

IQR – Interquartile range

* All models were checked for parallel lines to see if an ordinal test for significance was appropriate.
**3 graded outcomes: none to mild, moderate and high, 2 patients’ data are based on only one completed questionnaire.
***Major comorbidity included cardiovascular disease, COPD, renal failure, diabetes and malignancy.
Table 2  Association between severity of PTSD symptoms (3 categories) and severity of depression symptoms and PTSD symptoms (n= 105).

Depression symptoms measured by BDI-II and defined as none to minimal, mild, moderate and severe. PTSD categories are ‘none to mild’ PTSD symptoms, ‘moderate’ PTSD symptoms and ‘severe’ PTSD symptoms.

<table>
<thead>
<tr>
<th>Depression Symptoms</th>
<th>Total (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimal</td>
</tr>
<tr>
<td>None to Low n (%)</td>
<td>56 (85%)</td>
</tr>
<tr>
<td>Moderate n (%)</td>
<td>14 (47%)</td>
</tr>
<tr>
<td>High n (%)</td>
<td>1 (9%)</td>
</tr>
</tbody>
</table>

* ANOVA test for significant differences in mean BDI-II scores between the three PTSD symptoms group <0.001.
**3-graded outcomes: ‘none to mild’, ‘moderate’ and ‘high’ PTSD symptoms, 2 patients’ data are based on only one completed questionnaire.

Predictive factors

Univariate analysis
From the general patient characteristic age was associated with the development of PTSD symptoms, showing that older patients were less likely to report PTSD symptoms. The APS score, indicating initial severity of disease was not significantly predictive of the PTSD symptoms. A more complicated course of the disease, as signified by longer ICU, longer ventilation time, longer admission to the hospital or developing a disease-related morbidity within 6-months following the initial emergency laparotomy, were all associated with having more PTSD symptoms. The administration of hydrocortisone during ICU stay was not related to more PTSD symptoms. As well, patients who developed ARDS did not report more PTSD symptoms. These results are presented in Table 1.

Patients reporting traumatic memories and experiences of ICU/ hospital stay (nightmares, fear and panic, pain and difficulty breathing) were more likely to have PTSD symptoms (Table 3). The factor that included whether patients had experienced another trauma in their own life or whether a family or friend had experienced a trauma in the last 3 years was also a univariate predictor for PTSD symptoms (p=0.033).

Multivariate analysis
The final multivariate model included age, gender, length of ICU stay, disease-related morbidity during the 6-months follow-up, traumatic memories of the ICU or hospital stay, and other traumatic factors within the last three years (Table 4).
Instructions:

From section 1 of the nomogram, locate the age of the patient on the axis. Determine how many points the patient receives. Repeat this for each factor in section 1. Sum the points for all predictors.

From section 2, locate this sum score on the total points axis. Draw a line straight down to the bar graphs. These bars indicate the estimated probability of that patient having ‘no to mild’ PTSD symptoms (negative on both instruments), ‘moderate’ PTSD symptoms (positive on one instrument) and ‘high’ PTSD symptoms (positive on both instruments).
In our final model, increasing age was associated with a lower likelihood of PTSD symptomology (OR=0.74 per ten years increase in age, p=0.084). Gender was not predictive of PTSD symptoms (OR=0.90, p=0.822). Disease-related morbidity at the 6-months follow-up (OR=2.09, p=0.238) was no longer independently predictive of PTSD symptoms. Memories of the ICU/hospital stay (patients that reported some memories, OR=4.9, p<0.057, whilst patients that reported many memories, OR=55.5, p<0.001) were the most prominent independent risk factors for increased PTSD symptomology. Length of ICU stay was also significantly predictive in the development of PTSD symptomology in the multivariate model (OR=1.4, p=0.004). The proportional odds assumption was valid as indicated by a p-value of 0.694 for the test of parallel lines. Calibration of the model (closeness between predicted and observed probabilities) was good with a p-value of 0.987 for the goodness-of-fit test for ordinal models.

The relative strengths of these relationships are visualized in the nomogram (Figure 2). In this nomogram, one has the ability to use the information presented above to calculate per individual patient, the probability that this patient will score either no to mild, moderate or high PTSD symptoms according to the PTSS-10 and IES-R. Hence, using this nomogram allows for screening of patients at risk for PTSD symptomology following severe peritonitis.

A graphical impression of the model’s discriminative ability is shown in Figure 3. This figure shows that the mean risk score is significantly different between all three PTSD symptoms severity categories (p<0.001), although there is some overlap in the risk scores between patients from different categories of PTSD symptoms severity.

**Discussion**

The proportion of patients 12 months after peritonitis with ‘high-scoring’ PTSD symptomology was 10% and the percentage of ‘medium scoring’ patients (28%) was in line with earlier studies measuring PTSD symptoms in critically ill patients who had been admitted to ICU12,16,30,51, whilst the prevalence of PTSD recorded in the general population varies between 0.9% and 2.9% (the ESEMeD study)11,53,54. The proportion of patients after peritonitis that reported symptoms of depression according to the BDI-II was comparatively low. Only 7% of patients reported moderate depression and 5% of patients reported severe depression symptoms. As expected the presence of PTSD symptoms and depression symptoms were strongly correlated. However, the overlap between the two psychological symptoms profiles was not as high as we had expected, as this patient group clearly suffers more from PTSD expressed symptoms than depression symptoms.

In the clinical setting, there is a continuing debate whether to intervene in the more acute peri-traumatic psychological processes or in a later phase, when symptoms or prodromes of PTSD are observed. By better understanding which factors play an important role in the development of PTSD, we can better prevent PTSD symptoms in high-risk patients and decide when best to intervene. The aim of our predictive model is that it can be used by treating physicians; following the acute episode and phase of secondary peritonitis in which survival and physical recovery are the main concerns, to recognize high-risk PTSD patients. This relatively simple model can aid the surgeon for instance during the first outpatient visit in determining which patients are at higher risk for the development of PTSD.
The development of PTSD symptoms was, however, predominantly related to a more complicated course of secondary peritonitis. Longer ICU and hospital stays and major disease-related morbidity during the 6-months follow-up were associated with more PTSD symptoms. In concordance with earlier studies, the strongest predictor of having PTSD symptoms following abdominal sepsis was having traumatic memories and experiences during their initial hospital or ICU stay. These results suggest that the presence of traumatic memories is one of the most relevant aspects for the development of PTSD-related symptoms. Earlier studies also found that subjective interpretation of the intensive care experience emerged as a consistent predictor of adverse emotional outcome, in both the short and the long term.

Age plays a critical role in the development of PTSD symptoms. Younger patients are much more likely to develop and report PTSD symptoms. This relationship confirms the results of an earlier, retrospective study of a different cohort of patients 4 to 10 years following hospital admission for severe peritonitis. These findings suggest that older patients are more able to adapt to the limitations that are associated with experiencing such a major disease, most likely because they have already experienced a comorbid

### Table 3

<table>
<thead>
<tr>
<th>Traumatic memories of the ICU or hospital stay</th>
<th>None to mild (n=64)</th>
<th>Moderate (n=30)</th>
<th>High (n=11)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nightmares</td>
<td>39%</td>
<td>61%</td>
<td>82%</td>
<td>0.002</td>
</tr>
<tr>
<td>Fear and panic</td>
<td>24%</td>
<td>61%</td>
<td>100%</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Pain</td>
<td>67%</td>
<td>70%</td>
<td>82%</td>
<td>0.002</td>
</tr>
<tr>
<td>Difficulty breathing</td>
<td>33%</td>
<td>76%</td>
<td>100%</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Categories of traumatic memories</th>
<th>None (0)</th>
<th>Moderate (1 through 4)</th>
<th>Severe (more than 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None (0)</td>
<td>41%</td>
<td>50%</td>
<td>9%</td>
</tr>
<tr>
<td>Moderate (1 through 4)</td>
<td>7%</td>
<td>47%</td>
<td>47%</td>
</tr>
<tr>
<td>Severe (more than 4)</td>
<td>0%</td>
<td>18%</td>
<td>82%</td>
</tr>
</tbody>
</table>

‡2 patients not included in final analysis due to missing data on traumatic memories during ICU or hospital stay.
illness and health-related problems. In contrast to some other studies, gender did not play a role in the development of PTSD symptoms.

Patients with abdominal sepsis suffering from ARDS did not report more PTSD symptomology than those without ARDS. In earlier ICU studies, ARDS patients have reported considerable PTSD symptoms. In our cohort of abdominal sepsis patients we found different predictive factors for PTSD than those found in the ARDS patients. Secondary peritonitis in itself may have been severe enough with ICU admission and extended mechanical ventilation days to cause PTSD symptoms, therefore the added risk by ARDS may be moot. Lack of power may also be factor because the percentage of patients developing ARDS in this study was modest.

We did not find an association between hydrocortisone administration during ICU stay and PTSD symptoms within this peritonitis cohort, as has been demonstrated for other critically ill patient groups. Hydrocortisone was not a protective in developing PTSD symptoms, whereas other studies have found that administration of hydrocortisone during ICU can lead to a reduction in PTSD symptoms after discharge. In this study corticosteroid use during only the first 14 days of ICU was included in our analyses. The effect of prolonged use of hydrocortisone or late-stage use during conditional adrenal insufficiency cannot be excluded.

### Table 4 Association between severity of PTSD symptoms and patient, disease operative and postoperative characteristics and other traumatic experiences following peritonitis in a multivariate analysis

<table>
<thead>
<tr>
<th></th>
<th>Final Model (n=105)*</th>
<th>OR</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
</tr>
<tr>
<td>Age per 10 year increase</td>
<td>0.74</td>
<td>0.53</td>
<td>1.04</td>
<td>0.084</td>
</tr>
<tr>
<td>Female</td>
<td>0.9</td>
<td>0.94</td>
<td>2.3</td>
<td>0.822</td>
</tr>
<tr>
<td>Length of ICU stay (log2 transformed)</td>
<td>1.4</td>
<td>1.1</td>
<td>1.7</td>
<td>&lt;0.003</td>
</tr>
<tr>
<td>Major disease-related morbidity during the 6-month follow-up (including index hospital admittance)</td>
<td>2.1</td>
<td>0.61</td>
<td>7.11</td>
<td>0.238</td>
</tr>
<tr>
<td>Traumatic memories during ICU / hospital stay</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate (1 through 4)</td>
<td>4.9</td>
<td>0.95</td>
<td>24.9</td>
<td>0.058</td>
</tr>
<tr>
<td>Severe (&gt; 4)</td>
<td>55.5</td>
<td>9.4</td>
<td>328.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Other trauma within the last three years</td>
<td>2.4</td>
<td>0.94</td>
<td>6.3</td>
<td>0.085</td>
</tr>
</tbody>
</table>

*This multivariate ordinal analysis included a test for parallel lines (p=0.694)
‡2 patients not included in final analysis due to missing data on traumatic memories during ICU or hospital stay
Figure 3  Distribution of total points from nomogram (risk score) for the prediction of the severity of PTSD symptoms with use of the risk factors taken from the multivariate ordinal model

PTSD categories are graded according to severity; none to mild (negative on both instruments), moderate (positive on one instrument), severe (positive on both instruments).
Unfortunately due to the acute and life-threatening nature of secondary peritonitis, it was not possible to collect baseline information on PTSD or data on earlier psychological disorders. However, as recommended in a recent review by Griffiths and colleagues, to account for possible earlier traumas we have considered information pertaining to comorbid diseases, as well, we have collected data on other non-disease related traumatic events that had occurred within the last three years. These other non-disease related events were indeed associated with having more PTSD symptoms, and altered the initial odds ratios of the other factors to the extent that we considered it a moderate confounder. Timing also plays an important issue in collecting data on PTSD symptoms in critically ill and ICU patients. We chose the time period for the recording of PTSD symptoms to be 12 months for a very specific reason; in this severely ill patient group, we did not want to record patient recovery. Past studies have shown that critically ill patients only develop PTSD symptoms after their physical recovery period has passed. In most cases of PTSD symptoms following ICU stay, the development of PTSD symptoms has had a delayed onset.

In this study we have tried to learn from two questionnaires, one commonly used and validated in particular for critically ill patients, together with one of the most frequently used screening instruments for PTSD, the Impact of Events Scale–Revised. The prevalence of PTSD symptomology in the present study was based on a whether or not a patient scored above the cut-offs of the IES-R and the PTSS-10. We employed the two questionnaires as complementary for the detection of PTSD symptoms and not to compare results deducted from both questionnaires separately.

Combining the results of both questionnaires in our analysis was anticipated to lead to more robust assessment into the factors associated with more PTSD symptoms. Especially since these two questionnaires have many questions that do not overlap, combining the results might increase the likelihood of correctly determining the level of PTSD symptoms in patients and will be more useful in screening patients following ICU stay, and potentially reducing biases due to instrument variation. Although both questionnaires aim to measure the presence of PTSD symptoms, the concordance in classification of patients by the two instruments was not perfect with 30 patients (28%) being positive on one questionnaire but not on the other. This is a reflection of the difficulty of measuring PTSD based on questionnaires. It also means that both questionnaires are informative in their own right, and that combining both instruments will lead to more robust ordering of patients based on their level of PTSD symptoms. Therefore we combined both instruments and used ordinal regression techniques to examine which factors are related to higher levels of PTSD symptoms. These self-report questionnaires are frequently used, but the diagnostic values of these instruments in relation to a DSM-IV diagnosis are still being researched and discussed. Although some studies have reviewed the diagnostic value of the questionnaires, in general these studies were methodologically limited.

Although this study is based on information collected from two validated questionnaires, it did not include a clinical interview, SCID, for establishing a definite DSM-IV criteria diagnosis of PTSD. In clinical psychology, such an interview is recommended to determine an actual diagnosis for PTSD and/
or depression. However, considering the number of patients with moderate to high PTSD symptoms, questionnaire designs are more useful, as they screen patients who may score high on a self-report questionnaire\(^\text{64}\), and patients can therefore be correctly referred to an appropriate mental healthcare giver\(^\text{37}\). In a clinical setting, especially in the ICU, questionnaires may be the only available instruments to measure PTSD symptoms\(^\text{58}\). We assessed traumatic experiences during ICU/hospital stay based on the patients’ recollections (after 1 year). It is likely that the patients’ perceived traumatic experiences may have contributed to the development of PTSD symptoms, but it is also possible that having PTSD symptoms has influenced the perception of their stay. Future studies should aim to prospectively quantify traumatic experiences during or shortly after ICU stay to draw more causal conclusions, even though this might be difficult in patients with such a lengthy recovery period\(^\text{12,58,62,63}\).

Health related-quality of life (HR-QoL) was closely monitored in this trial (4 questionnaires in one year), and therefore patients may have received some psychological attention than they would not have in routine clinical care. We may therefore see a small underestimation in the proportion of patients reporting PTSD symptoms\(^\text{64}\).

The nomogram should primarily be seen as a tool to visualize and compare the predictive strength of different factors. Before using the nomogram to actually predict PTSD symptomology in clinical practice, it requires external validation of the model in different sets of patients with secondary peritonitis.

In conclusion, 10% of peritonitis patients report ‘high’ PTSD symptomology and another 28% moderate PTSD symptoms. Factors that were related to more PTSD symptoms included younger age, suffering from traumatic memories of the period of hospitalization, and length of ICU stay. Knowledge about these predictive factors is required to increase awareness and to develop tailored early treatment options for these high-risk patients. Our nomogram may assist in identifying these ‘high’ risk patients for PTSD.
References

(23) Granja C, Dias C, Costa-Pereira A, Sarmento A. Quality of life of survivors from severe sepsis and septic shock may be similar to that of others who survive critical illness. Crit Care. 2004;8:R91-R98.


Chapter 11

Kimberly R Boer, Johannes B Reitsma, Oddeke van Ruler, Arnold A van Emmerik, Mirjam A Sprangers, Marja A Boermeester, Corianne A de Borgie

Submitted
Abstract

**Introduction:** We examined the interrelationship between symptoms of post-traumatic stress disorder (PTSD) and health related-quality of life (HR-QoL) in patients surviving abdominal sepsis (severe peritonitis).

**Methods:** We prospectively measured HR-QoL (using the EQ-5D and a disease-specific questionnaire for peritonitis) and symptoms of PTSD (using the Post-Traumatic Stress Syndrome 10-Questions Inventory, PTSS-10 and Impact of Event Scale-Revised, IES-R) at 12 months following the initial surgery.

**Results:** 107 patients (79%) out of 135 eligible survivors responded. Elevated levels of PTSD symptoms were present in 41 (38%) of the patients, as they scored above the cutoff of 35 for the PTSS-10 or 24 for the IES-R. 'Moderate to high-scoring' PTSD patients (n=41) reported poorer HR-QoL (mean EQ-VAS 64.8, $p<0.001$) than 'none to low-scoring' patients (mean EQ-VAS 75.3). Moderate to high-scoring PTSD patients were approximately 3 times more likely to report problems with mobility (OR=3.5, 95% CI 1.2 to 10.5) daily activities (OR=6.2, 95% CI 2.1 to 18.3), pain (OR=2.8, 95% CI 1.1 to 7.0) and mood (OR=6.1, 95% CI 2.2 to 17.0). On the disease-specific instrument, SP-QoL, moderate to high scoring PTSD patients reported significantly more abdominal problems, worries, body image and sexuality problems, as well as problems specifically related to still having an enterostomy ($p<0.01$). Differences between the two PTSD symptom score groups remained apparent even after adjusting for differences in patient and postoperative characteristics, including age, gender, length of ICU stay and disease-related morbidity.

**Conclusion:** Abdominal sepsis is a severe condition that may trigger PTSD symptoms in a considerable portion of patients as well as reduce their HR-QoL. PTSD is related to both the mental health and physical HR-QoL dimensions, which is not be explained by differences in clinical severity. These results indicate the need for interventions targeting post-traumatic stress after critical illness and emergency surgery.
Introduction

Severe peritonitis (or abdominal sepsis) is an acute medical condition that can be acquired both nosocomial and in the community and is associated with a considerable mortality rate\(^1\,^2\,^3\). The primary treatment is an emergency operation to eliminate the source of infection\(^4\,^5\). These patients often need lengthy Intensive Care Unit (ICU) stays, repeated surgical procedures and long recovery times, exposing them to considerable stressors during both their initial illness and during their prolonged recovery. This combination of an acute, as well as, life-threatening event renders these patients more susceptible to developing post-traumatic stress disorder (PTSD) symptoms\(^6\). In two recent studies, we found a high prevalence of PTSD symptoms in patients with secondary peritonitis both after one year and more than 10 years after the initial illness\(^7\,^8\). Increased levels of PTSD symptoms have also been reported by patients admitted to the ICU in general and specifically by patients with Acute Respiratory Distress Syndrome (ARDS), patients undergoing cardiac surgery and patients with sepsis\(^9\,\,^10\,\,^11\).

PTSD symptoms have been shown to play an intricate role in health related-quality of life (HR-QoL), including reduced social and emotional functioning\(^12\,\,^13\,\,^14\). Patients with PTSD symptoms are known to report high levels of depressive symptoms\(^16\,\,^19\). PTSD symptoms may even negatively influence patients’ physical recovery\(^6\,\,^20\). Earlier ARDS, oncology and transplant studies described that patients reporting PTSD symptoms also report major impairments in some dimensions of HR-QoL, whereas patients without PTSD symptoms had HR-QoL scores that were in the range of the general population\(^21\,\,^22\,\,^23\,\,^24\). Although the relationship between PTSD and QoL has been studied for several patient groups, presently no data exists on how PTSD symptoms and HR-QoL relate together after abdominal sepsis. Therefore the aim of this paper is to examine this relationship in patients 12-months following their initial emergency laparotomy for abdominal sepsis.

Methods

Study Design

Our study was embedded in an ongoing peritonitis (RELAP) trial evaluating two surgical strategies for peritonitis. Patients were enrolled and followed between December 2001 and August 2006 in 2 academic and 7 regional teaching hospitals in The Netherlands.

Patients

Patients were eligible if they had a clinical diagnosis of secondary peritonitis, had an APACHE II score >10 and were between 18 and 80 years old. Further details of the RELAP trial can be found elsewhere\(^25\).
Instruments

We studied PTSD and HR-QoL in this trial population 12 months after initial emergency laparotomy. To measure HR-QoL, we used two generic self-report questionnaires, the Euroqol-Five Dimensions (EQ-5D) in combination with the Euroqol-Visual Analogue Scale (EQ-VAS). These questionnaires were recently recommended as the instrument of choice in critical care studies. The EQ-5D has response ranges from 0 (no problems) 1 (some problems), 2 (many problems), whereas the EQ-VAS score ranges from 0 (worst overall HR-QoL) to 100 (best HR-QoL). We also used a disease-specific HR-QoL questionnaire specifically designed for patients with secondary peritonitis, the SP-QoL. This questionnaire includes disease specific subscales for abdomen, nutrition, worries, body image, sexuality, defecation and stoma-related problems, and has been described in detail elsewhere. Each SP-QoL subscale has been linearly converted to scores that range from 0 (best HR-QoL) to 100 (worst HR-QoL).

To determine PTSD symptoms two standardized questionnaires were administered: the Impact of Event Scale-Revised (IES-R), the golden standard questionnaire for symptoms of PTSD, as well as the Post-traumatic Stress Syndrome 10-Questions Inventory (PTSS-10) which is particularly suited for ICU patients. PTSD symptoms were considered present if the cumulative score on the PTSS-10 (range 10 to 70) was 35 or above and/or if the total score on the IES-R (range 0 to 66) was above 24, as presented in an earlier study.

Data collection

Preoperative characteristics, peri-operative findings, and postoperative data on morbidity and the course of the disease were prospectively collected for all included patients. All questionnaires were distributed by mail to patients who survived at least 12 months, with a reminder by phone after two weeks if there was no response. After 1 month of non-response a new questionnaire with a reminder letter was sent.

Statistical analysis

Patients who scored above the cut-off scores for the IES-R (≥24) and/or the PTSS-10 (≥35) were considered ‘moderate- to high-scoring patients’, indicating a probable increased level of PTSD symptoms. Patients who scored below these cut-off values were considered ‘none to low-scoring patients’, indicating no increased level of PTSD symptoms.

Baseline characteristics between the two groups (‘moderate- to high-scoring patients’ vs. ‘none to low-scoring patients’) were compared using \( \chi^2 \), Student t test or Mann-Whitney U test. Logistic regression was used to model whether or not patients reported problems on the various dimensions of the EQ-5D. Linear regression was used to examine the EQ-VAS and SP-QoL scores on the various subscales between the two groups.

These analyses were initially carried out in a univariate way by only including the variable indicating whether patients were ‘moderate- to high-scoring’ or ‘none to low-scoring’ on PTSD symptoms. In the next step, a multivariate model was made. In this model we adjusted for characteristics that have been
mentioned in the literature for their possible association with either PTSD symptoms or HR-QoL or both \cite{7,9,10,15}. These factors included age, gender, comorbidity, severity of disease measured by the Acute Physiology and Chronic Health Evaluation (APACHE II) and length of ICU stay. Odds ratio with 95% confidence intervals (95% CI) were used to express the strength of association in case of binary outcomes, whereas absolute differences in scores were used when modeling continuous outcomes.

Results

Questionnaire response

A total of 107 (82%) out of 132 eligible patients responded to the questionnaires. Reasons reported by patients for not reporting PTSD questionnaires included: being severely ill (n=5); mentally handicapped (n=4), refusal (n=11); and unknown (n=4).

Forty-one (38%) patients reported scores above the cut-off value on one (n=30) or both (n=11) questionnaires; these two groups have been combined and have been labeled ‘moderate- to high-scoring’ PTSD patients. Sixty-six (62%) patients scored below the cut-off value on both the IES-R and the PTSS-10 and therefore scored few PTSD symptoms; this group is referred to as ‘none to low-scoring’ patients (Table 1).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Baseline and clinical follow-up characteristics of patients scoring ‘none to low’ and ‘moderate to high’ symptoms of PTSD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients scoring</td>
</tr>
<tr>
<td></td>
<td>‘none to low’ PTSD</td>
</tr>
<tr>
<td></td>
<td>symptoms N = 67</td>
</tr>
<tr>
<td>Age mean (sd)</td>
<td>66.7 (12)</td>
</tr>
<tr>
<td>Male n (%)</td>
<td>35 (52%)</td>
</tr>
<tr>
<td>Comorbidity present n (%)</td>
<td>37 (55%)</td>
</tr>
<tr>
<td>APACHE II mean (sd)</td>
<td>15.2 (3.4)</td>
</tr>
<tr>
<td>Hospital stay median (IQR)</td>
<td>27 (18 to 47)</td>
</tr>
<tr>
<td>Number of patients admitted to ICU n (%)</td>
<td>57 (85%)</td>
</tr>
<tr>
<td>ICU stay median (IQR)</td>
<td>7 (4 to 12)</td>
</tr>
<tr>
<td>Readmitted n (%)</td>
<td>36 (54)</td>
</tr>
<tr>
<td>Disease-related major morbidity at 6 months n (%)</td>
<td>24 (36)</td>
</tr>
</tbody>
</table>

Factors indicated in **bold** are significant at p<0.1 and were added to the multivariate model.

*Mann Whitney U test was used due to non-parametric data
Baseline characteristics

Baseline characteristics, peritonitis details and postoperative factors differed between the two PTSD symptom score groups, with the ‘moderate- to high-scoring’ group having a lower age, lower APACHE II scores and longer ICU stays (Table 1). There were no other differences found between the ‘moderate- to high-scoring’ patients and ‘none to low-scoring’ patients.

In Tables 2 and 3, differences on the EQ-5D, EQ-VAS and SP-QoL are presented for the univariate analyses as well as the multivariate analyses. We show results from the multivariate analysis, which includes patient characteristics from the literature (age, gender, APACHE II score and comorbidity) and aspects from the post-operative course (represented by length of ICU stay, Table 1).

EQ-5D: In the multivariate model, patients with moderate to high PTSD symptoms reported significantly more problems than the patients reporting none to low PTSD symptoms for the following EQ-5D domains: mobility (OR = 3.5, 95% CI 1.2 to 10.5); daily activities (OR = 6.2, 95% CI 2.1 to 18.3); pain (OR = 2.8, 95% CI 1.1 to 7.0) and mood problems (OR = 6.1, 95% CI 2.2 to 17.0; Table 2, Figure 1).

### Table 2: Univariate and multivariate analyses of problems on each of the five dimensions of the EQ-5D for patients scoring ‘none to low’ and ‘moderate to high’ on PTSD symptoms

<table>
<thead>
<tr>
<th>EQ-5D</th>
<th>n</th>
<th>Patients scoring 'none to low' PTSD symptoms N = 67</th>
<th>Patients scoring 'medium to high' PTSD symptoms n = 41</th>
<th>Unadjusted Odds Ratio (95% CI)</th>
<th>Adjusted Odds Ratio from Multivariate model * (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>105</td>
<td>44%</td>
<td>56%</td>
<td>1.7 (0.74 to 3.7)</td>
<td>3.5 (1.2 to 10.5)</td>
</tr>
<tr>
<td>Self-care</td>
<td>106</td>
<td>16%</td>
<td>31%</td>
<td>2.3 (0.89 to 5.8)</td>
<td>2.4 (0.83 to 7.1)</td>
</tr>
<tr>
<td>Daily activities</td>
<td>106</td>
<td>42%</td>
<td>74%</td>
<td>4.0 (1.7 to 9.6)</td>
<td>6.2 (2.1 to 18.3)</td>
</tr>
<tr>
<td>Pain</td>
<td>105</td>
<td>40%</td>
<td>66%</td>
<td>2.9 (1.2 to 6.5)</td>
<td>2.8 (1.1 to 7.0)</td>
</tr>
<tr>
<td>Mood</td>
<td>106</td>
<td>16%</td>
<td>62%</td>
<td>8.1 (3.3 to 20.3)</td>
<td>6.1 (2.2 to 17.0)</td>
</tr>
</tbody>
</table>

*Adjusted mean differences controlled in multivariate models for age, gender, APACHE II score and comorbidity and length of ICU stay.
EQ-VAS: The ‘moderate- to high-scoring’ patients had a mean VAS score that was 10.4 (95% CI 3.9 to 17.0) points lower than the ‘none to low-scoring’ patients (Table 3) in the multivariate model. The ‘none to low-scoring’ patients scored mean VAS scores of 75.3 (SD 14.9), which is comparable to that of a healthy Dutch general population sample (aged 60 to 69) who reported a mean EQ-VAS score of 79.7 (15.9)32.

SP-QoL: In the multivariate analysis ‘moderate to high scoring’ patients reported significantly more abdomen-related problems (difference in means= 12.2, 95% CI 3.8 to 20.6), more overall worries (difference in means=16.0, 95% CI 7.9 to 24.0), more problems related to body image (difference in means=26.0, 95% CI 15.0 to 37.0), more problems related to sexuality (mean difference=38.1 95% CI 21.5 to 54.6) and more enterostomy-related problems (difference in means=21.2, 95% CI 2.0 to 40.4) than the ‘none to low-scoring’ patients (Table 3).

<table>
<thead>
<tr>
<th>Outcome</th>
<th>n Patients scoring 'none to low' PTSD symptoms</th>
<th>n Patients scoring 'medium to high' PTSD symptoms</th>
<th>Unadjusted mean difference (95% CI)</th>
<th>Adjusted mean differences from the multivariate model * (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ-VAS</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAS score+</td>
<td>103</td>
<td>75.3 (14.9)</td>
<td>64.8 (14.4)</td>
<td>10.4 (4.6 to 16.5)</td>
</tr>
<tr>
<td>SP-QoL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdomen</td>
<td>102</td>
<td>12.2 (15.6)</td>
<td>23.0 (24.1)</td>
<td>10.8 (3.0 to 18.6)</td>
</tr>
<tr>
<td>Nutrition</td>
<td>105</td>
<td>7.8 (13.1)</td>
<td>12.8 (16.9)</td>
<td>5.0 (-0.82 to 10.9)</td>
</tr>
<tr>
<td>Worries</td>
<td>103</td>
<td>19.3 (16.7)</td>
<td>36.6 (21.8)</td>
<td>17.4 (9.8 to 24.9)</td>
</tr>
<tr>
<td>Body Image</td>
<td>105</td>
<td>21.2 (24.0)</td>
<td>52.5 (27.9)</td>
<td>31.3 (21.1 to 41.5)</td>
</tr>
<tr>
<td>Sexuality</td>
<td>78</td>
<td>23.2 (32.3)</td>
<td>56.8 (35.9)</td>
<td>33.6 (18.1 to 49.1)</td>
</tr>
<tr>
<td>Defecation</td>
<td>65</td>
<td>12.7 (11.4)</td>
<td>16.6 (12.7)</td>
<td>3.9 (-2.3 to 10.0)</td>
</tr>
<tr>
<td>Stoma-related problems</td>
<td>37</td>
<td>23.8 (19.6)</td>
<td>47.6 (26.9)</td>
<td>23.8 (8.3 to 39.3)</td>
</tr>
</tbody>
</table>

* Adjusted mean difference controlled in multivariate models for age, gender, APACHE II score, comorbidity, and length of ICU stay
+ The Dutch reference population (aged 60 to 69) has a mean EQ-VAS score of 79.7 (±15.9) 32
Discussion

Among patients surviving abdominal sepsis, those with moderate to high PTSD symptom scores reported significantly worse HR-QoL scores than ‘none to low-scoring’ patients. These worse scores were found on all domains of the EQ-5D, except self-care. These data suggest that survivors developing PTSD symptoms have both significant functional limitations and psychological difficulties. Patients without elevated levels of PTSD symptoms symptomology reported scores on the EQ-VAS comparable to that of the Dutch general population.

The link between PTSD and depression is well documented and it was therefore to be expected that patients with increased PTSD symptoms reported more problems on the mood dimension of the EQ-5D and poorer scores on the worries, body image and sexuality subscales of the SP-QoL. It was, however, unexpected to find such marked differences between the two groups for daily activities and pain on the EQ-5D. Patients with ‘moderate to high’ PTSD symptoms were three times more likely to report pain problems and limitations in daily activities than patients with ‘none to low’ PTSD symptoms. Large differences were also found on the abdomen and stoma-related problems subscales. These findings indicate that the physical aspects of patient HR-QoL and PTSD symptoms are intrinsically related to each other. One explanation may be that in PTSD patients many psychobiological systems that may be involved in pain perception and physical morbidity are deregulated. Similarly, in earlier studies cancer patients were found to suffer from poorer HR-QoL when having PTSD symptoms, also in the physical domains. In contrast, Deja and colleagues found differences between patients with high scores and those with low scores on PTSD symptomology with respect to mental health, but not on physical dimensions in critically ill patients.
Other factors, such as age, gender, length of ICU stay, associated with either PTSD or HR-QoL or both did not explain for the large differences found between these groups. In particular for gender this is surprising considering the well-documented higher PTSD rates for women. Specifically, we expected that the dissimilarity between the two groups in age and in length of ICU stay, which are both independent predictors of PTSD and HR-QoL, would at least in part account for differences in HR-QoL. The results from the adjusted multivariate models indicated that this was the case.

The use of two questionnaires to define PTSD symptoms is contentious. However, as we have discussed in detail in our earlier paper, we feel this allows for more information pertaining to the PTSD symptoms and allows for better screening of patients with potential PTSD problems. There is general consensus that more information is urgently needed about the relationship between PTSD questionnaires and the structured clinical interview for diagnosis (SCID) to determine questionnaires clinical and diagnostic value. Although some studies have reviewed the diagnostic value of the questionnaires, in general these studies had small sample sizes and were methodologically limited.

PTSD symptoms are a common sequel after recovery from critical illness. The long-term consequences of critical illness are not only detrimental for the individual, but also have cost implications for society, including diminished HR-QoL causing reduced ability to return to work and enjoy recreational activities. Mental disorders are also important determinants of work role disability and quality of life, often outnumbering the impact of common chronic physical disorders. Treating PTSD, therefore, may improve HR-QoL as well as reverse psychobiological abnormalities.

Unfortunately, at this stage it is not possible to determine whether HR-QoL affects PTSD or the other way around. The present study measured PTSD at only one time point. Longitudinal measurement of PTSD symptoms, for example at 3-, 6-, 9- and 12-months follow-up, as done with HR-QoL in this study might have been helpful in determining the direction of the relationship between the PTSD symptoms and HR-QoL and been more elucidating to the causal relationship between HR-QoL and PTSD symptoms. Also, in this patient group, due the acuteness of the illness it was and is not possible to obtain either HR-QoL information or information related to PTSD symptoms prior to onset of the disease. It is clear however that it is imperative to study PTSD symptoms together with HR-QoL in critically ill, surgery, cancer and ICU patients.

Although we do not know what the causal relationship is between PTSD and HR-QoL, treating PTSD symptoms could reduce HR-QoL problems. Therefore, future implications for improving treatment of survivors of abdominal sepsis could include identification of high-risk patients who could benefit from further evaluation, psychological intervention and support. Long-term follow-up at outpatient clinics needs to be re-shaped to address these long-term effects, including HR-QoL and PTSD.

In conclusion, these results suggest that severe peritonitis is a trauma similar to other more accepted definitions of trauma that can trigger the development of PTSD symptoms. PTSD symptoms occur frequently in this patient group and have a substantial negative impact on HR-QoL, both psychological and physical components. Differences found in PTSD symptomology and HR-QoL cannot be explained by differences in severity of abdominal sepsis. It is therefore important to be aware of these symptoms in all patients with abdominal sepsis, as they deserve early attention.
References

Chapter 12
General discussion
Summary
Nederlandse samenvatting

“Say not, ‘I found the truth’, but rather, ‘I have found a truth.’”

Kahlil Gibran (The Prophet)
General discussion

The debate about which surgical strategy planned or on-demand relaparotomy, to use when treating patients with secondary peritonitis following emergency laparotomy dates as far back as the call for a randomized trial comparing both strategies by the European Surgical Infection Society in 1990.1,2. Although several institutions have completed retrospective cohort studies reviewing this issue, until now there has not been a successful randomized controlled trial evaluating these surgical strategies3-6. The RELAP trial presented in this dissertation aimed to end - at least in part - this debate, by comparing the on-demand strategy to the planned relaparotomy surgical strategy in a multicenter randomized controlled trial.

Based on the evidence from the RELAP trial, we have seen that patients who received the on-demand relaparotomy strategy did not have a significantly lower rate of death or peritonitis-related morbidity but had substantial reductions in the number of relaparotomies, healthcare utilization and societal costs compared to patients treated with the planned relaparotomy strategy7. This was true regardless of the initial severity of disease. Hence the on-demand strategy has been shown to be the simpler strategy with significant reduction in healthcare utilization and costs. The RELAP trial was necessary to develop safe, evidence-based and cost-effective surgical practice8,9.

The primary strength of a randomized control trial is its rigorous design, where the risk of random or systematic (bias) error is minimized and thus the risk of making an incorrect conclusion about the efficacy of a treatment is also minimized10,11. During the RELAP trial several challenges arose specific to surgical trials that made this process more difficult7,10,12. Designing and conducting the RELAP trial to evaluate these surgical strategies included careful planning and strict protocolization of study procedures and trial endpoint definitions, but also included some compromises and creative design solutions13.

Methodological or logistical issues included, to name but a few, difficulties in patient sampling and difficulties in patient accrual. Issues related to surgical expertise and attitudes (patient and surgeon acceptance of surgical randomization) and standardization of the surgical procedure (variability in surgical proficiency/technique) also arose7. Issues such as double-blinding and a placebo-controlled trial, which often arise in surgical trials, were not problems in the RELAP study, as we presented a randomized study comparing surgical strategies and not surgical techniques11,14. In the case of comparison of treatment strategies the primary interest lies in the effect of the strategies as a whole15.

Health related-quality of life (HR-QoL) can also be an important factor in clinical decision-making. In the RELAP trial we collected randomized data from both generic and disease-specific HR-QoL questionnaires. We showed that HR-QoL did not vary according to surgical treatment strategy. In both surgical treatment arms, secondary peritonitis patients continued to report poor HR-QoL even 12 months following emergency surgery.
A limitation to this dissertation is that we have presented several statistically significant differences both between HR-QoL, surgical strategies and over time. However, in HR-QoL research it presently remains unclear when and to what extent differences in HR-QoL are clinically relevant, and how these differ from statistical differences. The term “effect size”, referring to clinical relevant difference, was coined by Cohen in 1988\(^{16}\), and has been discussed extensively in the literature\(^ {17}\). As yet there is no conclusion on how best to examine clinical relevance in HR-QoL in surgical patients (for example, minimal important differences, effect sizes, standard error of measurement or anchor-based estimates), and when this should affect clinical decision-making and patient benefit\(^ {18-22}\). In clinical trials, one concentrates on the relative differences between treatment arms; but no clear standard has been set on what is a clinically relevant improvement. Questions for the future of HTA studies should be: at what point is clinical decision-making made based on HR-QoL data, and is this quantifiable?\(^ {22}\)

From the RELAP trial, we have seen that strategy type does not determine differences in HR-QoL, but that at 6- and 12-months follow-up patients continue to report poor HR-QoL. Therefore we have subsequently considered which factors do affect HR-QoL. The factors we found to be most associated with poor HR-QoL at 6 months following initial surgery were gender, age, comorbidity at study entry, length of hospital stay and severe disease-related morbidity. As none of these factors are modifiable or adjustable our options to use this information to improve HR-QoL have been limited. Nevertheless, these results remain informative, as HR-QoL data can play an important role in better understanding the burden of disease from the patient’s perspective\(^ {23-25}\). Following the acute onset and high risk of mortality and patient survival, understanding the patients' perspectives during their lengthy recovery periods can help improve communication between doctors and patients, thereby enabling patients to adjust their expectations accordingly and better preparing them for their recovery process.

In this study we have presented a great deal of information pertaining to HR-QoL in patients with secondary peritonitis, using both generic and disease-specific tools to complement each other. Generic questionnaires are helpful in comparing different patient groups across different diseases, for example abdominal sepsis patients versus other ICU-admitted patients. Disease-specific questionnaires are more useful for determining detailed HR-QoL issues, specific to patients with a particular disease. These questionnaires should be more sensitive to changes within the same patient group, both for comparative trials and for evaluating HR-QoL improvement over time\(^ {26,27}\).

Based on the results from the generic Euroqol pain dimension, which did not show improvement during the 12-months follow-up, we recommend that the disease-specific secondary peritonitis questionnaire (SP-QoL) be adjusted to include a more detailed pain subscale for future use. This pain subscale should be similar to the ones in other European Organization of Research and Treatment of Cancer (EORTC) questionnaires (www.eortc.org).

An issue that we have not yet explored in this study group is one that many researchers are presently concentrating on: The relationship between patients’ HR-QoL and their survival. Several studies have shown that HR-QoL is a strong predictor of survival, primarily in cancer patients, with patient survival
improving positively correlated with better HR-QoL. In more acute situations this is more difficult to
determine and may have less implications for survival. Future studies should focus on whether in
the acute surgical field HR-QoL is a contributor to patient survival, and determine the potential clinical
consequences of such a finding.

Health technology assessment (HTA) is an important factor in aiding decision-making in healthcare,
therefore the RELAP trial also included healthcare cost estimates from a societal perspective. The
total healthcare costs associated with these surgical strategies varied with different assumptions that
were considered using sensitivity analyses, but the relative differences between the two strategies
remained stable. With approximately 21% less costs generated by the on-demand relaparotomy
strategy, this indicates that choosing and optimizing the on-demand strategy could reduce future costs
significantly.

The approximate societal costs (direct and indirect) of general sepsis patients ranges from 3.6 billion
to 7.7 billion euros yearly in Germany only, with ICU costs responsible for the majority of these costs. Estimates indicate that sepsis patients' ICU stays account for 25-33% of all ICU costs in Western Europe. In the RELAP trial, the most important component of the total direct healthcare costs was also ICU stay (often involving mechanical ventilation). These figures suggest that a strategy that
succeeds in reducing healthcare utilization, and ICU stay in particular, could achieve a substantial cost
reduction at this level.

Therefore, the next challenge in treating patients with secondary peritonitis is to optimize the on-
demand surgical strategy as a means of reducing disease-related mortality. For some patients, clinical
deterioration based on abdominal pathology is a clear indication that a relaparotomy is necessary. For
other patients clinical improvement is apparent, guaranteeing that re-intervention is not needed. However, for patients that fall into neither of these categories, where the patient's situation is not clear
cut, the decision-making process to determine if and when a surgeon should re-intervene is more
difficult. Re-operating on patients in this group unnecessarily may lead to avoidable clinical
deterioration, whereas not re-intervening in patients, who need it in a timely fashion, may mean
missing the opportunity to get the infection under control (infection source control). Presently no standard criteria are available regarding if and when to perform a relaparotomy; this is
often left to the surgeon's discretion. In chapter 5 of this manuscript, we have seen that none of
the commonly used existing ICU scoring systems can adequately support surgical decision-making. In
a retrospective cohort of secondary peritonitis patients, the difficulties in determining which variables
will best support surgeons in their decision-making were discussed at length.

Future studies therefore should concentrate on better monitoring of patients and creating better
decision-making tools for the selection of patients for a timely relaparotomy. Sequential longitudinal
scoring systems should be better examined for their usefulness in deciding when to reoperate. Other
indicators, such as cultures, patient characteristics, disease-related factors and postoperative
characteristics should also be prospectively examined. Finally, the availability of diagnostic tests has
strengthened the effectiveness of the on-demand strategy. CT scanning, for instance, can be a useful diagnostic tool for selecting patients with ongoing abdominal infection, although the diagnostic value of these tests should also be further evaluated for more timely and accurate patient selection.

The primary concern of the surgeon for abdominal sepsis patients is survival, with continued interest in disease-related morbidity and health-related quality of life. Recently, several psychological aspects of recovery have also become increasingly more important for surgeons and other caregivers in these acutely ill patients.

To put this in perspective, approximately 90% of all people will undergo a (potentially) traumatic event fulfilling A1 criteria as defined in the Diagnostic and Statistical Manual of Mental Disorders, edition IV (DSM-IV) during their lifetime. Whilst most people will undergo a natural recovery, a proportion of these people will report post-traumatic stress disorder (PTSD) symptoms following their traumatic event. In patients with critical illness, represented by heterogeneous groups of patients admitted to the ICU, the proportion of patients reporting PTSD symptoms ranged from 11 to 64%. Also, ICU patients report PTSD symptoms that can include intrusive recollections (re-experiencing the trauma in flashbacks, memories, or nightmares); avoidant and numbing symptoms (including diminished emotions and avoidance of situations that are reminders of the traumatic event); and hyperarousal (including increased irritability, exaggerated startle reactions, or difficulty sleeping or concentrating). Secondary peritonitis patients report PTSD symptoms in similar numbers to those of other ICU patients, ranging from 24 to 38% (as seen in chapters 9 and 10). In fact, secondary peritonitis patients may be particularly vulnerable for developing PTSD symptoms due to the acuteness and severity of the disease, which is frequently followed by a lengthy ICU stay and a long recovery period that often includes multiple surgical and non-surgical interventions. PTSD symptoms in secondary peritonitis patients have shown to be highly correlated with both depression symptoms and a reduction in HR-QoL. This does not only have consequences for the individual but also has implications for society due to reduced ability to return to work and a continuing financial burden for the healthcare system.

Timing plays an important issue in collecting data on PTSD symptoms in critically ill and ICU patients. In most critically ill patients discharged from the ICU, a natural recovery time is to be expected. Only a select group of patients will not undergo this natural recovery with PTSD symptoms developing, often with delayed onset. Critically ill patients often only develop PTSD symptoms after their physical recovery period has passed. In a retrospective study, we saw that patients ranging from 4 to 10 years following illness secondary peritonitis reported PTSD symptoms. In our prospective study, presented in chapter 10, we chose a 12-months the time period following secondary peritonitis to record PTSD symptoms for a very specific reason. In this severely ill patient group, we did not want to record patient recovery, nor ignore the necessity to allow patients a natural recovery time.

In diagnosing PTSD, the gold standard is considered the structured clinical interview. This is, however, not always feasible, as the ICU does not lend itself to the use of extensive clinical interviews. In general in the literature, we have come across a gap in knowledge concerning the diagnostic value of the questionnaires in comparison to DSM-IV diagnosis based on a structured clinical interview.
studies have tried to determine the diagnostic value of these questionnaires in critically ill patients, but
to date, these studies have been methodologically limited (mainly due to small sample sizes). In this dissertation, we therefore measured PTSD symptoms employing the most commonly used and validated PTSD questionnaires in critically ill patients. But our studies lacked a structured interview to determine a differential DSM-IV diagnosis. By determining PTSD symptomology by questionnaires, we hoped to be able to extrapolate the number of patients who would be diagnosed with PTSD according to the DSM-IV in a structured clinical interview. We feel that future studies should concentrate on better diagnostic tools in the ICU, specifically on the diagnostic value of the questionnaire in relation to the gold standard clinical interview taking into account feasibility issues presented by the ICU.

In chapter 10, we concentrated on identifying factors associated with the development of PTSD symptomology and developed a prognostic model to help identify patients with a high risk of PTSD. This is important because it avoids burdening all patients with in-depth and distressing psychological assessment interviews. Therefore providing a prognostic model to enhance the ability of the surgeon to make pre-selection of high-risk patients is important, although this model cannot be accepted into clinical practice before externally validated.

From this prognostic model, we have seen that the strongest predictor of PTSD symptoms is the actual experience of the patient whilst in the ICU, confirming the results of many earlier studies. Nightmares, panic, fear and suffocation are among only a few of the reported memories of patients during ICU stay. Therefore, new research should include prospective collection of traumatic memories of the ICU as, clearly, to prevent PTSD symptoms we need to have a better idea of the actual ICU experience of the patient. Collecting these data from patients whilst still admitted to the ICU can be troublesome, as asking patients about their experience whilst they are too ill to comprehend or not yet ready to discuss these issues may actually increase PTSD symptoms. In a recently published study on PTSD we have seen that early debriefing of patients on the subject of PTSD can run the risk of actually increasing symptoms and a diagnosis of PTSD. Therefore, we suggest for future studies, that ICU diaries, that are maintained by both the patient's contact person and/or the nursing staff during the patients ICU stay, be used to obtain these data. In addition, it has been argued that delirium is also associated with the development of PTSD symptoms. This has not yet been substantiated, therefore information pertaining to delirium during the ICU stay should be collected.

The primary focus of new research should be reducing PTSD symptoms in patients following critical illness; however, there are different schools of thought on how best to do this. Some argue that these measures should include medication following ICU stay or administration of hydrocortisone therapy during ICU stay or new sedation methods (daily sedation interruption). Others argue that less invasive measures should be used, such as the use of the ICU diary or debriefing. Communication and support to patients both during ICU stay and following ICU stay could also show reductions in all aspects of patient-related outcomes. Pattison argues that facilitating communication, explaining care and rationalizing interventions, ensuring patients are oriented as to time and place, reassuring patients about transfer, providing patients, where possible, with information about critical care before admission.
are practices that may have beneficial effect on patient care\textsuperscript{55-73,84}. Many researchers hypothesize that the ICU should offer follow-up services, which could potentially help patients come to terms with their experiences of critical illness and provide the opportunity for them to access further intervention if desired\textsuperscript{55-73,84}. Evidenced-based information booklets informing patients and relatives about patient transfer from ICU to general ward, have been shown to enhance verbal communication\textsuperscript{85,86}. But there are no RCT evidence-based data to support whether post-ICU programs or information booklets are also effective methods in reducing PTSD symptoms\textsuperscript{56}.

Although many researchers feel that increased social support may help in reducing PTSD symptoms in patients, the problem may be compounded by the fact that a substantial percentage of the relatives of ICU patients also report increased PTSD symptoms\textsuperscript{87-93}. In itself this may reduce the patients’ social support network as some relatives even report more PTSD symptoms than the patients themselves\textsuperscript{93}. Therefore future prevention methods should not only concentrate on improving the patients’ ICU experiences and the long-term psychosocial effects, but should also concentrate on reducing PTSD symptoms in relatives.

Evidence from randomized data on the prevention and/or treatment of PTSD following abdominal sepsis and ICU stay are sorely needed\textsuperscript{59,65}, as it is still unclear at this point which prevention and treatment options would be the most effective. It may even be possible that these treatments are not mutually exclusive, but that a combination of methods would be most effective.
Summary

In the general introduction, presented in chapter 1, we discuss current concepts in severe peritonitis, such as prevalence, diagnosis, treatment, as well as the overall burden to the patient and societal costs. We also introduce evidence-based medicine and health technology assessment in surgical trials and in particular the RELAP trial comparing surgical strategies for the treatment of patients with severe peritonitis.

Part I: In search of an evidence-based surgical approach

In chapter 2 of this dissertation we reported the results of the first randomized controlled trial comparing two commonly applied surgical treatment strategies in patients with abdominal sepsis after their initial emergency laparotomy: planned relaparotomy and relaparotomy on-demand. This trial revealed that mortality and morbidity were comparable between the two strategies, but that the on-demand relaparotomy strategy showed a substantial reduction in relaparotomies, healthcare utilization and medical costs. The on-demand patients had shorter median ICU stays (7 vs. 11 days; p=0.001), shorter median hospital stays (27 vs. 35 days; p=0.008) and less direct medical costs (per patient costs were reduced by 23% using the on-demand strategy, also indicating a less complicated course of disease).

In chapter 3 no differences between the two surgical strategies were found in health related-quality of life (HR-QoL) on either the generic or disease-specific questionnaires. Therefore in this trial we have demonstrated that HR-QoL should not play an important role in decision-making between the two surgical strategies. Over time HR-QoL improved only moderately for the generic EQ-VAS and some dimensions of the EQ-5D and disease-specific SP-QoL questionnaire. Symptom reductions in the self-care and daily activities dimensions were most prominent, whilst changes in the pain and mood dimensions were minimal. In line with clinical improvement and recovery, more changes over time were expected in HR-QoL, whilst several dimensions showed no improvement at all over a 12-months follow-up including the sexuality, body image and mood subscales.

In chapter 4 we reported a detailed account of the economic evaluation from the RELAP trial. Across the whole study population consisting of patients with different clinical conditions, resource utilization and associated costs generated by treatment and follow-up of secondary peritonitis are substantially lower for an on-demand strategy than for a planned relaparotomy strategy. For the whole study period, the mean total costs per patient associated with the on-demand relaparotomy strategy were €65,768 versus €83,450 with the planned relaparotomy strategy (absolute difference €17,682; 95% CI €5,062 to €29,004; relative difference 21%). Of these total costs, 75% were generated during the index admission, of which 45 to 55% were ICU costs. As the on-demand relaparotomy strategy leads to substantially lower costs than the planned relaparotomy strategy without compromising patient outcomes, we therefore conclude that the on-demand strategy should be the preferred surgical strategy.
Part II: Optimizing the on-demand strategy

In chapter 5, we evaluated the usefulness of APACHE II score to predict patient outcome (mortality and complicated course) and to select surgical strategy. The APACHE II score was useful in stratifying patients with abdominal sepsis according to their risk for mortality and morbidity, even in the present clinical setting with improvements in treatment and ICU care. All patients with relatively mild peritonitis (APACHE II score of 10 or lower) were safely treated by the on-demand strategy outside the trial. Furthermore, we found no evidence that disease severity as measured by the APACHE II score was an effect modifier of the treatment in the RELAP trial. This means that in the future both patients with mild peritonitis and patients with severe peritonitis can be safely treated by the on-demand strategy in the future.

In Part I of this dissertation we concluded that the on-demand strategy should be the recommended strategy for the future surgical treatment of secondary peritonitis patients. In order to optimize patient outcome using the on-demand strategy to treat abdominal sepsis it is essential to include timely and adequate selection of patients for relaparotomy after the initial emergency (index) laparotomy. Currently, there are no specific scoring systems aimed at selecting patients that require a relaparotomy. Therefore, in chapter 6 several existing scoring systems, originally designed to predict mortality in critically ill patients, for their predictive value in selecting patients requiring a relaparotomy or reintervention. However, none of these widely available scoring systems was of clinical value in identifying patients with abdominal sepsis who may require a relaparotomy. Therefore, paying careful attention to the dynamic nature of the underlying pathophysiological changes, we see a real need to develop more specific tools that can assist surgeons in their daily, potentially hourly monitoring of these patients after their initial emergency operation.

Part III: Health related-quality of life in patients with secondary peritonitis

In chapter 7, we compared the HR-QoL of patients surgically treated for secondary peritonitis to that of a general population, and prospectively identified factors associated with worse HR-QoL. HR-QoL was significantly poorer on all dimensions in peritonitis patients than in a healthy reference population. Peritonitis characteristics at initial presentation were not associated with poorer HR-QoL at 6-months follow-up. A more complicated course of the disease was most predictive for reduced HR-QoL at 6-months follow-up.

In chapter 8, we presented a newly compiled disease-specific quality of life questionnaire (SP-QoL) for patients with secondary peritonitis to obtain a better, more complete profile of illness-related problems and patient recovery. Questions of the SP-QoL were compiled from several European Organization for Research and Treatment of Cancer (EORTC) modules including the QLQ-CR38, the QLQ-OES18 and the QLQ-PAN26. We examined the psychometric performance of this newly constructed questionnaire.
based on systematic selection and compilation of existing items. The SP-QoL has shown to have high acceptability, to be internally consistent and to be adequately reliable and valid. Therefore with minor adjustments (specifically including a new pain subscale, e.g. from one of the EORTC modules and using individual defecation items instead of a 7-items subscale), we recommend this questionnaire for future use in comparing treatment strategies for patients with secondary peritonitis.

Part IV: Post-traumatic stress disorder in patients with secondary peritonitis

In chapter 9 we presented a retrospective study, initially designed to determine whether abdominal sepsis was a condition that could trigger post-traumatic stress disorder (PTSD) symptoms. This study was also performed to evaluate whether the ICU is an environment which particularly triggers PTSD symptoms in patients more so than the general hospital ward. In this study we have determined that patients with secondary peritonitis admitted to the ICU were at significantly greater risk for having PTSD symptoms, even after adjusting for baseline differences, in particular, age. ICU patients were more likely to report PTSD symptoms than patients not admitted to the ICU. As well, both younger patients as well as females were also more likely to report PTSD symptoms. In total almost a quarter of the surviving patients receiving surgical treatment for secondary peritonitis developed long term PTSD symptoms (4 to 10 years).

In chapter 10 we carried out a prospective analysis of PTSD symptoms in abdominal sepsis patients one year after their emergency operation. This study determined the extent to which patients, who survived abdominal sepsis, suffered from symptoms of PTSD and depression. It also identified potential risk factors for PTSD symptoms. PTSD symptom scores were reported by 38% of patients, with only 5% of these patients reporting severe depression symptoms. Factors associated with increased PTSD symptoms in a multivariate ordinal regression model were similar to those found in the retrospective study; younger age, length of ICU stay and having some or many traumatic memories of the ICU or hospital stay 12 months earlier. A nomogram was constructed that showed good discrimination. This nomogram may assist in identifying patients at increased risk for developing symptoms of PTSD, but needs external validation.

In chapter 11 we examined the interrelationship between symptoms of PTSD and HR-QoL in patients surviving severe peritonitis. We have seen that approximately 12 months following initial emergency surgery, patients with elevated levels of PTSD symptoms reported poorer HR-QoL on both the generic HR-QoL questionnaires (EQ-VAS and EQ-5D), as well as on the disease-specific instrument, SP-QoL. These differences between the two PTSD symptom score groups remained apparent even after adjusting for differences in patient and postoperative characteristics, including age, gender, length of ICU stay and disease-related morbidity. PTSD is related to both the mental health and physical HR-QoL dimensions, which could not be explained by differences in clinical severity.
In chapter 12, we discussed the challenges we faced in designing and running the RELAP trial, and how we resolved and documented these. We also discussed future ideas to optimize the on-demand surgical treatment of secondary peritonitis. Several important factors in determining HR-QoL in surgical patients were discussed, and finally we addressed some methodological issues in determining PTSD symptoms in these patients and possible methods by which PTSD symptoms could be reduced following ICU stay.
Samenvatting

Deel I: Op zoek naar een evidence-based chirurgische behandeling van peritonitis

Na de inleiding in hoofdstuk 1, beschrijven wij in hoofdstuk 2 van dit proefschrift de resultaten van de eerste gerandomiseerde studie (de RELAP-trial), waarin twee veelgebruikte chirurgische behandelstrategieën voor patiënten met abdominale sepsis (ernstige ontsteking van de buikholte) na een spoed-laparotomie met elkaar vergeleken worden: van tevoren geplande (planned) relaparotomie(ën) en een relaparotomie op indicatie (on-demand). Uit het onderzoek blijkt dat de mortaliteit en de morbiditeit niet significant verschillend zijn tussen de beide strategieën. In de groep patiënten waarbij de relaparotomie pas gedaan werd bij klinische verslechtering of uitblijven van verbetering (on-demand) was relaparotomie in slechts 31% van de gevallen niet nodig (t.o.v. 66% bij de geplande strategie). Ook werd er minder gebruik gemaakt van zorgvoorzieningen, en waren de medische kosten 23% lager. Tenslotte verbleven bij de on-demand strategie patiënten korter op de intensive care (7 versus 11 dagen; P=0,001), en korter in het ziekenhuis (27 versus 35 dagen; p=0,008).

In de studie beschreven in hoofdstuk 3 worden er geen verschillen geconstateerd in de gezondheidsgerelateerde kwaliteit van leven (KvL) tussen de twee chirurgische strategieën; niet op basis van de algemene, en niet op basis van de ziektespecifieke vragenlijsten. Daarmee toont deze studie aan dat de kwaliteit van leven geen belangrijke overweging zou moeten zijn bij de keuze tussen de twee strategieën. Op den duur verbeterde de kwaliteit van leven slechts matig voor wat betreft de EQ-VAS vragenlijst en enkele aspecten van de EQ-5D en SP-QoL vragenlijsten. De klachten verminderden het meest op de gebieden ‘zelfzorg’ en ‘dagelijkse activiteiten’; de veranderingen in ‘pijn’ en ‘stemming’ waren echter minimaal. Tevoren was de verwachting dat er een grotere verbetering in de kwaliteit van leven op de langere termijn zou optreden, welke meer in lijn zou zijn met de klinische verbetering en het herstel van de patiënten. Op bepaalde aspecten van de kwaliteit van leven vragenlijsten, zoals seksualiteit, lichaamsbeleving en stemming, werden over een follow-up-periode van 12 maanden helemaal geen verbeteringen geconstateerd.

In hoofdstuk 4 beschrijven we de economische evaluatie van de RELAP-trial in meer detail. Het gebruik van zorgvoorzieningen en de daaraan gerelateerde kosten van behandeling en follow-up van patiënten met secundaire peritonitis was aanzienlijk lager bij de on-demand strategie dan bij de geplande relaparotomie strategie. Over de gehele studieperiode genomen, waren de gemiddelde kosten per on-demand patiënt €65,768 versus €83,450 bij de patiënten met een geplande relaparotomie strategie (absoluut verschil €17,682; 95% BI €5,062 tot €29,004; relatief verschil 21%). Van de totale kosten bestond 75% uit kosten gemaakt tijdens de eerste opname, waarvan 45% tot 55% kosten voortvloeiend uit intensive care verblijf. Omdat de on-demand strategie leidt tot aanzienlijk lagere kosten dan de geplande strategie, zonder nadelige gevolgen voor de patiënt, concluderen we dat de on-demand strategie de chirurgische strategie van voorkeur zou moeten zijn.
In hoofdstuk 5 evalueren we de bruikbaarheid van de APACHE II score om de uitkomst voor de patiënt (mortaliteit en gecompliceerd beloop) te voorspellen en de chirurgische strategie te selecteren. De APACHE II score bleek een nuttig instrument om de patiënten met abdominale sepsis in te delen op basis van hun mortaliteits- en morbiditeitsrisico, zelfs in de huidige klinische setting met verbeteringen van behandeling en intensive care zorg. Alle patiënten met een relatief milde peritonitis (een buikvliesontsteking met een APACHE II score van 10 punten of lager) werden niet ingecludeerd in de RELAP-trial en werden buiten studieverband met de on-demand strategie behandeld. We vonden geen aanwijzingen dat de ernst van de ziekte (mild of juist zeer ernstig), zoals gemeten met de APACHE-score, de uitkomst van de behandeling in de RELAP-trial beïnvloedde. Dit betekent dat zowel patiënten met milde peritonitis als patiënten met zeer ernstige peritonitis in de toekomst veilig behandeld kunnen worden met de on-demand relaparotomie strategie.

In Deel 1 van dit proefschrift concluderen we dat de on-demand strategie de aangewezen strategie blijkt voor de behandeling van patiënten met secundaire peritonitis. Om de uitkomst voor de patiënt bij de behandeling van abdominale sepsis met de on-demand strategie te optimaliseren, is een tijdige en adequate selectie van de patiënten die een relaparotomie moeten ondergaan na de eerste (index) spoedlaparotomie, essentieel. Op dit moment bestaan er geen specifieke scoringssystemen gericht op het selecteren van deze patiënten. Daarom wordt in hoofdstuk 6 onderzocht of bestaande scoringssystemen, oorspronkelijk ontwikkeld om het mortaliteitsrisico van ernstig zieke patiënten te voorspellen, voorspellende waarde hebben voor het selecteren van patiënten die een relaparotomie of reinterventie nodig hebben. Deze scoringssystemen zijn de APACHE II score, de Mannheim Peritonitis Index (MPI), de Sequential Multiple Organ Dysfunction Score (MODS), de Sequential Organ Failure Assessment Score (SOFA) en de Simplified Acute Physiology Score (sAPS). Voor het onderzoeken van de voorspellende waarde maakten wij gebruik van de Area Under the Curves (AUC)-methode. Geen van deze veelgebruikte scoringssystemen bleek echter klinisch geschikt om te identificeren welke patiënten met abdominale sepsis een relaparotomie nodig hebben. Daarom constateren wij dat er een grote behoefte bestaat aan het ontwikkelen van meer specifieke methoden, die – rekening houdend met de dynamische aard van de onderliggende pathofysiologische processen – chirurgen kunnen ondersteunen in hun dagelijkse, of nog frequentere, besluitvorming bij deze patiënten na hun initiële spoedoperatie.

Deel III: Gezondheidsgerelateerde kwaliteit van leven in patiënten met secundaire peritonitis

In hoofdstuk 7 vergelijken we de gezondheidsgerelateerde kwaliteit van leven van patiënten die operatief behandeld waren voor secundaire peritonitis, met die van de algemene bevolking. Ook bepaalden we prospectief factoren die gerelateerd waren aan een slechtere kwaliteit van leven. De kwaliteit van leven was op alle dimensies significant slechter in peritonitis-patiënten dan in een
gezonde controlepopulatie. Kenmerken van de peritonitis bij de initiële presentatie waren niet voorspellend voor een slechtere kwaliteit van leven op 6 maanden. Een meer gecompliceerd beloop van de ziekte was het meest voorspellend voor de verminderde kwaliteit van leven op 6 maanden.

In hoofdstuk 8 presenteren we een nieuw samengestelde ziektespecifieke kwaliteit-van-levenvragenlijst (SP-QoL) voor patiënten met secundaire peritonitis. Deze vragenlijst werd ontwikkeld om een beter, meer compleet overzicht van ziektegerelateerde problemen en herstel van patiënten te krijgen. De SP-QoL werd samengesteld uit een aantal modules van de Europese Organisatie voor Onderzoek en Behandeling van Kanker (EORTC), inclusief de QLQ-CR38, de QLQ-OES18 en de QLQ-PAN26. Wij onderzochten de psychometrische kwaliteiten van deze nieuw samengestelde vragenlijst, die was gebaseerd op een systematische selectie en samenvoeging van bestaande items. De SP-QoL blijkt een hoge aanvaardbaarheid te hebben, heeft een goede interne consistentie, en is voldoende betrouwbaar en valide. Daarom raden we aan deze vragenlijst – zij het met enkele kleine aanpassingen (in het bijzonder het toevoegen van een nieuwe pijn-subschaal, bijvoorbeeld een van de EORTC-modules, en het gebruikmaken van individuele defecatie-items in plaats van een 7-item subschaal) – in de toekomst te gebruiken bij het vergelijken van behandelingsstrategieën voor patiënten met secundaire peritonitis.

Deel IV: Post-traumatische stress stoornis bij patiënten met secundaire peritonitis

In hoofdstuk 9 presenteren we een retrospectieve studie die de hypothese onderzoekt dat abdominale sepsis een dermate ernstige, levensbedreigende aandoening is dat deze ziekte een post-traumatische stress stoornis (PTSS) tot gevolg kan hebben. Daarnaast werd de studie uitgevoerd om de hypothese te testen dat de intensive care een omgeving is die wellicht vaker PTSS-symptomen veroorzaakt dan een normale ziekenhuisomgeving. In deze studie onderzochten we de lange-termijn prevalentie (4-10 jaar) van PTSS-symptomen bij patiënten na een operatieve behandeling voor secundaire peritonitis; in totaal ontwikkelde bijna een kwart van de overlevende patiënten symptomen van PTSS. Patiënten die opgenomen waren op de intensive care, hadden een groter risico op het ontwikkelen van PTSS-symptomen, zelfs na correctie voor verschillen in uitgangsrisico, in het bijzonder de leeftijd. Intensive care patiënten meldden vaker PTSS-symptomen dan patiënten die niet opgenomen waren geweest op de intensive care. Oudere patiënten en mannen hadden minder vaak symptomen van PTSS.

In hoofdstuk 10 voeren we een prospectieve analyse uit van PTSS-symptomen in patiënten met abdominale sepsis een jaar na hun spoedoperatie. Met deze studie bepaalden we de mate waarin patiënten die de abdominale sepsis overleefden, lijden aan symptomen van PTSS en depressie. Ook wilden we potentieel risicofactoren voor het ontwikkelen van PTSS-symptomen identificeren. Door 38% van de patiënten werden PTSS-symptomen gemeld. Slechts 5% (95% BI: 2 to 12) van de patiënten meldden klachten van ernstige depressie. De factoren die in een multivariaat ordinaal regressiemodel geassocieerd bleken met meer PTSS-symptomen, waren gelijk aan de symptomen die we vonden in de
retrospectieve studie: lagere leeftijd, duur van het verblijf op de intensive care, en het hebben van traumatische herinneringen aan het verblijf op de intensive care of in het ziekenhuis. We stelden een nomogram op dat goed in staat bleek te discrimineren tussen de drie categorieën van ernst van PTSS (geen tot mild, matig, ernstig). Dit nomogram zou patiënten kunnen identificeren met een verhoogd risico op het ontwikkelen van PTSS, maar moet nog extern worden ge valideerd.

In hoofdstuk 11 onderzochten we de relatie tussen symptomen van PTSS en kwaliteit van leven in patiënten die de abdominale sepsis (ernstige peritonitis) overleefden. We zagen dat ongeveer 12 maanden na de initiële spoedoperatie, patiënten met meer symptomen van PTSS een slechtere kwaliteit van leven meldden op zowel de algemene gezondheidsgerelateerde kwaliteit-van-leven-vragenlijsten (EQ-VAS en EQ-5D) als de ziektespecifieke vragenlijst (SP-QoL). Deze verschillen tussen de twee PTSS-scoregroepen bleven bestaan, ook na correctie voor verschillen in patiënteneigenschappen (zoals leeftijd en geslacht) en postoperatieve eigenschappen (zoals duur van het verblijf op de intensive care en ziektegerelateerde morbidity). PTSS is gerelateerd aan zowel geestelijke gezondheid als fysieke gezondheidsgerelateerde kwaliteit-van-leven dimensies, en kon niet verklaard worden door verschillen in ernst van de peritonitis.

In hoofdstuk 12 bespreken we de uitdagingen die we tegenkwamen bij het ontwikkelen en uitvoeren van de RELAP-trial, en hoe we deze uitdagingen hebben opgelost en geregistreerd. Tevens bespreken we ideeën om de chirurgische on-demand strategie voor peritonitis in de toekomst te optimaliseren. Een aantal belangrijke factoren bij het vaststellen van de kwaliteit van leven in chirurgische patiënten wordt besproken. Ten slotte stellen we enige methodologische aspecten van het vaststellen van PTSS-symptomen in deze patiënten aan de orde, en mogelijke methoden om het risico op het ontwikkelen van PTSS-symptomen na verblijf op de intensive care te verminderen.
References


(8) Terwee CB, Bot SD, de Boer MR et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34-42.


(64) Bleeker SE, Moll HA, Steyerberg EW et al. External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol.* 2003;56:826-832.


Appendices

“If a man will begin with certainties, he shall end in doubts; But if he will be content to begin with doubts, he shall end in certainties.”

Francis Bacon (The Advancement of Learning, 1605)
Appendix 1 Disease-related morbidity

A  Disease-related major morbidity needing readmission and conservative treatment but not surgery

- **Fistula**: nonanatomical connection between hollow organ and cutis or between two hollow organs
- **Wound dehiscence / incisional hernia with obstruction**: full-thickness discontinuity in abdominal wall with bulging of abdominal content
- **Abscess needing percutaneous drainage**: pus-containing non-preexisting cavity confirmed by positive Gram-stain or culture
- **Renal failure**: urine production < 500 mL/24h with rising level of blood urea nitrogen and creatinine combined with dehydration (decreased circulating volume with elevated hematocrit needing intravenous rehydration) based on inadequate oral intake, nausea/vomiting, or both (only when needing readmission).
- **Myocardial infarction** (electrocardiogram and enzyme changes suggestive of myocardial infarction or needing admission to coronary car unit), **pulmonary embolus** (ventilation-perfusion mismatch on lung scintigraphy), or **cerebrovascular accident** (ischemic or nonischemic with persistent paresis or paralysis without previous history)
- **Gastric or duodenal bleeding**: needing endoscopic treatment or embolisation therapy
- **Respiratory failure** due to pneumonia, pleural effusion, or pulmonary edema and needing oxygen therapy or mechanical ventilation
- **Urosepsis**: urinary tract infection with positive urine and blood cultures and circulatory shock

B  Disease-related major morbidity needing surgical intervention during first admission or readmission

- **Incisional hernia**: full-thickness discontinuity in abdominal wall with bulging of abdominal contents with or without obstruction with disabling concerns interfering with daily activities
- **Bowel obstruction or herniation due to intra-abdominal adhesions**: diagnosis must be confirmed during surgery
- **Burst abdomen**: complete midline or transverse discontinuity in abdominal wall
- **Abdominal compartment syndrome**: intra-abdominal hypertension >25 mmHg with tense abdomen and with increasing respiratory failure, renal failure, or both; measured by the urinary bladder pressure method (modified Burch criteria)
- **Fistula**: nonanatomical connection between intestine and cutis or between two hollow organs
- **Intra-abdominal bleeding**: only when septic bleeding after index laparotomy or relaparotomy or surgical bleeding after relaparotomy but not after index laparotomy
- **Intra-abdominal hematoma** needing surgical evacuation
- **Perforation** of visceral organ confirmed at surgery
• **Anastomotic leakage**: anastomotic leak on contrast imaging needing surgery or contrast-enhanced computed tomography scan, confirmed at relaparotomy

• **Ischemia or necrosis of a visceral organ**: critically reduced blood flow to an intra-abdominal organ causing tissue loss, confirmed at pathological examination

• **Enterostomy dysfunction** due to prolaps, stenosis, or retraction

• **Gastric or duodenal ulcer bleeding** needing intervention of any type
Appendix 2 Nederlandse ziektespecifieke kwaliteit van leven (KvL) vragenlijst (SP-QoL)

De volgende vragen in deze lijst gaan over hoe het met u en uw ziekte gaat. We vragen naar uw gezondheid, uw gevoelens en uw sociale activiteiten in de afgelopen week. Soms zeggen patiënten dat ze klachten of problemen hebben. Wilt u aangeven, in welke mate u de volgende klachten of problemen gedurende de afgelopen week heeft ervaren?

<table>
<thead>
<tr>
<th>Gedurende de afgelopen week:</th>
<th>helemaal niet</th>
<th>een beetje</th>
<th>nogal</th>
<th>heel erg</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Hoofd pijn</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>2 Heeft u een onprettig gevoel in uw buik gehad?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>3 Had u een opgeblazen gevoel in uw buik?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>4 Had u snel een vol gevoel?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>5 Had u problemen met het eten van vast voedsel?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>6 Had u problemen met drinken?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>7 Was u beperkt in het soort voedsel dat u kon eten ten gevolge van uw ziekte of behandeling?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>8 Was u beperkt in de hoeveelheid voedsel die u kon eten ten gevolge van uw ziekte of behandeling?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>9 Smaakten voedsel en drank anders dan u gewend was?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>10 Bent u afgevallen?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>11 Heeft u zich zorgen gemaakt dat uw gewicht te laag is?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>12 Heeft u zich slap gevoeld in uw armen en benen?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>13 Hoe belastend is uw ziekte of behandeling voor u geweest?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>14 Maakte u zich zorgen over uw gezondheid in de toekomst?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>15 Bent u beperkt geweest in uw dagelijkse activiteiten?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>16 Voelde u zich lichamelijk minder aantrekkelijk ten gevolge van uw ziekte of behandeling?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>17 Voelde u zich minder vrouwelijk/mannelijk ten gevolge van uw ziekte of behandeling?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>18 Was u ontevreden met uw lichaam?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>19 Had u minder belangstelling voor seks?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
<tr>
<td>20 Heeft u minder plezier beleefd aan seks?</td>
<td>☐ 1</td>
<td>☐ 2</td>
<td>☐ 3</td>
<td>☐ 4</td>
</tr>
</tbody>
</table>
21 Heeft u op dit moment een stoma? (wilt u Nee of Ja omcirkelen?)

<table>
<thead>
<tr>
<th></th>
<th>Nee ➔ Wilt u de vragen 22 t/m 28 beantwoorden?</th>
<th>Ja ➔ Wilt u de vragen 22 t/m 28 overslaan, en de vragen 29 t/m 38 beantwoorden?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Alleen voor patiënten ZONDER stoma

<table>
<thead>
<tr>
<th>Gedurende de afgelopen week:</th>
<th>helemaal niet</th>
<th>een beetje</th>
<th>nogal</th>
<th>heel erg</th>
</tr>
</thead>
<tbody>
<tr>
<td>22 Heeft u overdag vaak ontlasting gehad?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 Heeft u ’s nachts vaak ontlasting gehad?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24 Heeft u loze aandrang (= aandrang zonder ontlasting) gehad?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 Heeft u onbedoeld ontlasting verloren?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26 Heeft u bloed bij uw ontlasting gehad?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27 Had u een moeilijke stoelgang?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28 Had u een pijnlijke stoelgang?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Alleen voor patiënten MET stoma

<table>
<thead>
<tr>
<th>Gedurende de afgelopen week:</th>
<th>helemaal niet</th>
<th>een beetje</th>
<th>nogal</th>
<th>heel erg</th>
</tr>
</thead>
<tbody>
<tr>
<td>29 Was u bang dat anderen uw stoma zouden kunnen horen?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 Was u bang dat anderen uw ontlasting zouden kunnen ruiken?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31 Maakte u zich zorgen over mogelijke lekkage van de stomazakjes?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32 Had u problemen met de verzorging van uw stoma?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33 Was uw huid rondom het stoma geïrriteerd?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34 Voelde u zich opgelaten door uw stoma?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35 Voelde u zich minder compleet vanwege uw stoma?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 3  Members of the Dutch Peritonitis Study Group

RELAP trial clinical centers and investigators of the Dutch Peritonitis Study Group All investigators are from Departments of Surgery unless specified Clinical Epidemiology and Biostatistics (E), Intensive Care (I) or Medical Psychology (MP).
O van Ruler MD; KR Boer MSc (E); JB Reitsma MD, PhD (E); CW Mahler MD; EA Reuland MSc;
JWO van Till MD; BC Opmeer PhD (E); PMM Bossuyt PhD (E); MJ Schultz MD, PhD (I);
MA Sprangers MD, PhD (MP); DJ Gouma MD, PhD; H Obertop MD, PhD; CAJM de Borgie MD, PhD (E);
MA Boermeester MD, PhD, Academic Medical Center, Amsterdam; EPh Steller MD, PhD;
P. Tanis MD, PhD; H Hart MD (I), St Lucas Andreas Hospital, Amsterdam; MF Gerhards MD, PhD;
M Guijt MD; HM Oudemans MD, PhD (I), Onze Lieve Vrouwe Gasthuis, Amsterdam;
K. Bosscha MD, PhD; E Ritchie MD; M Vermeer, Bosch Medical Centre, Den Bosch;
PW de Graaf MD, PhD; B van Etten MD, PhD; C Haazer (I); E Salm MD, PhD (I);
Reinier de Graaf Hospital, Delft; B Lamme MD, PhD; EJ Hesselink MD, PhD;
H Rommes MD, PhD (I), Gelre Hospitals, Lucas Hospital Apeldoorn; RJ Oostenbroek MD, PhD;
L te Velde MD; G Govaert MD; HH Ponssen MD (I), Albert Schweitzer Hospital, Dordrecht;
HG Gooszen MD, PhD; MK Dinkelman MD; LPH Leenen MD, PhD (I), University Medical Centre Utrecht;
EGJM Pierik MD, PhD; KWW Lansink MD; J Bakker MD, PhD (I), Isala Clinics, Zwolle.

Key staff and steering committee at coordinating center (AMC Amsterdam)
RELAP trial

O van Ruler (study coordinator and investigator), EA Reuland (data management),
CW Mahler (investigator), JB Reitsma (epidemiologist), CAJM de Borgie (epidemiologist),
KR Boer (quality of life investigator), BC Opmeer (economist),
MA Boermeester (surgeon, supervisor, project leader) from the Department of Surgery,
Academic Medical Center Amsterdam, The Netherlands.
“‘In that direction,’ the Cat said, waving its right paw round, ‘lives the Hatter: and in that,’ waving the other paw, ‘lives a March Hare. Visit either you like: they’re both mad.’ ‘But I don’t want to go among mad people,’ Alice remarked. ‘Oh you can’t help that,’ said the Cat, ‘we’re all mad here. I’m mad. You’re mad.’ ‘How do you know I’m mad,’ said Alice. ‘You must be,’ said the Cat, ‘or you wouldn’t have come here.’”

*Lewis Caroll (Alice’s Adventures in Wonderland, 1865)*
Dr. Marja A Boermeester, Department of Surgery, Academic Medical Center, Amsterdam

Dr. Corianne A de Borgie, Department of Clinical Epidemiology, Biostatistics, and Bioinformatics, Academic Medical Center, Amsterdam

Prof. dr. Patrick M Bossuyt, Department of Clinical Epidemiology, Biostatistics, and Bioinformatics, Academic Medical Center, Amsterdam

Dr. Arnold AP van Emmerik, Department of Psychology, University of Leiden

Dr. Michael F Gerhards, Department of Surgery, Onze Lieve Vrouwe Gasthuis, Amsterdam

Prof. dr. Hein G Gooszen, Department of Surgery, University Medical Center Utrecht, Utrecht

Prof. dr. Dirk J Gouma, Department of Surgery, Academic Medical Center, Amsterdam

Dr. Peter W de Graaf, Departement of Surgery, Reinier de Graaf Gasthuis, Delft

Dr. Bas Lamme, Department of Surgery, Gelre Hospital, Apeldoorn

Cecilia W Mahler, Department of Surgery, Academic Medical Center, Amsterdam

Dr. Brent C Opmeer, Department of Clinical Epidemiology, Biostatistics, and Bioinformatics, Academic Medical Center, Amsterdam

Dr. Johannes B Reitsma, Department of Clinical Epidemiology, Biostatistics, and Bioinformatics, Academic Medical Center, Amsterdam

E Ascelijn Reuland, Department of Surgery, Academic Medical Center, Amsterdam

Dr. Sophia E de Rooij, Department of Internal medicine, Geriatrics, Academic Medical Center, University of Amsterdam

Oddeke van Ruler, Department of Surgery, Academic Medical Center, Amsterdam

Prof. dr. Mirjam A Sprangers, Department of Medical Psychology, Academic Medical Center, University of Amsterdam

Dr. E Philip Steller, Department of Surgery, Sint Lucas Andreas Hospital, Amsterdam

JW Olivier van Till, Department of Surgery, Academic Medical Center, Amsterdam

Cagdas Unlu, Department of Surgery, Academic Medical Center, Amsterdam

Prof. dr. Margreeth B Vroom, Department of Intensive Care Medicine, Academic Medical Center, Amsterdam

Manon LW Ziech, Department of Surgery, Academic Medical Center, Amsterdam
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADL</td>
<td>Activities in Daily Life</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>APACHE II</td>
<td>Acute Physiology and Chronic Health Evaluation II</td>
</tr>
<tr>
<td>APS</td>
<td>Acute Physiology Score</td>
</tr>
<tr>
<td>ARDS</td>
<td>Acute Respiratory Distress Syndrome</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under the Curve</td>
</tr>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>DSM-IV</td>
<td>Diagnostic and Statistical Manual of Mental Disorders- Fourth Edition</td>
</tr>
<tr>
<td>DM</td>
<td>Diabetes Mellitus</td>
</tr>
<tr>
<td>EBM</td>
<td>Evidenced Based Medicine</td>
</tr>
<tr>
<td>EORTC</td>
<td>European Organisation for Research and Treatment of Cancer</td>
</tr>
<tr>
<td>EORTC-QL</td>
<td>European Organisation for Research and Treatment of Cancer Quality of Life group</td>
</tr>
<tr>
<td>QLQ-PAN26</td>
<td>EORTC module for Pancreatic cancer</td>
</tr>
<tr>
<td>QLQ-CR38</td>
<td>EORTC module for Colorectal cancer</td>
</tr>
<tr>
<td>QLQ-OES18</td>
<td>EORTC module Oesophageal cancer</td>
</tr>
<tr>
<td>EQ-VAS</td>
<td>Euroqol-Visual analogue scale</td>
</tr>
<tr>
<td>EQ-5D</td>
<td>Euroqol-Five-dimensions</td>
</tr>
<tr>
<td>ES</td>
<td>Effect Size</td>
</tr>
<tr>
<td>GEE</td>
<td>Generalized Estimating Equations</td>
</tr>
<tr>
<td>HR-QoL</td>
<td>Health Related-Quality of Life</td>
</tr>
<tr>
<td>HTA</td>
<td>Health Technology Assessment</td>
</tr>
<tr>
<td>IES-R</td>
<td>Impact of Event Scale–Revised</td>
</tr>
<tr>
<td>ICU</td>
<td>Intensive Care Unit</td>
</tr>
<tr>
<td>IQR</td>
<td>Inter-Quartile Range</td>
</tr>
<tr>
<td>MPI</td>
<td>Mannheim Peritonitis Scale</td>
</tr>
<tr>
<td>MODS</td>
<td>Multiple Organ Dysfunction Score</td>
</tr>
<tr>
<td>n</td>
<td>Number of patients</td>
</tr>
<tr>
<td>n.a.</td>
<td>Not applicable</td>
</tr>
<tr>
<td>NNT</td>
<td>Number Needed to Treat</td>
</tr>
<tr>
<td>OR</td>
<td>Odds Ratio</td>
</tr>
</tbody>
</table>
PCD  Percutaneous Drainage
PTSD  Post-traumatic Stress Disorder
PTSS-10  Post-traumatic Stress Syndrome 10-question Inventory

ROC  Receiver Operating Curve
RR  Relative Risk
RD  Relative Difference

SOFA  Sequential Organ Failure Assessment
SAPS II  Simplified Acute Physiology Score II
SCID  Standardized Clinical Interview Diagnosis
SD  Standard Deviation
SP-QoL  Secondary Peritonitis Quality of Life questionnaire
AB  SP-QoL Abdomen subscale
NU  SP-QoL Nutrition subscale
WO  SP-QoL Worries subscale
BI  SP-QoL Body-image subscale
SEX  SP-QoL Sexuality subscale
DEF  SP-QoL Defecation subscale
SRP  SP-QoL Stoma related problems subscale

US  Ultrasound

95% CI  95% Confidence Intervals
$\chi^2$  Chi-Square test
Acknowledgements

“If he is indeed wise he does not bid you enter the house of his wisdom, but rather leads you to the threshold of your own wisdom.”

Kahlil Gabran (The Prophet)
For every sentence, every paragraph and every chapter in this thesis, there are a group of people who hypothesized, discussed and often even negotiated to make them perfect. For the tears, sweat and blood that they have put into this book I am eternally grateful.

Professor Bossuyt, dear Patrick, I have thoroughly enjoyed working with you. I have sought your advice on many occasions about my future plans and to help guide me through the sometimes complex world of epidemiology, with so many choices to be made. One of my greatest strengths is also one of my greatest weaknesses; I am always interested in 50 things at once. Thankfully, you have helped me streamline my thoughts and concentrate on one topic at a time (well, maybe not one, but close). Thank you for keeping me on track.

Dr. Boermeester, dear Marja, to watch you in the surgical theater, taking on an acute surgery with three students at different levels around you, all hanging on your every word; as you with ease, teach, cut and heal, was like watching a director in her orchestra. You always took the time to teach me on my level. As well your interest in PTSD and HR-QoL is unparalleled in the surgical world. I am very proud to have you as my co-promoter. My respect and admiration for you is enormous and I am very thankful for the opportunity you have given me over the past 5 years and the vast information you have had the patience to share with me.

Dr. de Borgie, dear Corianne, I could not have imagined a better, more attentive companion to start at the KEBB. We have shared so many interests together; at work, our joint interest in psychological effects of critical illness; and in our free time, our love of art and an interest in medical ethics. We both love to multi-task – so sometimes it was not easy to concentrate on the work at hand; there were so many other interesting things to discuss. It was an absolute pleasure working with you and I hope we will soon have the time and opportunity to discuss the multi-facets of medical ethics.

Dr. Reitsma, dear Hans, I am not quite sure what my thesis would have looked like without you – but to be honest – I don’t even want to speculate. Your methodological input and critical approach to each paper was essential, but most importantly your ability to always make complicated methodological issues clinically relevant and practical is a true talent. I have learned so much from you in the last 3 years. You are the best clinical epidemiologist I have ever met. Thank you for sharing your brilliance with me and always challenging me. I will miss working with you as there is still so much for me to learn from you.

Professor dr. H.G. Gooszen, Professor dr. J.C.J.M. de Haes, Professor dr. H. Obertop, dr. M Olff, Professor dr. E.W. Steyerberg and and Professor dr. M.B. Vroom, thank you for your participation in my PhD committee, for taking the time to critically read my manuscript and for the opportunity to discuss your reservations during my public defense.
Dear Oddeke, thank you for being so easy to work with. Working this hard could never have been so much fun, if it were not for doing it with you (thank you for a million laughs about nothing I can never repeat in public). Thank you for being awake at 2 o’clock in the morning, so we could just finish one more article. I am truly honored to be standing on the podium after you, and I know you are going to be brilliant.

I would like to thank all the coauthors of this thesis, Prof. Gooszen, Prof. Gouma, Dr. de Graaf, Dr. Steller, thank you for this fruitful collaboration. Arnold, Cecilia, Brent, Bas, Ascelijn, Sophia, Olivier, Cagdas, Manon, without you all this work would not have been possible.

Dear Brent, thank you for being a stable factor in the RELAP trial. You have always been available for all of my questions and frequent panics; your support has been never-ending and highly appreciated.

Dear Mirjam, thank you for your insights into many of these papers. It has been irreplaceable, your comments making each paper stronger and more concrete. You were always able to see the broader perspective. I have enjoyed our long chats, and hope we have time to continue them in the future.

Dear Prof. Vroom, Neely van Os and Dave Dongelmans, I am so pleased that you have given me the opportunity to evaluate the use of diaries for the reduction of PTSD symptoms in ICU patients. It was such a pleasure to find clinicians who shared this interest. Towards the end of my thesis, when my energy was waning, you re-inspired me and made the ‘laatste loodjes’ more interesting and clinically more tangible.

Dear Marcel, I am sorry we did not get to complete my original ‘proefschrift’ together, but definitely not for lack of trying. I am so pleased we had that time to work together. You gave me a new perspective on the fundamentals of epidemiological research: never underestimate the importance of the quality of the study and the scientific value. Our exercise in sample size and patient recruitment has changed my career path for good; and inspired me to aim at strengthening research standards. But most of all, we had a lot of good laughs together.

To my colleagues at the KEBB department, first and foremost, Petra and Gre, oh .... my life-savers, I owe you so much ..... You are mijn steun en toeverlaat!! I will miss you both very much and will come by at regular intervals for hugs. Marije, Helene, Rebecca, Sandra and Willem, my roommates at the KEBB department, we have had many very interesting conversations over the years; politics, relationships, religion, you name it, we have discussed it. Thank you for always allowing me to express my frustrations and for sharing my excitements. Olga thank you for your kindness and gentle demeanor and laughing at my jokes when we both knew they were not really that funny; Anne, thank you for showing me the way on the 14th of December 2005, those were some good times we had preparing for your special day, and I miss you tons! Lotty, I look forward to and count on our philosophical conversations, I am so grateful for our friendship. Mariska, Marlies, Fleur, Nadine, Bart, Hans, Barbara and Joost, I hope we continue to have fun together and cannot wait to see you up on that podium soon!
Dear Boy and Mathieu, thank you for taking such good care of me in preparing this manuscript for press. Your knowledge, advice and quiet composure have surely kept me from lunacy.

Justin, thank you so much for your fantastic illustrations, made during a hectic schedule of exams, moving and working. I think you are amazingly talented and are going to be a fantastic doctor. To you and Kristin, my utmost admiration for moving to Amsterdam for this new city and new adventure, we are very glad you did.

To my great friends, those of you that are still speaking to me after easily five years of neglect. To the girls, Suzie, Sige, Lara thanks for your listening ears and your kind words of encouragement, and the occasional strong gin and tonic. To my London–New York connection - Alison, Annex, Audrey, Anna H, David, Barbar, Josh and Anna P, thanks for always believing in me.

Sonja, you have been the best friend anyone can have, an amazing support. Thank you for catching me when I began to fall or filling the Kleenex box when the tears poured out or shopped till we dropped when the mood called for it. And of course thank you for easily a million telephone calls over the last 5 years.

Nynke, I probably owe you a great deal of the sanity I still have left. I stood proudly next to you almost exactly one year ago, whilst the world crossed their legs as you gracefully and with full conviction defended your thesis. I can only hope to make you as proud of me on the 14th of November. Through thick and thin, we have become stronger researchers and friends for life.

Emilie, I am so pleased you have held my hand through this experience. I could not have chosen a better sister if I tried and I am so glad that moving back to the Netherlands has allowed us to grow close. You and Jan hold a very special place in our lives.

From the day mom ran after a bus to find Ms. Knowles to hand in my history project on time in 1994, to the night before this manuscript went to press, when we pulled an all-nighter to make sure every sentence was perfect. This manuscript is as much yours as it is mine. Mom and dad you mean the world to me. Dad, thank you for translating every abstract, questionnaire and paper from English to Dutch and back again; and mom thank you for reading every chapter in this book – at least twice – dotting every i and removing every hyphen, even in the acknowledgements. Without your love, your brilliant sense of humor, your never-ending support and a handy six-hour time difference, I would, simply, have never made it through this. There are no words – no tears – that can ever express how grateful I am to have you both.

Peter, my love, in every boekje, there is always a line at the end dedicated to the spouse – it usually starts with a promise of more time to do wonderful things together and that the hectic times are over – As you know, I cannot make this promise. Honestly being married to me means that the hectic times will never truly be over. But I know you know that. What I can promise you, are wonderful things, great laughs and a long loving future. Even though, I have run you in circles in the last year, you never once complained. I adore you, and I am so lucky to have you, you are the kindest, sweetest man I have ever met and I thank you for sharing your life with me.
List of publications

“Remember that only on paper has humanity yet achieved glory, beauty, truth, knowledge, virtue and ability.”

George Bernard Shaw


KR Boer, BC Opmeer, MA Sprangers, CW Mahler, JB Reitsma, CA de Borgie, MA Boermeester. The Psychometric Performance of a Disease Specific Health Related Quality of Life Questionnaire for Patients with Secondary Peritonitis: SP-QoL. *Quality of Life Research, in press.*


O van Ruler, MLW Ziech, KR Boer, EA Reuland, JB Reitsma, MA Boermeester. Usefulness of the Apache II scores to predict disease outcome and to evaluate safety of surgical strategy in patients with abdominal sepsis. *Digestive Surgery, submitted*  

O van Ruler, KR Boer, B Lamme, JB Reitsma, MA Boermeester, JB Reitsma. Failure of available scoring systems to predict ungoing infection in patients with abdominal sepsis after their initial emergency laparotomy. *submitted*  


Curriculum Vitae

TWO roads diverged in a yellow wood,
And sorry I could not travel both
And be one traveler, long I stood
And looked down one as far as I could
To where it bent in the undergrowth;

THEN took the other, as just as fair,
And having perhaps the better claim,
Because it was grassy and wanted wear;
Though as for that the passing there
Had worn them really about the same,

AND both that morning equally lay
In leaves no step had trodden black.
Oh, I kept the first for another day!
Yet knowing how way leads on to way,
I doubted if I should ever come back.

I shall be telling this with a sigh
Somewhere ages and ages hence:
Two roads diverged in a wood, and I
– I took the one less traveled by,
And that has made all the difference.

Robert Frost (1874–1963). The Road Not Taken, Mountain Interval, 1916
Born on October 2nd, 1976 in Haarlem, The Netherlands, Kimberly Boer moved to Harrington Park, NJ, a suburb of New York City, when she was one year old and lived there until 1986. In that year, her family returned to the Netherlands. She attended secondary school at the International School of the Hague, completing her International Baccalaureate in 1994. Again moving back to the United States, Kimberly attended New York University and then the University of South Florida, where she finished her Bachelors degree in Psychology. After working in the field of HIV-AIDS, at a non-profit organization and later at a CRO in New York City, she decided to continue her career in epidemiology. In 2001 she completed her MSc in Communicable Disease Epidemiology at the London School of Hygiene and Tropical Disease. During this time in London she resided in the London Goodenough Trust, allowing her the opportunity to travel to Western Kenya to complete her master’s thesis on malaria and anemia in schoolgirls. After a short stint at Columbia University working on HIV-AIDS research, she again returned to The Netherlands. In 2002, she joined the Academic Medical Center, Amsterdam and began work on her doctoral thesis on surgical interventions for patients with severe secondary peritonitis, at the Department of Clinical Epidemiology and Biostatistics under the supervision of Prof. P.M.M. Bossuyt, Dr. C.A.J.M. de Borgie, Dr. J.B. Reitsma and Dr. M.A. Boermeester. At the same time, she had the opportunity to work on research with many other clinical departments of the AMC, including the Departments of General Practice, Internal Medicine (Syncope Unit), Gastroenterology and Adult Intensive Care Unit. Presently, Kimberly works at the Department of Biomedical Research at the Royal Tropical Institute in The Netherlands, concentrating on diagnostic research and medical ethics in developing countries.
This dissertation concentrates on patients with severe secondary peritonitis (abdominal sepsis), a condition with high mortality and disease-related morbidity. Secondary peritonitis is a clinical diagnosis requiring an emergency surgical laparotomy to confirm diagnosis and to tailor surgical treatment. With an estimated incidence in the United States of 9 cases of secondary peritonitis per 1000 emergency hospital admissions and with extensive intensive care admissions and lengthy hospital stays, these patients represent a substantial cost to the healthcare system.

Patients surviving peritonitis also continue to report substantially reduced health-related quality of life and are often readmitted to hospital with disease-related complications in the first year following their initial acute illness and surgery. Therefore in this thesis, we concentrate on both the psychological and physiological recovery of these severely ill patients.