Finding sepsis, what we seek and what we find

Tanca Minderhoud

Finding sepsis

what we seek and what we find

Tanca Minderhoud

Tanca Minderhoud
Universiteit van Amsterdam

ISBN: 9789464375299

Paranimfen:
Kim Sigaloff
Claartje Thijs

Vormgeving en opmaak: Erik Elferink / Meneer E. illustratie en vormgeving Amsterdam

Illustratie omslag: Hanna de Haan, http://www.hannadehaan.nl

Organisatie: margreet@morganiseren.nl

Printing: Digiforce, Vianen

VRIJE UNIVERSITEIT

FINDING SEPSIS, WHAT WE SEEK AND WHAT WE FIND

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philosophy aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op woensdag 22 juni 2022 om 13.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door

Tannetje Cornelia Minderhoud

geboren te Wycombe, Verenigd Koninkrijk

promotoren: prof.dr. K.D.P.W.B. Nanayakkara
prof.dr. M.A. van Agtmael

promotiecommissie: prof.dr. J.M. Prins
dr. L. Vaughan
prof.dr. S.C.E. Klein Nagelvoort
prof.dr. P. Pickkers
dr. E. Sieswerda

Juliet:

'Tis but thy name that is my enemy;

Thou art thyself, though not a Montague.

What's Montague? It is nor hand, nor foot,

Nor arm, nor face, nor any other part

Belonging to a man. O, be some other name!

What's in a name? That which we call a rose

By any other name would smell as sweet

Willliam Shakespeare, Romeo & Juliet, Act II, scene II, Lines 38-47

Een lantarenpaal met takken

Dat is dus een boom

Kijken met je ogen dicht

Dat is dus een droom

Twee vingers in het stopcontact

Dat is dus stroom

Voor alles is een naam

De zon die 's nachts schijnt, heet maan

En een gele komkommer is banaan

Bram Vermeulen, Onzin

Introduction

Chapter 1

General introduction and outline of the thesis

GENERAL INTRODUCTION AND OUTLINE OF THE THESIS

In 2008 I encountered the sickest patient I had ever seen. I had only just graduated from medical school when a 55-year-old man entered the Emergency Department of the small hospital I was working in. He came in with confusion, general weakness and a low blood pressure. Before this episode, he had been in good health. I noticed tiny red marks all over his body, which later turned into obvious petechial hemorrhages. His lab showed multi-organ failure, and only a few hours later he was critically ill in the intensive care department. The diagnosis was sepsis by meningococcal infection. My supervisor lectured me on the importance of door-to-needle time for antibiotics[1]. The huge impact of sepsis and the rate of deterioration left me in fear and awe of this disease and triggered an interest that has ultimately resulted in this thesis.

THE HISTORY OF SEPSIS

The concept of sepsis (σηψις) has already been described in ancient Greece by Homer (~750 b.c.) and Hippocrates (~400 b.c). Stemming from σηπω (I rot), it was described as a general decay of the body. From a historical perspective, sepsis or septicaemia was used for patients that were severely ill due to infections, although for decades it was unclear what caused infections. With the discovery of bacteria and hand hygiene in the nineteenth century, the understanding of the pathogenesis of infections was much improved. The real game-changer, antibiotic treatment, was discovered by Alexander Fleming who published on penicillin in 1929. Despite these improvements in the treatment of infections, sepsis as a condition was not clearly delineated and several terms were used like septiceamia, endotoxemia and bacteremia. Many studies on the topic of sepsis focused on blood-stream infections, especially on the role of endotoxins in gram-negative bacteremia. Studies in the seventies and eighties found that early empirical therapy reduced the mortality of gram-negative bacteremia, however these were predominantly small studies, and quite a few of these studies focused on neutropenic patients[2].

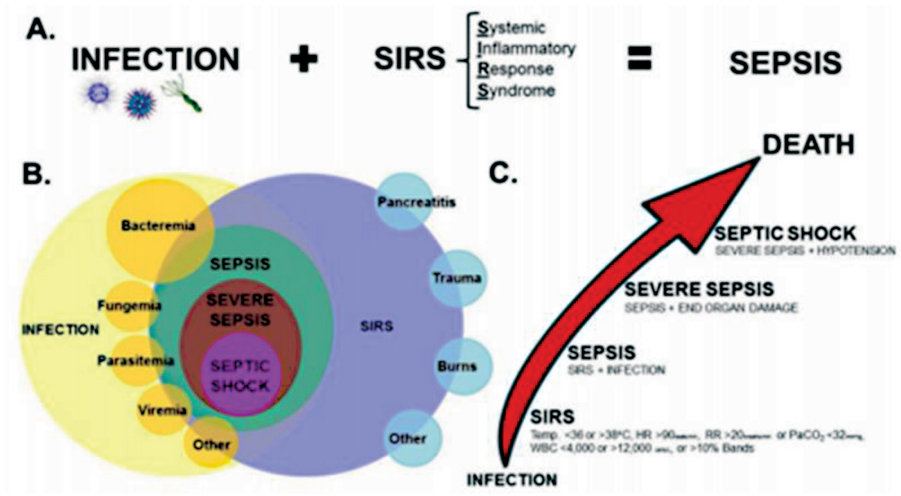
In 1992 a first international conference was held on the topic of sepsis. At this conference, consensus was reached to define sepsis as a systemic inflammatory response to an infection[3], named the Sepsis-I definition. Sepsis was further categorized in sepsis (SIRS + suspected infection), severe sepsis (organ dysfunction) and septic shock. The systemic inflammatory response and the sepsis classifications are illustrated in figure 1.1. After this conference, a lot of research was conducted to gain more understanding in the pathogenesis of sepsis. In 2001, the sepsis definition was refined to incorporate the threshold values for organ damage, and it was called sepsis-II.

[1]Kumar A, Roberts D, Wood KE, Light B, Parrillo JE, Sharma S, et al. Duration of 10 hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. Crit Care Med. 2006 Jun;34(6):1589-96.

[2]Calandra T, Cometta A. Antibiotic therapy for gram-negative bacteremia. Infect Dis Clin North Am. 1991 Dec;5(4):817-34. PMID: 1783770.

[3]Bone RC, Balk RA, Cerra FB, Dellinger RP, Fein AM, Knaus WA, Schein RM, Sibbald WJ. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. Chest. 1992 Jun;101(6):1644-55. doi: 10.1378/chest.101.6.1644. PMID: 1303622.

Figure 1.1 Earlier Conceptual View and Definition of Systemic Inflammatory Response Syndrome (SIRS), Sepsis, Severe Sepsis and Septic Shock



Author: Matthew J. Delano and Peter A. Ward, Immunol Rev. 2016 November; 274(1):330-353., who adapted a figure from Bone RC et al, Cehst.1992,101:1644-55

SCIENTIFIC SUCCESS AND THE SURVIVING SEPSIS CAMPAIGN

A landmark trial was published in 2001 by Rivers[4], in which a bundle of interventions was implemented for patients with sepsis were admitted to the intensive care unit. This bundle consisted of intravenous fluids based on measurements such as lactate and the central venous pressure. The bundle was named Early Goal Directed Therapy (EGDT). Patients treated with the bundle had a much lower in-hospital mortality of 30.5 %, compared with 46.5% in the usual care group. To improve sepsis recognition and outcome, a worldwide campaign titled the Surviving Sepsis Campaign (SSC) was launched in 2001. In 2004, the SSC produced the first consensus guidelines[5], recommending EGDT and early antibiotic treatment. In 2006, another very important trial was published by Kumar[6] who found that every hour that antibiotic therapy was delayed in patients with septic shock, mortality increased with 7.6%. This study underpinned the SSC recommendation for early antibiotic treatment.

[4] Rivers E, Nguyen B, Havstad S, et al. Early Goal-Directed Therapy Collaborative Group. Early goal-directed therapy in the treatment of severe sepsis and septic shock. N Engl J Med. 2001 Nov 8;345(19):1368-77. PMID: 11794169.

[5] Dellinger RP, Carlet JM, Masur H, et al; Surviving Sepsis Campaign Management Guidelines Committee. Surviving Sepsis Campaign guidelines for management of severe sepsis and septic shock. Crit Care Med. 2004 Mar;32(3):858-73. Erratum in: Crit Care Med. 2004 Jun;32(6):1448. Dosage error in article text. Erratum in: Crit Care Med. 2004 Oct;32(10):2169-70. PMID: 15090974.

[6] Kumar et al Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. Crit Care Med. 2006 Jun;34(6):1589-96. doi: 10.1097/01.CCM.0000217961.75225.E9. PMID: 16625125.

Given the mortality rate of sepsis, and the large effect size of early antibiotics and EGDT, great urgency was felt to implement the recommendations of the Surviving Sepsis campaign. Following the SSC, many emergency departments implemented sepsis screening using the SIRS criteria. However, the SIRS criteria were criticized from the introduction, with many in the field asserting that any person who ran up a flight of stairs might meet 2 of the SIRS criteria[7,8]. The focus on early antibiotic treatment also led to concern on antibiotic overuse[9]. Often, it was perceived, clinicians would err on the side of caution. However, without an accepted alternative, SIRS was accepted as the best-practice.

NEGATIVE TRIALS AND CONTROVERSY – TIME FOR A NEW DEFINITION

The recommendations of the SSC were widely implemented, but not all components of the EGDT were regarded as equally effective. In 2014 and 2015, three large trials (ARISE, ProCESS and ProMISe) were published that did not find a mortality benefit using EGDT. The exact timing of antibiotic treatment was also disputed. Two large retrospective trials found a significant effect of early antibiotic treatment, but the effect was strongest in patients with shock[10,11]. This led critics to dispute if the results of early antibiotic treatment in patients with shock should be extrapolated to patients without shock. The only prospective, randomized trial on antibiotics, the Phantasi[12] trial, also did not find benefit of early antibiotic treatment. However, it was disputed if the candidates included in these trials were really similar to the patients in the trials of Rivers and Kumar. It was argued that the SIRS-based sepsis definitions were too broad, leading to poor patient selection and overuse of antibiotics. Due to the observed difficulties of the SIRS-based definition, the definition of sepsis was replaced in 2016. In the sepsis III paper, sepsis was defined as life-threatening organ dysfunction caused by a dysregulated host response to infection[13]. This new definition was widely adopted and the SIRS criteria were abandoned.

SEPSIS III – NEW CHALLENGES

After the new definition was launched, many clinicians and researchers had to re-invent their approach of sepsis. With the new definition, it was unclear how to approach the topic of sepsis in the Emergency department. Should we identify patients based on (suspected) infection or based on organ dysfunction/ threat to life? How can we accurately identify patients with sepsis in the Emergency Department? And how can we avoid overuse of antibiotics?

[7] Vincent Dear SIRS, I'm sorry to say that I don't like you.. Crit Care Med. 1997 Feb;25(2):372-4. PMID: 9034279.
[8] Marshall J. Both the disposition and the means of cure: "Severe SIRS," "sterile shock," and the ongoing challenge of description. Crit Care Med. 1997 Nov;25(11):1765-6. PMID: 9366746.
[9] Fitzpatrick F, Tarrant C, Hamilton V, et al Sepsis and antimicrobial stewardship: two sides of the same coin BMJ Quality & Safety 2019;28:758-761.
[10] Liu VX, Fielding-Singh V, Greene JD, Baker JM, Iwashyna TJ, Bhattacharya J, Escobar GJ. The Timing of Early Antibiotics and Hospital Mortality in Sepsis. Am J Respir Crit Care Med. 2017 Oct 1;196(7):856-863. doi: 10.1164/rccm.201609-1848OC. PMID: 28345952; PMCID: PMC5649973.
[11] Seymour CW, Gesten F, Prescott HC, Friedrich ME, Iwashyna TJ, Phillips GS, Lemeshow S, Osborn T, Terry KM, Levy MM. Time to Treatment and Mortality during Mandated Emergency Care for Sepsis. N Engl J Med. 2017 Jun 8;376(23):2235-2244. doi: 10.1056/NEJMoa1703058. Epub 2017 May 21. PMID: 28528569; PMCID: PMC5538258.
[12] Alam N, Oskam E, Stassen PM, et al. ; Prehospital Antibiotics against Sepsis Trial Investigators and the ORCA (Onderzoeks Consortium Acute Geneeskunde) Research Consortium the Netherlands. Prehospital antibiotics in the ambulance for sepsis: a multicentre, open label, randomised trial. Lancet Respir Med 2018; 6:40–50.
[13] Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016;315(8):801–810. doi:10.1001/jama.2016.0287

To answer these questions, we first need to understand what the SEPSIS III definition consist of, and how the components are related to each other. In a background paper on the sepsis definition in 2016, Derek Angus[14] unraveled the process of the development and interpretation of sepsis definitions. In this article it was stated that "sepsis is a function of four variables linked in a causal pathway, with, from left to right, one conditional upon the other". In the article the definition was written as a logic statement (see box 1.1).

Box 1.1 Sepsis defenition as logic statement

$$sepsis = f(threat\ to\ life \mid organ\ dysfunction \mid dysregulated\ host\ response \mid infection)$$

If we want to accurately recognize sepsis, we can start by examining how accurately we measure the variables that it consists of. Then, we must add up the components and asses the causal pathway to decide if the patient has sepsis.

SEPSIS AND ITS COMPONENTS, A CASE

As pointed out above, the SEPSIS III definition consists of four components linked in a causal pathway. How do we measure these components in clinical practice and how do we assess the causal pathway? To illustrate this, a case, Sheila, is presented in box 1.2. In order to sepsis, we would like to find out if Sheila has an infection (1), a dysregulated host response (2), organ dysfunction (3) and if her situation is life threatening (4). If all these components are present and they are causally related, than her illness is consistent with the concept of sepsis[15]. In the next paragraphs we will discuss how these components are defined and handled in the clinical process.

Box 1.2

Sheila comes to the emergency department a few days after visiting her grandchildren. During her visit, one of the children turned ill and vomited a few times. One day after the meeting, Sheila starting vomiting, and progressed to have diarrhea for 3 days in a row. She has a history of hypertension and tension headache.

She has been taking her blood pressure medication (perindopril) and ibuprofen for her headache.
In the emergency department, Sheila has a low blood pressure, 80/40 mmHg and a high pulse of 110 beats per minute. She is not quite alert.

Her MEWS is 4 and her qSOFA score is 2. Her blood tests show evidence of an acute kidney injury, as well as signs of mild inflammation (normal leucocyte count with a moderately elevated CRP).

[14] Angus DC, Seymour CW, Coopersmith CM, Deutschman CS, Klompas M, Levy MM, Martin GS, Osborn TM, Rhee C, Watson RS. A Framework for the Development and Interpretation of Different Sepsis Definitions and Clinical Criteria. Crit Care Med. 2016 Mar;44(3):e113-21. doi: 10.1097/CCM.0000000000001730. PMID: 26901559; PMCID: PMC4765912.
[15] Angus DC, Seymour CW, Coopersmith CM, et al. A Framework for the Development and Interpretation of Different Sepsis Definitions and Clinical Criteria. Crit Care Med. 2016;44(3):e113-e121.

Infection

The definition of an infection, or infectious disease is an invasion and multiplication of a pathogenic organisms in (a part of) the body[16]. Many organisms can cause diseases, such as viruses, bacteria, yeast, funghi and parasites. To determine if a person has an infection, we ideally want to identify the organism causing it. Several methods are used to demonstrate the presence and invasion of these organisms: molecular methods (such as polymerase chain reaction, PCR), direct methods (such as cultures), and serological methods. In clinical practice, we often rely on cultures, but cultures usually take 24-72 hours to yield a result. Cultures also have limited sensitivity, since in some infectious diseases, the majority of patients have negative cultures. The most important example of this is pneumonia. In about 50% of the patients with pneumonia, the causative organism is not found[17]. This is the reason why pneumonia is usually defined by radiological (x-ray) abnormalities, combined with clinical features. Therefore, clinicians need to assess clinical, radiological and microbiological data to establish if a patient has an infection. Clinicians are used to analyzing multiple sources of information to come to a diagnosis. For the purpose of this thesis, we were interested in how accurately we diagnose infection.

Organ dysfunction

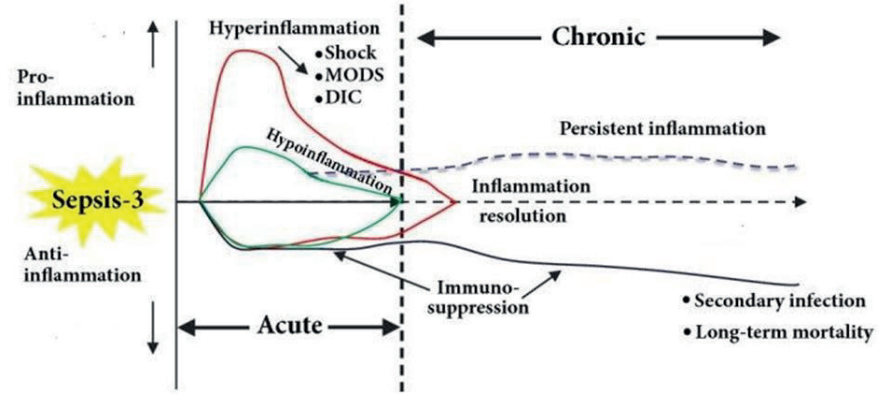
Organ dysfunction is hard to define because every organ has its own type of dysfunction. Some patients might present with delirium, which is indicative of malfunctioning of the central nervous system, whereas others present with widespread endothelial damage and an activated coagulation cascade (diffuse intravasal coagulation) and yet others present with hypotension or oliguric kidney injury. For every type of organ failure, we have to find a cut-off to differentiate function from dysfunction. To complicate things, the organ dysfunction could have been pre-existing, or could have been caused by other mechanisms than infection. Paraphrasing Tolstoy it seems that " in short, healthy bodies all function alike; every dysfunctional body is dysfunctional in its own way."[18] Despite these challenges, organ dysfunction does correlate with mortality. Several scoring systems, based on organ dysfunction are used successfully to predict the chance of mortality. In intensive care units (ICU), a score was developed and validated to measure organ dysfunction, to predict the risk of death. This score, the Sequential Organ Function Assesment (SOFA)[19] has been validated en used for many years. However, this score is unsuitable for use in the emergency department (ED), since it requires many measurements that are not available in the ED. With the introduction of the SEPSIS III definition, a new score was also introduced, named the quick SOFA (qSOFA). The qSOFA was introduced as a simple tool for risk stratification, outside of the ICU. However, many other scoring systems, some specifically for infections, some generic, were already used to risk-stratify patients in the Emergency Department. In the first two chapters of this thesis, we will examine how accurate these scores identify patients in the ED who are at risk for adverse outcomes (death, ICU admission).

Summarizing, organ dysfunction is usually measured by using scoring systems, we just have to find out which score has the best accuracy for our ED patients.

Dysregulated host response

In reaction to a pathogen invading our body, the immune system will start a response to fight off the invading organism. Early on, it was postulated that it in some cases, it was not the pathogen that killed the host, but rather the overzealous reaction of the immune system that caused the illness and demise of the host. The immune system consists of the innate immune response and the adaptive response. The innate response, when triggered, has components that stimulate inflammation (pro-inflammatory respons), but simultaneously has components that dampen the inflammatory response (the anti-inflammatory response). The current concept is that in sepsis, the balance between pro-inflammatory factors and anti-inflammatory factors is lost. A overshoot of the innate response leads to hyperinflammation. Clinical signs of these hyperinflammation are thought to consist of shock, multi-organ dysfunction and dysregulated clotting (diffuse intravasal coagulation). The concept of dysregulated host response is illustrated in figure 1.2. The dysregulated host response has been the target of many studies and treatments and is rapidly evolving[20] but has also been the source of dispute since many trials that sought to treat this response failed. In a critical article in 2018, Alcock, disputed the notion that the found

Figure 1.2 Adapted model of sepsis-3



In the acute phase of sepsis, the host inflammatory response to an infection is heterogeneous. In some patients, the classical picture of hyperinflammation (red line) is observed, whereas others have signs of hypoinflammation (green line). In the chronic phase of sepsis, persistent inflammation and immunosuppression can cause secondary infection and long-term mortality.
Abbreviations: DIC, disseminated intravascular coagulation; MODS, multiple organ dysfunction syndrome; SOFA, sequential organ failure assessment.
This figures was adapted and simplified from. Ding R, Meng Y, Ma X. The Central Role of the Inflammatory Response in Understanding the Heterogeneity of Sepsis-3 et al. Biomed Res Int. 2018;2018:5086516. Published 2018 Jun 7.

[16] M Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases
[17] Cilloniz, Catia et al. "Microbial Etiology of Pneumonia: Epidemiology, Diagnosis and Resistance Patterns." International journal of molecular sciences vol. 17,12 2120. 16 Dec. 2016, doi:10.3390/ijms17122120
[18] Leo Tolstoy: Anna Karenina
[19] Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, Reinhart CK, Suter PM, Thijs LG. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive Care Med. 1996 Jul;22(7):707-10. doi: 10.1007/BF01709751. PMID: 8844239.

immune responses are in fact dysregulated, and argued that it is more likely that these responses are physiological[21].
Despite a large body of literature on the dysregulated host response, translation to bedside medicine has proved difficult. In usual care, the dysregulated host response it is not currently measured in a uniform and clinically relevant way. Since this thesis was focused on bedside medicine, this element was not the subject of further study in this thesis.

Threat to life

The item of ''threat to life" in the definition of SEPSIS III was a new component, but it was hardly discussed in the paper proposing the new definition[4], or in subsequent papers discussing the definition[5]. In the original paper, it was explained that in lay terms, sepsis is a life-threatening condition that arises when the body's response to an infection injures its own tissues and organs. It seems therefore to function as a threshold. Only if the infection, leading to organ dysfunction has led to a 'threat to life', only then can we call it sepsis. In both the original study, and in following articles, it is unclear how we should measure 'threat to life'. However since qSOFA was designed and tested to predict mortality, it becomes clear that the purpose of the item of threat to life is to confer that a certain degree of organ dysfunction is required for the diagnosis of sepsis. From this it follows that the elements of organ dysfunction and threat to life are connected, and these items will be discussed together in this thesis.

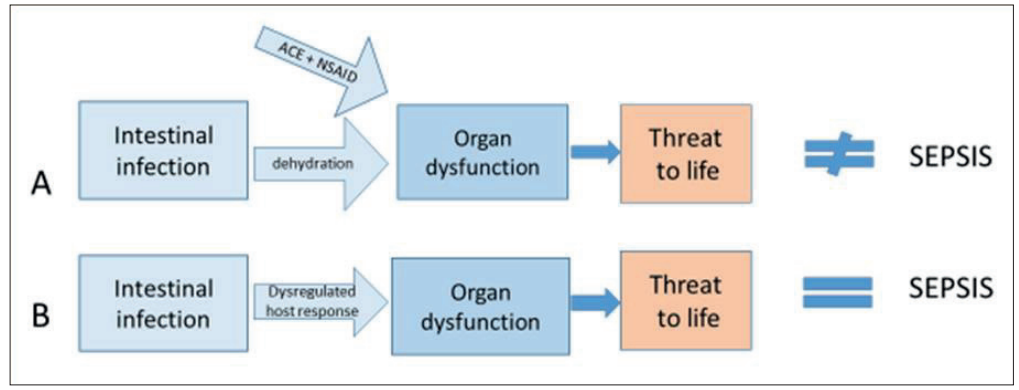
1

EVALUATION OF SHEILA'S CASE

Now that we know the background of the components of sepsis, let's apply this knowledge to Sheila's case. Sheila's situation can be modeled in several ways. Some physicians might argue that it is most likely that she has a self-limiting intestinal infection, leading to dehydration due to ongoing losses, which, in combination with perindopril and ibuprofen has led to kidney dysfunction and low blood pressure (Figure 1.3, model A). This in turn, leads to a life-threatening situation.
Another physician may hypothesize that the intestinal infection has led to an imbalance in the immune system, the dysregulated host response. Unfortunately, we cannot measure this respons, but we can measure the resulting organ dysfunction. In Sheila's case the hypothesized dysregulation has injured the kidneys and has led to distributory shock, thus resulting in a life-threatening situation. (Figure 1.3, model B).

[20] Wiersinga WJ, Leopold SJ, Cranendonk DR, van der Poll T. Host innate immune responses to sepsis. Virulence. 2014 Jan;5(1):36-44. DOI: 10.4161/viru.25436.
[21] Alcock J. The Emperor Has No Clothes? Searching for Dysregulation in Sepsis. J Clin Med. 2018 Aug 29; 7(9):247. doi: 10.3390/jcm7090247. PMID: 30158480; PMCID: PMC6162833.

Figure 1.3 Model of Sheila's situation



TREATMENT OF SEPSIS AND THE PROBLEM OF HETEROGENEITY

Does it matter if we follow model A, or model B? It is quite likely that both models lead to a good outcome. However, if it we would like to include Sheila in a trial, with the target of treating a dysregulated host-response, it would be important that she indeed has a dysregulated host-response. Similarly, if we include her in a trial on early antibiotic treatment, it is relevant if she has a bacterial intestinal infection, and not a viral infection.
If we include the patients without the targeted disease in a trial, we risk finding an effect, that is not really present (type 1 error) or not finding an effect that is present (type 2 error). For example, assume that model A is true, and that Sheila is dehydrated, and responds well to fluids. She might be included in a trial on sepsis, which studies a treatment bundle consisting of both early antibiotic treatment and standardized fluid resuscitation. Sheila would have a favourable outcome, because she does respond well to fluids. If many patients like her would be included in the trial, it might show a significant beneficial effect of the treatment bundle for suspected sepsis. However, if the trial is flawed it might erroneously lead to the conclusion that all patients like Sheila need antibiotic treatment.

The negative outcomes of the large EGDT trials (ARISE, ProCESS and ProMISe) and early antibiotic therapy (Phantasi) were in fact attributed to heterogeneity and poor patient selection using the SIRS-based sepsis definition. The new SEPSIS-III definition was introduced to improve patient selection and homogeneity. But is it really more accurate? How can we use the SEPSIS-III definition in the ED?

The main goal of this thesis is to find out in how accurate we are in recognizing sepsis. If we are wrong about sepsis in only a small percentage of the cases, the effect on conclusions will be small. But if we are wrong in in one-third or even half of the cases this has a profound impact on the validity of our studies. In this thesis, we tried to find out how accurate we can recognize sepsis in the ED, by studying the underlying components: infection, organ dysfunction and threat to life.

1

AIMS AND OUTLINE OF THE THESIS

In 2012, we started to evaluate the sepsis protocol that was implemented at the time in the Albert Schweitzer hospital (Dordrecht). The central question I started with in 2012 was: how many of the patients treated for sepsis in the Emergency Department, actually have sepsis? However, after the introduction of the sepsis-3 definition, we realized we needed more insight in the components that sepsis consist of, to answer the bigger question.

Part one of the thesis studies scoring systems that are used to measure organ dysfunction, with the aim of identifying patients that are in a life-threatening situation, in the Emergency department. In *chapter 2* we performed a narrative review of the diagnostic performance of risk-stratification scores in the Emergency Department. In this review, we included risk-stratification scores that were studied for the whole ED population, but we separately described the studies that focused on patients with infections.
In *chapter 3*, the results are described of our prospective multicenter study, where we compared the qSOFA score to other risk stratification scores and the SIRS criteria for early risk stratification in the Emergency department. We chose to study the first set of vital parameters to calculate the risk scores, because treatment decisions such as antibiotics, fluids are often made early in the ED visit. Which score can we use best for this setting?

Part two of this thesis explores several research questions in relation to diagnosing infection in the emergency department. In *chapter 4*, a retrospective evaluation of the (suspected) sepsis patients in the Albert Schweitzer hospital is described. Of all the patients that were treated with antibiotics for suspected sepsis, we studied how many had objective signs of a bacterial infection.
In *chapter 5*, all patients that were tested for influenza in the influenza season of 2017/2018 were studied. Identifying viral infections accurately when patients are admitted is very relevant because it influences the correct treatment, but also because it influences decisions on placement on the ward, such as allocation of isolation rooms. By chart review, we scored predefined clinical symptoms to see if clinical symptoms could be used to distinguish patients with influenza from those without influenza.
In *chapter 6*, we studied if procalcitonin can help to rule out bacteremia in patients in with suspected infections in the Emergency Department. Based on chapter 5, and on our clinical experience, we hypothesized that a viral infection affects the likelihood of bacteremia in patients with suspected infection in the ED. In a cohort of patients in whom the physician apparently doubted between a viral infection and a bloodstream infection, we measured procalcitonin and observed the results. We then studied the likelihood of bacteremia in patients with and without viral infection, and studied the optimal cut-off of procalcitonin to exclude bacteremia in these groups.

In *Chapter 7* a new approach on diagnosing infection was sought. In the ED, many patients have extensive testing for infections including blood cultures. However, blood cultures are often negative, contaminated, and even if they are positive often do not impact clinical care. In this chapter we studied a machine learning algorithm to predict positive blood cultures.

Part three of this thesis discusses the broader implications of sepsis. In *chapter 8* we studied the quality of life in patients with sepsis, that were included in the Phantasi trial. The Phantasi trial was designed to test if prehospital antibiotic treatment could improve the outcome in patient with suspected sepsis. Approximately 1 month after discharge, patients included in the study were sent a survey on the quality of life (SF-36). The results were compared to age-matched controls and are described in the article.

Chapter 9 consists of a letter to the editor. In 2020, an article in Lancet, discussing the global burden of disease, concluded that an estimated 48·9 million (95% uncertainty interval 38·9–62·9) incident cases of sepsis were recorded worldwide in 2017[22]. About 1 in 5 deaths globally, was ascribed to sepsis, with the majority of deaths in sub-Sahar Africa. In our letter, we discuss if naming their condition sepsis can really benefit these patients. Originally this letter was named; What's in a name, referring to Romeo & Juliet, by Shakespeare. Juliet wonders if we call a rose by any other name, would it not smell as sweet? In the letter, we wonder if it really changes anything for our patients if we attach the name sepsis to deaths caused by diarrhea, pneumonia, malaria and HIV.

[22] Rudd, Kristina E et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study The Lancet, Volume 395, Issue 10219, 200 - 211

Part 1

Identifying threat to life

Chapter

2

Prognostic value of early warning scores in the emergency department (ED) and acute medical unit (AMU): A narrative review

Eur J Intern Med. 2017 Nov;45:20-31

R.S. Nannan Panday

T.C. Minderhoud

N. Alam

P.W.B. Nanayakkara

ABSTRACT

Background

A wide array of early warning scores (EWS) have been developed and are used in different settings to detect which patients are at risk of deterioration. The aim of this review is to provide an overview of studies conducted on the value of EWS on predicting intensive care (ICU) admission and mortality in the emergency de- partment (ED) and acute medical unit (AMU).

Methods

A literature search was conducted in the bibliographic databases PubMed and EMBASE, from inception to April 2017. Two reviewers independently screened all potentially relevant titles and abstracts for eligibility.

Results

42 studies were included. 36 studies reported on mortality as an endpoint, 13 reported ICU admission and 9 reported the composite outcome of mortality and ICU admission. For mortality prediction National Early Warning Score (NEWS) was the most accurate score in the general ED population and in those with respiratory distress, Mortality in Emergency Department Sepsis score (MEDS) had the best accuracy in patients with an infection or sepsis. ICU admission was best predicted with NEWS, however in patients with an infection or sepsis Modified Early Warning Score (MEWS) yielded better results for this outcome.

Conclusion

MEWS and NEWS generally had favourable results in the ED and AMU for all endpoints. Many studies have been performed on ED and AMU populations using heterogeneous prognostic scores. However, future stud- ies should concentrate on a simple and easy to use prognostic score such as NEWS with the aim of introducing this throughout the (pre-hospital and hospital) acute care chain.

INTRODUCTION

Studies have consistently shown that clinical deterioration of hospi- talized patients is often preceded by changes in vital signs up to 6 to 24 h before an adverse event [1–4]. However, these changes in vital signs were often underreported or were disregarded[5]. In an effort to prevent these adverse events, several systems have been developed to identify patients that are most likely to deteriorate. By 2015 over 36 Early Warning Systems (EWS) were developed with variable success and rate of implementation [6]. Hospitals worldwide use different EWS and due to the large number of EWS models it might be difficult to determine which EWS is most suitable for different settings in the acute care chain. Most of the studies have investigated the value of EWS in clinical wards. However, large scale studies investigating the value of EWS on top of triage systems in the emergency department (ED) to timely detect deterioration is lacking.

Some groups have developed separate scoring systems specifically designed for medical patients in the ED. An example is the Rapid Emer- gency Medicine Score (REMS) that was introduced in 2004 [7]. In addition, some scores have been developed to be used specifically in certain patient groups. For example the Mortality in Emergency Department Sepsis score (MEDS) for patients with an infectious disease [8], and CURB-65 (acronym for Confusion, Urea nitrogen, Respiratory rate, Blood pressure and age ≥ 65) for patients with pneumonia[9].

Different types of EWS have also been used in the Acute Medical Unit (AMU), which is a department that has been implemented in several countries to optimize care for acutely admitted medical patients [10]. This department is essentially a multi-disciplinary gateway between the emergency department and the ward of the hospital caring for acutely admitted during the first 72 h. However patients can only be admitted from the outpatients setting to the AMU, but not from a ward to the AMU[11,12].

As specialists in (acute) internal medicine, we see our patients both in ED, the AMU and the clinical wards. A uniform scoring system in the acute care chain would be preferable instead of using separate scores for ED, the AMU and the clinical wards. For the purpose of this review, we decided to provide an overview of the prognostic value of various (early warning) scores in predicting mortality or ICU admission that have been studied in medical patients in the ED and in the AMU.

MATERIAL AND METHODS

Study identification/search strategy

To identify all relevant publications, we performed a systematic search on 15 April 2017 in the bibliographic databases PubMed and EMBASE from inception to April 2017. Search terms included controlled terms from MeSH in PubMed, EMtree in EMBASE as well as free text terms. Search terms expressing "EWS" were used in combination with search terms comprising "emergency department" and terms for "Acute Medical Unit" and "adults". The search strategy can be found in appendix 1. Subsequently, the references of identified articles were manually searched for relevant publications.

Eligibility criteria

The literature search generated a total of 1651 references. Hereafter, all titles were re- viewed to determine which studies were possibly con- ducted in the ED and/or AMU and involved the use of Early Warning Scores (EWSs). Thereafter, abstract selection was performed. Studies were selected if they met the following inclusion criteria: (i) the study was a retrospective or prospective observational study; (ii) the study population consisted of patients (16 years and older) at the ED or AMU; (iii) the study used the predictive value of EWS as a primary or secondary outcome; (iv) The predictive value of the EWS was studied for mortality, intensive care admission or a composite outcome of these. Studies were excluded if: (i) they were conducted exclusively on patients from disciplines other than internal medicine; (ii) unclear when the first assessment of EWS was performed (iii) first assessment of EWS was done after the ED or AMU; (iv) if the aim of the study was to determine whether implementation of an EWS led to an improvement in patient mortality and/or ICU admission; (v) no full text was available (in English); (vi) certain publication types such as editorials, letters, legal cases, interviews, posters etc.

Study selection and data collection

Two reviewers (RN and TM) independently screened all potentially relevant articles and abstracts for eligibility. Where required, the full ar- ticle was checked for eligibility criteria. Differences in opinion were re- solved through deliberation. All retrieved abstracts were in English. The authors extracted data from the different studies independently and inserted this in a standardised worksheet, which contained among others the following: study type, inclusion period, study setting, study group size, type of EWS used, the (primary) outcome variables and the conclusions of the study (Flowchart Fig.2.1).

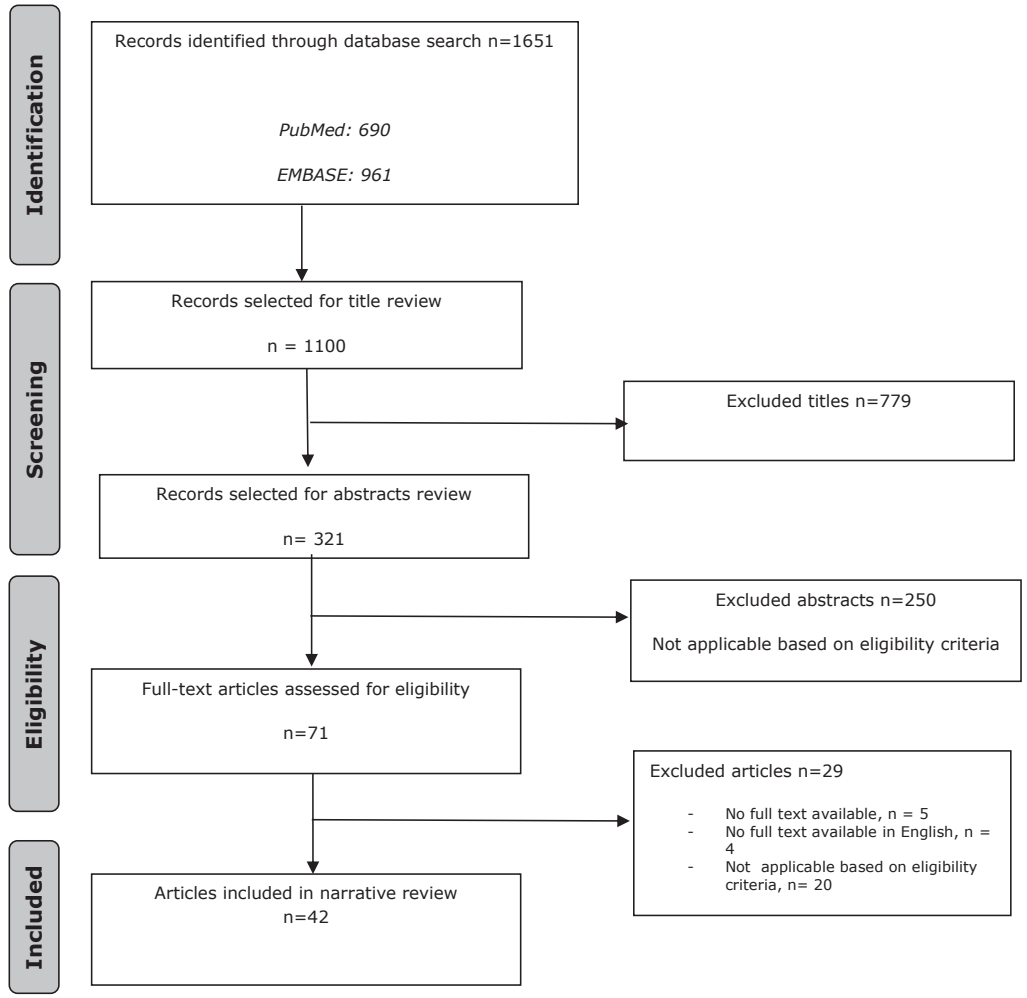
Quality assessment data

Study quality was assessed with the Quality In Prognostic Studies (QUIPS) tool which has been recommended by the Cochrane Prognosis Method Group [13]. The QUIPS consists of six quality criteria: (i) study participation; (ii) study attrition; (iii) prognostic factor measurement; (iv) outcome measurement; (v) study confounding and (vi) statistical analysis and reporting. These six quality criteria can each be scored in either low bias, moderate bias or high bias and thus a conclusion regarding the bias and quality of each individual study can be determined objectively.

Data presentation

Area under the receiver operator characteristic (AUROC) is de- scribed using the standard terms, where AUROC 0.6–0.7 is considered poor, 0.7–0.8 fair, 0.8–0.9 good and N 0.9 is excellent [14] and will as such be applied in our description of the results. In the case that AUROC is not reported, we will use the p value of either the odds ratio (OR), hazard ratio (HR) or relative risk (RR) to determine whether there is a strong correlation between the prognostic score and the out- come. A p value b 0.05 will be considered a significant correlation. A de- tailed overview of the results of the prognostic studies with their respective follow-up moments can be found in Table 2.4 (study outcomes).

Figure 2.1 Flowchart of study selection



RESULTS

Study characteristics

Of the total 1651 citations that were found, 71 articles were assessed for eligibility of which 42 studies were found suitable for inclusion (Figure 2.1 Flowchart). The study characteristics of these 42 studies can be found in Table 2.1 (study characteristics). A total of 25 different types of EWS were identified in these 42 articles. The most frequently used prognostic scores were the Modified Early Warning Score (MEWS), which was applied in 19 studies [15-33], and the National Early Warning Score (NEWS), which was used in 12 studies [18,20,34–43]. Nine studies used the REMS [16,18,19,26,44–48] and seven studied MEDS [18,21, 25, 26,30,33,46]. Several variations of the EWS were used, with slight modifications such as adding

age, adding laboratory values or different cutoff values. An overview of scores and the information needed to cal- culate them, is provided in Table 2.2 (overview of scores).

Study quality

All 42 included studies were assessed using the QUIPS tool. Accord- ing to this tool 18 studies [16,19,20,26–28,30,32,37,39,40,42–46,49,50] were found to have a low risk of bias and were thus regarded to be of high quality, 22 studies [17,18,21–25,29,31,33–36,38,41,47,48,51–55] had a moderate risk of bias and were considered to be of moderate quality and 2 studies [15,56] had a high risk of bias and therefore were considered to be of low quality. The subdomains that were most at risk of high bias were study attrition (n = 9) mainly due to inadequate reporting of data on patient follow-up and study confounding due to in- complete reporting on (possible) confounders (n = 14). Table 2.3 (Risk of bias) provides an overview of the included studies with their respective QUIPS scores.

Prognostic score performance

Due to the heterogeneity of the included studies, results will be pre- sented in three groups: studies that included the general ED population, studies that only included patients with a possible infection or sepsis and studies that specifically included patients who had either a community acquired pneumonia or respiratory distress. In these three groups of patients a distinction will be made for the prognostic value in three domains, which are: mortality, ICU admission and a composite outcome which consists of mortality and ICU admission. Mortality will be split up in in-hospital mortality, short-term (0–30 days) and long-term mortal- ity (N 30 days). A detailed overview of the outcomes of each study can be found in table 2.4 (study outcomes).

Table 2.1 Study characteristics

Author Year	Country	Study Design	Inclusion Period	Study group size	Age, Years, median (IQR/TR), mean ± SD
Abott 2015	United Kingdom	Retrospective	March 2013 – April 2013	431	60.9 ± 22.4
Abott 2016	United Kingdom	Prospective	March 2013 – April 2013	310	63 ±21.8
Alam 2015	Netherlands	Prospective	January 2013 – February 2013	274	U
Armagan 2008	Turkey	Prospective	April 2007 – August 2007	309	57.1 ± 15.3
Barlow 2006	United Kingdom	Retrospective	November 2001 - April 2002 and November 2002 - April 2003	419	74 (TR 16-98)
Bilben 2016	Norway	Prospective	July 2014 - November 2014	246	70.5 (IQR 60-80)
Bulut 2014	Turkey	Prospective	October 2011 - April 2012	2000	61.4 ± 18.9
Burch 2008	South Africa	Prospective	U	790	43 (TR 16-89)
Cattermole 2009	Hong Kong	Prospective	April 2006 – May 2006	330	61.3 ± 20.6
Cattermole 2014	Hong Kong	Prospective	November 2008 - January 2009	234	65.8 ± 18.1
Cetinakya 2016	Turkey	Prospective	February 2015 - June 2015	616	74.3 ± 6.9
Churpek 2017	United States	Retrospective	November 2008 - January 2016	18523	58 ± 18.0
Cildir 2012	Turkey	Prospective	August 2009 - February 2011	230	U
Corfield 2014	United Kingdom	Retrospective	March 2009 - May 2009	2003	72 (IQR 59-81)
Delgado-Hurtado 2016	United States	Retrospective	January 2014 - May 2015	2147	56 (IQR 38-73)
Duckitt 2007	United Kingdom	Prospective	October 2005 - November 2005	1102	U
Dundar 2016	Turkey	Prospective	January 2014 - February 2014	671	75 ± 11
Fairclough 2009	United Kingdom	Prospective	January 2004, January 2005 and January 2006	300	73 (TR 16-99)
Geier 2013	Germany	Prospective	August 2012 - September 2012	125	68.3 ± 18

Table 2.1 (continued)

Author Year	Country	Study Design	Inclusion Period	Study group size	Age, Years, median (IQR/TR), mean ± SD
Ghanem-Zoubi 2011	Israel	Prospective	February 2008 - April 2009	1072	74.7 ± 16.1
Goodacre 2005	United Kingdom	Retrospective	1996 to 2001	5583	63.4 (U)
Groarke 2008	Ireland	Prospective	U	225	64.7 ± 19.1
Ha 2015	Vietnam	Prospective	U	1746	65.9 ± 17.0
Heitz 2010	United States	Retrospective	2005	280	56 (IQR 41-73)
Hilderink 2015	Netherlands	Retrospective	August 2009 - August 2010	600	64.6 ± 17.6
Ho 2013	Singapore	Retrospective	November 2006 - December 2007	1024	U
Hodgson 2016	United Kingdom	Retrospective	March 2012 - February 2014	2361	74 (IQR 67-82)
Howell 2007	United States	Prospective	December 2003 - September 2004	2132	61 (IQR 44.5-77)
Huggan 2015	Singapore	Prospective	May 2011 - June 2011	398	64.6 ± 20.2
Innocenti 2017	Italy	Retrospective	June 2008 - April 2016	742	75.0 ± 14.0
Jo 2016	South Korea	Retrospective	September 2014 - October 2014	4624	57.7 ± 18.9
Köksal 2016	Turkey	Prospective	U	502	62 (TR 18-102)
Nickel 2016	Denmark	Retrospective	October 2008 - Febuary 2009	1201	62.7 ± 18.8
Olsson 2003	Sweden	Prospective	November 1995 - November 1996	1027	70 ± 18.1
Olsson 2004	Sweden	Prospective	October 1995 - November 1996	11751	61.9 ± 20.7
Paterson 2006	United Kingdom	Prospective	October 2013 - November 2013	435	69 (IQR 43-79)
Perera 2011	Sri Lanka	Prospective	June 2009	242	49.4 ± 18.7
Prytherch 2010	United Kingdom	Retrospective	May 2006 - June 2008	39992	67.7 (U)

Table 2.1 (continued)

Sbiti-Rohr 2016	Switzerland	Retrospective	October 2008 - March 2009	925	73 (IRQ 59-82)
Singer 2017	United States	Retrospective	January 2014 – March 2015	22530	54 ± 21.0
Smith 2013	United Kingdom	Retrospective	2006-2008	35585	67.7 (U)
Vorwerk 2009	United Kingdom	Retrospective	January 2006 and January 2007	307	69.7 (U)

IQR = interquartile range, TR = total range, SD= standard deviation, U= unknown

General ED population

Mortality

22 studies [15–17,20,22–24,28,31,36,40,41,43–45,47,48,50–52,54,56] were conducted on mortality as an outcome value and used a total of 12 prognostic scores. MEWS was the most applied method of assessment (n = 8)[15–17,22–24,28,31], followed by REMS (n = 5)[16,44,45,47,48] and NEWS (n = 4)[36,40,41,43]. 12 of the 22 studies assessed in-hospital mortality[15–17,22–24,40,44,47,52,54,56], ten short-term mortality[16,28,31,36,40,41,43,45,48,50] and two studies investigated long-term mortality[41,48].

All studies that used the MEWS as a predictor of in-hospital mortal- ity[15–17,22 24,40,44,47,52,54,56] found it to be an acceptable method of mortality prediction with AUROCs ranging from 0.707 to 0.891. Studies that did not use the AUROC as the outcome varible for the MEWS all found a significant correlation for this outcome (p b 0.05)[15,17,22,24,52,56]. REMS was the second most frequent method of assessment of inhospital mortality and found fair to excellent results (AUROC 0.70 to 0.911)[16,44,47]. Other methods of assessment that were used for inhospital mortality were: Vitalpac Early Warning Score (ViEWS), Rapid Acute Physiology Score (RAPS), Groarke's EWS, Worthing Physiological Score (WPS), Standardised Early Warning Score (SEWS), acute physiology and chronic health evaluation II score (APACHE II) and the NEWS, of which the APACHE II yielded the highest (AUROC 0.901)[23,40,44,47,52,54,56].

Studies that were conducted on short-term mortality[28,31,36,40,41, 43,45,48,50,51] most frequently used the NEWS (n = 4)[36,40,41,43], followed by the MEWS (n = 2)[28,31] and the REMS (n = 2)[45,48]. The NEWS had fair to excellent results (AUROC 0.768–0.94) [36,40,43], which improved when combined with lactate (AUROC 0.96)[40]. Combining NEWS with d-dimer was also beneficial for mortality prediction [41]. Studies that applied the MEWS found fair to good predictive value (AUROC 0.71–0.846)[28,31], which was in line with the results of the findings of the REMS (AUROC 0.712 and HR 1.34)[45,48]. Less frequently used EWS were WPS, ViEWS and the GAP (acronym for Glasgow coma scale, Age and systolic blood Pressure). Of these, the ViEWS had the best prognostic performance (AUROC 0.888)[31,45,50].

Table 2.2 Overview of prognostic scores

Name	Data required	Range
MEWS Modified Early Warning Score	Pulse, respiratory rate, temperature, urinary output, blood pressure, AVPU	0-17
SEWS Standardised Early Warning Score	Pulse, respiratory rate, temperature, blood pressure, spO2, AVPU	0-18
NEWS National Early Warning score	Pulse, respiratory rate, temperature, blood pressure, spO2, oxygen supplemental, AVPU	0-20
VIEWS Vitalpac Early Warning score	Pulse, respiratory rate, temperature, blood pressure, spO2, oxygen supplemental, AVPU	0-21
WPS Worthing Physiological Score	Respiratory rate, pulse, blood pressure, temperature, spO2, AVPU	0-14
REMS Rapid Emergency Medicine Score	Age, blood pressure, blood pressure, Heart rate, Respiratory rate, spO2, GCS	0-26
qSOFA quick Sequential Organ Function Assessment	Blood pressure, respiratory rate, mental status,	0-3
RAPS Rapid Acute Physiology Score	MAP, HR, Respiratory rate, GCS	0-16
RTS Revised trauma score	Blood pressure, respiratory rate GCS,	0-15
GAP (Glasgow- Age- Systolic BLOOD PRESSURE)	GCS, age, blood pressure	3-21
CCI Charlson Comorbidity index	Age, comorbid conditions.	0-6
CURB-65 (Confusion, Urea, Respiratory Rate, Blood pressure, Age)	Mental status, urea, respiratory rate, blood pressure, age (>65)	0-5
CRB-65 (Confusion, Respiratory rate, Blood pressure, Age)	Mental status, respiratory rate, blood pressure, age (>65)	0-4
APACHE II	Vital parameters, GCS, laboratory results.	0-44
PIRO (Predisposition, Insult, Response, Organ dysfunction)	Combination of comorbidity, lab results, current physiological parameters	0-33
PSI Pneumonia Severity Index	Age, type of residence, laboratory values, vital parameters	0-395
MEDS (Mortality in Emergency Department Score)	Functional status, vital parameters, lab values	0-27
SIRS Systemic Inflammatory Response Syndrome	Vital parameters + lab values	0-4
SCC (Simple Clinical Score)	Based on ABCDEF parameters: A; age, airway, spO2 B; breathing (resp rate) C: blood pressure/ pulse D: stroke, altered mental status, pulse, E: ECG (abnormal ECG) F: fever	0-21
MEES Mainz Emergency Evaluation Score	GCS, HR, respiratory rate, ECG, level of pain, blood pressure, spO2	6-28
PEDS Prince of Wales Emergency Department Score	Blood pressure, GCS, glucose, HCO3, white cell count, metastatic cancer history	-2 - 58
THERM The Resuscitation Management Score	GCS, HCO3, blood pressure	0-37
ESI Emergency Severity Index	Clinical judgement, heart rate, respiratory rate, spO2	0-5
CREWS Chronic Respiratory Early Warning Score	Pulse, respiratory rate, temperature, blood pressure, spO2, Oxygen supplemental, AVPU	0-20

GCS= Glasgow Coma Scale, HR = Heart Rate, MAP = Mean Arterial Pressure

Two studies investigated long-term mortality [41,48]. Nickel et al.[41] concluded that a combination of NEWS and d-dimer could predict patients at risk of 1-year mortality, while similar results were found for the REMS in a study by Olsson et al. for 90-day, 1-year mortality and 4.7- year mortality [48].

ICU admission

Nine studies [15,16,22,23,28,36,40,43,52] were conducted on the predictive value of prognostic scores on ICU admission. Five studies used the MEWS [15,16,22,23,28], three studies used the NEWS [36,40, 43] or a derivate of it and one study included the REMS [16]. Groarke et al. used a self-developed EWS [52], while Dundar et al. included the ViEWS [23]. Three studies [15,22,23] found the ICU admission rate to increase significantly (p b 0.05) for every point the MEWS increased, while two studies [16,28] considered it to be a poor predictor with AUROC of 0.538 and 0.49 respectively. All included studies [36,40,43] found a strong correlation between the NEWS and ICU admission. In- crease in NEWS led to a significant increase in ICU admission [36], the addition of lactate improved the AUROC, increasing from 0.78 to 0.83 [40]. A study by Smith et al. even found the NEWS outperforming 33 other EWS that had previously been evaluated [43]. The EWS of Groarke [52] and the ViEWS [23] were both strongly correlated with ICU admission, while the REMS had a poor predictive value for this outcome (AUROC = 0.598) [16].

Composite outcome of ICU admission and mortality

Eight studies used a composite outcome of ICU admission and mortality [18,19,27,29,32,34,35,40.] Five of the studies included the MEWS [18,19,27,29,32], while four studies used a NEWS or a derivate [18,34, 35,40] as method of assessment and two included the REMS [18,19].

Studies that applied the MEWS in the general ED population found it to be a poor predictor for composite outcomes with AUROC for this value ranging from 0.668 to 0.680 and concluded a cut-off value of 5 for the MEWS to be optimal [27,29,32]. In patients admitted to the resuscitation room the MEWS was superior to the NEWS (AUROCs 0.730 - 0.761 vs 0.71) [18,19]. The NEWS outperforms the REMS with an AUROC of 0.696 in these patients in one study [19].

The NEWS was also used in the general ED population and deemed to be a fair predictor of the composite outcome in the general ED population (AUROC 0.74–0.79) [34,35,40]. Adding the lactate to the NEWS further improved the composite outcome to an AUROC of 0.84 [40]. Eight other EWS were used for the composite outcome, which were the Revised Trauma Score (RTS), APACHE II, Prince of Wales Emergency Department Score (PEDS), Mainz Emergency Evaluation Score (MEES), The Resuscitation Management Score (THERM), Simple Clinical Score (SCS), Patient At Risk Score (PARS) and MEDS. Of these eight EWS, The PEDS yielded the highest AUROC (0.909) [18,19,35].

Prognostic scores in patients with (suspected) infection or sepsis

Mortality

Ten studies [20,21,25,26,30,33,38,46,49,55] were identified that studied early warning scores in patients with suspected infection or sepsis. With the exception of two studies [38,55], all studied and compared multiple scoring systems. Six studies assessed in-hospital mortality [20,25,26,38,46,55], five short-term mortality [21,26,30,33,49], while one study [26] was conducted on long-term mortality.

The most common methods of prognostic outcome scoring were the MEDS (n = 7) [21,25,26,30,33,46,49] and MEWS (n = 6) [20,21,25,26,30,33]. Other common prognostic scores were Charlson Comorbidity Index (CCI) [21,25,30], REMS [26,46,49] and quick Sequential Organ Function Assessment (qSOFA) [20,30,55], which were each used in three studies.

In studies that investigated the in-hospital mortality, the MEWS (n= 3) [20,25,26] and the MEDS (n = 3) [25,26,46] were the most utilized methods of assessment. The prognostic value of the MEWS was poor to fair (AUROC 0.642–0.73) [20,25,26], while the MEDS performed better with fair to good prognostic value (0.73–0.871) [25,26,46]. The follow- ing eight prognostic scores were also investigated: CURB-65, Simple Clinical Score (SCS), Emergency Severity Index (ESI), CCI, Rapid Emer- gency Medicine Score (REMS), NEWS, Systemic Inflammatory Syn- drome (SIRS) and qSOFA tested in this group of patients [20,25,26,38, 46,55]. The REMS was the best performing prognostic score among them (AUROC 0.80) [46], thereby outperforming the qSOFA (AUROC 0.76) [20,55] which was specifically developed to detect mortality in pa- tients with an infection or sepsis [20].

Five studies assessed short-term mortality [21,26,30,33,49], all five of them included the MEDS as a prognostic score and found varying results, with the MEDS being a poor to good predictor (AUROC 0.674– 0.82). The MEWS was the second most common prognostic value for this outcome, four studies [21,26,30,33] included this prognostic score and found this score to be a poor to fair predictor of short-term mortality (AUROC 0.608–0.72). Other scores that were included in the studies on short-term mortality were: SCS, REMS, qSOFA, Sequential Organ Function Assessment (SOFA), CCI, APACHE II, PIRO (acronym for Predisposition, Insult response, organ dysfunction), CURB-65, RAPS and Near Patient lactate Test (NPT). The CURB-65 and REMS had the best results of these (AUROC 0.78 for both scores) [21,26,30,33,49].

Ghanem et al. [26] conducted the only study on long-term mortality and assessed this with the MEWS, SCS, MEDS and REMS. For 60-day mortality, the SCS was the most accurate tool (AUROC 0.76).

ICU admission

Corfield et al. and Singer et al. conducted studies on ICU admission predictability of the NEWS and qSOFA respectively. The predictive value was concluded to be poor (AUROC 0.67 vs 0.61) [38,55]. A study by Innocenti et al. [30] found a significant correlation between ICU ad- mission and each of the following EWS: MEWS, qSOFA, CCI, SOFA, APACHE II and MEDS. All EWS had a strong significant correlation with ICU admission in this group of patients (p b 0.05).

Composite outcome of ICU admission and mortality

Two studies were conducted on the composite outcome of ICU and mortality [20,38]. One study [20] investigated the composite outcome of ICU admission and mortality for the NEWS, MEWS, qSOFA and SIRS. The NEWS was the most accurate, while the SIRS was the least accurate (AUROC 0.72 vs 0.58) [20]. A study by Corfield et al. found similar results with the NEWS having an AUROC of 0.70 for this outcome [38].

Patients with community acquired pneumonia or respiratory distress

Four studies were conducted in subgroup of patients with community acquired pneumonia or respiratory distress [37,39,42,53]. All four studies reported on mortality, of which one study reported in-hospital mortality [39], two studies investigated short-term mortality [42,53] and another two studies [37,42] long-term mortality. Hodgson et al. investigated the value of in-hospital mortality for the NEWS and the Chronic Respiratory Early Warning Score (CREWS), and found the NEWS to be the superior of these scores (AUROC 0.74 vs 0.062) [39]. Short-term mortality studies [42,53] included a multitude of prognostic scores, namely: NEWS, Pneumonia Severity Index (PSI), CURB-65, CRB- 65, SIRS and SEWS of which the PSI was the best performing scores (AUROC 0.80) [42,53].

Sbiti-Rohr et al. [42] was the only study that investigated the value of prognostic scores for the prediction of long-term in this subgroup and found the PSI to be the most reliable predictor for 6-year mortality, when compared to the NEWS and CURB-65 (AUROC 0.79, 0.60 and 0.73 respectively). Sbiti-Rohr et al. also investigated the value of these three prognostic scores for ICU admission and found the NEWS to be outperforming both the CURB-65 and PSI for this outcome (AUROC 0.73, 0.64 and 0.64 respectively) [42].

DISCUSSION

We conclude that a wide array of prognostic scores is in use in different settings with a considerable heterogeneity in the used parameters. All studies that included AUROC for one or more outcome measures found AUROCs which were far greater than N 0.5, which is the cut-off for correlation that is reached by chance alone [14]. The majority of the studies that objectified the performance of EWS with either HR, OR, RR or p value found a strong significant correlation with their out- come variables.

Mortality was the most prevalent prognostic outcome, followed by ICU admission and the composite outcome of mortality and ICU admission. Patient population and time to follow-up greatly influenced the performance of EWS, with some scores reaching good to excellent AUROC in some populations, but only poor AUROC in others. For in-hospital mortality MEWS was the most reliable method of assessment in the general population,

Table 2.3 Risk of bias

Study	Overall risk of bias	Study participation	Study Attrition	Prognostic Factor measurement	Outcome measurement	Study confounding	Statistical analysis and reporting
Abott 2016	Moderate	L	M	L	L	M	L
Abott 2015	Moderate	L	M	L	L	H	M
Alam 2015	Moderate	M	M	L	L	M	L
Armagan 2008	High	M	H	L	L	H	H
Barlow 2007	Moderate	M	H	L	L	N/A	L
Bilben 2014	Low	L	N/A	L	L	L	L
Bulut 2014	Low	L	N/A	L	L	M	L
Burch 2008	Moderate	M	M	L	L	M	M
Cattermole 2009	Low	L	L	L	L	M	M
Cattermole 2013	Moderate	L	M	L	L	H	M
Cetinkaya 2016	Moderate	L	H	L	L	H	L
Churpek 2017	Low	L	L	L	L	N/A	L
Cildir 2012	Moderate	M	H	L	M	L	L
Corfield 2013	Moderate	M	M	L	L	M	M
Delgado-Hurtado 2016	Moderate	H	M	L	L	M	L
Duckitt 2007	Moderate	M	L	L	L	H	M
Dundar 2016	Moderate	M	H	L	L	H	L
Fairclough 2009	Moderate	M	H	L	L	M	H
Geier 2013	Moderate	L	H	L	L	H	L
Ghanem-Zoubi 2011	Low	L	L	L	L	N/A	L
Goodacre 2005	Low	M	L	L	L	L	L
Groarke 2008	Moderate	H	L	L	L	H	M
Ha 2015	Low	M	L	L	L	L	L
Heitz 2010	Low	M	L	L	L	L	L
Ho 2013	Low	M	L	L	L	N/A	L
Hodgson 2016	Low	L	H	L	L	L	L
Howell 2007	Low	L	M	L	L	N/A	L
Hilderink 2015	Low	L	L	L	L	N/A	L
Huggan 2015	Moderate	L	M	L	L	M	L
Innocenti 2017	Low	M	L	L	L	N/A	M
Jo 2016	Low	L	L	L	L	L	L
Koksal 2016	Moderate	M	L	L	L	H	L
Nickel 2016	Moderate	M	L	L	L	M	M
Olsson 2003	Moderate	M	H	L	L	M	L
Olsson 2004	Moderate	M	H	L	L	H	L
Paterson 2006	High	M	H	H	H	H	L
Perera 2011	Low	L	L	H	L	L	L
Prytherch 2010	Low	M	N/A	L	L	L	L
Sbiti-Rohr 2016	Low	L	L	L	L	M	L
Singer 2017	Moderate	M	M	L	M	M	H
Smith 2013	Low	L	N/A	L	L	H	L
Vorwerk 2009	Moderate	M	N/A	L	L	H	L

L low risk of bias, M moderate risk of bias, H high risk of Bias, N/A Not applicable

MEDS reached the highest AUROC in those with an infection and NEWS was the most reliable in those with community acquired pneumonia. Short-term mortality was best predicted by the NEWS in the general population, the MEDS in the population with an infection and the PSI in those with community acquired pneumonia. NEWS was the best predictor of long-term mortality in the general population and those with community acquired pneumonia, but was outperformed by the SCS in patients with an infection.

Innocenti et al. [30] studied the predictive value of multiple prognostic scores for ICU admission in patients with an infection. All prognostic scores, including the MEWS, had a significant correlation with ICU ad- mission. Although this study did not include it, other studies found that NEWS outperformed qSOFA [38,55] despite the fact that the qSOFA was specifically developed for the purpose of detecting deterioration in patients with an infection [20]. Moreover, NEWS had the strongest correlation of all prognostic scores for ICU admission in the general population and those with community acquired pneumonia. For the composite outcome NEWS outperformed the other included EWS in both the general population and in those with an infection or sepsis. Studies in patients with respiratory symptoms did not include composite outcomes.

EWS are developed with the purpose of detecting patient deteriora- tion and are espe- cially applied in the acute healthcare chain and should therefore be simple, reliable and comprehensible among all the healthcare providers in this chain. Therefore uniformity in the used EWS across all departments of the healthcare chain might be beneficial for the improvement of patient care. The ideal prognostic score should be easy to calculate, preferably without the need of laboratory results and should show good predictive value. Simple bedside systems such as RTS, CRB-65 or qSOFA are appealing due to their sim- plicity and ease of use, however it is difficult to combine both simplicity and accuracy, as this review shows that simple prognostic scores were outperformed by more elaborate scoring systems such as the NEWS and MEDS.

As this study shows, staffs in different hospitals and settings are like- ly to use different EWS and often multiple EWS are used in one hospital [57]. The fact that pre-hospital care- givers also have EWS which differ from the hospital causes further fragmentation. Recog- nizing this issue, an effort was made in the United Kingdom in 2015 to use one uniform EWS throughout the whole healthcare chain in order to reduce the chance of miscom- munication between the different disciplines and caregivers [57]. Although the results of this implementation are not yet known, it is likely to increase the detection of patients that are at risk chain and therefore optimizing adequate treatment.

Although EWS recognize patients which are at risk of deterioration, they are unable to aid in the prevention of these adverse events. More-over, EWS are scored on the basis of vital parameters and a study by Baker et al.[58] has shown that recognition and treat- ment of deteriorating vital parameters does not necessarily lead to improved patient outcomes. Recognition and treatment of systolic hypotension at admission was the only factor that significantly improved patient survival, while treating other factors such as GCS and tachycardia did not im- prove patient outcome. This emphasizes that EWS

can predict adverse outcomes, but unfortunately carry few implications regarding improvement of patient outcome.

This study has several strengths. To the best of our knowledge this is the first narrative review that describes the prognostic value of different prognostic scores in the acute setting of the ED and AMU. Previous work has been conducted on the effects of implementing an EWS on patient outcome in ED and ward patients [59], but this was aimed at finding if implementation of an EWS led to better clinical outcomes instead of the prognostic value. No studies have evaluated the prognostic value of EWS itself in both the ED and AMU. Furthermore, we have described the outcomes of all included EWS in each of the studies included in this review as to give an objective overview of the usability of these EWS in the ED and AMU.

Despite these strengths, there are also some limitations. First, this study only assessed the value of prognostic stores on the ED and AMU and thereby excluded multiple studies that have been conducted on the ward. Secondly, most of the studies in this review did not describe if the measurement of early warning scores was done electronically. Previous studies have shown that early warning studies are often incomplete or miscalculated and that electronic measurement leads to less underscoring and improvement of the accuracy of EWS values and reduced mortality [6,43,50,60–62].

Thirdly, due to the heterogeneity in sample size, follow-up time and number of studies that tested each individual EWS, a definitive conclusion as to which prognostic score should be used cannot be reached.

Fourth, the majority of the studies included in this review were conducted in developed countries. Previous studies have shown that EWS which perform well in developed countries, may not necessarily do so in developing countries [63]. The hospitalized population in developing countries differs from those in developed countries and resources are limited. This might hamper the implementation of certain EWS such as the NEWS and ViEWS, which depend on scoring oxygen supplementation. Oxygen supplementation is not always available for patients in developing countries, thus making these EWS less reliable in such conditions [64]. A recent large scale study by Moore et al. [65] derived a new EWS called the Universal Vital Assessment (UVA) score, which took these factors in account and found it to outperform the MEWS and qSOFA in developing countries.

Table 2.4 Study outcomes

Study, Year	Study population	EWS assessed	Moment of assessment	Endpoints	AUROC/RR/OR/HR/ correlation for ICU admission*	AUROC/RR/OR/HR/ correlation for mortality *	AUROC/RR/OR/HR/ correlation for composite outcome of ICU admission and mortality *	Overall risk of bias according to QUIPS
Abott 2016	General ED patients	NEWS in combination with blood gas variables	AMU	Composite of mortality and ICU admission within 2 days	N/A	N/A	AUROC NEWS 0.74 OR NEWS 1.48 OR NEWS + lactate: 1.19 OR NEWS + glucose: 1.02 OR NEWS + base excess: 1.13	Moderate
Abott 2015	General ED patients	NEWS and PARS	AMU	Composite of mortality and ICU admission within 2 days	N/A	N/A	OR NEWS: 1.54 OR PARS: 1.42	Moderate
Alam 2015	General ED patients	NEWS	ED	ICU admission and 30-day mortality	Correlation NEWS: p=0.003	AUROC NEWS: 0.768	N/A	Moderate
Armagan 2008	General ED patients	MEWS	ED	ICU admission, ED mortality and in-hospital mortality	OR MEWS: 1.95	OR MEWS ED: 35.13 OR MEWS In-hospital: 14.81	N/A	High
Barlow 2007	ED patients with suspected CAP	CURB-65, CRB-65, SIRS and SEWS	ED	30-day mortality	N/A	AUROC CURB-65: 0.78 AUROC CRB-65: 0.73 AUROC SIRS: 0.68 AUROC SEWS 0.64	N/A	Moderate
Bilben 2014	ED patients with respiratory distress	NEWS	ED	90-day mortality	N/A	AUROC NEWS: 0.809	N/A	Low
Bulut 2014	General ED patients	MEWS, REMS	ED	ICU admission and in-hospital mortality	AUROC MEWS: 0.538 AUROC REMS: 0.589	AUROC MEWS: 0.630 AUROC REMS: 0.707	N/A	Low
Burch 2008	General ED patients	MEWS	ED	In-hospital mortality	N/A	Correlation MEWS: p<0.001	N/A	Moderate
Cattermole 2009	ED Patients admitted to resuscitation room	RTS, REMS, MEWS, APACHE II and PEDS	ED	Composite of mortality and ICU admission within 7 days	N/A	N/A	AUROC RTS: 0.748 AUROC REMS: 0.696 AUROC MEWS: 0.761 AUROC APACHE II: 0.733 AUROC PEDS: 0.909	Low

Table 2.4 (continued)

Study, Year	Study population	EWS assessed	Moment of assessment	Endpoints	AUROC/RR/OR/HR/ correlation for ICU admission*	AUROC/RR/OR/HR/ correlation for mortality *	AUROC/RR/OR/HR/ correlation for composite outcome of ICU admission and mortality *	Overall risk of bias according to QUIPS
Cattermole 2013	ED Patients admitted to resuscitation room	MEES, PEDS, THERM, MEWS, NEWS, REMS, SCS and MEDS	ED	Composite of mortality and ICU admission within 7 days	N/A	N/A	AUROC MEES: 0.75 AUROC PEDS: 0.75 AUROC THERM: 0.84 AUROC MEWS: 0.73 AUROC REMS: 0.70 AUROC SCS: 0.70 AUROC MEDS: 0.59	Moderate
Cetinkaya 2016	General ED patients > 65 years	ViEWS in combination with lactate	ED	24-hour mortality	N/A	AUROC ViEWS with lactate: 0.872	N/A	Moderate
Churpek 2017	ED patients with suspected infection	qSOFA, SIRS, MEWS and NEWS	ED	In-hospital mortality, composite outcome of in-hospital mortality and ICU stay	N/A	AUROC qSOFA: 0.69 AUROC SIRS: 0.65 AUROC MEWS: 0.73 AUROC NEWS: 0.77	AUROC qSOFA: 0.62 AUROC SIRS: 0.58 AUROC MEWS: 0.68 AUROC NEWS: 0.72	Low
Cildir 2012	ED patients with suspected sepsis	MEDS, MEWS and CCI	ED	5-day and 28-day mortality	N/A	5-day AUROC MEDS: 0.716 5-day AUROC MEWS: 0.625 5-day AUROC CCI: 0.637 28-day AUROC MEDS: 0.772 28-day AUROC MEWS: 0.608 28-day AUROC CCI: 0.647	N/A	Moderate
Corfield 2013	ED patients with suspected sepsis	NEWS	ED	ICU admission, in-hospital mortality and composite outcome of in-hospital mortality and ICU stay	AUROC NEWS: 0.67	AUROC NEWS: 0.70	AUROC NEWS: 0.70	Moderate

Table 2.4 (continued)

Study, Year	Study population	EWS assessed	Moment of assessment	Endpoints	AUROC/RR/OR/HR/ correlation for ICU admission*	AUROC/RR/OR/HR/ correlation for mortality *	AUROC/RR/OR/HR/ correlation for composite outcome of ICU admission and mortality *	Overall risk of bias according to QUIPS
Delgado-Hurtado 2016	General ED patients	MEWS	ED	ICU admission and in-hospital mortality	Correlation MEWS: p<0.0001	Correlation MEWS: p<0.0001	N/A	Moderate
Duckitt 2007	General ED patients	WPS	ED	In-hospital mortality	N/A	AUROC WPS: 0.74	N/A	Moderate
Dundar 2016	General ED patients ≥ 65 years	MEWS and ViEWS	ED	ICU admission and in-hospital mortality	Correlation MEWS: p<0.0001 Correlation ViEWS: p<0.0001	AUROC MEWS: 0.891 AUROC ViEWS: 0.900	N/A	Moderate
Fairclough 2009	General ED patients	MEWS	AMU	In-hospital mortality	N/A	RR MEWS >4: 7.8	N/A	Moderate
Geier 2013	ED patients with suspected sepsis	ESI, MEWS, MEDS, CCI	ED	In-hospital mortality	N/A	AUROC ESI: 0.617 AUROC MEWS: 0.642 AUROC MEDS: 0.871 AUROC CCI: 0.673	N/A	Moderate
Ghanem-Zoubi 2011	ED patients with suspected sepsis	MEWS, SCS, MEDS and REMS	ED	In-hospital mortality, 30-day mortality and 60-day mortality	N/A	In-hospital AUROC MEWS: 0.69 In-hospital AUROC SCS: 0.77 In-hospital AUROC MEDS: 0.73 In-hospital AUROC REMS: 0.77 30-day mortality AUROC MEWS: 0.67 30-day mortality AUROC SCS: 0.77 30-day mortality AUROC MEDS: 0.75 30-day mortality AUROC REMS: 0.76 60-day mortality AUROC MEWS: 0.65 60-day mortality SCS: 0.76 60-day mortality MEDS: 0.74 60-day mortality REMS: 0.74	N/A	Low

Table 2.4 (continued)

Study, Year	Study population	EWS assessed	Moment of assessment	Endpoints	AUROC/RR/OR/HR/ correlation for ICU admission*	AUROC/RR/OR/HR/ correlation for mortality *	AUROC/RR/OR/HR/ correlation for composite outcome of ICU admission and mortality *	Overall risk of bias according to QUIPS
Goodacre 2005	General ED patients	RAPS and REMS	ED	In-hospital mortality	N/A	AUROC RAPS: 0.64 AUROC REMS: 0.74	N/A	Low
Groarke 2008	General ED patients	Self-developed EWS	AMU	ICU admission and in-hospital mortality	OR self-developed EWS: 3.35	OR self-developed EWS: 2.19	N/A	Moderate
Ha 2015	General ED patients	REMS and WPS	ED	30-day mortality	N/A	AUROC REMS: 0.712 AUROC WPS: 0.797	N/A	Low
Heitz 2010	General ED patients	MEWS	ED	Composite outcome of ICU admission and 24-hour mortality	N/A	N/A	AUROC MEWS: 0.668	Low
Hilderink 2015	ED patients with suspected sepsis	MEDS, CURB-65, APACHE II, RAPS and REMS	ED	28-day mortality	N/A	AUROC MEDS: 0.82 AUROC CURB-65: 0.78 AUROC APACHE II: 0.71 AUROC RAPS: 0.72 AUROC REMS: 0.78	N/A	Low
Ho 2013	General ED patients	MEWS	ED	ICU admission and 30-day mortality	AUROC MEWS: 0.49	AUROC MEWS: 0.71	N/A	Low
Hodgson 2016	ED patients with suspected exacerbation COPD	NEWS, CREWS and S-NEWS	AMU	In-hospital mortality	N/A	AUROC NEWS: 0.74 AUROC AUROC CREWS: 0.62 AUROC S-NEWS: 0.62	N/A	Low
Howell 2007	ED patients with suspected sepsis	MEDS, REMS and CURB-65	ED	In-hospital mortality	N/A	AUROC MEDS: 0.85 AUROC REMS: 0.80 AUROC CURB-65: 0.79	N/A	Low
Huggan 2015	General ED patients	MEWS	ED	Composite outcome of in-hospital mortality and ICU admission	N/A	N/A	HR MEWS >4: 5.50	Moderate

Table 2.4 (continued)

Study, Year	Study population	EWS assessed	Moment of assessment	Endpoints	AUROC/RR/OR/HR/ correlation for ICU admission*	AUROC/RR/OR/HR/ correlation for mortality *	AUROC/RR/OR/HR/ correlation for composite outcome of ICU admission and mortality *	Overall risk of bias according to QUIPS
Innocenti 2017	ED patients with suspected sepsis	MEWS, qSOFA, CCI, SOFA, APACHE II, MEDS and PIRO	ED	ICU admission and 30-day mortality	Correlation MEWS: p<0.0001 Correlation qSOFA: p<0.0001 Correlation SOFA: p<0.0001 Correlation CCI: p=-0.0003 Correlation APACHE II: p<0.0001 Correlation MEDS: p<0.0001 Correlation PIRO: p<0.0001	AUROC MEWS: 0.662 AUROC qSOFA: 0.625 AUROC SOFA: 0.696 AUROC CCI: 0.596 AUROC APACHE II: 0.756 AUROC MEDS: 0.674 AUROC PIRO: 0.646	N/A	Low
Jo 2016	General ED patients	NEWS and NEWS-L	ED	ICU admission, in-hospital mortality, 2-day mortality, 7-day mortality and composite outcome of 2-day mortality and ICU admission	AUROC NEWS: 0.78 AUROC NEWS-L: 0.83	In-hospital AUROC NEWS: 0.84 2-day AUROC NEWS: 0.94 7-day AUROC NEWS: 0.91 In-hospital AUROC NEWS-L: 0.87 2-day AUROC NEWS-L: 0.96 7-day AUROC NEWS-L: 0.94	AUROC NEWS: 0.79 AUROC NEWS-L: 0.84	Low
Koksal 2016	General ED patients	MEWS and GAP	ED	28-day mortality	N/A	AUROC MEWS: 0.846 AUROC GAP: 0.821	N/A	Moderate
Nickel 2016	General ED patients	NEWS in combination with d-dimer	ED	30-day and 1-year mortality	N/A	30-day OR NEWS $\geq$3 with elevated d-dimer vs NEWS $\geq$3 with normal d-dimer : 3.7 1-year OR NEWS $\geq$3 with elevated-dimer vs NEWS $\geq$3 with normal d-dimer : 4.1	N/A	Moderate
Olsson 2003	General ED patients	RAPS, REMS and APACHE II	ED	In-hospital mortality	N/A	AUROC RAPS: 0.870 AUROC REMS: 0.911 AUROC APACHE II: 0.901	N/A	Moderate

Table 2.4 (continued)

Study, Year	Study population	EWS assessed	Moment of assessment	Endpoints	AUROC/RR/OR/HR/ correlation for ICU admission*	AUROC/RR/OR/HR/ correlation for mortality *	AUROC/RR/OR/HR/ correlation for composite outcome of ICU admission and mortality *	Overall risk of bias according to QUIPS
Olsson 2004	General ED patients	REMS	ED	7-day mortality, 30-day mortality, 90-day mortality, 1-year mortality and 4.7 years mortality	N/A	7-day REMS HR: 1.34 30-day REMS HR: 1.30 90-day REMS HR: 1.29 1-year REMS HR: 1.26 4.7 years REMS HR: 1.26	N/A	Moderate
Paterson 2006	General ED patients	SEWS	ED	In-hospital mortality	N/A	Correlation SEWS: p<0.001	N/A	High
Perera 2011	General ED patients	MEWS	ED	Composite outcome of in-hospital mortality and ICU admission	N/A	N/A	AUROC MEWS: 0.68	Low
Prytherch 2010	General AMU patients	ViEWS vs 33 other EWS	AMU	24-hour mortality	N/A	AUROC ViEWS: 0.888 AUROC 33 other EWS: 0.803 to 0.850	N/A	Low
Sbiti-Rohr 2016	ED patients with suspected CAP	NEWS, PSI and CURB-65	ED	ICU admission, 30-day mortality, 180-day mortality and 6-year mortality	AUROC NEWS: 0.73 AUROC PSI: 0.64 AUROC CURB-65: 0.64	30-day AUROC NEWS: 0.65 30-day AUROC PSI: 0.80 30-day AUROC CURB-65: 0.72 180-day AUROC NEWS: 0.62 180-day AUROC PSI: 0.76 180-day AUROC CURB-65: 0.69 6-year AUROC NEWS: 0.60 6-year AUROC PSI: 0.79 6-year AUROC CURB-65: 0.73	N/A	Low
Singer 2017	ED patients with suspected infection	qSOFA	ED	ICU admission and in-hospital mortality	AUROC qSOFA: 0.61	AUROC qSOFA: 0.76	N/A	Moderate
Smith 2013	General AMU patients	NEWS vs 33 other EWS	AMU	24-hour mortality and ICU admission	AUROC NEWS: 0.857 AUROC 33 other EWS: 0.570 to 0.827	AUROC NEWS: 0.894 AUROC 33 other EWS: 0.813 to 0.858	N/A	Low
Vorwerk 2009	ED patients diagnosed with suspected sepsis	MEDS, MEWS, NPT	ED	28-day mortality	N/A	AUROC MEDS: 0.82 AUROC MEWS: 0.72 AUROC NPT: 0.62	N/A	Moderate

* Where available area under receiver operating characteristic (AUORC) are presented, otherwise relative risk (RR), odds ratios (OR) or hazard ratios (HR) are shown. If none of these values was available, a significance level of correlation between the outcome and the prognostic score is presented.

Abbreviations: AMU= Acute Medical Unit, APACHE-II= Acute Physiology And Chronic Health Evaluation II, CAP= Community Acquired Pneumonia, CCI= Charlson Comorbidity Index, CRB-65= acronym for Confusion, Urea nitrogen, Respiratory rate, Blood pressure and age ≥ 65, CREWS= Chronic Respiratory Early Warning Score, CURB-65= acronym for Confusion, Urea nitrogen, Respiratory rate, Blood pressure and age ≥ 65 years, ED= Emergency Department, ESI= Emergency Severity Index, EWS= Early Warning Score, GAP= Glasgow, Age, Systolic blood Pressure, MEDS= Mortality in the Emergency Department Sepsis score, MEES= Mainz Emergency Evaluation Score, MEWS= Modified Early Warning Score, N/A = Not Applicable, NEWS= National Early Warning Score, NEWS-L= NEWS combined with lactate, PARS= Patient At Risk Score, PEDS= Prince of wales Emergency Department Score, PIRO= acronym for Predisposition, Infection, Response and Organ failure, qSOFA= quick Sequential Organ Function Assessment, PIRO= Predisposition, Insult Response, Organ dysfunction, PSI= Pneumonia Severity Index, NPT= Near Patient lactate Test, QUIPS = Quality In Study Prognosis tool, RAPS= Rapid Acute Physiology Score, REMS= Rapid Emergency Medicine Score, RTS= Revised Trauma Score, S-NEWS = Salford National Early Warning Score, SCS= Simple Clinical Score, SEWS= Standardised Early Warning Score, SIRS= Systemic Inflammatory Response Syndrome, SOFA= Sequential Organ Function Assessment, THERM= The Resuscitation Management score, ViEWS= Vitalpac Early Warning Score, WPS= Worthing Physiological Score

** Where available area under receiver operating characteristic (AUORC) are presented, otherwise relative risk (RR), odds ratios (OR) or hazard ratios (HR) are shown. If none of these values was available, a significance level of correlation between the outcome and the prognostic score is presented.*

AMU= Acute Medical Unit, APACHE-II= Acute Physiology And Chronic Health Evaluation II, CAP= Community Acquired Pneumonia, CCI= Charlson Comorbidity Index, CRB-65= acronym for Confusion, Urea nitrogen, Respiratory rate, Blood pressure and age > 65, CREWS= Chronic Respiratory Early Warning Score, CURB-65= acronym for Confusion, Urea nitrogen, Respiratory rate, Blood pressure and age > 65 years, ED= Emergency Department, ESI= Emergency Severity Index, EWS= Early Warning Score, GAP= Glasgow, Age, Systolic blood Pressure, MEDS= Mortality in the Emergency Department Sepsis score, MEES= Mainz Emergency Evaluation Score, MEWS= Modified Early Warning Score, N/A = Not Applicable, NEWS= National Early Warning Score, NEWS-L= NEWS combined with lactate, PARS= Patient At Risk Score, PEDS= Prince of wales Emergency Department Score, PIRO= acronym for Predisposition, Infection, Response and Organ failure, qSOFA= quick Sequential Organ Function Assessment, PIRO= Predisposition, Insult Response, Organ dysfunction, PSI= Pneumonia Severity Index, NPT= Near Patient lactate Test, QUIPS = Quality In Study Prognosis tool, RAPS= Rapid Acute Physiology Score, REMS= Rapid Emergency Medicine Score, RTS= Revised Trauma Score, S-NEWS = Salford National Early Warning Score, SCS= Simple Clinical Score, SEWS= Standardised Early Warning Score, SIRS= Systemic Inflammatory Response Syndrome, SOFA= Sequential Organ Function Assessment, THERM= The Resuscitation Management score, ViEWS= Vitalpac Early Warning Score, WPS= Worthing Physiological Score

CONCLUSION

Evidence regarding the prognostic value of EWS in the ED and AMU is hampered by the use of a large array of similar, yet slightly different scoring systems. A reliable comparison among these systems is therefore difficult. The ideal EWS is both simple and accurate, with few chances of calculation errors and can be used in the whole acute care chain. No prognostic score can impeccably detect all patients at risk of an adverse outcome. Therefore EWS should not replace clinical judgement, but should instead be used as a complementary method of judgement which assists clinicians in their decision making. Future studies should be conducted on the development and implementation of a simple but accurate EWS for the whole acute healthcare chain, which might lead to improved care in both developing and developed countries. These studies should be conducted with at least the NEWS in the general population and the MEDS in those with an infection as reference score, as we have found these scores to be best performing prognostic scores.

REFERENCES

1 Hillman KM, Bristow PJ, Chey T, Daffurn K, Jacques T, Norman SL, et al. Duration of life-threatening antecedents prior to intensive care admission. Intensive Care Med 2002;28:1629-34.
2 Schein RM, Hazday N, Pena M, Ruben BH, Sprung CL. Clinical antecedents to in-hospital cardiopulmonary arrest. Chest 1990;98:1388–92.
3 Franklin C, Mathew J. Developing strategies to prevent inhospital cardiac arrest: analyzing responses of physicians and nurses in the hours before the event. Crit Care Med 1994;22:244-7.
4 Kause J, Smith G, Prytherch D, Parr M, Flabouris A, Hillman K, et al. A comparison of antecedents to cardiac arrests, deaths and emergency intensive care admissions in Australia and New Zealand, and the United Kingdom-the ACADEMIA study. Resus- citation 2004;62:275-82.
5 McQuillan P, Pilkington S, Allan A, Taylor B, Short A, Morgan G, et al. Confidential in- quiry into quality of care before admission to intensive care. BMJ 1998;316:1853-8.
6 Jarvis S, Kovacs C, Briggs J, Meredith P, Schmidt PE, Featherstone PI, et al. Can binary early warning scores perform as well as standard early warning scores for discriminating a patient's risk of cardiac arrest, death or unanticipated intensive care unit admission? Resuscitation 2015;93:46-52.
7 Olsson T, Terent A, Lind L. Rapid Emergency Medicine score: a new prognostic tool for in-hospital mortality in nonsurgical emergency department patients. J Intern Med 2004;255:579-87.
8 Shapiro NI, Wolfe RE, Moore RB, Smith E, Burdick E, Bates DW. Mortality in Emer- gency Department Sepsis (MEDS) score: a prospectively derived and validated clinical prediction rule. Crit Care Med 2003;31:670-5.
9 Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, et al. Defin- ing community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax 2003;58:377-82.
10 van Galen LS, Lammers EM, Schoonmade LJ, Alam N, Kramer MH, Nanayakkara PW. Acute medical units: The way to go? A literature review. Eur J Intern Med 2017;39: 24-31.
11 Reid LE, Dinesen LC, Jones MC, Morrison ZJ, Weir CJ, Lone NI. The effectiveness and variation of acute medical units: a systematic review. Int J Qual Health Care 2016; 28:433-46.
12 Scott I, Vaughan L, Bell D. Effectiveness of acute medical units in hospitals: a systematic review. Int J Qual Health Care 2009;21:397-407.
13 Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med 2013;158:280-6.
14 MacDonald N, Ahmad T, Mohr O, Kirk-Bayley J, Moppett I, Hinds CJ, et al. Dynamic preload markers to predict fluid responsiveness during and after major gastrointes- tinal surgery: an observational substudy of the OPTIMISE trial. Br J Anaesth 2015; 114:598-604.
15 Armagan E, Yilmaz Y, Olmez OF, Simsek G, Gul CB. Predictive value of the modified Early Warning Score in a Turkish emergency department. Eur J Emerg Med 2008;15: 338-40.
16 Bulut M, Cebicci H, Sigirli D, Sak A, Durmus O, Top AA, et al. The comparison of modified early warning score with rapid emergency medicine score: A prospective multicentre observational cohort study on medical and surgical patients presenting to emergency department. Emerg Med J 2014;31: 476-81.
17 Burch VC, Tarr G, Morroni C. Modified early warning score predicts the need for hospital admission and inhospital mortality. Emerg Med J 2008;25:674-8.
18 Cattermole GN, Liow EC, Graham CA, Rainer TH. THERM: the Resuscitation Manage- ment score. A prognostic tool to identify critically ill patients in the emergency department. Emerg Med J 2014; 31:803-7.
19 Cattermole GN, Mak SK, Liow CH, Ho MF, Hung KY, Keung KM, et al. Derivation of a prognostic score for identifying critically ill patients in an emergency department resuscitation room. Resuscitation 2009;80:1000-5.
20 Churpek MM, Snyder A, Han X, Sokol S, Pettit N, Howell MD, et al. Quick sepsis-re- lated organ failure assessment, systemic inflammatory response syndrome, and early warning scores for detecting clinical deterioration in infected patients outside the intensive care unit. Am J Respir Crit Care Med 2017;195:906-11.
21 Cildir E, Bulut M, Akalin H, Kocabas E, Ocakoglu G, Aydin SA. Evaluation of the modified MEDS, MEWS score and Charlson comorbidity index in patients with community acquired sepsis in the emergency department. Intern Emerg Med 2013;8:255-60.
22 Delgado-Hurtado JJ, Berger A, Bansal AB. Emergency department Modified Early Warning Score association with admission, admission disposition, mortality, and length of stay. J Community Hosp Intern Med Perspect 2016;6:31456.
23 Dundar ZD, Ergin M, Karamercan MA, Ayranci K, Colak T, Tuncar A, et al. Modified Early Warning Score and VitalPac Early Warning Score in geriatric patients admitted to emergency department. Eur J Emerg Med 2016;23:406–12.

24 Fairclough E, Cairns E, Hamilton J, Kelly C. Evaluation of a modified early warning system for acute medical admissions and comparison with C-reactive protein/albumin ratio as a predictor of patient outcome. Clin Med (Lond) 2009;9:30-3.
25 Geier F, Popp S, Greve Y, Achterberg A, Glockner E, Ziegler R, et al. Severity illness scoring systems for early identification and prediction of in-hospital mortality in patients with suspected sepsis pre senting to the emergency department. Wien Klin Wochenschr 2013;125:508-15.
26 Ghanem-Zoubi NO, Vardi M, Laor A, Weber G, Bitterman H. Assessment of disease severity scoring systems for patients with sepsis in general internal medicine departments. Crit Care 2011;15.
27 Heitz CR, Gaillard JP, Blumstein H, Case D, Messick C, Miller CD. Performance of the maximum modified early warning score to predict the need for higher care utilization among admitted emergency department patients. J Hosp Med 2010;5:E46-2.
28 Ho le O, Li H, Shahidah N, Koh ZX, Sultana P, Hock Ong ME. Poor performance of the modified early warning score for predicting mortality in critically ill patients presenting to an emergency department. World J Emerg Med 2013;4:273-8.
29 Huggan PJ, Akram F, Er BH, Christen LS, Weixian L, Lim V, et al. Measures of acute physiology, comorbidity and functional status to differentiate illness severity and length of stay among acute general medical admissions: a prospective cohort study. Intern Med J 2015;45:732-40.
30 Innocenti F, Tozzi C, Donnini C, De Villa E, Conti A, Zanobetti M, et al. SOFA score in septic patients: incremental prognostic value over age, comorbidities, and parameters of sepsis severity. Intern Emerg Med 2017.
31 Koksal O, Torun G, Ahun E, Sigirli D, Guney SB, Aydin MO. The comparison of modified early warning score and Glasgow coma scale-age-systolic blood pressure scores in the assessment of nontraumatic critical patients in Emergency Department. Niger J Clin Pract 2016;19:761-5.
32 Perera YS, Ranasinghe P, Adikari AM, Welivita WD, Perera WM, Wijesundara WM, et al. The value of the Modified Early Warning Score and biochemical parameters as predictors of patient outcome in acute medical admissions a prospective study. Acute Med 2011;10:126-32.
33 Vorwerk C, Loryman B, Coats TJ, Stephenson JA, Gray LD, Reddy G, et al. Prediction of mortality in adult emergency department patients with sepsis. Emerg Med J 2009; 26:254-8.
34 Abbott TE, Torrance HD, Cron N, Vaid N, Emmanuel JA. Single-centre cohort study of National Early Warning Score (NEWS) and near patient testing in acute medical admissions. Eur J Intern Med 2016; 35:78-82.
35 Abbott TE, Vaid N, Ip D, Cron N, Wells M, Torrance HD, et al. A single-centre observational cohort study of admission National Early Warning Score (NEWS). Resuscitation 2015;92:89-93.
36 Alam N, Vegting IL, Houben E, van Berkel B, Vaughan L, Kramer MH, et al. Exploring the performance of the National Early Warning Score (NEWS) in a European emergency department. Resuscitation 2015;90:111–5.
37 Bilben B, Grandal L, Sovik S. National Early Warning Score (NEWS) as an emergency department predictor of disease severity and 90-day survival in the acutely dyspneic patient - a prospective observational study. Scand J Trauma Resusc Emerg Med 2016;24:80.
38 Corfield AR, Lees F, Zealley I, Houston G, Dickie S, Ward K, et al. Utility of a single early warning score in patients with sepsis in the emergency department. Emerg Med J 2014;31:482-7.
39 Hodgson LE, Dimitrov BD, Congleton J, Venn R, Forni LG, Roderick PJA. validation of the National Early Warning Score to predict outcome in patients with COPD exacerbation. Thorax 2017;72:23–30.
40 Jo S, Yoon J, Lee JB, Jin Y, Jeong T, Park B. Predictive value of the National Early Warn- ing Score-Lactate for mortality and the need for critical care among general emer- gency department patients. J Crit Care 2016;36:60-8.
41 Nickel CH, Kellett J, Cooksley T, Bingisser R, Henriksen DP, Brabrand M. Combined use of the National Early Warning Score and D-dimer levels to predict 30-day and 365-day mortality in medical patients. Resuscitation 2016;106:49–52.
42 Sbiti-Rohr D, Kutz A, Christ-Crain M, Thomann R, Zimmerli W, Hoess C, et al. The National Early Warning Score (NEWS) for outcome prediction in emergency department patients with community-acquired pneumonia: results from a 6-year prospective cohort study. BMJ Open 2016;6:e011021.
43 Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation 2013;84:465-70.
44 Goodacre S, Turner J, Nicholl J. Prediction of mortality among emergency medical admissions. Emerg Med J 2006;23:372-5.
45 Ha DT, Dang TQ, Tran NV, Vo NY, Nguyen ND, Nguyen TV. Prognostic performance of the Rapid Emergency Medicine Score (REMS) and Worthing Physiological Scoring system (WPS) in emergency department. Int J Emerg Med 2015;8:18.
46 Howell MD, Donnino MW, Talmor D, Clardy P, Ngo L, Shapiro NI. Performance of severity of illness scoring systems in emergency department patients with infection. Acad Emerg Med 2007;14:709-14.
47 Olsson T, Lind L. Comparison of the rapid emergency medicine score and APACHE II in nonsurgical emergency department patients. Acad Emerg Med 2003;10:1040-8.
48 Olsson T, Terent A, Lind L. Rapid Emergency Medicine Score can predict long-term mortality in non surgical emergency department patients. Acad Emerg Med 2004; 11:1008-13.
49 Hilderink MJ, Roest AA, Hermans M, Keulemans YC, Stehouwer CD, Stassen PM. Predictive accuracy and feasibility of risk stratification scores for 28-day mortality of patients with sepsis in an emergency department. Eur J Emerg Med 2015;22:331-7.
50 Prytherch DR, Smith GB, Schmidt PE, Featherstone PI. ViEWS-towards a national early warning score for detecting adult inpatient deterioration. Resuscitation 2010; 81:932-7.
51 Cetinkaya HB, Koksal O, Sigirli D, Leylek EH, Karasu O. The predictive value of the modified early warning score with rapid lactate level (ViEWS-L) for mortality in patients of age 65 or older visiting the emergency department. Intern Emerg Med 2016.
52. Groarke JD, Gallagher J, Stack J, Aftab A, Dwyer C, McGovern R, et al. Use of an admission early warning score to predict patient morbidity and mortality and treatment success. Emerg Med J 2008; 25:803-6.
53 Barlow G, Nathwani D, Davey P. The CURB65 pneumonia severity score outperforms generic sepsis and early warning scores in predicting mortality in community-ac- quired pneumonia. Thorax 2007;62: 253-9.
54 Duckitt RW, Buxton-Thomas R, Walker J, Cheek E, Bewick V, Venn R, et al. Worthing physiological scoring system: derivation and validation of a physiological early- warning system for medical admissions. An observational, population-based single-centre study. Br J Anaesth 2007;98:769-74.
55 Singer AJ, Ng J, Thode Jr HC, Spiegel R, Weingart S, Quick SOFA. Scores Predict Mortality in Adult Emergency Department Patients With and Without Suspected Infec- tion. Ann Emerg Med 2017;69: 475-9.
56 Paterson R, MacLeod DC, Thetford D, Beattie A, Graham C, Lam S, et al. Prediction of inhospital 6;6:281-4.
57 RCoPo. L. National Early Warning Score (NEWS) - standardising the assessment of acute-illness severity in the NHS - report of a working party; 2012.
58 Baker T, Schell CO, Lugazia E, Blixt J, Mulungu M, Castegren M, et al. Vital signs directed therapy: improving care in an intensive care unit in a low-income country. PLoS One 2015;10:e0144801.
59 Alam N, Hobbelink EL, van Tienhoven AJ, van de Ven PM, Jansma EP, Nanayakkara PW. The impact of the use of the Early Warning Score (EWS) on patient outcomes: a systematic review. Resuscitation 2014;85:587-94.
60 Pullinger R, Wilson S, Way R, Santos M, Wong D, Clifton D, et al. Implementing an electronic observation and early warning score chart in the emergency department: a feasibility study. Eur J Emerg Med 2016.
61 Smith AF, Oakey RJ. Incidence and significance of errors in a patient 'track and trigger' system during an epidemic of Legionnaires' disease: retrospective casenote analysis. Anaesthesia 2006;61:222-8.
62 Wilson SJ, Wong D, Clifton D, Fleming S, Way R, Pullinger R, et al. Track and trigger in an emergency department: an observational evaluation study. Emerg Med J 2013; 30:186-91.
63 Wheeler I, Price C, Sitch A, Banda P, Kellett J, Nyirenda M, et al. Early warning scores generated in developed healthcare settings are not sufficient at predicting early mortality in Blantyre, Malawi: a prospective cohort study. PLoS One 2013;8:e59830.
64 Opio MO, Nansubuga G, Kellett J. Validation of the VitalPAC Early Warning Score (ViEWS) in acutely ill medical patients attending a resource-poor hospital in sub-Sa- haran Africa. Resuscitation 2013; 84:743-6.
65 Moore CC, Hazard R, Saulters KJ, Ainsworth J, Adakun SA, Amir A, et al. Derivation and validation of a universal vital assessment (UVA) score: a tool for predicting mor- tality in adult hospitalised patients in sub-Saharan Africa. BMJ. Glob Health 2017;2.

Chapter

3

A prospective, observational study of the performance of MEWS, NEWS, SIRS and qSOFA for early risk stratification for adverse outcomes in patients with suspected infections at the emergency department

Acute Medicine 2021;20(2):116-124.

T. C. Minderhoud

P. Mohammadi

R. Dekker

V. Brown

T. Attaye

S. J. Huisman

A. A. Hettinga-Roest

P. W.B. Nanayakkara

A prospective, observational study of the performance of MEWS, NEWS, SIRS and qSOFA for early risk stratification for adverse outcomes in patients with suspected infections at the emergency department. Kaoutar Azijli & TC Minderhoud (shared first), Parisa Mohammadi, Rachelle Dekker, Vanessa Brown , Tamana Attaye, Sebastiaan J. Huisman, Asselina A. Hettinga-Roest, Prabath W.B. Nanayakkara

Risk stratification

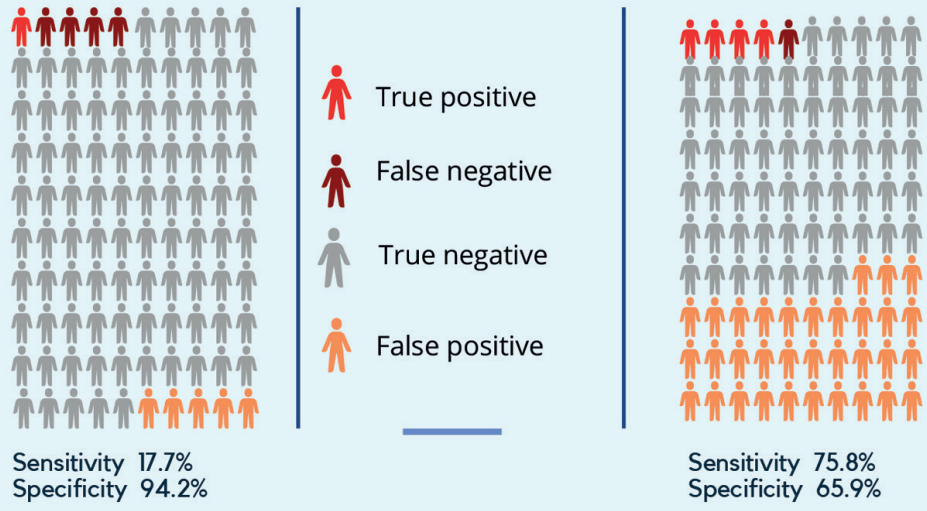
What is the best score for the ED?

We performed a prospective study of 1328 patients with suspected infection in the Emergency department.

Inclusion: suspected systemic infection
Methods: first set of vital signs
Outcome: 30 days-mortality

62/ 1328 died (4.7%)

Results



The NEWS had the highest AUROC of 0.740 (95% CI 0.682-0.798) for predicting 30-day mortality, compared to qSOFA, MEWS and SIRS. It also had the highest sensitivity of 76.8 % at the cut-off NEWS ≥5. qSOFA ≥ 2 had the lowest sensitivity but the highest specifity, as illustrated above.

Conclusion

- Among patients presenting to the ED with suspected infection, early risk stratification with NEWS (cut-off of ≥5) is more sensitive for prediction of mortality than qSOFA ≥ 2 , MEWS ≥3 or SIRS≥2, with adequate specificity.
- The alternative cut off point of qSOFA≥ 1 increased the sensitivity to 69.4% (CI 56.4-80.4), with a specificity of 66.1% (95% CI 63.4-98.7). In some settings, qSOFA≥ 1 can be a suitable and easy to calculate alternative.

3

ABSTRACT

Background

Many patients with suspected infection are presented to the emergency Department. Several scoring systems have been proposed to identify patients at high risk of adverse outcomes.

Methods

We compared generic early warning scores (MEWS and NEWS) to the (SIRS) criteria and quick Sequential Organ Failure Assessement (qSOFA), for early risk stratification in 1400 patients with suspected infection in the ED. The primary study end point was 30-day mortality.

Results

The AUROC of the NEWS score for predicting 30-day mortality was 0.740 (95% Confidence Interval 0.682-0.798), higher than qSOFA (AUROC of 0.689, 95% CI 0.615- 0.763), MEWS (AUROC 0.643 (95% CI 0.583-0.702) and SIRS (AUROC 0.586, 95%CI 0.521 - 0.651). The sensitivity was also highest for NEWS≥ 5 (sensitivity 75,8% specificity of 67,4%).

Conclusion

Among patients presenting to the ED with suspected infection, early risk stratification with NEWS (cut-off of ≥5) is more sensitive for prediction of mortality than qSOFA, MEWS or SIRS, with adequate specificity.

INTRODUCTION

Severe infection, leading to organ dysfunction is a public health concern affecting millions of people around the world each year[1, 2]. A study in the United States by Wang [3] estimated that approximately 21% of the patients that present to the Emergency Depart-
ment have (severe) infections, and approximately 3 % have associated organ dysfunction. This suggests that systemic infections are a frequent cause of ED presentation, but only a smaller proportion of these infections actually lead to sepsis and mortality. To determine if a patient, coming in with suspected infection, suffers from sepsis, organ function has to be assessed. However, while some patients have overt organ dysfunction at presentation, others have organ dysfunction that can only be detected by laboratory tests (taking time), whereas others will not develop organ dysfunction at all.
It will therefore not always be possible to establish if organ dysfunction is present early in the process, however decisions regarding management (such as early antibiotic treatment) will have to be made early. To enable early decision making, an early risk assessment in patients suspected of infection is needed.

Over the past years, several scoring systems have been proposed to assess the risk of death or an adverse outcomes in Emergency Department patients. Some were specifically designed for patients suspected of infection or sepsis, such as the quick Sequential Organ Failure Assessment (qSOFA)[1], and the Systemic Inflammatory Response Syndrome (SIRS). The latter was not designed as a risk stratification tool, but was introduced as a diagnostic criterion. However, over the years several studies have used it as risk stratification tool [4-6]. In the Netherlands SIRS was routinely used by emergency departments for flagging patients with possible sepsis up to 2016, as this was suggested in a government program introduced in 2008 [7].

Since changing the definition of sepsis in 2016, many Emergency Departments struggled how to proceed with early risk stratification in patients suspected of infection. In the UK, the NICE guideline on sepsis (2016) suggested the use of a generic early warning score to assess the risk of death in patients at risk of sepsis[8]. The most commonly used Early Warning Scores are the Modified Early Warning Score (MEWS) and the National Early Warning Score (NEWS). A review of early warning scores in 2018 found that the National Early Warning score (NEWS) was the most accurate score in the general ED population for prediction of mortality [9]. The Surviving Sepsis campaign introduced qSOFA as a risk stratification tool, also in 2016. It was not prospectively validated in the ED setting at the time. Studies published in the meantime have found conflicting results regarding the accuracy of qSOFA in ED populations [4, 5, 10-13].

The aim of this study was to compare the predictive value of early risk stratification scores (qSOFA, SIRS, MEWS and NEWS) at the primary assessment for predicting death or the need for ICU admission in adults presenting to the ED with suspected infection.

METHODS

Design

This study was a multi-centre prospective observational study and was performed in the ED in two general and one university hospital in the Netherlands. We included patients between November 2016 and May 2018 in the university hospital (Amsterdam University Medical Center, location VUmc) and between November 2017 and January 2018 in the two general hospitals.

Study population

Patients ≥18 years who visited the ED with a suspected infection (who were admitted to the hospital or discharged from the ED) were prospectively included. Both medical and surgical patients were included. Patients that could not complete follow-up (i.e. tourists) were excluded. Patients were included during peak hours (10 am to 8 pm) by trained researchers.

Ethics

All patients, or their proxies, signed written informed consent forms prior to inclusion.

Variables

All clinical data were measured upon arrival at the ED and registered in a secure database, Castor (https://www.castoredc.com) compliant with Good Clinical Practice guidelines. The following data were collected at the time of arrival: baseline characteristics and the vital signs. The NEWS, MEWS, qSOFA and SIRS scores were calculated in the database using the first set of vital signs. A detailed overview of variables included in these scores can be found in the supplement (S3.1). Information regarding medical history, organ dysfunction and laboratory tests were extracted from the electronic medical record (EMR) in the ED. After 30 days, the record was checked for the clinical endpoints death and/or ICU admission, revisitation and culture results.

Definitions

Suspected infection was flagged by the (trained) triage nurse or treating physician or both
at presentation at the Emergency Department. This was based on the reason for referral (referral by General Practitioner, or paramedics, if available) and the first impression at triage or primary assessment. This broad inclusion criterion was chosen as it reflects current practice, because infection is suspected in many different clinical situations. Since the signs of infection are diverse, any more specific criteria might lead to selection bias and poor external validity. For example, up to 33% of patients with bacteremia in the ED may present without fever [14].

End point

The primary study end point was the 30-day mortality. Secondary endpoints were the need of ICU admission and a composite outcome measure of adverse outcome (30-day mortality and ICU admission).

STATISTICS

All continuous variables with a normal distribution are expressed as percentages, means and standard deviations (SD) and non-normally distributed variables as median with interquartile range (IQR). Differences between groups were assessed using independent T-testing for normally distributed variables and Mann-Whitney U for non-normally distributed variables. Chi-square or Fisher's exact test were used for categorical variables. Receiver Operating Characteristics (ROC) were constructed and the sensitivity and specificity for predefined cut-offs was calculated. We used the cut-offs that were standard at the hospitals, during the time of inclusion, (qSOFA≥ 2, SIRS ≥ 2, MEWS ≥ 3, NEWS of ≥ 5) for the primary comparison of sensitivity, specificity, as well as negative predictive value (NPV) and positive predictive value (PPV) [15, 16]. Because several cut-offs have been proposed in the past, we also calculated the test specifics at alternative cut-offs to see if this could lead to optimization of sensitivity and specificity and PPV and NPV. A p-value ≤0.05 was considered to be statistically significant for all analyses. An AUROC from 0.6 to 0.7, 0.7 to 0.8, 0.8 to 0.9, and 0.9 or higher were respectively considered as poor, adequate, good, and excellent [17].

Missing values

In some cases, the set of vital signs in the EMR were incomplete. In case of missing values, we assumed a normal result to calculate the risk stratification scores. During the inclusion of patients, we found that the respiratory rate was often poorly recorded, which has been observed before in literature [18]. As this might influence the accuracy of the scores, we performed a sensitivity analysis to see how this affected the AUROC. All data analyses were performed using SPSS software (SPSS version 22.0, IBM, Armonk, NY, USA)

Results

Between November 2016 and May 2018 1401 patients were included in the study as is shown in Figure 1. Due to incomplete data and withdrawal of consent, 73 patients were excluded from the final analysis.
In total 1328 Patients were included in the final analysis. At day 30, the survival rate was 95.3% (n=1266) and 4.7% (n=62) had died. In-hospital 30-day mortality was 3.5 % (N=46), out-of-hospital 30 day mortality was 1.2 % (N=16). Of all included patients, 63 (4.7%) were admitted to the ICU. 17 (27%) of the patients admitted to the ICU died. The composite outcome of ICU and/or death was reached by 108 patients (8.1%).
The inclusion process is illustrated in Figure 3.1.
Baseline characteristics are shown in table 3.1. Non-survivors were significantly older and had a higher Charlson Comorbidity index (CCI).

Figure 3.1 Inclusion process

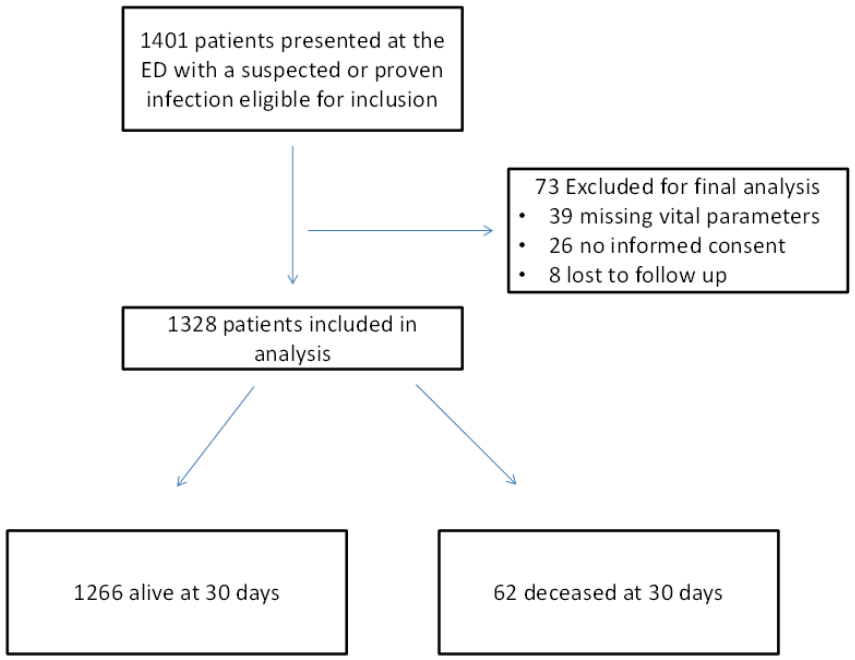


Table 3.1 Baseline characteristics of Study Participants

	All patients N=1328	Survivors N=1266	Nonsurvivors N =62
Demographics			
Age, mean (SD)	62.62 (2.3)	62.1 (18.0)	73.0 (13.5)
Male, n (%)	694 (5.3)	655 (51.5)	39 (62.9)
Admission, n (%)	963 (72.5)	903 (71.3)	60 (96.8)
Admission IC, n (%)	63 (4.7)	46 (3.6)	17 (27.4)
Readmission, n (%)	160 (12.0)	151 (11.9)	9 (14.5)
Co-morbidities, n (%)[a]			
COPD (GOLD >2) n (%)	135 (10.5)	127 (10.0)	8 (12.9)
Astma	62 (4.8)	60 (4.8)	2 (3.2)
Heart failure	92 (7.1)	74 (5.8)	18 (29.0)
Renal insufficiency	133 (10.3)	125 (9.9)	8 (6.0)
Current Malignancy	284 (22.0)	260 (20.5)	24 (38.7)
CCI (data complete N=994), mean (SD)	2.8 (2.2)	2.8 (2.2)	3.5 (2.5)
Laboratory analysis on admission [b]			
White bloodcell count (10^9/L, mean (SD)	11.9 (7.6)	11.8(7.5)	14.2 (9.8)
Lactate (mmol/L), mean (SD)	1.6 (1.0)	1.5 (1.0)	2.2 (1.4)
Platelets (*10^9/L), mean (SD)	260 (119)	259 (116)	278 (168)
Creatinine (µg/L), mean (SD)	109.0 (95.6)	107.7 (95.3)	135.1 (98.1)
CRP (mg/L), mean (SD)	104.6 (102.2)	102.4 (101.3)	147.9 (110.7)
True positive BC/ total BC (%)	94/694 (13.5%)	88/659 (13 %)	6/35 (17%)
Positive scores at primary survey			
MEWS ≥ 3, n(%)	560 (42,2)	521 (41.2)	39 (62.9)
NEWS ≥5, n (%)	481 (36,2)	434 (34.3)	47 (75.8)
qSOFA ≥2, n (%)	67 (5.0)	57 (4.5)	10 (16.1)
SIRS ≥2, n (%)	758 (57.1)	711 (56.2)	47 (75.8)

MAIN OUTCOME

The accuracy of the qSOFA, SIRS, MEWS and NEWS was assessed with AUROC curves. Results are shown in Table 3.2. The curves are added in the supplement S3.2. The NEWS score showed an adequate AUROC of 0.740 (95% CI 0.682- 0.789) for mortality. qSOFA scored an AUROC of 0.689 (95%-CI 0.615 - 0.763) for prediction of mortality, followed by MEWS with an AUROC 0.643 (95% CI 0.583 – 0.702) and SIRS, AUROC 0.586 (95%-CI 0.521 – 0.651). For predicting ICU admission NEWS had an adequate AUROC of 0.775 (95%-CI 0.719 - 0.831), followed by MEWS with an AUROC of 0.731 (95%-CI 0,671-0,797) and SIRS with an AUROC of 0.722 (95%-CI 0,625-0,764). qSOFA scored the lowest AUROC of 0.702 (95%-CI 0.625 - 0.764).

For the composite outcome, NEWS had an adequate AUROC of 0.763 (95% CI 0.719-0.808), followed by MEWS with an AUROC of 0,697 (95%-CI 0.648-0.746). qSOFA had a AUROC of 0.691 (95%CI 0.638-0.744), and SIRS an AUROC of 0.667 (0.618-0.717). The AUROC curves can be found in the Supplement (S3.2).

Table 3.2 AUROC of risk-stratification scores, ranked high-to-low

Outcome: mortality		
Score	**30-day mortality**	**95 % Confidence interval**
NEWS	0.740	(0.682-0.798))
qSOFA	0.687	(0.618-0.756)
MEWS	0.643	(0.583-0.702)
SIRS	0.586	(0.521-0.651)
Outcome: ICU admission		
NEWS	0.775	(0.719-0.831)
MEWS	0.731	(0.671-0.797)
SIRS	0.722	(0.681-0.790)
qSOFA	0.702	(0.625-0.764)
Composite: Death/ICU admission		
NEWS	0.763	(0.719-0.808)
MEWS	0.697	(0.648-0.746)
qSOFA	0.691	(0.638-0.744)
SIRS	0.667	(0.618-0.717)

qSOFA: quick Sequential Organ Failure; SIRS Systemic Inflammatory Response Syndrome; MEWS: Modified Early Warning Score; NEWS: National Early Warning score
CI: confidence interval.

Sensitivity, specificity, negative predictive value and positive predictive value

30-day mortality

Table 3.3 shows the sensitivity, specificity and the predictive value of the scores with various cut-off values. NEWS ≥ 5 had sensitivity of 75.8 % (95%-CI 63.3-85.8) and a reasonable specificity of 65.9% (95%-CI 63.2-68.5). The PPV was 9,8 (95%-CI 8.5-11.3). NEWS≥ 5 had also the highest NPV of 98,2% (95%-CI 97.3-98.9).

Of the 62 patients that died within 30 days, 51 had a negative qSOFA score. qSOFA≥2 had a very low sensitivity for predicting 30-day mortality of 17.7 % (95%-CI, 9.20-29.5), but the highest specificity of 94.2% (95%-CI 92.8-95.5) and PPV of 13,1% (95%-CI 7.7-21.2).

The MEWS score at a cut-off point of 3 or above had a sensitivity of 72.6% (95%-CI 59.8 - 83.2%) with a specificity of 54,9% (95%-CI 52.1-57.7 %). The PPV and NPV were 5.6% (95%-CI 3.4 - 9.0) and 95.5% (95%-CI 94.9-96.0) respectively.

Alternative cut-off points for NEWS and MEWS are displayed in table 3, but did not improve sensitivity. The alternative cut off point of qSOFA≥ 1 increased the sensitivity to 69.4% (CI 56.4-80.4), with a specificity of 66.1% (95% CI 63.4-98.7). The PPV was 9,19.1 (95% CI 7.7-10.7) with a NPV of 97.8 (95% CI 96.8-98.5)

Table 3.3 Diagnostic Performance for Prediction of 30-day Mortality Across Different Score Thresholds

	All patients N=1328	Survivors N=1266	Nonsurvivors N =62
Demographics			
Age, mean (SD)	62.62 (2.3)	62.1 (18.0)	73.0 (13.5)
Male, n (%)	694 (5.3)	655 (51.5)	39 (62.9)
Admission, n (%)	963 (72.5)	903 (71.3)	60 (96.8)
Admission IC, n (%)	63 (4.7)	46 (3.6)	17 (27.4)
Readmission, n (%)	160 (12.0)	151 (11.9)	9 (14.5)
Co-morbidities, n (%)[a]			
COPD (GOLD >2) n (%)	135 (10.5)	127 (10.0)	8 (12.9)
Astma	62 (4.8)	60 (4.8)	2 (3.2)
Heart failure	92 (7.1)	74 (5.8)	18 (29.0)
Renal insufficiency	133 (10.3)	125 (9.9)	8 (6.0)
Current Malignancy	284 (22.0)	260 (20.5)	24 (38.7)
CCI (data complete N=994), mean (SD)	2.8 (2.2)	2.8 (2.2)	3.5 (2.5)
Laboratory analysis on admission [b]			
White bloodcell count (10^9/L, mean (SD)	11.9 (7.6)	11.8(7.5)	14.2 (9.8)
Lactate (mmol/L), mean (SD)	1.6 (1.0)	1.5 (1.0)	2.2 (1.4)
Platelets (*10^9/L), mean (SD)	260 (119)	259 (116)	278 (168)
Creatinine (µg/L), mean (SD)	109.0 (95.6)	107.7 (95.3)	135.1 (98.1)
CRP (mg/L), mean (SD)	104.6 (102.2)	102.4 (101.3)	147.9 (110.7)
True positive BC/ total BC (%)	94/694 (13.5%)	88/659 (13 %)	6/35 (17%)
Positive scores at primary survey			
MEWS ≥ 3, n(%)	560 (42,2)	521 (41.2)	39 (62.9)
NEWS ≥5, n (%)	481 (36,2)	434 (34.3)	47 (75.8)
qSOFA ≥2, n (%)	67 (5.0)	57 (4.5)	10 (16.1)
SIRS ≥2, n (%)	758 (57.1)	711 (56.2)	47 (75.8)

qSOFA, quick Sequential Organ Failure; SIRS, Systemic Inflammatory Response Syndrome; MEWS, Modified Early Warning Score; NEWS, National Early Warning score; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval.
Outcomes for sensitivity, specificity, PPV and NPV are shown as percentage with the 95% confidence interval.

ICU admission

Table 3.4 shows that the SIRS score has the highest sensitivity for predicting ICU admission of 92.0 % (95% CI 82.4-97.4) with a specificity of 44,6% (95% CI 41.8-47.4). NEWS ≥ 5 had a sensitivity of 79.4 (95% CI 67.3-88.5) with a specificity of 65,9 % (95% CI 63.2-68.5). qSOFA ≥ 2 had a sensitivity of 15.9 % (95% CI 7.9-27.3) with a specificity of 94.1 (95% CI 92.7-95.4). qSOFA with a cut-off of 1 had a sensitivity of 73.0 % (95% CI 60.4-83.4) with a specificity of 66,9% (95% CI 64.2-69.5).

Table 3.4 Diagnostic Performance for Prediction of IC admission Across Different Score Thresholds

Score/ threshold	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
NEWS ≥5	79.4 (67.3-88.5)	65.9 (63.2-68.5)	10.6 (9.3-12.1)	98.2(97.5-99.0)
NEWS ≥ 7	57.1 (44.1-69.5)	78.9 (76.5-81.1)	12.1 (9.7-14.9)	97.3 (96.5-98.0)
qSOFA **≥1**	73.0 (60.4-83.4)	66.9 (64.2-69.5)	10.1 (8.7-11.7)	98.0 (97.0- 98.7)
qSOFA **≥2**	15.9 (7.9-27.3)	94.1 (92.7-95.4)	12.1 (6.9-20.1)	95.7 (95.2-96.1)
MEWS **≥** 3	76.2 (63.8-86.0)	59.6 (56.8- 62.3)	8.73 (7.6-10.0)	98.0 (96.9-98.7)
MEWS **≥ 5**	42.9 (30.5-56.0)	85.1 (83.0-87.0)	12.7 (9.6-16.7)	96.7 (95.9-97.3)
SIRS **≥2**	92.0 (82.4-97.4)	44.6 (41.8-47.4)	7.77(7.2-8.4)	99.1 (97.9-99.6)

qSOFA, quick Sequential Organ Failure; SIRS, Systemic Inflammatory Response Syndrome; MEWS, Modified Early Warning Score; NEWS, National Early Warning score; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval.
Outcomes for sensitivity, specificity, PPV and NPV are shown as percentage with the 95% confidence interval.

Composite outcome Death/ICU

Results for all scores are displayed in table 3.5. For the composite outcome, NEWS ≥5 had a sensitivity of 76.9 % and a specificity of 67.4 %. Only SIRS ≥2 showed higher sensitivity but with a specifity of only 45.2%. qSOFA ≥ 2 had a sensitivity of 15.7 % and a specifity of 94.5. qSOFA cut-off ≥1 showed a sensitivity of 69.4% and a specifity of 67.5%.

Table 3.5 Diagnostic Performance for Prediction of death/ ICU Across Different Score Thresholds

Score/ threshold	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
NEWS ≥5	76.9 (67.8-84.4)	67.4 (64.7-70.0)	17.3 (15.5-19.2)	97.1 (95.9-97.9)
NEWS ≥ 7	55.1 (45.2-64.8)	79.8 (77.5-82.1)	19.3 (16.4-22.7	95.3 (94.3-96.2)
qSOFA **≥1**	69.4 (59.8-78.0)	67.5 (64.8-70.1)	15.9 (14.0-18.0	96.2 (95.0-97.0)
qSOFA **≥2**	15.7 (9.5 -24.0)	94.5 (93.1- 95.7)	20.2 (13.4-29.4)	92.7 (92.1-93.2)
MEWS **≥** 3	68.5 (58.9-77.1)	60.2 (57.4-62.9)	13.2 (11.6-15.0)	95.6 (94.2-96.6)
MEWS **≥ 5**	32.4 (23.7-42.1	85.2 (83.0-87.0)	16.2 (12.5-20.8)	93.4 (92.6-94.2)
SIRS **≥2**	82.4 (73.9-89.1)	45.2 (42.4-48.0)	11.7 (10.7-12.8)	96.7 (95.1-97.8)

qSOFA, quick Sequential Organ Failure; SIRS, Systemic Inflammatory Response Syndrome; MEWS, Modified Early Warning Score; NEWS, National Early Warning score; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval.
Outcomes for sensitivity, specificity, PPV and NPV are shown as percentage with the 95% confidence interval.

Sensitivity Analysis
Of all the vital parameters, the respiratory rate was missing most often (136 of 1328 records). Closer analysis of these cases revealed that of these patients, one patient died, one (other) patient was admitted to the ICU, thus a total of two patients reached the composite outcome (death/ICU admission). The assumption of normal values might have more influence on the calculation of the qSOFA score, as this score consists only of 3 parameters. Therefore, we performed a ROC analysis of the sample, excluding the patients with missing respiratory rates. The area under the curve was slightly different, but the ranking stayed the same. NEWS scored the highest AUROC of 0.729 for 30-day mortality, 0.759 for ICU admission and 0.751 on the composite outcome. The curves and values of AUROC can be found in the Supplement, S3.3 and S3.4.

DISCUSSION

Our prospective observational study in patients with suspected infection demonstrates that the NEWS at a cut-off of ≥5, measured during the primary assessment, has the highest sensitivity for predicting the risk of death or IC admission compared to qSOFA, SIRS, MEWS scores.

This finding is in line with the findings of several other (retrospective) studies. Churpek et al.[10] who retrospectively studied 30,677 ward and ED patients and found a comparable AUROC of 0.77 for NEWS and of 0.69 for qSOFA. The study of Goulden et all showed similar results [4]. This study was also a retrospective cohort study which included patients who were presented to the ED or medical admissions unit with suspected sepsis.
An observational prehospital study by Silcock also showed similar results for the combined outcome for NEWS comparing it with qSOFA [19].

Our study studied the early warning scores in a prospective cohort. In the ED, decisions regarding treatment and admission often have to be made early and within a short time frame. For that reason, we need a sensitive score to identify patients at risk for adverse outcome as early as possible in the ED. Ideally this score also has a high negative predictive value in a population with low incidence of the outcome of interest in this case mortality. In our study, NEWS consistently showed the highest AUROC for prediction of mortality, ICU admission. At a cut-off of ≥5 the sensitivity was the highest of all scores tested for mortality, coupled with a high negative predictive value. Therefore, we find it the most suitable score to identify patients at risk of adverse outcomes early in the ED presentation.

Strengths and weaknesses

We decided to include patients with suspected infection. This suspicion was based on early clinical judgment of the treating physician or triage nurse. This reflects daily practice, where decisions on management, such as antibiotic treatment are often made during the primary assessment of the patient. Some studies investigated the predictive validity of qSOFA, including only patients in whom the suspicion of infection was based upon body fluid culture and or administration of antibiotics [2, 10, 12, 20] However this method can limit external validity since risk stratification is usually done before the decision to give empirical antibiotics. A recent article of the NEJM illustrated how hard it is to make treatment decisions on patients with suspected infection with organ dysfunction[21]
The strength of our broad inclusion criterion is the generalizability. However, it has also resulted in inclusion of a very heterogeneous population of patients with suspected systemic infection in the ED. This explains the low rates of mortality and ICU admission in contrast to other studies. However, this represents the diversity of patients that are daily seen in the ED, and in whom decisions about treatment and admission have to be made.

The findings of our study must be interpreted in light of several limitations. First this is an observational study which is subject to errors in data collection and entry. Second, the most severely ill patients, who directly presented to the critical-care room were difficult to include because the informed consent could not be obtained. This has probably affected the results, in particular the PPV and NPV which are influenced by the low incidence of the outcome measure. However, in daily practice, we need risk assessment scores most in the patients that are moderately ill. It is not difficult to identify those patients that are obviously severely ill, for example, overtly hypotensive. It is much harder to spot those with a normal blood pressure, but already elevated respiratory rate, and at risk of deterioration in the next 24 hours.

Another important limitation is that we included a convenient sample of patients, that presented between 10AM to 8PM. We missed a portion of patients who presented during the evening and night. However, we think our results are still valuable, especially since an earlier analysis of the patient flow in our hospital showed that most patient presentations in our ED occurs between 10 AM and 9 PM[22].

IN CONCLUSION

In conclusion, we found that NEWS with a cut-off of 5 provides the best balance between sensitivity, specificity and NPV for early identification of high-risk patients with suspected infection in the Emergency Department. As NEWS is a general predictive score already being used in many ED`s, we recommend using it for early risk stratification in patients with suspected or proven infections at the Emergency Department.

REFERENCES

1 Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). Jama. 2016;315(8):801-10.
2 Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, et al. Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). Jama. 2016;315(8):762-74.
3 Wang HE, Jones AR, Donnelly JP. Revised National Estimates of Emergency Department Visits for Sepsis in the United States. Critical care medicine. 2017;45(9):1443-9.
4 Goulden R. qSOFA, SIRS and NEWS for predicting inhospital mortality and ICU admission in emergency admissions treated as sepsis. Emergency medicine journal. 2018;35(6):345-9.
5 Askim A, Moser F, Gustad LT, Stene H, Gundersen M, Asvold BO, et al. Poor performance of quick-SOFA (qSOFA) score in predicting severe sepsis and mortality - a prospective study of patients admitted with infection to the emergency department. Scandinavian journal of trauma, resuscitation and emergency medicine. 2017;25(1):56.
6 Gonzalez Del Castillo J, Julian-Jimenez A, Gonzalez-Martinez F, Alvarez-Manzanares J, Pinera P, Navarro Bustos C, et al. Prognostic accuracy of SIRS criteria, qSOFA score and GYM score for 30-day-mortality in older non-severely dependent infected patients attended in the emergency department. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2017;36(12):2361-9.
7 program) VvDNIfH-s. Voorkomen van lijnsepsis en behandeling van ernstige sepsis 2009.
8 (NICE) NIfhaE. Sepsis: recognition, diagnosis and early management (2016) NICE guideline NG51.
9 Nannan Panday RS, Minderhoud TC, Alam N, Nanayakkara PWB. Prognostic value of early warning scores in the emergency department (ED) and acute medical unit (AMU): A narrative review. European journal of internal medicine. 2017;45:20-31.
10 Churpek MM, Snyder A, Han X, Sokol S, Pettit N, Howell MD, et al. Quick Sepsis-related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients outside the Intensive Care Unit. American journal of respiratory and critical care medicine. 2017;195(7):906-11.
11 Singer AJ, Ng J, Thode HC, Jr., Spiegel R, Weingart S. Quick SOFA Scores Predict Mortality in Adult Emergency Department Patients With and Without Suspected Infection. Annals of emergency medcine. 2017;69(4):475-9.
12 April MD, Lantry JH. Prognostic Accuracy of Quick Sequential Organ Failure Assessment Among Emergency Department Patients Admitted to an ICU. Annals of emergency medicine. 2017;70(4):600-1.
13 Song JU, Sin CK, Park HK, Shim SR, Lee J. Performance of the quick Sequential (sepsis-related) Organ Failure Assessment score as a prognostic tool in infected patients outside the intensive care unit: a systematic review and meta-analysis. Critical care (London, England). 2018;22(1):28.
14 Seigel TA, Cocchi MN, Salciccioli J, Shapiro NI, Howell M, Tang A, et al. Inadequacy of temperature and white blood cell count in predicting bacteremia in patients with suspected infection. The Journal of emergency medicine. 2012;42(3):254-9.
15 McGinley A, Pearse RM. A national early warning score for acutely ill patients. BMJ (Clinical research ed). 2012;345:e5310.
16 van Galen LS, Dijkstra CC, Ludikhuize J, Kramer MH, Nanayakkara PW. A Protocolised Once a Day Modified Early Warning Score (MEWS) Measurement Is an Appropriate Screening Tool for Major Adverse Events in a General Hospital Population. PloS one. 2016;11(8):e0160811.
17 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29-36.
18 Hogan J. Why don't nurses monitor the respiratory rates of patients? British journal of nursing (Mark Allen Publishing). 2006;15(9):489-92.
19 Silcock DJ, Corfield AR, Staines H, Rooney KD. Superior performance of National Early Warning Score compared with quick Sepsis-related Organ Failure Assessment Score in predicting adverse outcomes: a retrospective observational study of patients in the prehospital setting. European journal of emergency medicine : official journal of the European Society for Emergency Medicine. 2019;26(6):433-9.
20 de Groot B, Stolwijk F, Warmerdam M, Lucke JA, Singh GK, Abbas M, et al. The most commonly used disease severity scores are inappropriate for risk stratification of older emergency department sepsis patients: an observational multi-centre study. Scandinavian journal of trauma, resuscitation and emergency medicine. 2017;25(1):91.
21 Mi MY, Klompas M, Evans L. Early Administration of Antibiotics for Suspected Sepsis. The New England journal of medicine. 2019;380(6):593-6.
22 Vegting IL, Alam N, Ghanes K, Jouini O, Mulder F, Vreeburg M, et al. What are we waiting for? Factors influencing completion times in an academic and peripheral emergency department. The Netherlands journal of medicine. 2015;73(7):331-40.

Part 1 / Chapter 3
Supplementary materials

Table S3.1 Components of qSOFA, SIRS, MEWS and NEWS

Parameter	qSOFA	SIRS	MEWS	NEWS
Respiratory rate	X	X	X	X
Systolic blood pressure	X		X	X
Mental status	X		X	X
Heart rate		X	X	X
Temperature		X	X	X
Oxygen saturation				X
Use of supplemental oxygen				X
White blood cell count		X		

qSOFA, quick Sequential Organ Failure; SIRS, Systemic Inflammatory Response Syndrome; MEWS, Modified Early Warning Score; NEWS, National Early Warning score.

Table S3.2 ROC curves for qSOFA, SIRS, MEWS and NEWS score

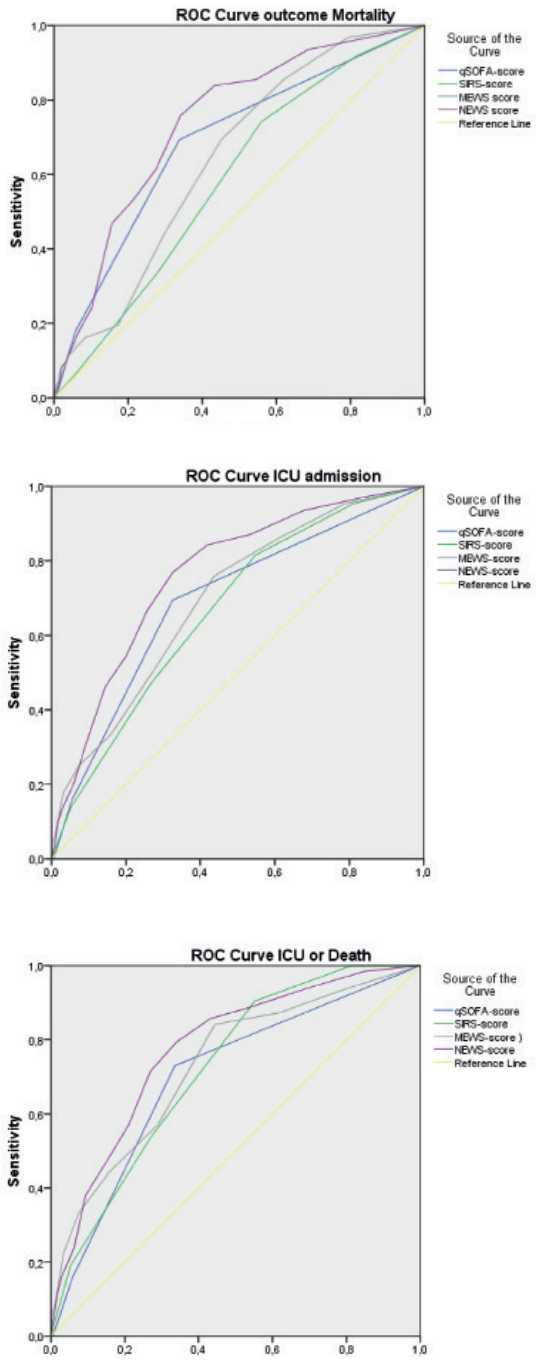


Table S3.3 Missing analysis - ROC curves (patients with missing respiratory rate excluded)

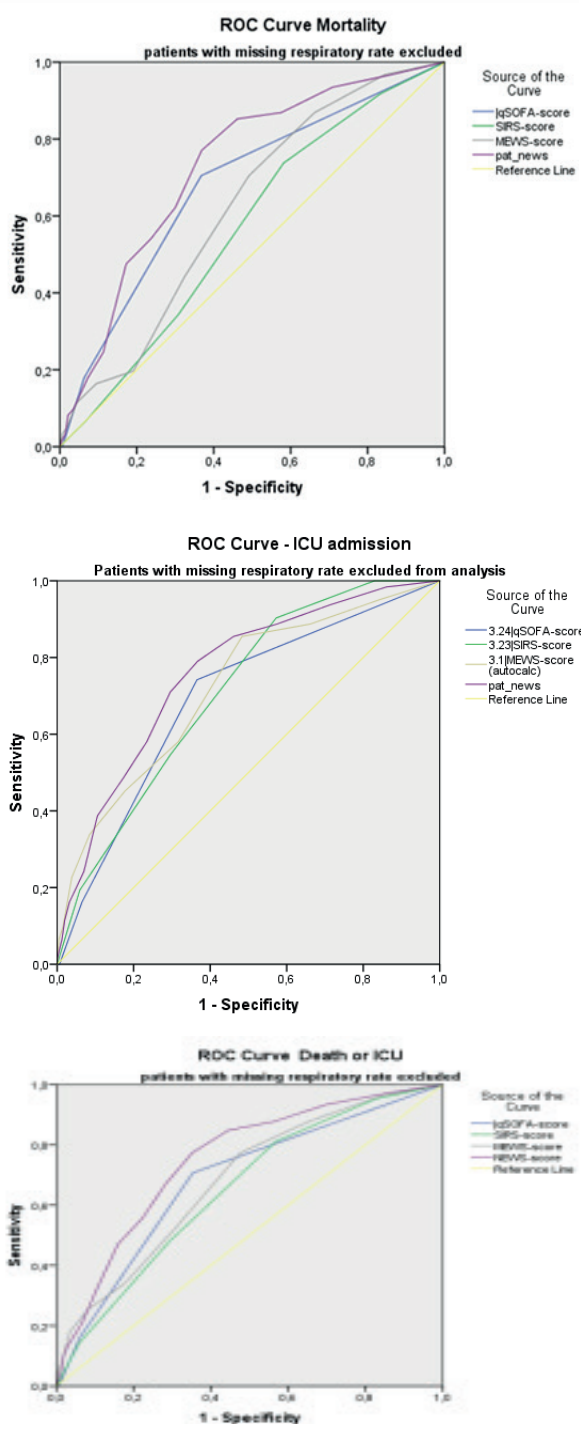


Table S3.4 AUROC of risk-stratification scores, ranked high-to-low, sensitivity analysis

Outcome: mortality	**95 % CI**	
score	**AUROC**	**95%CI**
NEWS	0.729	0.670- 0.789
qSOFA	0.679	0.609 -0.748
MEWS	0.620	0.558- 0.683
SIRS	0.568	0.501- 0.634
Outcome: ICU admission		
NEWS	0.759	0.701- 0.818
MEWS	0.721	0.656- 0.787
SIRS	0.709	0.653- 0.765
qSOFA	0.693	0.627- 0.759
Composite: Death/ICU admission		
NEWS	0.751	0.705- 0.797
qSOFA	0.683	0.630- 0.737
MEWS	0.682	0.632- 0.733
SIRS	0.652	0.601- 0.703

qSOFA: quick Sequential Organ Failure; SIRS Systemic Inflammatory Response Syndrome; MEWS: Modified Early Warning Score; NEWS: National Early Warning score; CI: confidence interval.

Part 2

Diagnosing infections in the Emergency Department

Chapter 4

Microbiological outcomes and antibiotic overuse in Emergency Department patients with suspected sepsis

Neth J Med. 2017 Jun;75(5):196-203

T.C. Minderhoud

C. Spruyt

S. Huisman

E. Oskam

S.C.E. Schuit

M.D. Levin

ABSTRACT

Objective
To study the presence of bacterial disease and antibiotic use in patients in the emergency department (ED) included in the local sepsis protocol.

Methods
An observational retrospective cohort study. Adults aged > 18 years, presenting to the ED of a large teaching hospital, from 1 January to 1 June 2011, with more than two SIRS criteria and a clinical suspicion of sepsis were included.

Results
Bacterial disease was suspected or confirmed in only 71% of all the patients with suspected sepsis (2008 definition) and consequently treated with antibiotics. Most of these patients (58%) suffered from systemic inflammatory response syndrome (SIRS) without signs of organ dysfunction, hypotension or hypoperfusion. Despite absence of bacterial disease in 29% of the patients after rigorous diagnostics, median antibiotic treatment in this group was still seven days (IQR 4-10).

Conclusions
Standard sepsis detection using SIRS criteria and clinical suspicion identified patients with suspected or confirmed bacterial disease in 71% of the cases. A significant proportion of patients were exposed to prolonged antibiotic use without proof of bacterial disease. This study illustrates the difficulties in correctly identifying bacterial disease and sepsis, and shows that overuse of antibiotics may be the consequence.

INTRODUCTION

Over the last decade, sepsis has been increasingly recognised as a major cause of death. After Rivers' publication and the start of the Surviving Sepsis Campaign, early detection and treatment has become a general endeavour with a special focus on the early administration of antibiotics.[1] It should be noted that current evidence regarding early treatment with antibiotics was founded on studies including only patients with severe sepsis (mostly needing ICU treatment) or septic shock.[1-4] However, in the effort to avoid delays in identifying sepsis many emergency departments (EDs) have started using the criteria of systemic inflammatory response syndrome (SIRS) and a clinical suspicion of infection as a way to screen their patients. The difficulty in identifying severe sepsis is that testing for organ damage (for example: renal function) takes time, whereas the recommendation is to treat severe sepsis within one hour (2012 Sepsis Guidelines[5]). Under the presumption that this waiting time may harm the patient, patients are increasingly treated with antibiotics without awaiting (all) the test results.

The Surviving Sepsis Guidelines of 2004 recommended administration of antibiotics in patients with severe sepsis or septic shock. The 2008 guideline did not give guidance as to how patients should be screened, but the 2012 guideline recommended screening of potentially infected seriously ill patients. The proposed instrument to screen for severe sepsis was an instrument based on SIRS criteria and clinical judgment, which was, however, only validated in ICU patients.[3]
From the very start, the SIRS criteria were criticised as defining condition for sepsis and recent publications refuelled the discussion.[6-8] In February 2016, sepsis was redefined, removing the SIRS criteria and adding that the term sepsis had to be reserved for patients with severe organ dysfunction. The term 'severe' sepsis was dismissed and replaced by sepsis-3 and septic shock.[9] However, SIRS criteria are still in use in clinical practice.

This study was undertaken to evaluate current practice and study the likelihood of bacterial infection in patients treated for sepsis in the ED according to the SIRS criteria.
To address issues regarding antimicrobial stewardship, duration of antibiotic therapy was also evaluated.

METHODS

Study design and setting
A retrospective analysis was conducted using a cohort of consecutive adult patients presenting to the ED of the Albert Schweitzer Hospital (a large teaching hospital in Dordrecht, the Netherlands) from 1 January to 30 June 2011.

Inclusion
Patients diagnosed with sepsis (2008 definition) according to two or more SIRS criteria and a clinical suspicion of an infection, as assessed by the resident or ED physician, who received antibiotics upon admission were eligible.

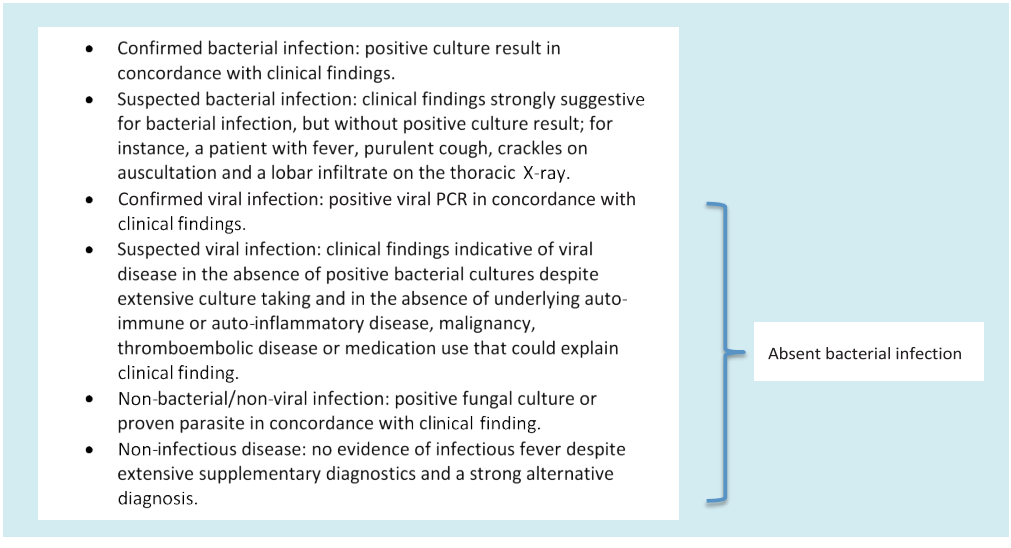
Data collection

Data regarding vital parameters were extracted from the standard protocol form if completely filled out. All additional and missing data were extracted by hospital chart review. If more than one measurement of parameters was done in the ED, the most aberrant measurement was used in the analysis (the lowest BP recorded in the ED, or the highest respiratory rate or pulse). The primary investigator, as well as authors Spruyt and Huisman, performed the data extraction. The primary investigator then checked the data and in case of doubt regarding the primary outcome measures the case was discussed between the primary investigator and Dr. Levin until consensus was reached. Consensus was reached in all 37 patients that were discussed.

Definitions

The definitions for SIRS, sepsis, severe sepsis, sepsis-induced hypotension and septic shock were derived from the guidelines of the Surviving Sepsis Campaign in 2008.
For the sake of clarity, in this article sepsis conform the old criteria will be referred to as sepsis and if referring to the new definition, sepsis-3 will be used.
The criteria to define confirmed or suspected bacterial infection were derived from an article by Limper,[10] as displayed in figure 4.1. In addition to these criteria another criterion, as found in previous literature, was added for further clarification of suspected bacterial disease: 'Clinically documented infection: presence of gross purulence or an abscess (anatomical and/or by imaging and/or histological evidence), which may not be microbiologically documented if the culture remains sterile due to antibiotic therapy.'[11]

Figure 4.1 Definitions of groups of infection, derived from Limper[10]



Outcome measures

Primary outcome measures were the number of patients with confirmed or suspected bacterial infection as assessed by the primary investigator using predefined criteria (stated above) and days of antibiotic use in these patients. Secondary outcome measures were severity of sepsis, rate of ICU admission, and mortality.

Data analysis

Results are expressed as mean ± standard deviation (SD), or median ± inter-quartile range (IQR) depending on normality of the data. Comparison between patients with and without bacterial infection was performed using the χ^2 test for categorical variables and the Students t-test (equal variances) or nonparametric Mann-Whitney U test for continuous data with non-normality. Statistical analyses were carried out using IBM SPSS 22.0.0 for OSX (SPSS, Chicago, IL). Missing data were excluded list-wise.

Ethics

The local institutional ethics review board approved the study design and a waiver for the retrieval of informed consent was obtained.

Table 4.1 Baseline characteristics

	Confirmed / suspected bacterial infection (n = 191)	Absent bacterial infection (n = 78)	P-value
Age, years (± SD)	67 (± 17,8)	61 (± 18,8)	,009^
Males, %	89 (46,8%)	37 (46,8%)	,900*
Comorbid conditions			
Immune deficiency	43 (22,6%)	24 (30,4%)	,155*
Current malignancy	28 (14,7%)	11 (13,9%)	,906*
Liver cirrhosis	2 (1.1 %)	2 (2.5 %)	,351*
Renal insufficiency	30 (15,8%)	6 (7,6%)	,080*
Congestive heart failure	8 (4,2%)	5 (6,3%)	,446*
Respiratory disease (COPD)	28 (14,8%)	19 (24,1%)	,060*
Laboratory findings			
C-reactive protein (CRP) day 0	151 (± 132,0)	66 (± 66,97)	,000^
CRP maximum (first 72 hrs)	212 (± 124,1)	99,7 (± 74,0)	,000^
Bilirubin	16,0 (± 10,5)	17,7 (± 38,8)	,571^
Creatinine	95 (± 54,7)	85,3 (± 39,1)	,143^
Lactate	2,31 (± 2,89)	1,94 (± 1,19)	,292^
Bacterial outcomes			
Positive blood cultures	51 (26,8%)		
Sepsis severity			
Sepsis	111 (58.1%)	55 (70,5%)	,058*
Severe sepsis	57 (30,0%)	20 (26,3%)	,511*
Sepsis-induced hypotension	18 (9,5%)	2 (2,6%)	,057*
Septic shock	5 (2,6%)	1 (1,3%)	,506*
Length of stay (days) (median (IQR)	8 (5-11,7)	6 (3-10)	,006§
ICU admission %	8 (10,1%)	17 (8,9 %)	
Duration of antibiotic treatment (median + IQR)	10 (7-14)	7 (4-10)	,000§
Mortality	4 (5%)	17(8,9%)	,295*

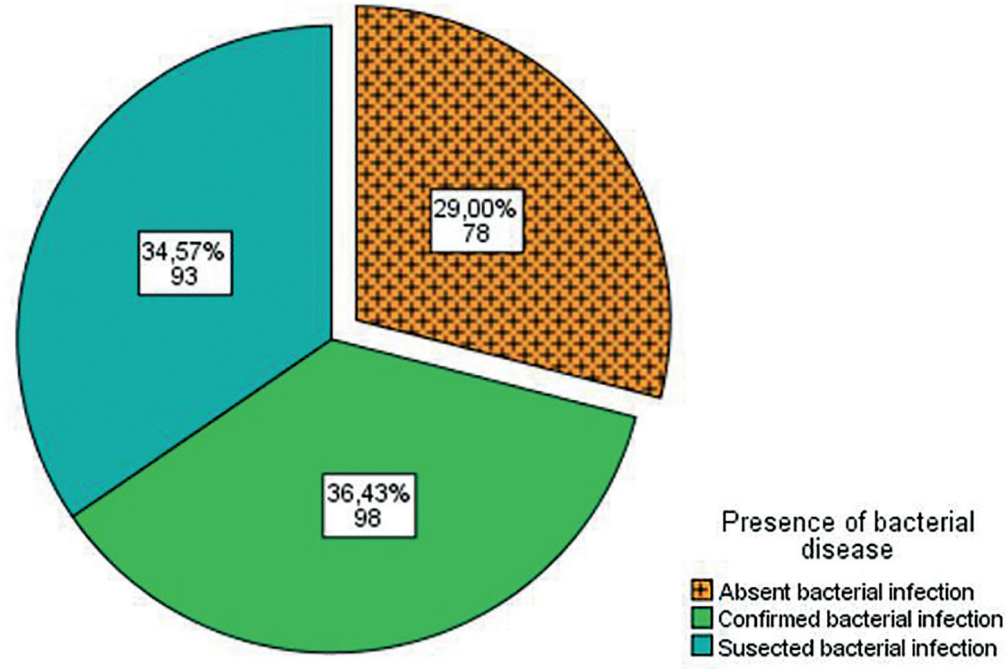
*Continuous data are presented as mean ± standard deviation, unless otherwise mentioned. Categorical data as number (percentages); *chi square, ^t-test, § Mann-Whitney U test.*

RESULTS

Patients and likelihood of bacterial infection

From 1 January 2011 to 30 June 2011, a total of 269 patients were diagnosed with sepsis (2008 definition) in the ED and received antibiotic treatment in the ED. Table 4.1 shows the baseline characteristics and main outcome measures. Retrospective analysis of clinical signs, cultures and other investigations using predefined criteria[10] (figure 4.1) showed a confirmed bacterial infection in 98 (36%) patients, of whom 51 patients had bacteraemia. In addition, 93 patients (35%) were classified as suspected bacterial disease without microbiological proof. A total of 78 patients (29%) did not have objective evidence of bacterial disease. Amongst them 21 suffered from proven or suspected viral infection. Figure 4.2 illustrates the proportions.

Figure 4.2 Presence of bacterial disease



Severity of illness

In total 71% of patients, identified with sepsis in the ED, were likely to have bacterial infection. In the group with bacterial infection, the largest proportion (58%) fulfilled criteria for sepsis, 30% fulfilled criteria for severe sepsis, and only 9.5% showed sepsis-induced hypotension. A small percentage (3%) suffered from septic shock. This means that the

largest proportion of the patients identified with bacterial infection (58%) would probably not fulfil the current sepsis-3 definition, although mental status was not documented reliably in all patients.

Factors associated with patients without bacterial infection

As shown in table 4.2 no significant differences in sex or comorbidity between patients with and without bacterial disease were established. The patients with bacterial disease were significantly older compared with the group without bacterial disease (p = 0.014). C-reactive protein at day 0 and day 3 was significantly higher in patients with bacterial disease than those without (p < 0.001 in both cases).

Table 4.2 SIRS criteria and bacterial infection

Mean (± SD)	Confirmed / suspected bacterial infection (n = 191)	Absent bacterial infection (n = 78)	P-value
Temperature (continuous)	38,99	38,89	,475^
Normothermia (36-38 °C)	n =2 (2,5%)	n = 17 (8,9%)	,061*
Leucocytes count (mean ± SD)	13,6 (± 7,2)	10,5 (± 5,4)	,001^
Blood pressure (MAP)	91,9 (± 17,3)	97,7 (± 15,3)	,013^
Systolic BP	129,3 (± 25,3)	133,5 (± 19,7)	,205^
Diastolic BP	73,3 (± 17,8)	79,7 (± 16,1)	,007^
Respiratory rate	24,6 (± 6,7)	23,6 (± 6,6)	,279^
Pulse	107,8 (± 18,7)	115,6 (± 22,9)	,004^
SIRS criteria			
≤3	n = 132	n = 62	,018*
4	n = 74	n = 17	

*^ Students t-test, * Pearson c2*

The number of SIRS criteria was significantly associated with the presence of bacterial disease. The odds ratio of having a bacterial infection was 2.32 (CI 1.3-4.3) if all four SIRS criteria were met in comparison with ≤ 3 criteria. The mean arterial pressure was significantly lower in the group with proven infection (p = 0.012), even though this was not reflected in the systolic blood pressure, but rather in the diastolic blood pressure. Leucocyte count was significantly higher in the group with bacterial infection (p = 0.001). Unexpectedly, patients with bacterial disease had a lower pulse than patients without a bacterial disease (p < 0. 001). Taking into account the severity of sepsis, the patients with more severe forms of sepsis (severe sepsis, sepsis-induced hypotension or septic shock) were significantly (p = 0.046) more likely to have bacterial infection compared with the group with sepsis alone.

Alternative diagnosis in patients without bacterial infection

To further understand how patients become misdiagnosed as possible sepsis we carefully studied the alternative diagnoses in the residual group (table 4.3). The most frequent alternative diagnosis was exacerbation of chronic obstructive pulmonary disease (COPD) with or without viral respiratory infection (n = 26). Congestive heart failure (n = 7), neutropenic fever (n = 5), pulmonary embolism (n = 4) and viral pneumonia due to H1N1 influenza (n = 4) were the most prevalent alternative diagnoses, apart from quite a large group (n = 15), in which no clear diagnosis was made.

Table 4.3 Final diagnosis in patients without bacterial infection

Characteristics of patients in sepsis protocol without bacterial infection

Clinical severity of sepsis	N = 78	Final diagnosis	N
Sepsis	55	Arrhythmia / congestive heart failure	5
		Exacerbation of COPD or upper respiratory infection (viral)	2 1
		Neutropenic fever	3
		Pulmonary embolism	3
		Malignancy (tumour-related fever)	2
		Epstein-Barr virus infection	1
		Meningitis (viral)	1
		Pericarditis (viral)	1
		Unclear diagnosis / insufficient information	1 2
Severe sepsis	20	Neutropenic fever	2
		Fever in immunocompromised host (not neutropenic)	3
		H1N1 infection / pneumonia	3
		Viral hepatitis (Epstein-Barr virus)	1
		Exacerbation of COPD / upper respiratory infection	5
		Pulmonary embolism	1
		Congestive heart failure	2
		Unclear diagnosis / insufficient information	3
Sepsis-induced hypotension	2	H1N1 pneumonia	1
		Multi organ failure in a patient with new diagnosis of aggressive lymphoma (DLBCL) and history of mRCC	1
Septic shock	1	Diabetic keto-acidosis	1

Met onder de table de tekst
COPD = chronic obstructive pulmonary disease; H1N1 = influenza of subtype H1N1; DLBCL = diffuse large B-cell lymphoma; mRCC = metastatic renal cell carcinoma

Antibiotic use

Data regarding duration of antibiotic treatment were available for 251 patients. In the remaining patients, data could not be retrieved, for example due to transfer to another hospital. The median duration of antibiotics for all patients was 9 days (IQR 3-15), but media[11] days (IQR 7-14) in patients with bacterial infection and 7 days (IQR 4-10 days) in patients without bacterial infections as displayed in table 4.4.

Table 4.4 Antibiotic duration in subgroups

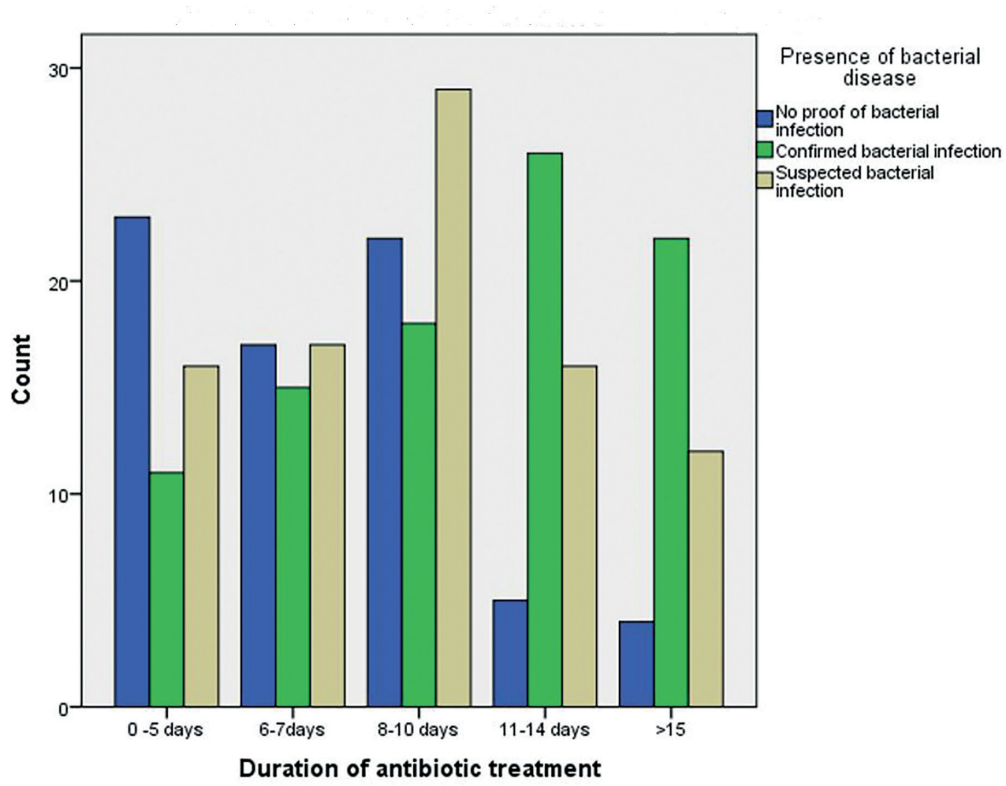
n = 251	Duration of antibiotic treatment (median, days)	IQR
Absent bacterial infection (n = 72)	7	4-10
Suspected / confirmed bacterial infection (n = 179)		
Pulmonary (n =88)	10	7-11
Abdominal (n = 25)	7	5-12
Urinary tract (n = 39)	10	7-15
Soft tissue / skin infection (n = 15)	13	8-14
Endocarditis (n = 4)	35	N/A
Joints / bone infection (n = 3)	9	N/A
CNS (epidural abscess) (n = 1)	84	N/A
PM line associated infection (n = 1)	42	N/A
Ear-nose-throat (n = 1)	8	N/A
Bacteraemia with unknown focus (n = 2)	16 (mean)	N/A

Total n = 251

IQR = interquartile range; CNS = central nervous system; PM = pacemaker; N/A not applicable

The most frequent infection was respiratory infection, which was treated for a median of 10 days. This is remarkable as evidence has shown that shorter treatments are safe and effective.[12-14] Antibiotics were stopped in the first 5 days in only 23 (32%) of the patients without bacterial infection, see figure 4.3 for more information. In this group antibiotic duration was significantly longer (p = 0.037) in patients with COPD in relation to patients with other comorbidities (current malignancy, congestive heart failure, liver cirrhosis, chronic renal insufficiency). Median duration was shortest in patients with confirmed or suspected viral disease (median 3 and 4 days, IQR 1-10.5 and 1.5-7.5).

Figure 4.3 Distribution of antibiotic duration among subgroups



DISCUSSION

This study demonstrates that in almost 30% of the patients with suspected sepsis in the ED no objective evidence of bacterial disease could be found. This puts patients at risk of overtreatment with antibiotics. This finding is in concordance with an earlier report of patients admitted to the ICU with a diagnosis of sepsis,15 in whom no evidence of bacterial infection could be found in 13% and only a possible infection could be established in 30%. In spite of the improved outcome of patients treated early with antibiotics for severe sepsis or septic shock, this antibiotic overtreatment in patients with sepsis is a very important finding and often underreported. It is of paramount importance to establish that patients included in sepsis research (on clinical suspicion) are in fact suffering from an infectious disease. Future research will have to report infectious outcomes in detail, to enable correct interpretation and extrapolation of the results.

Antibiotic treatment within the hour

The most important intervention in severe sepsis treatment in the last decades, next to fluid treatment, has been the emphasis on early antibiotic treatment. The problem in every ED, however, is that signs and symptoms of severe sepsis can be deceiving or occult.

Postponing antibiotic treatment whilst awaiting basic test results (i.e. kidney function, chest X-ray) does not fit well within the one-hour target which has been outlined by the sepsis guidelines. The benefit of early antibiotic treatment has been established in suspected sepsis patients admitted to the ICU.[1-4,16-18] Two other studies showed benefit of early antibiotic treatment in ED patients but selected only patients with sepsis and organ dysfunction or patients with hypotension/ hyperlactataemia (lactate > 4 mmol/l).[19-21] However the largest group identified by our screening did not have organ dysfunction, and only about 10% needed ICU care. This means that more than half of our patients could have awaited basic test results (which might have raised suspicion of alternative diagnoses), thus allowing more time to consider if antibiotic treatment is really indicated. In pneumonia, studies have shown that treatment within four hours is safe.[22] This leaves more than enough time for at least a chest X-ray and lab results to come in.

Antibiotic treatment in the ED within the hour should generally be reserved for critically ill patients, patients deteriorating quickly, or specific patient groups such as neutropenic patients. Future research will hopefully guide us further as to which risk-stratification score (Modified Early Warning Score (MEWS), National Early Warning Score (NEWS) or Quick Sequential Organ Failure Assessment (qSOFA) is most helpful with identifying patients at risk of deterioration or death.

Duration of antibiotic therapy

Overall the duration of antibiotic therapy was long in our cohort. This may reflect local standards or may be because our patients were selected in 2011.
Of concern, patients in our cohort without evidence for bacterial disease were treated with antibiotics for a median duration of 7 days, pointing to overuse. Antibiotic treatment was stopped in the first 5 days in only 32% of the patients with negative culture results. Several reasons for the prolonged use of antibiotics can be suggested.

1. The ED presumptive diagnosis of sepsis makes it hard to stop antibiotics despite negative cultures. This could be due to cognitive errors such as the tendency to stick to first impressions (anchoring error) and the tendency to stick to prior diagnoses (confirmation bias) despite conflicting evidence.
2. The large number of patients suffering from COPD in this subgroup, in whom antibiotic treatment is often given despite negative cultures. Even so, evidence is mounting that shorter regimens are safe for bronchitis and pneumonia.12-14
3. Clinical improvement of patients after admission and starting antibiotics.
4. Fear of undiagnosed bacterial disease by physician or patient.
5. Fear of inducing antimicrobial resistance if antibiotics are stopped prematurely. This is a theoretical problem which is hard to prove or refute in practice. Though widespread, it has been challenged over recent years. New research in the area of pneumonia shows that shorter treatment regimens are safe without signs of inducing microbial resistance.12-14 A review in 2016, looking at de-escalation of antimicrobials, concluded that de-escalation appears safe and effective for certain conditions, but calls for further, high-quality, research.23 All in all, de-escalation seems safe, and if antibiotics are used for too long for fear if inducing resistance, this might actually constitute antibiotic overuse.

Limitations

Limitations of the study are its retrospective character and the single-cohort design in a single hospital. Another point of concern is the allocation of patients to groups suffering from proven/ suspected or no bacterial disease. It has been pointed out before that many patients suffering from a bacterial infection (i.e. pneumonia) may not have positive culture results. A patient suffering from urosepsis may have negative cultures due to prior treatment initiated by the primary care physician. In these patients, it is hard to determine in retrospect if they were truly suffering from bacterial disease. We have put a lot of effort into accurately determining the correct group for each patient, but in some cases it is inevitable that discussion will always remain. However, as this reflects daily practice it does not reduce our concerns of overtreatment and the protracted duration of antibiotic use.

Future investigations evaluating the sepsis campaign or regarding screening in the ED should report microbiological outcomes and include overuse and possible harm of antibiotics as endpoint to avoid a singular focus on benefits of early sepsis treatment.

Relevance and recommendations

With the new sepsis-3 definition, treatment within one hour based only on SIRS criteria cannot be substantiated. However, it is still difficult to know which patients in the ED have to be treated within the hour. qSofa was introduced as an instrument to identify patients with sepsis who are likely to fare poorly and should thus be treated early with broad-spectrum antibiotics.9 This is an important step forward. However, it was noted that early treatment should not be limited to patients with a positive qSofa. Several reports have been made since, but acceptance of qSOFA is not universal. One investigation showed poor sensitivity of 63% for qSOFA in the ED population.24 The same report found that the NEWS was the most accurate tool in predicting in-hospital and ICU mortality. In the UK, use of NEWS is mandatory and qSOFA has not been implemented. Since the best way to identify a septic patient in the ED is still under discussion, this study offers valuable information regarding the use of SIRS criteria.
With respect to antibiotic duration and de-escalation, the current guidelines recommend daily reconsideration of antibiotic therapy. Unfortunately, only a few studies have been performed regarding the safety of early de-escalation in patients outside the ICU. More research is needed in the area of de-escalation in suspected sepsis patients.

CONCLUSION

Sepsis detection in the ED is a continuous challenge. This study shows that early recognition of sepsis using SIRS criteria leads to over identification of sepsis. More than half of the patients suspected of sepsis would probably not fulfil the current sepsis-3 definition, and almost 30% did not have objective evidence of a bacterial infection. In some of the patients without bacterial infection, awaiting basic tests might have confirmed an alternative diagnosis and antibiotic treatment could have been avoided.

In a significant proportion of patients, empiric therapy was justified but with a median duration of therapy of seven days de-escalation should have been much more rigorous.

REFERENCES

1 Rivers E, Nguyen B, Havstad S, et al. Early Goal-Directed Therapy Collaborative Group: Early goal-directed therapy in the treatment of severe sepsis and septic shock. N Engl J Med. 2001;345:1368-77.
2 Kumar A, Roberts D, Wood KE, et al. Duration of 10 hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. Crit Care Med. 2006;34:1 589-96
3 Moore LJ, Jones SL, Kreiner LA, et al. Validation of a screening tool for the early identification of sepsis. J Trauma. 2009; 66:1539-46.
4. Barie PS, Hydo LJ, Shou J, Larone DH, Eachempati SR. Influence of antibiotic therapy on mortality of critical surgical illness caused or complicated by infection. Surg Infect (Larchmt). 2005;6:41-54.
5 Dellinger RP, Levy MM, Rhodes A, et al. The Surviving Sepsis Campaign Guidelines Committee including The Pediatric Subgroup Surviving Sepsis Campaign: International Guidelines for Management of Severe sepsis and septic Shock, 2012. Intensive Care Med. 2013;39:165-228.
6 Vincent JL. Dear SIRS, I'm sorry to say that I don't like you… Crit Care Med. 1997;25:372-4.
7 Kaukonen KM, Bailey M, Pilcher D, et al. Systemic inflammatory response syndrome criteria in defining severe sepsis. N Engl J Med. 2015;372:1629-38.
8 Beesley SJ, Lanspa MJ. Why we need a new definition of sepsis. Ann Transl Med. 2015;3:296.
9 Shankar-Hari M, Phillips GS, Levy ML, et al. Developing a New Definition and Assessing New Clinical Criteria for Septic Shock For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016;315:775-8.
10 Limper M, Eeftinck Schattenkerk D, de Kruif MD, et al. One year epidemiology of fever at the Emergency Department. Neth J Med. 2011;69:124-8.
11 Alberti C, Brun-Buisson C, Burchardi H, et al. Epidemiology of sepsis and infection in ICU patients from an international multicentre cohort study. Intensive Care Med. 2002;28:108-21.
12 Fagalas ME, Avgeri SG, Matthaiou DK, Dimopoulos G, Siempos I. Short- versus long-duration antimicrobial treatment for exacerbations of chronic bronchitis: a meta-analysis. J Antimicrob Chemother. 2008;62:442-50.
13 Li JZ, Winston LG, Moore DH, Bent S. Efficacy of Short-course antibiotic regimens for community-acquired pneumonia: A Meta-analysis. Am J Med. 2007;120:783-90.
14 El Moussaoui R, de Borgie CAJM, van den Broek P, et al. Effectiveness of discontinuing antibiotic treatment after three days versus eight days in mild to moderate-severe community acquired pneumonia: randomised, double blind study. BMJ. 2006;332:1355.
15 Klein Klouwenberg PMC, Cremer O, van Vught LA, et al. Likelihood of infection in patients with presumed sepsis at the time of intensive care unit admission: a cohort study. Crit Care. 2015;19:319.
16 Castellanos-Ortega A, Suberviola B, Garcia-Astudillo LA, et al. Impact of the surviving sepsis protocols on hospital length of stay and mortality in septic shock patients: results of a three-year follow-up quasiexperimental study. Crit Care Med. 2010;38:1036-43.
17 El Solh AA, Akinnusi ME, Alsawalha LN, et al. Outcome of septic shock in older adults after implementation of the sepsis ''Bundle''. J Am Geriat Soc. 2008;56:272-8.
18 Puskarich MA, Trzeciak S, Shapiro MI, et al. Association between timing of antibiotic administration and mortality from septic shock inpatients treated with a quantitative resuscitation protocol. Crit Care Med. 2011;39:2066-71
19 Gaieski DF, Mikkelsen ME, Band RA, et al. Impact of time of antibiotics on survival in patients with severe sepsis or septic shock in whom early goal-directed therapy was initiated in the emergency department. Crit Care Med. 2010;38: 1047-53.
20 Kennedy M, Joyce N, Howell MD, Mottley JL, Shapiro NI. Identifying infected emergency department patients admitted to the hospital ward at risk of clinical deterioration and Intensive care unit transfer. Acad Emerg Med. 2010;17:1080-5.
21 The Australasian Resuscitation in Sepsis Evaluation (ARISE) Investigators and the Australian and New Zealand Intensive Care Society (ANZICS) Adult Patient Database (APD) Management Committee. The outcome of patients with sepsis and septic shock presenting to emergency departments in Australia and New Zealand. Crit Care Resusc. 2007;9:8-18.
22 Wiersinga WJ, Bonten MJ, Boersman WG, et al, for SWAB/NVALT. Guidelines Community-acquired Pneumonia. November 2011. Available through www.swab.nl
23 Ohji G, Doi A, Yamamoto S, Iwata K. Is de-escalation of antimicrobials effective? A systematic review and meta-analysis. Int J Infect Dis. 2016;49:71-9.
24 Churpek MM, Snyder A, Han X, et al. Quick Sepsis-related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients outside the Intensive Care Unit. Am J Respir Crit Care Med. 2017;195:906-11.

4

Chapter 5

Acute admission of patients with suspected influenza. Can we identify high-risk patients based on clinical symptoms?

submitted BMC Emergency Medicine

T.C. Minderhoud

A. Mers

M. Jonges

R. van Mansfeld

M. A. van Agtmael

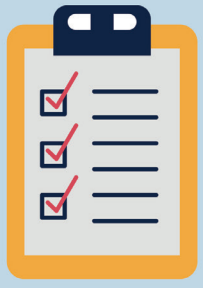
P. W.B. Nanayakkara

M.K. Bomers

INFLUENZA

ED IDENTIFICATION AND ISOLATION

During influenza pandemics, many patients present in the Emergency Department with possible influenza. We performed a retrospective analysis of the 2017–2018. In this cohort of 419 patients, 148 (35%) tested positive for influenza.

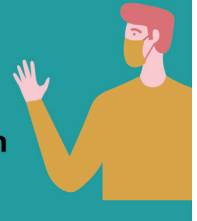


A DECISION RULE WAS IMPLEMENTED TO IDENTIFY HIGH RISK INFLUENZA PATIENTS IN 2017

Three items were assessed
- Acute onset
- Upper respiratory tract symptoms
- Fever with systemic symptoms

If a patient had ALL three items, a patient was high risk and isolated pre-emtively.

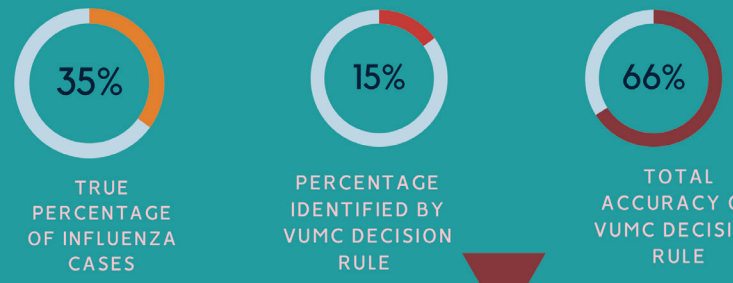
Acording to the VUmc decision rule, only 1/10 patients was high risk



Clinical complaints

Among influenza patients the most common complaints were coughing (87,8%), fever (73.6%), and dyspnea (56.1%). In patients without influenza, the most common complaints were coughing (64.9%), dyspnea (57.6%) and fever (50.9%).

Clinical decision rule results



Conclusion

The clinical decison rule used in our emergency room to guide isolation in possible influenza has a high specificity of 94 % and leads to restrictive use of isolation rooms. However, due to its low sensitivity of 15,9 % for identifying patients with influenza, in-hospital transmission of influenza is unacceptably high. Clinical symptoms are unreliable to identify influenza, and rapid, molecular 24/7 influenza testing is the most efficient solution for correct allocation of isolation measures in patients with suspected influenza.

ABSTRACT

Objective

A clinical decision rule (CDR) was introduced in our tertiary care hospital to use isolation measures for influenza more effectively. The aim of this study is to evaluate the accuracy our clinical decision rule.

Methods

We conducted a retrospective cohort study of all admitted patients who were tested for influenza in the 2017-2018 influenza season. Symptoms at admission were extracted from the electronic health record, as was the influenza PCR result. Performance of our CDR was evaluated by calculating sensitivity, specificity and overall accuracy.

Results

Of all patients tested, 35,5 % were influenza positive. The CDR had a sensitivity of 15,9% (95% CI 10.33 – 22.84), specificity of 94.3 (95% CI 90.8 – 96.8) and an accuracy of 66,4 (95%CI 61.6 -71.0) %.

Conclusion

The CDR used in our emergency room to guide isolation in possible influenza has a high specificity of 94 % and leads to restrictive use of isolation rooms. However, due to its low sensitivity of 15,9 % for identifying patients with influenza, in-hospital transmission of influenza is a risk. Clinical symptoms are unreliable to identify influenza, and rapid, molecular 24/7 influenza testing seems the most efficient solution for correct allocation of isolation measures in patients with suspected influenza.

INTRODUCTION

Background

Each winter an influenza epidemic occurs in Europe. In 2017-2018 the influenza-epidemic in the Netherlands lasted 18 weeks, starting December 11th 2017 and ending April 15th 2018. An annual survey by the National institute for Public Health and Environment (RIVM)[1] shows that in 2017-2018 approximately 900.000 people got infected. An estimated 16.000 people were admitted into hospitals due to influenza-like illness.

Importance

In many hospitals, the yearly influenza epidemic leads to dilemmas on testing and isolation. On the one hand, we want to prevent transmission of influenza in the hospital. On the other hand, the number of isolation rooms is restricted. For optimal use of available isolation rooms, we want to avoid isolation in patients who turn out to be influenza negative. With the yearly surge of patients needing admission, and a concomitant increase of influenza isolation, ED crowding, lack of admission capacity are often seen during the influenza season[2,3].

Several strategies have been proposed to optimize diagnosis and the use of isolation measures: clinical decision rules and rapid molecular testing. A recent review [4] of literature published up till 2017 showed good sensitivity and specificity of rapid molecular testing. Though the rapid testing decreases turn-around times, in many hospitals it is not available out of office hours. In our hospital we implemented an influenza protocol in 2017 to optimize care for suspected influenza patients. In an effort to diminish unnecessary use of isolation rooms, it was decided to presumptively isolate only high-risk patients. These patients were admitted in a single room in droplet isolation, conform infection prevention guidelines [5]. Low risk patients were admitted with standard hygiene measures on a ward. After definitive diagnosis, influenza positive patients were admitted in a single isolation room or, depending on availability and influenza strain, were admitted in cohort isolation. In the same season, turnaround time for rapid molecular testing between office hours was improved with results available within 2 hours including Saturday and Sunday [6]. As a consequence, in case a low-risk patient unexpectedly did have influenza, improving the time to diagnosis would allow less time for in-hospital transmission.

Goals of this investigation

To evaluate this strategy, the cohort of 2017-2018 was studied retrospectively. The primary outcome was how accurate the clinical decision rule was at identifying influenza in patients needing admission. Secondary, we also studied if other combinations of symptoms could improve accuracy.

MATERIALS AND METHODS

Study design and setting

We conducted a retrospective cohort study of the 18 weeks in which the 2017/2018 influenza epidemic took place. The cohort comprised all adult patients who were tested for influenza virus and admitted to the Amsterdam University Medical Center, location Vrije Universiteit Medical Center (VUmc). The research received an approval from the

Medical Ethics Board of the VU University medical center. A waiver for informed consent was granted. Our approach and reporting follows STROBE guidelines.

Patient selection
The patient data were derived from the electronic medical records after an automated search for all adult (>18 years) patients tested for influenza virus and admitted during this period. All patients whose PCR test date was more than five days after their hospital admission, were excluded from the study. Age, sex, and clinical measurements including body temperatures, and laboratory findings were also extracted automatically from the electronic patient file.

Data abstraction
The clinical symptoms included in this study were extracted from the physicians electronic notes by manual search. Data abstraction was done by a bachelor student of Health Sciences (Amber Mers- AM). Based on literature research, AM and a specialist in Infectious Diseases, Marije Bomers (MB) determined which clinical symptoms were relevant for prediction of influenza, and they were defined in a coding log. After data extraction and coding the first 40 patients, AM and MB checked the coding, and discussed any problems in the coding process. Any issues were resolved by discussion, determining a definitive version of the coding log. Tanca Minderhoud (TM) checked the accuracy by checking random cases afterwards. Symptoms scored included: acute onset of symptoms, fever, headache, arthralgia, myalgia, coughing, sneezing, rhinitis, sternal pain and a sore throat. All these complaints were scored separately. An acute onset was defined as the start of symptoms within a time frame of 72 hours [8] Fever is defined as either a body temperature of 38.0 degrees Celsius or a history of fever as mentioned in the medical records.

VUmc Clinical Decision Rule
The VUmc clinical decision rule consists of 3 items, and was based on the case definition in the Influenza Guideline of the National institute for Public Health and Environment [9]. If patients fulfill all 3 items, the patient was deemed high risk. If the patient has less than 3 items, he is classified as low risk. The items are 1) acute onset of the symptoms, 2) a fever with at least one of the following systemic symptoms: headache, arthralgia or myalgia, and 3) respiratory symptoms: coughing, sneezing, rhinitis, sternal pain and/or a sore throat.

Outcome
The primary outcome measure was the result of the influenza test. Influenza testing was performed by Cepheid Xpert Flu A/B/RSV XC assay. The secondary outcome was the efficacy of allocation of isolation. Influenza isolation comprises of admission in a separate room and use of a protective gown, mask and gloves.

Primary Data analysis
All continuous variables with a normal distribution are expressed as percentages, means and standard deviations (SD). Non-normally distributed variables are expressed as median with interquartile range (IQR). Differences between groups were assessed using independent T-testing for normally distributed variables and Mann-Whitney U for non-normally distributed variables. Chi-square or Fisher's exact test were used for categorical variables. Sensitivity and specificity and Rand accuracy were calculated using crosstables. (Rand) accuracy is defined as the proportion of true results (true positives and true negatives) among the total number of cases examined[10].
To study if we had missed important clinical factors in our cohort, that were associated with a positive influenza test, a binary logistic regression was performed. The model was reduced using the back-ward selection method. All variables with a p-value above 0.05 were excluded from the model.

Missing values and extreme values
In ten people, clinical complaints could not be assessed due to lack of a history, for example due to reduced consciousness. In case of missing clinical complaints, patients were classified as low risk. One of these patients turned out to have influenza. NAll data analyses were performed using SPSS software (SPSS version 24.0, IBM, Armonk, NY, USA)

RESULTS

Characteristics of study subjects
A total of 419 patients that were tested for influenza in the season of 2018-2019 were identified. Of these 419 patients, 148 (35,3%) were influenza positive. The clinical signs, radiological and laboratory results of patients with and without influenza are shown below in table 5.1. The most common complaint in influenza patients were coughing (87,8%), fever (73.6%), and dyspnea (56.1%). In patients without influenza, the most common complaints were coughing (64.9%), dyspnea (57.6%) and fever (50.9%).

Testing properties of VUmc CDR and Dugas CDG
The clinical decision rule flagged 38 patients (9%) of all of the patients as high risk. Of those patients identified as high risk, 23 (60,5%) tested positive for influenza. Of the 381 (91%) patients who were classified as low risk, 125 (32,8%) tested positive in subsequent testing. Table 5.2 displays the performance of the single and combined parameters of the clinical decision rule. For respiratory signs the sensitivity was 91.8% (95%CI 86.08 - 95.68) with a specificity of 26.9% (95%CI 21.64 – 32.67) and total accuracy was 50.0% (95% CI 45.1-55.0) . For two out of three symptoms, the difference in performance were minimal. The combination fever with one or more systemic symptoms and an acute onset showed a sensitivity of 15.9% (95%CI 10.33 – 22.84) with a specificity of 89% and an accuracy of 63.0%. Adding respiratory signs to this, (conform the current CDR) did not alter sensitivity but did raise the specificity to 94.3% (95%CI 90.8 – 96.8). Rand accuracy was 66.4 % (95%CI 61.6 – 71.0).

Table 5.1 Baseline characteristics of patients with and without influenza

	Influenza (N = 148)	**No Influenza (N = 271)**
Age, mean (SD)	63.4 (19.0)	61.5 (17.3)
Female N (%)	73 (49.3%)	125 (46.1 %)
Systemic symptoms[1] N (%)		
Myalgia	27 (18.2)	34 (12.5)
Arthralgia	3 (2.0)	8 (3.0)
Shivering	23 (15.5)	41 (15.1)
Headache	26 (17.6)	36 (13.3)
Fever N (%)	109 (73.6)	138 (50.9)
Respiratory signs[1] N (%)		
Coughing	130 (87.8)	176 (64.9)
Rhinitis	34 (23.0)	51 (18.8)
Sternal pain	4 (2.7)	7 (2.6)
Sore throat	19 (12.8)	27 (10.0)
Dyspnea N (%)	83 (56.1)	156 (57.6)
Acute onset N (%)	64 (43.2)	87 (32.1)
X-ray findings N (%) [2]		
No pneumonia	91(61.5)	150 (55.4)
Possible pneumonia	33 (22.3)	67 (24.7)
Pneumonia	16(10.8)	34 (12.5)
Death (during admission) N (%)	15 (10.1)	26 (9.6)
Charlson Comorbidity Index, median (IQR)	4 (3-6)	4 (2-6)
CRP , mean (SD)[3]	67.7 (84.7)	90.5 (103.8)

N: Number, IQR: InterQuartile Range, SD: Standard deviation, CDR: Clinical Decision Rule.
1. Outcome is missing for 8 influenza patients, and 2 non-influenza patients. A combination of symptoms can be present in patients.
2. Outcome is missing for 8 influenza patients and 20 non-influenza patients.
3. Outcome is missing for 9 influenza patients and 24 non-influenza patients.

Table 5.2 Properties of VUmc clinical decision rule and of its components

Descriptions	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Rand Accuracy (95% CI)
Respiratory signsτ	91.8 (86.1 – 95.7)	26.9 (21.6 – 32.7)	41 (38.9 – 43.1)	85.5 (76.9 – 91.3)	50.0 (45.1 – 55.0)
Fever	74.8 (67.0-81.6)	47.5 (41.4-53.6)	43.4 (39.8-47.1)	77.8 (72.1-82.7)	57.1 (52.1-61.9)
Fever AND respiratory signs	67.8 (59.6-75.3)	64.1 (58.0-70.0)	50.4 (45.5-55.3)	78.7 (74.2-82.6)	65.4 (60.6-70.0)
Fever OR respiratory signs	99.3 (96.2-99.8)	10.3 (6.9-14.6)	37.4 (36.4-38.4)	96.5 (79.3-99.5)	41.5 (36.6-46.4)
Clinical Decision Rule (Fever*, acute onset, respiratory signs)	15.9 (10.3 – 22.8)	94.3 (90.8 – 96.8)	60.5 (45.4 – 73.9)	67.1 (65.3-68.7)	66.4 (61.6 – 71.0)

** Fever accompanied by one of more systemic symptoms, as per definition in the VUmc CDR.*
τ Respiratory signs: coughing, sneezing, rhinitis, sternal pain and/or a sore throat.
95% CI: 95 percent Confidence Interval PPV: Positive Predictive Value, NPV: Negative Predictive Value

LIMITATIONS

Our study has several limitations. First of all, our data are from a single influenza season with a relative high number of influenza B infections. Symptoms of influenza can vary according to the strain and the season, and therefore the results of our CDR and of the Dugas CDG could be different in other seasons. However, this variability in clinical symptoms probably also means that any clinical score will show variable performance over the years, as it is impossible to modify the score every year. This probably illustrates an inherent downside of using clinical symptoms in any type of predictive score.

Another limitation is that our data were from one center, and that the introduction of the CDR could have led to selection bias. It is quite likely that clinicians would have been inclined to think of influenza most in patients that would have symptoms that are included in the CDR. Since we only investigated patients in whom physicians already suspected influenza, our data cannot answer the question which patients we should select for influenza testing. Another limitation of the single center data is that our hospital ED is visited by a high number of patients undergoing specialized oncologic and hematologic care. In our sample the median comorbidity index (CCI) is 4, indicating considerable comorbidity. Signs and symptoms of influenza are more likely to be atypical in patients with advanced age or comorbidity. A more general hospital might have a different casemix, limiting the generalizability of our data.

A third limitation is the high prevalence of influenza in our sample. This could be due to selecting only the patients that were to be admitted, or that physicians only tested patients that were otherwise seen as 'high risk', for example due to being older or immunocompromised. In these patients, atypical presentations of common diseases are more prevalent. This would have affected the test parameters, especially the PPV and NPV. However, this is also precisely the group of patients that needs admission for influenza.

DISCUSSION

We investigated if we could improve allocation of isolation measures in patients with possible influenza awaiting test results by using a clinical decision rule to identify high risk patients. The goal was to minimize both inappropriate isolation of non-influenza patients and ward admittance of undiagnosed influenza patients.

In our evaluation, we found a very low sensitivity of our CDR for influenza of 15,9 % (95%CI 10.33 – 22.84) albeit with a high specificity of 90.8 (85.28 – 94.76). The accuracy was 66.4% (95%CI 61.6-71.0), which means that 66% of the patients were true positive or true negative. Compared to using no decision rule, in many hospitals it is custom to isolate all patients with suspicion of influenza. Given that 35% tested positive, if you use isolation preemptively, 65% of the patients in our sample would always be wrongly allocated. By changing the default, the CDR increased the prior chance of correct isolation of 35% to 66%. If all patients that might have influenza were not isolated (pending test-results) than the prior chance of having it right would have been 65%, and adding the CDR brought it up to 66.4%. The downside of this approach is the large number of false negatives. Only a 38 (9%) of 419 patients with possible influenza of patients were admitted directly in isolation because the CDR identified them as high risk. But following the definitive test result, another 125 (29.8%) had to be isolated. Though our CDR decreased the use of isolation measures, it was at the risk of increasing (possible) transmission to other patient and healthcare workers. The risk of transmission was limited in duration, due to short turnaround times and opening of the laboratory on Saturdays and Sundays.

An alternative approach to solve the dilemmas around isolation and influenza transmission is to implement 24/7 rapid viral testing. The review[4] in 2019 assessing accuracy of rapid testing also investigated the clinical impact of rapid molecular testing. It found high heterogeneity in design and outcome of the studies, but concluded that there was high quality evidence that length of stay was reduced. Two other studies, published in 2019, were not included in this review. Lankelma, in a Dutch study [11] reported positive effects on hospital flow and length of stay and also reported that implementation was cost-effective. Youngs [12] showed a reduction of hospital-acquired influenza after implementation of POCT, but cost-effectiveness was not reported. Some studies have also reported a reduction in antibiotic use after implementation of rapid testing [13, 14] but this effect was not confirmed in the review by Vos et al. However, due to the COVID-19 pandemic, rapid viral testing upon admission has become standard of practice in many hospitals for patients with respiratory complaints or fever.

In summary, the CDR used in our hospital in 2017 to guide isolation in suspected influenza patients was effective in reducing isolation. But due to a low sensitivity of 15,9 %, the risk of in-hospital transmission of influenza was high. In light of these limitations, 24/7 rapid molecular testing for patients, suspected of influenza and needing admission, is the best solution to ensure prudent use of isolation rooms and to avoid nosocomial transmission. It should be kept in mind that clinical symptoms have limited value in predicting the presence of viral disease, especially in older patients and patients with comorbidity. Due to COVID-19, rapid viral testing has already become common practice in the Emergency Department.

5

REFERENCES

1 RIVM. Annual report Surveillance of influenza and other respiratory infections in the Netherlands: winter 2017/2018. https://www.rivm.nl/bibliotheek/rapporten/2018-0049.pdf 2019.
2 Schilling PL, Campbell DA, Jr., Englesbe MJ, Davis MM. A comparison of in-hospital mortality risk conferred by high hospital occupancy, differences in nurse staffing levels, weekend admission, and seasonal influenza. Medical care. 2010;48(3):224-32.
3 Glatman-Freedman A, Kaufman Z, Stein Y, Sefty H, Zadka H, Gordon B, et al. Influenza Season Hospitalization Trends in Israel: A Multi-Year Comparative Analysis 2005/2006 Through 2012/2013. Journal of hospital medicine. 2017;12(9):710-6.
4 Vos LM, Bruning AHL, Reitsma JB, Schuurman R, Riezebos-Brilman A, Hoepelman AIM, et al. Rapid molecular tests for influenza, respiratory syncytial virus, and other respiratory viruses: a systematic review of diagnostic accuracy and clinical impact studies. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2019.
5 Control CfD. Prevention Strategies for Seasonal Influenza in Healthcare Settings. 2018.
6 M. Jonges ADH, B. Blok-Schuurs, M.K. Bomers, R. van Mansfeld. Rapid influenza diagnostics supports effective and efficient patient isolation. 2019;Presented at the European Society for Clinical Virology Conference september 2019.
7 Dugas AF, Hsieh YH, Lovecchio F, Moran GJ, Steele MT, Talan DA, et al. Derivation and Validation of a Clinical Decision Guideline for Influenza Testing in Four U.S. Emergency Departments. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2019.
8 Boivin G, Hardy I, Tellier G, Maziade J. Predicting influenza infections during epidemics with use of a clinical case definition. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2000;31(5):1166-9.
9 dr. A. Meijer RCI. Influenza Richtlijn https://lci.rivm.nl/richtlijnen/influenza 2012 (Dutch)
10 Metz CE. Basic principles of ROC analysis. Seminars in nuclear medicine. 1978;8(4):283-98.
11 Lankelma JM, Hermans MHA, Hazenberg E, Macken T, Dautzenberg PLJ, Koeijvoets K, et al. Implementation of point-of-care testing and a temporary influenza ward in a Dutch hospital. The Netherlands journal of medicine. 2019;77(3):109-15.
12 Youngs J, Marshall B, Farragher M, Whitney L, Glass S, Pope C, et al. Implementation of influenza point-of-care testing and patient cohorting during a high-incidence season: a retrospective analysis of impact on infection prevention and control and clinical outcomes. The Journal of hospital infection. 2019;101(3):276-84.
13 Tillekeratne LG, Bodinayake CK, Nagahawatte A, Vidanagama D, Devasiri V, Arachchi WK, et al. Use of Rapid Influenza Testing to Reduce Antibiotic Prescriptions Among Outpatients with Influenza-Like Illness in Southern Sri Lanka. The American journal of tropical medicine and hygiene. 2015;93(5):1031-7.
14 Falsey AR, Murata Y, Walsh EE. Impact of rapid diagnosis on management of adults hospitalized with influenza. Archives of internal medicine. 2007;167(4):354-60.
15 Vegting IL, Alam N, Ghanes K, Jouini O, Mulder F, Vreeburg M, et al. What are we waiting for? Factors influencing completion times in an academic and peripheral emergency department. The Netherlands journal of medicine. 2015;73(7):331-40.
16 Nesher L, Tsaban G, Dreiher J, Rolston KVI, Ifergane G, Shemer Y, et al. The impact of incorporating early rapid influenza diagnosis on hospital occupancy and hospital acquired influenza. Infection control and hospital epidemiology. 2019;40(8):897-903.

Chapter 6

Optimizing the cut-off of procalcitonin to rule out bacteremia in patients with suspected viral infection.

accepted for publication JACEP Open

K. Azijli

T. C. Minderhoud

C. J. de Gans

A.W.E. Lieveld

P.W.B. Nanayakkara

Procalcitonin to exclude bacteremia

Optimizing cut-off by using pre-test probability

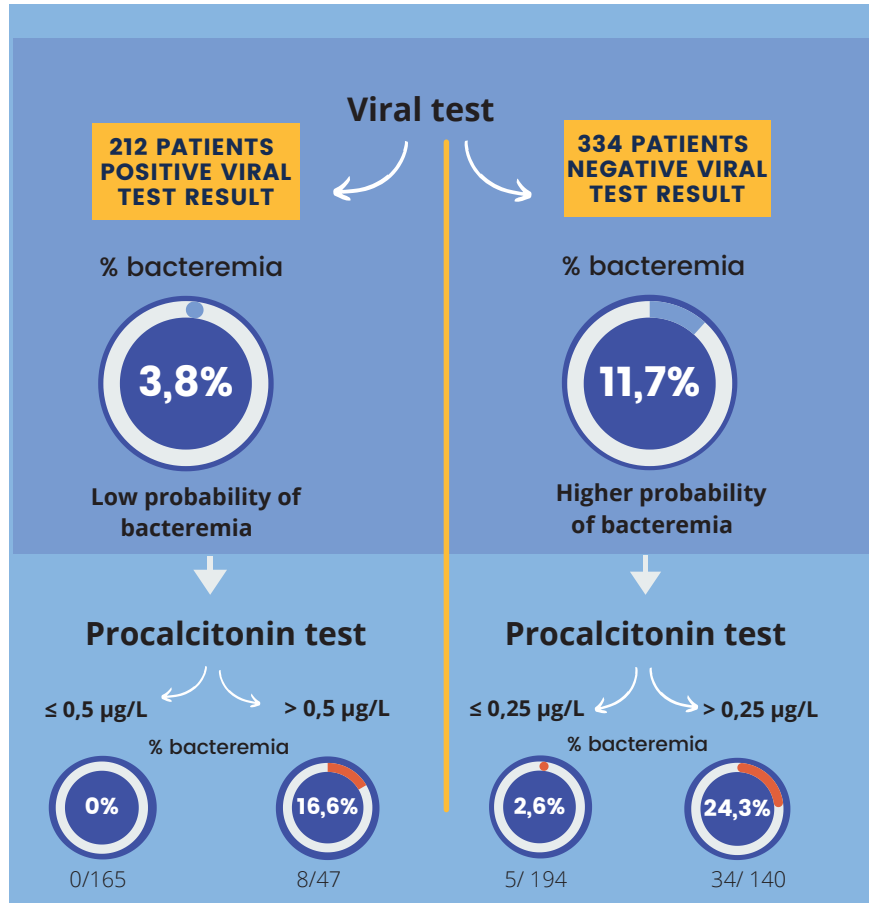
Many patients in the Emergency Department (ED) present with suspected infection. It is hard to determine which patients has a bacterial, viral infection or co-infection, especially during epidemics of viral infections. It is essential to start the right treatment. We studied the optimal use of procalcitonin to exclude bacteremia in retrospect.

Can we optimize the PCT cut-off based on the result of viral testing?

546 ED patients with suspected infection

All underwent
- **rapid viral testing**
- **blood cultures**
- **procalcitonin (PCT) test**

Rate of bacteremia: 8,6% (47/546)
Rate of positive viral test: 39,3 % (212/546)



Conclusion: In patients with suspected infection, viral testing can be used to estimate the probability of bacteremia. In patients with a low probabilty of bacteremia, a pct of ≤ 0.5 µg/L ruled out bacteremia, with a sensitivity of 100% ((63.1-100) and a LR of zero. In patients with a high probability, pct of ≤ 0.25 µg/L can be used, with a negative LR of 0.20.

ABSTRACT

Objective

During the winter, many patients present with suspected infection, that could be a viral or a bacterial (co-)infection. The aim of this study is to investigate whether the optimal use of procalcitonin is different in patients with and without proven viral infections for the purpose of excluding bacteremia. We hypothesize that when a viral infection is confirmed, this lowers the probability of bacteremia, and therefore influences the appropriate cut-off of procalcitonin.

Methods

This observational study was conducted in the Emergency Department of an academic medical center, in The Netherlands, in the winter seasons of 2019 and 2020. Adults (>18 years) with suspected infection and in whom a blood culture, procalcitonin and a rapid viral PCR test was performed were included.

Results

A total of 546 patients were included of which 47 (8,6%) had a positive blood culture. Procalcitonin had an area under the curve (AUROC) of 0.85 (95% CI 0.80-0.91) for prediction of bacteremia. In patients with a proven viral infection (N=212), f PCT <0.5µg/L had a positive likelihood ratio of 5.21 (95% CI 3.93 -6.90) and a negative likelihood ratio of 0.0 to exclude bacteremia. In patients without a viral infection, the procalcitonin cut-off point of <0.25 µg/L showed a LR(+) of 2.43 (95% CI 2.00-2.96) and LR (-) of 0.20 (95% CI 0.09 -0.45).

Conclusion

In patients with a viral infection, a procalcitonin of < 0.50 µg/L ruled out bacteremia. However in patients without a viral cause of the complaints, the most appropriate cut-off was <0.25 µg/L.

Background

Annually more than 20% of adult emergency department (ED) visits occur due to severe infections.[1] The most frequent presenting symptoms are fever and respiratory complaints. It is difficult to distinguish between a viral and bacterial cause of these complaints based on clinical symptoms since the complaints in viral and bacterial disease show great overlap[2,3]. During the winter season in the Netherlands, this clinical dilemma is encountered more often as the incidence of viral infections rises. This is usually caused by the annual influenza epidemic, however in 2020, this epidemic was curtailed by the COVID-19 pandemic. Rapid diagnosis of viral infections has become easier due to the availability of accurate and rapid PCR test.[4] However, despite a positive viral test, clinicians often prescribe antibiotics[5,6], since viral infections predispose patients to bacterial (co-)infection especially in the elderly and mortality in those cases are higher.[7-9]
The rate of bacterial co-infection in viral infections is highly variable. In influenza bacterial co-infection rates varying from 2-65% have been reported. 7 A recent meta-analysis re-0ported a rate of bacterial co-infection in patients with SARS-COV-2 at presentation of 3.5 %.[10] The rate of bacterial co-infections with other viruses such as rhinovirus or RSV have been mostly reported from intensive care units, which is hard to extrapolate to emergency department patients.[11] Estimated rates of unnecessary antibiotic use at the emergency department are between 30-60%, and it has been described as the most preventable cause of antibiotic resistance.[12-14]. A recent report by the World Health Organization (WHO) found that antibiotic resistance could lead to a significant increase in economic costs, and 10 million annual deaths globally by 2050 without a sustained effort to contain it[15].

To reduce antibiotic use and identify bacterial (co-)infection more accurately, biological markers such as procalcitonin (PCT) have been used but with conflicting results.[16,17]
One of the largest studies in the Cochrane meta-analysis of procalcitonin showed high negative predictive value (NPV) of 91.9% for exclusion of a bacterial co-infection in patients admitted to the intensive care unit (ICU) with influenza, despite a high prevalence of co-infection.[18]
Procalcitonin has been used with several cut-offs (< 0,10 µg/L, < 0,25 µg/L, < 0,5 µg/L). In a review of procalcitonin algorithms, it was advised to take the pre-test likelihood of bacterial infection into account, to choose the appropriate cut-off.[19-21].
The aim of this study is to investigate if the procalcitonin cut-off to exclude bacteremia should be different in patients with and without confirmed viral infections, presenting during a viral epidemic or pandemic. We hypothesize that if a viral infection is found, this lowers the probability of bacteremia, and therefore influences the appropriate cut-off of procalcitonin.

METHODS

Setting

This was a prospective study performed on the Emergency Department of Amsterdam University Medical Centre (location VUmc) in The Netherlands, during two separate winter seasons. The first inclusion period was from January 2019 to April 2019, the second inclusion period was from January 2020 until April 2020, coinciding with the start of the COVID pandemic.
The study was approved by the local medical ethics committee, a waiver for informed consent was obtained.

Participants

All patients 18 years and older in whom a blood culture and a viral test were ordered in the ED were included. This was a consecutive sample of patients.

STUDY PROTOCOL

In our hospital, testing for influenza, was based on the case definition of the National institute for Public Health and Environment (RIVM) which includes fever and respiratory symptoms[22]. Testing for the novel SARS-COV-2 was based on the clinical case definition of the National institute for Public Health and Environment (RIVM) in march 2020. A blood culture was drawn when a bacterial (co-)infection was suspected by the treating physician. In all included patients PCT was determined in the blood sample that was drawn for other biochemical tests. No additional sample was needed for this study. Procalcitonin levels were not actively communicated to the treating physicians. Patients were treated according to standard care. Patients were excluded if a blood culture, procalcitonin or viral test was not available.

Test methods

Bacteremia was defined as true positive blood cultures. All blood cultures are processed with the BACTEC system (Becton Dickinson). Contaminated cultures were assessed according to pre-established criteria [National Nosocomial Infection Surveillance (NNIS) parameters and surveillance criteria for blood stream infection] and considered negative[23]. Procalcitonin was measured using the Elecsys BRAHMS PCT assay. Viral testing was performed by Cepheid Xpert Flu A/B/RSV XC assay in 2019, and by the Cepheid Xpert SARS-CoV-2/Flu/RSV assay in 2020. If more extensive viral testing was deemed indicated by the treating physician, the result was used in our analysis. All patients with a positive viral test were considered to have a confirmed viral infection.

Data management

Data of patients were gathered by using chart review in EPIC and were entered in a clinical data management platform, Castor EDC[24], in compliance with Good Clinical Practice regulations.

Outcome

The primary outcome was the diagnostic accuracy of PCT to exclude bacteremia at the pre-defined cut-offs of < 0,10 µg/L, < 0,25 µg/L, < 0,5 µg/L, overall, and in subgroups based on results of viral testing. The diagnostic accuracy of PCT was defined as the sensitivity, specificity, positive/negative predictive value, Likelihood ratio (LR) and the area-under-the-curve (AUC) in excluding bacteremia.
Secondary outcomes were the difference between incidence of bacteremia in patients with and without viral infection, and the use of antibiotics.
For exploratory purposes, the time to onset of complaints, and general patient characteristics of patients with a low procalcitonin (< 0,5 µg/L) and a true positive blood culture were reported.

Statistical analysis

All data analyses were performed in SPSS version 26. Normally distributed continuous variables were expressed by their mean and standard deviation (SD). Continuous variables that were not normally distributed were expressed by their median and interquartile range (IQR). Comparison of continuous values, not normally distributed was done with a non-parametric test (Mann Whitney). Analysis of proportions was done with chi-square test (dichotomous variables) . Receiver operating characteristic (ROC) curves were drawn to assess the overall diagnostic accuracy of PCT by calculating the area under the curve (AUC) with confidence intervals (CI). Sensitivity, specificity, negative predictive value and positive predictive value and the LR were calculated using MedCalc version 19.4 (MedCalc Software, Ostend, Belgium).

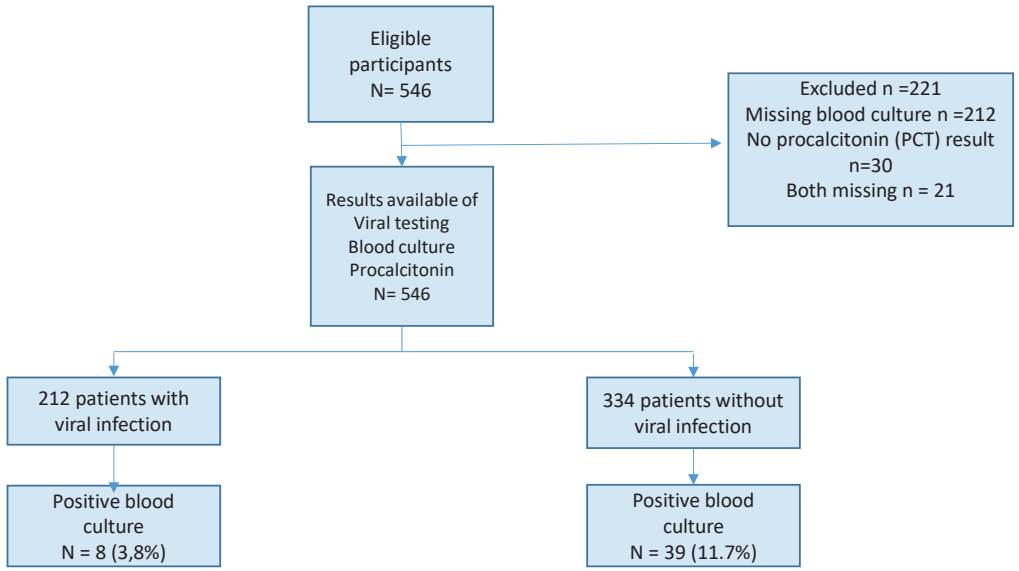
Sensitivity analysis

Due to the pandemic, SARS-COV-2 infection heavily influenced the group of patients with a viral infection, since 109 of the 210 patients with viral infection had COVID-19. Therefore, we performed an additional analysis of the AUROC of procalcitonin in the patients with a viral infection, but excluding the patients with COVID-19.

Reporting

Reporting was done in concordance with the guideline for Standards for Reporting diagnostic Accuracy (STARD) 2015.[25]

Figure 6.1 Diagram of flow of participants through the study



RESULTS

Population

During the study period a total of 767 patients presented with suspected infection at the emergency department. A total of 221 patients were excluded because of missing data; they lacked either blood culture (n=212), PCT (n=30) or both values. Figure 6.1 displays the diagram of flow of participants. The final study population included 546 patients. Patient demographics are described in table 6.1. Forty seven patients (8.6%) had a true positive blood culture. Fourteen blood cultures were classified as contaminated. Details of the contaminated cultures are described in supplementary table S6.1. Of the 546 patients, 212 patients (38.8%) had a viral infection. Infection with sars-cov-2 was the most frequently diagnosed viral infection (N= 109), followed by influenza A/B (N=62). More details of viral test results can be found in supplementary table S6.2.

Overall diagnostic value of procalcitonin (PCT) for bacteremia

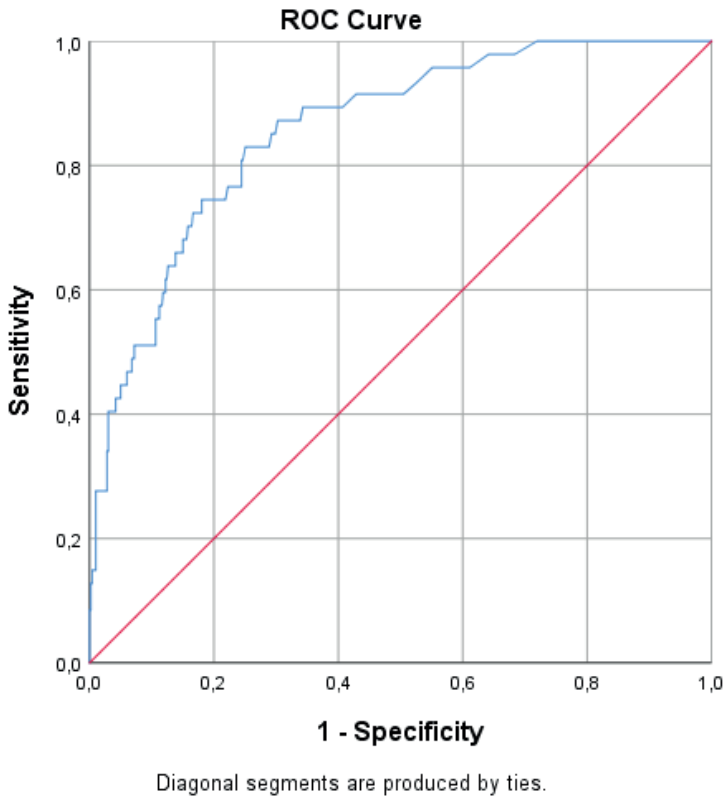
The median PCT in the total study population (n=546) was 0.15 µg/L (IQR 0.06-0.57). The area under the curve (AUROC) was 0.86 (95% CI 0.81-0.91) and is displayed in figure 6.2. The sensitivity, specificity, PPV and NPV, and likelihood ratios, at the pre-specified cut-offs are presented in table 6.2.

Table 6.1 Demographics

General characteristics	Study population n = 546
Gender	
Female, N (%)	245 (44.9)
Age, mean (SD)	64.0 (17.0)
Admission	
Hospital admission, n (%)	363 (66.5)
ICU admission, n (%)	45 (8.2)
Readmission (30-days), n (%)	47 (8.7)
Comorbidities	
Respiratory disease, n (%)	89 (16.4)
Diabetes mellitus, n (%)	94 (17.3)
Vital parameters	
Respiratory rate, mean (SD)	21.7 (7.2)
Oxygen saturation levels, mean(SD)	95.3 (3.9)
Temperature °C, mean (SD)	37.4 (1.3)
MEWS score (median with IQR)	2.0 (2-4)
Laboratory tests (median with IQR)	
CRP (mg/l)	55.0 (17.0-115)
PCT (ng/ml)	0.15 (0.06-0.57)
Positive SARS-CoV-2 PCR (%)	106 (19.5)
Positive blood culture (%)	47 (8.6)
30-days mortality n (%)	63 (11.5)

IQR: interquartile range, SD standard deviation, MEWS: modified Early Warning Score, CRP: c-reactive protein, PCT: procalcitonin.

Figure 6.2 ROC curve of procalcitonin for the outcome of bacteremia, total group



Area Under the Curve *0.857* *(95% CI 0.805 - 0.910)*

Table 6.2 Diagnostic accuracy PCT in predicting positive blood culture in total study population

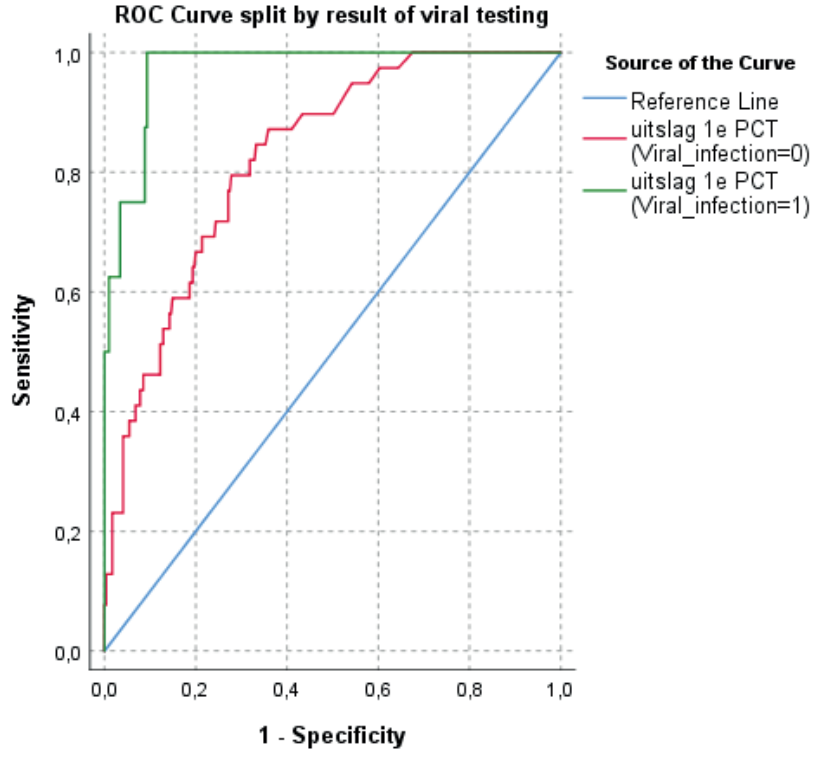
Groups	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR + (95% CI)	LR- (95%CI)
PCT < 0.10	**95.7 (85.5-99.5)**	**38.9 (34.6-43.4)**	**12.9 (15.9-18.6)**	**99.0 (96.1-99.7)**	**1.57 (1.43-1.72)**	**0.11 (0.03-0.43)**
PCT < 0.25	89.4 (76.9-96.5)	65.7 (61.3-69.8)	25.7 (22.8-28.8)	97.90 (95.3-99.0)	2.61 (2.23-3.05)	0.16 (0.07-0.37)
PCT < 0.50	76.6 (62.0-87.7)	77.6 (73.9-81.3)	24.5 (20.5-28.9)	97.2 (95.5-98.3)	3.44 (2.74-4.32)	0.30 (0.18-0.51)

All values are percentages, with the 95% confidence interval between brackets.
**PPV= positive predictive value, NPV= negative predictive value, LR + = positive likelihood ratio, LR-= negative likelihood ratio.*

Probability of bacteremia in patients with and without viral infection
Eight (3.8%) of the 212 patients with a confirmed viral infection had bacteremia. In the patients without a viral infection, 39 of the 334 (11.7%) had bacteremia. This was a significant difference, p-value 0.001 (chi-square test)

Diagnostic accuracy of procalcitonin in patients with a proven viral infection
In patients with a viral infection (N=212), median PCT was 0.14 µg/L (IQR 0.07-0.34). The AUROC of PCT for prediction of bacteremia was 0.97 (95% CI 0.94-1.00), displayed in figure 6.3. All eight patients with a viral infection and bacteremia had a procalcitonin value of ≥ 0.5 µg/L. The sensitivity, specificity, PPV and NPV, and likelihood ratios, at the pre-specified cut-offs are presented in table 6.3. In the group with a viral infection, the negative likelihood ratio was zero (95% CI N/A), with a positive LR of 5.21 (95% CI 3.93-6.90).

Figure 6.3 ROC curve of procalcitonin for the outcome of bacteremia, split by viral infection



Area under ROC curve
No viral infection (red) *0,826 (96% CI 0. 764- 0.887)*
Viral infection (green) *0. 972 (96% CI 0.942 - 1.002)*

Table 6.3 Diagnostic accuracy of PCT in predicting positive blood culture in study population with positive viral test (low-pre-test probability)

Groups	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR + (95% CI)	LR- (95%CI)
PCT < 0.10	100 (63.1-100)	34.0 (27.5-41.0)	5.65 (5.14-6.20)	100	1.51 (1.37-1.67)	0.00
PCT < 0.25	100 (63.1-100)	68.0 (61.1-74.4)	11.0 (9.17-13.1)	100	3.12 (2.56-3.82)	0.00
PCT < 0.50	100 (63.1-100)	81.2 (75.1-86.3)	17.4 (13.7-21.9)	100	5.21 (3.93 -6.90)	0.00

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

Diagnostic accuracy in patients without viral infection.

In patients without confirmed viral infection (n = 334), the median PCT was 0.13 (IQR 0.05-0.48). Of the 39 patients in this group with a positive blood culture, 28 had a procalcitonin ≥ 0.5 µg/L. Of the other 11 patients with positive bloodcultures , 2 had a procalcitonin < 0.10 µg/L. Three had a procalcitonin ≥ 0.10 µg/L but <0.25 µg/L, and 6 had a procalcitonin ≥ 0.25 µg/L but <0.5 µg/L. The AUROC of PCT for the outcome of bacteremia was 0.83 (95% CI 0.76-0.88). The sensitivity, specificity, PPV and NPV, and likelihood ratios are presented in table 6.4.

Table 6.4 Diagnostic accuracy of PCT in predicting positive blood culture in study population with negative viral test (high pre-test probability).

Groups	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR + (95% CI)	LR- (95%CI)
PCT < 0.10	94.9 (82.7-99.4)	42.0 (36.4-48.0)	17.8 (16.1-19.6)	98.4 (94.1.-99.6)	1.64 (1.45-1.85)	0.12 (0.03-0.47)
PCT < 0.25	87.2 (72.6-95.7)	64.1 (58.3-69.6)	24.3 (21.0-28.3)	97.4 (94.3-98.9)	2.43 (2.00-2.96)	0.20 (0.09 -0.45)
PCT < 0.50	71.8 (55.1-85.0)	75.6 (70.3-80.4)	28.0 (22.7-34.0)	95.3 (92.4-97.1)	2.94 (2.22-3.90)	0.37 (0.23-0.62)

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

Patients with low procalcitonin and positive blood cultures

Eleven patients had a procalcitonin < 0,5 µg/L but a true positive blood culture. None had a proven viral infection. The patient characteristics are summarized in S3. None of these patients died within 30 days of this episode. Five patients had bacteremia and a procalcitonin < 0.25 µg/L. The time between onset of complaints and the first laboratory measurement of procalcitonin (Δ time) was less than 24 hours in all of these patients.

Use of antibiotics

Of the total of 543, 288 patients (53.0%) received antibiotics at the emergency department, including 37 patients with a positive blood culture. In patients with a confirmed viral infection, 107 of the 210 (50.9%) patients were treated with empirical antibiotics in the emergency department. Of these 107 patients, 74 (69,1%) had a PCT < 0.5 µg/L. Of the patients without a viral infection (n=333), 180 (54,0%) patients were treated with empirical antibiotics in the Emergency Department. Of these patients, 71 (39.4%) had a PCT < 0.25 µg/L.

The most frequent administered antibiotic class were cefalosporines intravenous (n= 154, 51,8%). The median duration of given antibiotics was 6 (IQR 5-7) days.

Sensitivity analysis excluding SARS-Cov-2 patients

In 103 patients with a viral infection, but not sars-cov-2 the AUROC of PCT for prediction of bacteremia was 0.96 (95% CI 0.90-1.00). The difference was not statistically significant, when compared to the AUROC of the total population of patients with a viral infection of 0.96 (95% CI 0.93-0.995).

DISCUSSION

In this study, we found that in patients with a confirmed viral infection, bacteremia can be safely ruled out if the procalcitonin is < 0.5 µg/L, with a 100% sensitivity and a negative LR of zero. However, in the patients without a viral infection, the cut-off of < 0.5 µg/L resulted in a low sensitivity of 71.8 % (95% CI 51.9-81.9) and only a moderate negative LR 0.37 (95% CI 0.23-0.62). In this group, the cut-off value of PCT < 0.25 had better results, with a fair sensitivity of 85.4 % (95% CI 55.1-85.0) and a negative LR of 0.20 (95% CI 0.09 -0.45).

In an earlier epidemiological study at our institution, the rate of bacteremia in the ED was 11.2%.26 In our study population, the patients with confirmedviral infection had a significantly lower rate of bacteremia of 3.8%. The patients without viral infection however, had a comparable rate of bacteremia of 11.7%. In line with Bayes theorem we used this knowledge of prior probability to optimize the use of procalcitonin, illustrated in the infographic (page 112,113).

If procalcitonin, dependent on probability of bacteremia, had been used as a guide in this population to initiate antibiotics, antibiotic use could have been reduced considerably. In the group with a viral infection, antibiotic use could have been decreased from 50.9% to 21.9%. However in the group without viral infection, a cut-off of < 0.25 µg/L would have resulted in a much smaller reduction of 54% to 42%.

What this study adds to the current knowledge is the results of (rapid) viral testing in the ED, can be used to estimate the prior probability of bacteremia. And, if a viral infection is confirmed, the patient thus has a low prior probability of bacteremia, a procalcitonin of < 0,5 µg/L can be used to exclude bacteremia. If patients do not have a viral infection, the cut-off of < 0,25 µg/L is more appropriate. However, in patients with a complaint of less than 24 hours duration, a procalcitonin < 0,25 µg/L does not rule out bacteremia. Five patients with positive blood cultures would have been missed if we had used PCT < 0.25 µg/L as a guide to initiate antibiotics. All of this patients presented within 24 hours after the onset of symptoms. Earlier studies have shown that procalcitonin can rise within

2-4 hours, but it can take 8-24 hours to reach high values.[19, 20] Therefore, when the patients present early after onset of symptoms the procalcitonin measurements should be repeated to ensure it is representative, in line with previous recommendations.[19]

LIMITATIONS

An important limitation of this study is that this study focuses on bacteremia. Blood cultures are the gold standard for detecting bacteria in blood. However, a negative blood culture is no guarantee that there is no bacterial infection. Patients with pneumonia have low rates of bacteremia.[27] In addition, patients previously treated with antibiotics before presentation at the emergency department might have lower yield of blood culture.[26] However, positive blood cultures are a clinically relevant outcome that most clinicians would act upon instantly. Bacteremia is related to adverse outcomes and death.[28] In this cohort, 30-day mortality was clearly higher in patients with positive blood cultures (16.3 %) than in patients without bacteremia (11.1%). Results of sputum or urine cultures are less time critical. In addition, it is easier to correlate blood cultures with clinical illness than sputum or urine cultures. Sputum cultures and urine cultures are harder to interpret, as they may represent colonization or asymptomatic bacteriuria, rather than an infection [29-31].
In our cohort, very few sputum samples were retrieved (N=25). Therefore, we did not include this in our outcome.
Another limitation is that it was done in a single-center in the Netherlands, with a specific case-mix. First of all, due to a strong primary care system, almost all patients in the Netherlands are seen first by a GP, who do not refer uncomplicated cases of (viral) infection. Because of this selection, patients presenting to our ED probably represent a subgroup of patients with more severe illness. This is reflected in the high number of admissions in our cohort (68%). Secondly, the population of our tertiary ED has a higher proportion of patients with malignancies and immunosuppressive treatments. This might influence the probability of bacteremia, and limit external validity.
Although all patients underwent influenza A/B/ RSV, and in 2020 sars-cov-2 testing, we did not limit our analysis to patients with influenza, RSV and SARS-Cov-2. We pooled all viral infections that were found, because we expected any viral explanation for infectious complaints would make the chance of also having bacteremia smaller. In SARS-Cov-2 infection it has become clear that the incidence of bacterial (co)infection is very low.
The sensitivity analysis, excluding the SARS-cov-2 patients, did not meaningfully change our results. However the results should be confirmed in a larger sample of patients, with particular attention for the incidence of bacteremia in patients with influenza.

Summary

In summary, procalcitonin has a good overall diagnostic accuracy for ruling out bacteremia. Rapid viral tests can be used to choose the most appropriate cut-off of procalcitonin in the Emergency Department, because patients with a virus infection have a low prior probability of bacteremia. In patients with a proven viral infection a procalcitonin of < 0.5 µg/L ruled out bacteremia with a negative LR of zero. In patients without a viral infection, the lower cut-off of <0,25 µg/L was more appropriate.

REFERENCES

1 Wang HE, Jones AR, Donnelly JP. Revised National Estimates of Emergency Department Visits for Sepsis in the United States. Critical care medicine. 2017;45(9):1443-1449.
2 Angeles Marcos M, Camps M, Pumarola T, et al. The role of viruses in the aetiology of community-acquired pneumonia in adults. Antiviral therapy. 2006;11(3):351-359.
3 Huijskens EGW, Koopmans M, Palmen FMH, van Erkel AJM, Mulder PGH, Rossen JWA. The value of signs and symptoms in differentiating between bacterial, viral and mixed aetiology in patients with community-acquired pneumonia. Journal of medical microbiology. 2014;63(Pt 3):441-452.
4 Vos LM, Bruning AHL, Reitsma JB, et al. Rapid Molecular Tests for Influenza, Respiratory Syncytial Virus, and Other Respiratory Viruses: A Systematic Review of Diagnostic Accuracy and Clinical Impact Studies. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2019;69(7):1243-1253.
5 Havers FP, Hicks LA, Chung JR, et al. Outpatient Antibiotic Prescribing for Acute Respiratory Infections During Influenza Seasons. JAMA Network Open. 2018;1(2):e180243-e180243.
6 Langford BJ, So M, Raybardhan S, et al. Antibiotic prescribing in patients with COVID-19: rapid review and meta-analysis. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2021;27(4):520-531.
7 Klein EY, Monteforte B, Gupta A, et al. The frequency of influenza and bacterial coinfection: a systematic review and meta-analysis. Influenza and other respiratory viruses. 2016;10(5):394-403.
8 Metersky ML, Masterton RG, Lode H, File TM, Jr., Babinchak T. Epidemiology, microbiology, and treatment considerations for bacterial pneumonia complicating influenza. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 2012; 16(5):e321-331.
9 McCullers JA. Insights into the interaction between influenza virus and pneumococcus. Clinical microbiology reviews. 2006;19(3):571-582.
10 Langford BJ, So M, Raybardhan S, et al. Bacterial co-infection and secondary infection in patients with COVID-19: a living rapid review and meta-analysis. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2020;26(12):1622-1629.
11 Cawcutt K, Kalil AC. Pneumonia with bacterial and viral coinfection. Current opinion in critical care. 2017;23(5):385-390.
12 Minderhoud TC, Spruyt C, Huisman S, Oskam E, Schuit SCE, Levin MD. Microbiological outcomes and antibiotic overuse in Emergency Department patients with suspected sepsis. The Netherlands journal of medicine. 2017;75(5):196-203.
13 May L, Cosgrove S, L'Archeveque M, et al. A call to action for antimicrobial stewardship in the emergency department: approaches and strategies. Annals of emergency medicine. 2013;62(1):69-77.e62.
14 Denny KJ, Gartside JG, Alcorn K, Cross JW, Maloney S, Keijzers G. Appropriateness of antibiotic prescribing in the Emergency Department. The Journal of antimicrobial chemotherapy. 2019;74(2): 515-520.
15 WHO. NO TIME TO WAIT: SECURING THE FUTURE FROM DRUG-RESISTANT INFECTIONS. 2019.
16 Kamat IS, Ramachandran V, Eswaran H, Guffey D, Musher DM. Procalcitonin to Distinguish Viral From Bacterial Pneumonia: A Systematic Review and Meta-analysis. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2020;70(3):538-542.
17 Schuetz P, Wirz Y, Sager R, et al. Procalcitonin to initiate or discontinue antibiotics in acute respiratory tract infections. Cochrane Database of Systematic Reviews. 2017(10).
18 Rodríguez AH, Avilés-Jurado FX, Díaz E, et al. Procalcitonin (PCT) levels for ruling-out bacterial co-infection in ICU patients with influenza: A CHAID decision-tree analysis. Journal of Infection. 2016; 72(2):143-151.
19 Schuetz P, Albrich W, Mueller B. Procalcitonin for diagnosis of infection and guide to antibiotic decisions: past, present and future. BMC medicine. 2011;9:107.
20 Schuetz P, Falsey AR. Procalcitonin in patients with fever: one approach does not fit all. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2018;24(12):1229-1230.
21 Schuetz P, Chiappa V, Briel M, Greenwald JL. Procalcitonin Algorithms for Antibiotic Therapy Decisions: A Systematic Review of Randomized Controlled Trials and Recommendations for Clinical Algorithms. Archives of internal medicine. 2011;171(15):1322-1331.
22 dr. A. Meijer RCI. Influenza Richtlijn https://lcirivmnl/richtlijnen/influenza. 2012.
23 Hall KK, Lyman JA. Updated review of blood culture contamination. Clinical microbiology reviews. 2006;19(4):788-802.
24 Castor EDC. Castor Electronic Data Capture. https://castoredc.com. Published 2019. Accessed August 28, 2019.

25 Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ (Clinical research ed). 2015;351:h5527.
26 Nannan Panday RS, Wang S, van de Ven PM, Hekker TAM, Alam N, Nanayakkara PWB. Evaluation of blood culture epidemiology and efficiency in a large European teaching hospital. PloS one. 2019;14(3): e0214052-e0214052.
27 Kennedy M, Bates DW, Wright SB, Ruiz R, Wolfe RE, Shapiro NI. Do Emergency Department Blood Cultures Change Practice in Patients With Pneumonia? Annals of emergency medicine. 2005;46(5): 393-400.
28 Nielsen SL. The incidence and prognosis of patients with bacteremia. Danish medical journal. 2015; 62(7).
29 Gunasekaran J, Saksena R, Jain M, Gaind R. Can sputum gram stain be used to predict lower respiratory tract infection and guide empiric antimicrobial treatment: Experience from a tertiary care hospital. Journal of microbiological methods. 2019;166:105731.
30 Lentino JR, Lucks DA. Nonvalue of sputum culture in the management of lower respiratory tract infections. Journal of clinical microbiology. 1987;25(5):758-762.
31 Werman HA, Brown CG. Utility of urine cultures in the emergency department. Annals of emergency medicine. 1986;15(3):302-307.

Part 2 / Chapter 6
Supplementary materials

Table S6.1 Details of contaminated cultures

Bacterial pathogen found	PCT values (ng/L)
Rothia mucilaginosa (n=1)	0.62
Stahylococcus capitis (n= 2)	0.02, 0.07
Staphylococcus pettenkoferi (n= 2)	2.11, 0.06
Staphylococcus hominis (n = 5)	1.89, 0.11, 2.15, 0.24, 0.17
Acinetobacter johnsonii and Chryseobacterium hominis (n=1)	0.04
Staphylococcus epidermidis and Kocuria species (n=1)	6.22
Streptococcus oralis (n=1)	0.08

Table S6.2 Results of virology diagnostics

Viral diagnosis	Number of patients
Influenza A	56
Influenza B*	4
SARS-cov-2	54
Respiratory Syncytial Virus (RSV)	10
Human Metapneumovirus *	10
Coronavirus (not sars-cov-2)	6
Rhinovirus	4
Parainfluenza	3
Other (boca,virus, adenovirus, EBV)	4

**In one patient both influenza B and Human Metapneumovirus was detected*

Table S6.3 Characteristics of patients with a positive blood culture & PCT ≤0,5 ng/L

BC result	Age	Time onset to first measurement	PCT ng/L	WBC (10^9)	CRP (mg/l)	MEWS	Blood pressure mmHg	Final diagnosis
E.coli	82	6-12 hrs	0.07	13.7	9	2	100/60	Urosepsis
E.coli	93	6-12 hrs + 10 hrs*	0.02 26.7 *	2.7 15.7*	< 2.5 22 *	3	149/73	E. coli bacteremia
S. aureus	51	unclear	.32	12.0	273	4	142/82	S. aureus bacteremia due to osteomyelitis

Chapter 7

Predicting blood culture outcomes in the emergency department using machine learning.

accepted for publication BMJ Open

A.W. Boerman

M. Schinkel

L. Meijerink

E.S. van den Ende

L.C.A. Pladet

M.G. Scholtemeijer

J. Zeeuw

A.Y. van der Zaag

T.C. Minderhoud

P. Elbers

W.J. Wiersinga

R. de Jong

M.H.H. Kramer

P.W.B. Nanayakkara

ABSTRACT

Background

Blood cultures (BCs) are the gold standard for detecting bacteraemia, a condition with high morbidity and mortality. The threshold for collecting BCs is low, resulting in a low yield. Machine learning models may help identify patients with low bacteraemia risk and reduce the number of unnecessary BCs and also antibiotic use at the emergency department (ED).

Objective

To develop machine learning models to predict BC outcomes in an ED setting.

Design

Retrospective observational study.

Setting

ED of a large teaching hospital in the Netherlands between September 1st 2018 and June 24th 2020.

Patients

Adult patients from whom BCs were collected in the ED.

Measurements

The primary outcome was the performance of two machine learning models to predict bacteraemia in ED patients, defined as at least one true positive BC collected at the ED. Data in the electronic health record, available at the end of the ED visits were used.

Results

In 4885 out of 51399 ED visits (9.5%), BCs were collected. In 598/4885 (12.2%) visits, at least one of the BCs was true positive. Both the gradient boosted tree model and logistic regression model showed good performance in predicting BC results with AUROCs of 0.77 (95%-CI = 0.73-0.82) in the test sets.

Limitations

Not all data was uniformly registered in the electronic health records and only a proportion of all data was used.

Conclusion

Both machine learning models can accurately identify patients with low risk of bacteraemia at the ED and may be useful to reduce unnecessary BCs, antibiotic use and associated healthcare costs. Implementation studies are necessary to investigate the potential clinical benefits.

INTRODUCTION

Over 20% of adult emergency department (ED) visits occur due to serious infections[1]. Current diagnostic modalities cannot sufficiently distinguish between bacterial and non-bacterial disease during an early stage of the diagnostic workup[2]. However, timely distinction between bacterial and non-bacterial disease can reduce unnecessary diagnostic tests and treatment with antibiotics. In case of a bacteraemia (bloodstream infection), blood cultures (BCs) are the gold standard test. Unfortunately, turnaround times of BC results of 24-72 hours make these cultures unhelpful for timely diagnosis of bacterial infections at the ED. Accurate and early identification of patients with a high or low risk of bacteraemia may be a first step to help to distinguish bacterial from non-bacterial disease early.

Bacteraemia is associated with high morbidity and mortality, which makes missing a possible bacteraemia very harmful[3]. Therefore, physicians order BCs frequently and the overall BC yields are low[2]. Around 11-15% of collected BCs are positive and studies show that up to half of those are false positives through contamination[4–7]. These contaminated BCs can lead to unnecessary downstream diagnostics, antibiotic overuse, and increased hospital length of stay[8,9]. Currently, we are unable to recognize patients with low risk of bacteraemia, in which we could safely withhold BC testing and even antibiotics.

Machine learning already has significant impact on healthcare. Machine learning models can use many data points from large numbers of patients to detect subtle patterns that may go unnoticed by health care professionals. These insights may support the swift assessment of a patient and selection of the appropriate diagnostic and treatment strategies. Complex situations, where multiple physiological mechanisms interact are perfect areas to investigate machine learning decision support[10]. The diagnostic workup of suspected bacterial infections is such an area.

In this paper, we aim to create machine learning models to predict BC outcomes in an ED setting which may help reduce unnecessary BCs and provide physicians with an additional tool to help decide whether or not antibiotic treatment is needed. A secondary aim of this study is to create models that can straightforwardly be implemented at the ED.

METHODS

Study setting

We performed a retrospective observational study on data from the electronic health records (EHR) of Amsterdam UMC, location VU University Medical Center, between September 1st 2018 and June 24th 2020. The VU University Medical Center is a large teaching hospital with an estimated 28.000 ED presentations annually. The study adhered to the "transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD)"[11]. The study protocol was assessed by the local Medical Ethics Review Committee (IRB number: IRB00002991; case: 2020.486). The need for informed consent was waived.

Population
We included all adult patients who presented to the ED and in whom at least one BC was taken during their ED stay because a bacterial infection was suspected on clinical grounds. We included patients of all medical specialties. Whenever a patient presented to the ED multiple times during the study period, each encounter was classified as a unique visit.

Data collection
All data that was available under local privacy regulations was extracted from the EHR. The data included demographic information, vital signs, laboratory results, and information about imaging procedures and administered medications in the ED. Data on comorbidities or medication usage at home were not available. We only used data that would be available before the end of the ED visit, which is the time when the prediction can potentially have clinical consequences on the use of BCs and initiation of antibiotic therapy. The data extracted from the EHR was further preprocessed to be used for predictive modelling. Details about preprocessing are described in the supplementary files (e-Methods and e-Tables 1-4).

Outcome
We aimed to predict bacteraemia, which was defined as at least one positive BC with a pathogenic microorganism collected during the ED visit.

AB and MS mapped all microorganisms to be pathogens or contaminants based on previous literature under supervision of WJW [2,4,5,12]. e-Table 5 lists all organisms that we classified as contaminants. Then, we assigned the most important result to a specific BC set (prioritizing positive over contamination over negative). Afterwards, the combination of all BC sets in a unique ED visit was mapped to represent a visit with growth of a clinically significant pathogen in at least one BC set (positive) or a visit with only negative or contaminated cultures (negative).

Model development and feature selection
We used all variables that were reported in over 10% of the ED visits as features. We also created indicator features for all variables to indicate whether this variable was measured or not. The dataset was randomly split into a training (75%) and test (25%) set for model development. We used median imputation except for some situations where imputation based on domain knowledge was used (see e-Table 6 for details). Median imputation is a practical and adequate solution for handling missing data in non-linear models. Furthermore, the combination of median imputation and indicator features as we used is also adequate for linear models, especially with data missing not at random [13]. Additional standard scaling around the mean was applied. We trained the models on the training set using the full set of features, since the used models are robust to unimportant features.

We used a gradient boosted tree model and a logistic regression model with L1 regularization. These different model classes are known to be suitable for our type of data, which is limited in size and of mixed type. We used gradient boosted trees as a powerful representative of tree-based models, which can uncover complex feature interdependencies and nonlinearities. We also used a simpler logistic regression for comparison, since its coefficients are easier to interpret.

Within the training set, a fivefold cross-validated grid search was performed to find the hyperparameters that optimize the model's performances. An overview of the pipeline from raw data to model can be found in the e-Methods section of the appendix.

Modelling was performed using Python version 3.7.9 (Python software foundation, http://www.python.org) and the Scikit-learn package (version 23.1).

Model evaluation
The model performances were tested using the Area Under curve of the Receiver Operating Characteristics (AUROC), together with the Area Under the Precision Recall Curve (AUPRC) since we had imbalanced outcome classes. We also reported Brier scores and F1-scores during cross-validation as well as on the test set. The model calibration is presented in calibration plots.

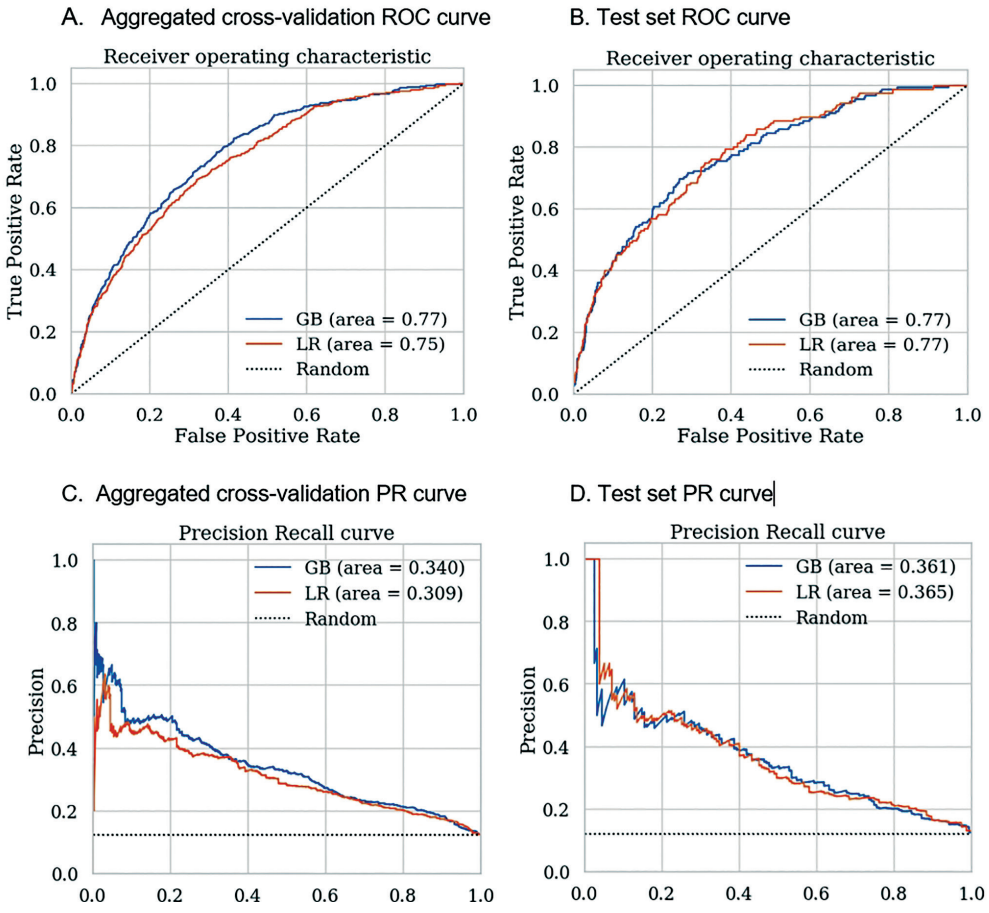
The model's output was the probability for the BC to be positive. To provide a clinically meaningful result, we report on two preselected probability thresholds that predict BCs to be positive above this threshold. Firstly, we show performances on the most optimal sensitivity-specificity threshold based on maximization of the sensitivity-specificity sum or minimization of the sensitivity-specificity difference [14]. These approaches are useful when omission errors (false negatives) should be avoided and provide a diagnostic test with the power to rule out a diagnosis [14,15]. Furthermore, we present model performances on a threshold that retains a sensitivity of 90%, which would lower the number of false negatives further at the expense of higher false positive rates.

Funding source
This project was funded by a research grant from the Dutch federation for acute internal medicine (NVIAG). The funder had no involvement in any part of the study.

RESULTS

Baseline characteristics
We identified 51.399 ED visits by 41.280 unique adult patients in the VU Medical Center between September 1st 2018 and June 24th 2020. One or more BC samples were taken in 4885 (9,5%) of those visits. In 598/4885 (12.2%) of those visits, at least one of the cultures was a true positive. In 254/4885 (5.2%) of the visits, at least one of the cultures was contaminated (later mapped to be negative). Overall, 4074/4885 (83.4%) visits had only truly negative cultures. Table 7.1 shows the baseline characteristics of the study population stratified by culture outcomes.

Figure 7.1 Receiver Operating Characteristic and Precision Recall curves for positive blood cultures in aggregated cross-validation sets and test set



GB = gradient boosted tree model; LR = Logistic regression model; Area = area under the curve; ROC = Receiver Operating Characteristics; PR = Precision Recall.

Predictive performance

The gradient boosted tree model's AUROC in the cross-validation (training) sets and internal test set were 0.77 (standard deviation = 0.03) and 0.77 (95%-CI = 0.73-0.82), respectively. The logistic regression model's AUROC in the cross-validation and internal test set were 0.75 (standard deviation = 0.02) and 0.77 (95%-CI = 0.73-0.82). The AUROCs of both models are shown in Figure 7.1. Table 7.2 shows the corresponding performance scores. The calibration plots are presented in e-Figure 1 of the appendix.

Table 7.1 Baseline characteristics of the study population stratified on blood culture outcomes

Characteristic	Negative cultures* (N=4287)	Positive cultures (N=598)	Total (N=4885)
Age, years			
Median (IQR)	66 (51-75)	70 (59-79)	66 (52-76)
Sex			
Male	56.3%	62.2%	57.0%
Modified Early Warning Score			
Median (IQR)	3 (2-4)	4 (2-5)	3 (2-4)
Missing (N)	2515	351	2866
Heart rate (beats per minute)			
Median (IQR)	94 (82-107)	100 (88-111)	95 (82-108)
Missing (N)	181	15	196
Systolic blood pressure (mmHg)			
Median (IQR)	91 (73-121)	86 (66-114)	90 (73-120)
Missing (N)	371	35	406
Respiratory rate (per minute)			
Median (IQR)	19 (15-23)	21 (16-25)	19 (15-24)
Missing (N)	1310	149	1459
Temperature (Celsius)			
Median (IQR)	37.8 (37.0-38.5)	38.1 (37.2-38.8)	37.8 (37.0-38.5)
Missing (N)	198	26	224
C-Reactive Protein (µmol/L)			
Median (IQR)	60 (25-134)	104 (39-216)	64 (25-144)
Missing (N)	132	23	155
White blood cell counts (10^9/L)			
Median (IQR)	10 (6.8-13.8)	11.9 (8.2-16.0)	10.2 (6.9-14.2)
Missing (N)	144	22	166
Thrombocyte counts (10^9/L)			
Median (IQR)	234 (174-311)	211 (149-273)	231 (171-307)
Missing (N)	593	105	698
Bilirubin (µmol/L)			
Median (IQR)	9 (6-13)	13 (8-22)	9 (6-14)
Missing (N)	1205	163	1368
Creatinine (µmol/L)			
Median (IQR)	82 (65-113)	105 (73-160)	84 (66-119)
Missing (N)	171	27	198
Length of ED stay (hours)			
Median (IQR)	4.3 (3.2-5.8)	4.7 (3.3-6.3)	4.4 (3.2-5.9)
Hospital admission			
Admitted	68.0%	84.6%	70.0%
30-day mortality			
Died	6.7%	11.5%	7.3%

IQR = Interquartile range; ED = Emergency department.
**Likely contaminants are classified as negative cultures in this table.*

Table 7.2 Performance metrics of both models in the aggregated cross-validation sets and the test set

Model	Modelling phase	AUROC	AUPRC	Brier score*	F1-score**
Gradient Boosted trees	Cross-validation mean	0.77 (SD = 0.03)	0.340	0.066	0.16
	Test	0.77 (95% CI: 0.73-0.82)	0.361	0.093	0.19
Logistic Regression	Cross-validation mean	0.75 (SD = 0.02)	0.309	0.098	0.14
	Test	0.77 (95% CI: 0.73-0.82)	0.365	0.092	0.15

SD = Standard Deviation; CI = Confidence Interval; AUROC = Area Under the curve of the Receiver Operating Characteristics; AUPRC = Area Under the Precision Recall Curve.
**The Brier score is a cost function that measures performance of probabilistic predictions. The score ranges from 0 to 1. The lower the score, the more accurate the prediction.*
***F1-scores present a balance between precision and recall. The lower the score, the more accurate the prediction.*

Feature importances

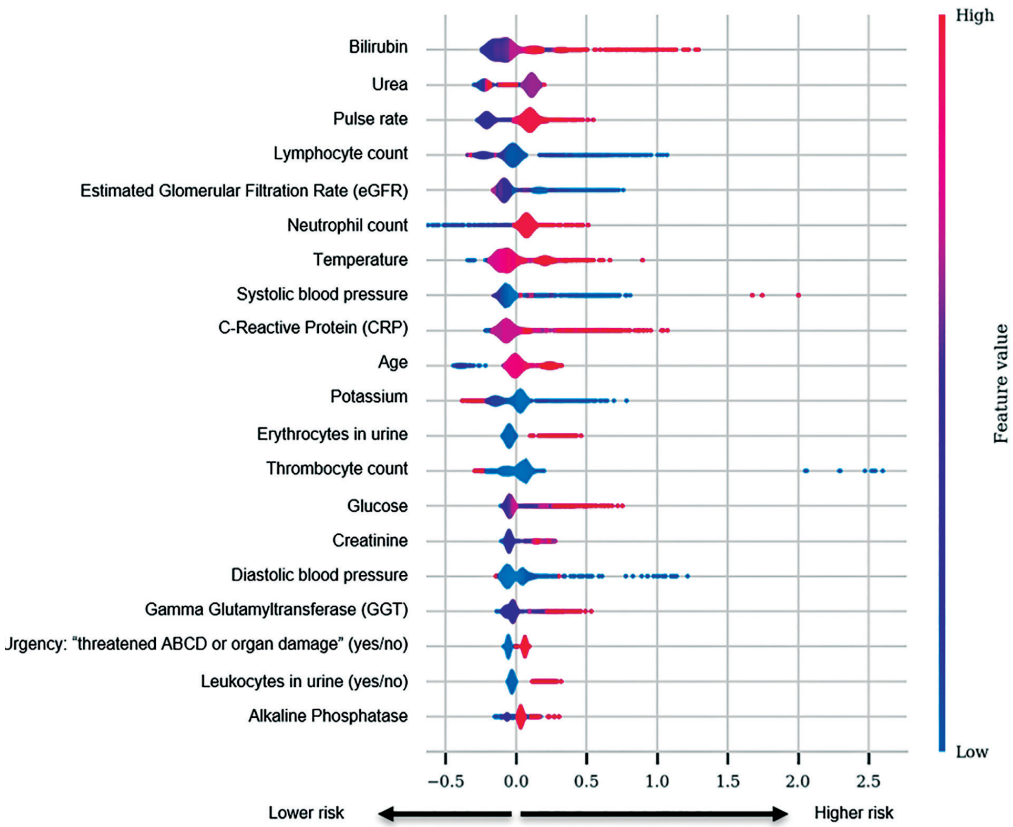
Gradient Boosted trees

Feature importances for non-linear tree based models only indicate the magnitude and not the directionality (positive/negative) of the effect. We present the feature contributions using Shapley additive explanation (SHAP) values, as depicted in Figure 7.2 [16]. These are distributions of local contributions per feature and per data point. Figure 7.2 shows the 20 most important features that drive predictions in the gradient boosted tree model (see e-Table 2-4 for the full lists of features). This model recognizes bilirubin values to be the strongest predictor of a positive BC. We see that high (red) bilirubin values are associated with a higher risk of a positive BC (right on the x-axis). Conversely, high (red) potassium levels are associated with a lower risk of a positive BC (left on the x-axis).

Logistic regression

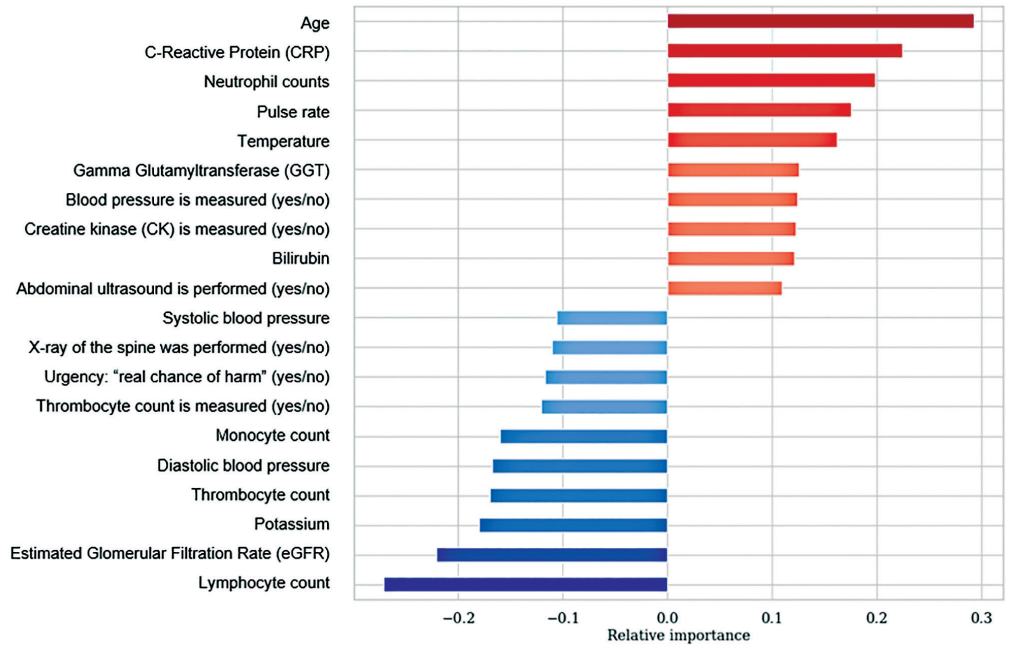
The 20 features with the largest absolute coefficients in the logistic regression model are presented in Figure 7.3. Age and lymphocyte counts are the strongest predictors. A high age is associated with a higher risk of a positive BC, whereas a high lymphocyte count is associated with a lower risk (see e-Table 7 for a full list of coefficients). Due to the imputation and the fact that physiological parameters are not strictly independent of each other, no valid estimation of the odds ratios can be provided.

Figure 7.2 SHAP-plot of feature importance in the gradient boosted tree model



The Shapley additive explanation (SHAP) values, present distributions of local contributions per data point per feature. The figure shows the 20 most important features in the gradient boosted tree model, in order of importance on the Y-axis. The relative effect of the feature on the risk of a positive blood culture is shown on the X-axis (right of 0.0 = increased risk, left of 0.0 = lower risk). The colours represent the actual values of the features themself. Blue depicts a low actual value of the feature while red depicts a high actual value. With yes/no features, no is depicted as a low value (blue) and yes as a high value (red).

Figure 7.3 Feature importances of the logistic regression model



The 20 most important features in the logistic regression model are shown. The features for which a high value is predictive of a positive BC are shown in blue and those predictive of a negative culture in red. The X-axis presents the relative importance of these features.

Thresholds

The models sensitivity and specificity depend on the probability threshold that is used to predict a positive or negative BC. Table 7.3 presents model performances for the optimal sensitivity-specificity threshold and a threshold that retains a sensitivity of 90%. The optimal threshold in the gradient boosted tree model would predict 69.1% of BCs in the test set to be negative, with a negative predictive value of over 94%. An extensive list of thresholds and corresponding performances in both sets can be found in e-Table 8 and 9.

Table 7.3 Performance metrics for both models at preselected thresholds in the aggregated cross-validation sets and the test set

Model and metric	Optimal sensitivity-specificity		Sensitivity retained at over 90%	
	Cross-validation (n=3608)	Test (n=1277)	Cross-validation (n=3608)	Test (n=1277)
Gradient boosted tree model				
Threshold for positive prediction	10%	12.5%	6%	5%
True Negative (n (%))	2119 (58.7%)	831 (65.1%)	1325 (36.7%)	360 (28.2%)
True Positive (n (%))	325 (9.0%)	104 (8.1%)	405 (11.2%)	144 (11.3%)
False Negative (n (%))	118 (3.3%)	51 (4.0%)	38 (1.1%)	11 (0.9%)
False Positive (n (%))	1046 (28.8%)	291 (22.8%)	1840 (51.0%)	762 (59.7%)
Sensitivity (%)	73.4	67.1	91.4	92.9
Specificity (%)	67	74.1	41.9	32.1
Positive predictive value (%)	23.7	26.3	18	15.9
Negative predictive value (%)	94.7	94.2	97.2	97
Logistic regression model				
Threshold for positive prediction	12.5%	10%*	6%	5%
True Negative (n (%))	2160 (59.9%)	676 (52.9%)	1120 (31.0%)	339 (26.5%)
True Positive (n (%))	302 (8.4%)	123 (9.6%)	412 (11.4%)	146 (11.4%)
False Negative (n (%))	141 (3.9%)	32 (2.5%)	31 (0.8%)	9 (0.7%)
False Positive (n (%))	1005 (27.9%)	446 (34.9%)	2045 (56.7%)	783 (61.3%)
Sensitivity (%)	68.2	79.4	93	94.2
Specificity (%)	68.2	60.2	35.4	30.2
Positive predictive value (%)	23.1	21.6	16.8	15.7
Negative predictive value (%)	93.9	95.5	97.3	97.4

**This is the only scenario where the optimal threshold would be different when based on the maximum sum of sensitivity and specificity or on a minimal difference between sensitivity and specificity. In this case, the threshold was chosen based on the maximum sum of sensitivity and specificity.*

Medication administered in the ED

In coming to the final models, we evaluated the effects of excluding different groups of features, such as medications given in the ED. Excluding all ED medication features led to comparable model performances (see e-Table 10 for details). When including the ED medication features, almost none provided predictive value, except for the administration of antibiotics (see e-Figures 7.2 and 7.3). Because this event may be associated with the physician's suspicion of bacteraemia, we decided to exclude ED medication features in order to retain a model that can augment physician decision-making instead of depending on it.

DISCUSSION

Short summary

We present two machine learning models that aim to predict the outcome of a BC that is drawn during an ED visit. Both models show comparably good performance in predicting BC results with AUROCs of 0.77 (95%-CI = 0.73-0.82) in the test sets. The models can identify patients in the ED with low risk for bacteraemia and can be useful to reduce unnecessary BCs and provide physician decision support on the necessity of antibiotic therapy.

In context of the literature

Many studies have aimed to identify factors associated with positive BCs or predict BC outcomes. A 2012 systematic review reported on 35 studies that evaluated the performance of clinical variables to detect bacteraemia [2]. Those clinical variables alone seemed insufficient to detect bacteraemia and further studies on this subject have focused on more advanced predictive models to detect bacteraemia. A 2015 systematic review presented fifteen machine learning models that predicted BC outcomes [17]. An additional few were published since [18-21].

The various studies on this subject have been conducted in different settings, where the reasons for drawing BCs vary. We focused on the ED setting, as the legacy of a probable diagnosis of infection at the ED greatly influences decision-making throughout the hospital stay, especially with regards to antibiotic treatment [22]. Based on the 2015 systematic review, only two other studies have been carried out fully in an ED setting [17,23,24]. Those models showed AUROCs of 0.75 and 0.74 in the test sets. The major difference with our study is that those earlier models were trained on data that were prospectively collected by researchers. This manual data collection resulted in few missing values, with 97.6% of laboratory data being available [23]. This will not occur in clinical practice and may lead to dramatic losses in predictive performance in implementation studies, when missing values need to be imputed in order to do any prediction. Therefore, these models have less potential for daily use in clinical practice and it will be difficult to implement them successfully.

Another aspect of the manual data collection in earlier studies is that predictors like the suspicion of endocarditis, which was an important predictor of BC outcomes, could be used [23]. This is very specific data that will rarely be available in the EHR, which again limits the translation to clinical practice and automation of the prediction within an EHR environment. As we illustrate here, the use of data that is not routinely captured in clinical practice is one of the key reasons why none of these prediction models have been implemented in clinical practice yet [17]. In contrast, our models are based on routinely collected clinical data that is available at the end of an ED visit and could be implemented straightforwardly in practice.

Most of the literature on BC predictions focuses on the ICU setting. Recent examples are models created by Roimi and colleagues and van Steenkiste et al [19,20]. Those models show excellent performances with AUROCs of up to 0.98 in the critical care setting.

These models are trained on temporal trends that have occurred over a period of at least 48 hours, in contrast with the short and heterogeneous ED visits during which patients are not constantly monitored and where time-series data is rarely captured. Also, the approaches as taken for most ICU models seem to be overfitting to the training data and will likely perform worse in an external validation. This is underscored in the model by Roimi and colleagues, in which the AUROC decreases from 0.92 to 0.60 during external validation [20].

Clinical value

The main clinical value of our predictive model lies in the ability to identify a population in which the chance of a positive BC is very low. The prediction can be made at the end of the ED visit and can identify patients in which we can safely withhold BC testing. Even in cases where BCs are already taken, there would be the option to not go through with the analyses, where most of the costs are made. We showed that we would be able to withhold BC draws or analyses in almost 70% of the population while still retaining a negative predictive value of over 94%.

Our algorithm also has added value with regards to treatment selection, especially in cases with high diagnostic uncertainty at the end of an ED visit. The BC outcome prediction can be used as decision support tool to decide whether or not antibiotic treatment is needed. Predictions of negative BCs can be an additional argument for withholding antibiotic treatment at that point and may help avoid unnecessary courses of empiric broad-spectrum antibiotics that can sometimes be given for several days due to delays in the turnaround time of BC results [25]. When a specific infection such as pneumonia is very likely, then antibiotic treatment will be initiated regardless of the BC draw. However, in these cases our algorithm can still be used to withhold unnecessary BC testing.

Another clinically relevant aspect of this study is that we were able to show that routine laboratory results are associated with positive BCs. A low lymphocyte count appears to be related to a positive BC. This association has been described in earlier studies, but this variable has not been included in bacteraemia prediction models up until now [26,27]. Bilirubin is another notably strong predictor of a positive BC. Elevated bilirubin levels have been observed in patients with sepsis, and it is included in prognostic scores, like the Sequential Organ Failure Assessment (SOFA) score for sepsis [28,29]. The association with positive BCs of other variables such as thrombocyte counts, temperature, blood pressure, heart rate and age is in line with previous studies [2,6,23].

Strengths

The main strength that distinguishes this work from what has been done before is the comprehensive pipeline from raw data to model. The preprocessing and feature engineering phases were conducted in collaboration with a machine learning scale-up company (Pacmed, the Netherlands), which has considerable experience with machine learning in healthcare. The strategy towards the selection of features and algorithms that were used to predict BC outcomes presents a significant improvement over currently accepted methods in the medical literature. Our pipeline used all available data so that the models themselves would decide on the importance of any feature.

With this approach, the models were not limited by the selection of features through current medical knowledge and had the potential to discover unknown associations with bacteraemia. Throughout preprocessing stages, we put emphasis on only using data that would routinely be available at the end of the ED visit, when the final treatment and admission decisions have to be made. This approach facilitates straightforward implementation of the models in clinical practice. Finally, we compare the results of the more complex gradient boosted tree model with a simpler logistic regression that is easier to understand for physicians, to improve the overall interpretability.

Limitations

There are several important limitations within this study. Firstly, defining a positive BC is difficult. Our definition of contamination, which was defined as BCs that grew pathogens that are generally considered contaminants, is in line with previous literature [2,4,5,12]. Nevertheless, it is still possible that samples, that were mapped as contamination, actually represented a true pathogen in that individual patient. However, the true positive rate of collected BC's in our population was somewhat higher than those described in previous literature [4,6,7,30]. This may be due to conservative mapping of pathogens to likely contaminants. A related limitation is that the model should not be used when a physician wants to detect a clinically relevant blood stream infection with pathogens that we considered to be contaminants, as with suspected central line-associated bloodstream infections (CLABSI). Our algorithm should be used as additive to the clinical pretest probability of bacteraemia, based on syndromes with a high likelihood of bacteraemia reported in earlier studies [31].

Another limitation of this study is that various potentially predictive variables could not be adequately extracted from the EHR system. Comorbidities, medication at home and placement of lines are not well documented within the EHR and this data would not be reliable enough to use in a prediction model. Furthermore, we were not able to use free-text data due to privacy concerns. Therefore, we could not use physician and nurse reports.

Future research

Our current study gives rise to several potential follow-up studies. Firstly, an implementation study should be done to study the clinical benefits derived from using such an algorithm to safely withhold BC draws or analyses in low-risk patients.

Additionally, there is a need to further improve the model performance, without losing sight of the usability of the model. For example, various studies have shown that procalcitonin can predict BC positivity with good performance [18,32]. We would be interested to see what the addition of procalcitonin and other novel biomarkers can do to our model performances. Another important step could be to include additional clinical information by using free-text data.

In conclusion, we created two machine learning models that predict BC outcomes in the ED with an AUROC of 0.77. The models are based on routinely captured clinical data and are well suited for implementation in clinical practice. The main value of these models lies in the ability to identify patients at low risk of bacteraemia, which can help reduce unnecessary BC testing and provides an additional tool to decide whether antibiotic treatment is needed. Based on the model predictions, we would be able to withhold BC testing in 70% of the population with few omission.

REFERENCES

1 Wang HE, Jones AR, Donnelly JP. Revised National Estimates of Emergency Department Visits for Sepsis in the United States. Crit Care Med [Internet]. 2017 Sep 1 [cited 2021 Feb 5];45(9):1443–9. Available from: /pmc/articles/PMC5609499/?report=abstract

2 Coburn B, Morris AM, Tomlinson G, Detsky AS. Does this adult patient with suspected bacteremia require blood cultures? JAMA - J Am Med Assoc [Internet]. 2012 Aug 1 [cited 2020 Aug 3];308(5): 502–11. Available from: www.jamaarchivescme.com

3 Goto M, Al-Hasan MN. Overall burden of bloodstream infection and nosocomial bloodstream infection in North America and Europe [Internet]. Vol. 19, Clinical Microbiology and Infection. Blackwell Publishing Ltd; 2013 [cited 2021 Feb 5]. p. 501–9. Available from: www.statcan.

4 Nannan Panday RS, Wang S, Van De Ven PM, Hekker TAM, Alam N, Nanayakkara PWB. Evaluation of blood culture epidemiology and efficiency in a large European teaching hospital. PLoS One. 2019 Mar 1;14(3).

5 Dargère S, Cormier H, Verdon R. Contaminants in blood cultures: importance, implications, interpretation and prevention [Internet]. Vol. 24, Clinical Microbiology and Infection. Elsevier B.V.; 2018 [cited 2020 Nov 30]. p. 964–9. Available from: https://pubmed.ncbi.nlm.nih.gov/29621616/

6 Jessen MK, Mackenhauer J, Hvass AMSW, Ellermann-Eriksen S, Skibsted S, Kirkegaard H, et al. Prediction of bacteremia in the emergency department. Eur J Emerg Med [Internet]. 2016 Feb [cited 2021 Feb 5];23(1):44–9. Available from: http://journals.lww.com/00063110-201602000-00010

7 Denny KJ, Sweeny A, Crilly J, Maloney S, Keijzers G. Is it time for a culture change? Blood culture collection in the emergency department. Emerg Med Australas [Internet]. 2018 Aug 1 [cited 2021 Mar 8];30(4):575–7. Available from: http://doi.wiley.com/10.1111/1742-6723.13122

8 Bates DW, Goldman L, Lee TH. Contaminant Blood Cultures and Resource Utilization: The True Consequences of False-Positive Results. JAMA J Am Med Assoc [Internet]. 1991 Jan 16 [cited 2021 Feb 5];265(3):365–9. Available from: https://jamanetwork.com/

9 Zwang O, Albert RK. Analysis of strategies to improve cost effectiveness of blood cultures. J Hosp Med [Internet]. 2006 Sep 1 [cited 2021 Feb 5];1(5):272–6. Available from: http://www.journalofhospitalmedicine.com/jhospmed/article/126938/improving-cost-effectiveness-blood-cultures

10 Schinkel M, Paranjape K, Panday RSN, Skyttberg N, Nanayakkara PWB. Clinical applications of artificial intelligence in sepsis: A narrative review. Comput Biol Med. 2019 Oct;103488.

11 Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. Ann Intern Med [Internet]. 2015 Jan 6 [cited 2021 Feb 22];162(1):55. Available from: http://annals.org/article.aspx?doi=10.7326/M14-0697

12 Harvey DJ, Albert S. Standardized definition of contamination and evidence-based target necessary for high-quality blood culture contamination rate audit [Internet]. Vol. 83, Journal of Hospital Infection. Elsevier; 2013 [cited 2020 Nov 30]. p. 265–6. Available from: http://www.journalofhospitalinfection.com/article/S0195670112003982/fulltext

13 Choi J, Dekkers OM, le Cessie S. A comparison of different methods to handle missing data in the context of propensity score analysis. Eur J Epidemiol [Internet]. 2019 Jan 15 [cited 2021 Feb 18];34(1):23–36. Available from: https://doi.org/10.1007/s10654-018-0447-z

14 Jiménez-Valverde A, Lobo JM. Threshold criteria for conversion of probability of species presence to either-or presence-absence.

15 Pewsner D, Battaglia M, Minder C, Marx A, Bucher HC, Egger M. Ruling a diagnosis in or out with "SpPIn" and "SnNOut": A note of caution [Internet]. Vol. 329, British Medical Journal. BMJ Publishing Group; 2004 [cited 2021 Feb 24]. p. 209–13. Available from: /pmc/articles/PMC487735/

16 5.10 SHAP (SHapley Additive exPlanations) | Interpretable Machine Learning [Internet]. [cited 2021 Feb 3]. Available from: https://christophm.github.io/interpretable-ml-book/shap.html

17 Eliakim-Raz N, Bates DW, Leibovici L. Predicting bacteraemia in validated models-a systematic review. Vol. 21, Clinical Microbiology and Infection. Elsevier; 2015. p. 295–301.

18 Ratzinger F, Haslacher H, Perkmann T, Pinzan M, Anner P, Makristathis A, et al. Machine learning for fast identification of bacteraemia in SIRS patients treated on standard care wards: a cohort study. Sci Rep [Internet]. 2018 Dec 1 [cited 2020 Aug 3];8(1):12233. Available from: www.nature.com/scientificreports

19 Van Steenkiste T, Ruyssinck J, De Baets L, Decruyenaere J, De Turck F, Ongenae F, et al. Accurate prediction of blood culture outcome in the intensive care unit using long short-term memory neural networks. Artif Intell Med. 2019 Jun 1;97:38–43.

20 Roimi M, Neuberger A, Shrot A, Paul M, Geffen Y, Bar-Lavie Y. Early diagnosis of bloodstream infections in the intensive care unit using machine-learning algorithms. Intensive Care Med [Internet]. 2020 Mar 1 [cited 2020 Aug 3];46(3):454–62. Available from: https://doi.org/10.1007/s00134-019-05876-8

21 Liu VX, Wiens J. 'No growth to date'? Predicting positive blood cultures in critical illness [Internet]. Vol. 46, Intensive Care Medicine. Springer; 2020 [cited 2020 Aug 3]. p. 525–7. Available from: https://doi.org/10.1007/s00134-019-05917-2

22 Charani E, Ahmad R, Rawson TM, Castro-Sanchèz E, Tarrant C, Holmes AH. The Differences in Antibiotic Decision-making Between Acute Surgical and Acute Medical Teams: An Ethnographic Study of Culture and Team Dynamics. Clin Infect Dis [Internet]. 2019 Jun 18 [cited 2021 Feb 18];69(1):12–20 Available from: https://academic.oup.com/cid/article/69/1/12/5174860

23 Shapiro NI, Wolfe RE, Wright SB, Moore R, Bates DW. Who Needs a Blood Culture? A Prospectively Derived and Validated Prediction Rule. J Emerg Med. 2008 Oct 1;35(3):255–64.

24 Tudela P, Lacoma A, Prat C, Mòdol JM, Giménez M, Barallat J, et al. Prediccin de bacteriemia en los pacientes con sospecha de infeccin en urgencias. Med Clin (Barc). 2010 Nov 20;135(15):685–90.

25 Sweeney TE, Liesenfeld O, May L. Diagnosis of bacterial sepsis: why are tests for bacteremia not sufficient? Expert Rev Mol Diagn [Internet]. 2019 Nov 2 [cited 2021 Feb 18];19(11):959–62. Available from: https://www.tandfonline.com/doi/full/10.1080/14737159.2019.1660644

26 de Jager CPC, van Wijk PTL, Mathoera RB, de Jongh-Leuvenink J, van der Poll T, Wever PC. Lymphocytopenia and neutrophil-lymphocyte count ratio predict bacteremia better than conventional infection markers in an emergency care unit. Crit Care [Internet]. 2010 Oct 29 [cited 2021 Feb 5];14(5). Available from: https://pubmed.ncbi.nlm.nih.gov/21034463/

27 Laukemann S, Kasper N, Kulkarni P, Steiner D, Rast AC, Kutz A, et al. Can We Reduce Negative Blood Cultures With Clinical Scores and Blood Markers? Results From an Observational Cohort Study. Medicine (Baltimore) [Internet]. 2015 Dec [cited 2021 Feb 5];94(49):e2264. Available from: http://journals.lww.com/00005792-201512080-00053

28 Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016 Feb;315(8):801.

29 Minemura M, Tajiri K, Shimizu Y. Liver involvement in systemic infection. World J Hepatol [Internet]. 2014 [cited 2021 Feb 22];6(9):632–42. Available from: /pmc/articles/PMC4179142/

30 Bates DW, Cook EF, Goldman L, Lee TH. Predicting bacteremia in hospitalized patients: A prospectively validated model. Ann Intern Med [Internet]. 1990 [cited 2021 Mar 8];113(7):495–500. Available from: https://pubmed.ncbi.nlm.nih.gov/2393205/

31 Fabre V, Sharara SL, Salinas AB, Carroll KC, Desai S, Cosgrove SE. Does This Patient Need Blood Cultures? A Scoping Review of Indications for Blood Cultures in Adult Nonneutropenic Inpatients. Clin Infect Dis [Internet]. 2020 Aug 22 [cited 2021 Feb 5];71(5):1339–47. Available from: https://academic.oup.com/cid/article/71/5/1339/5703622

32 Chirouze C, Schuhmacher H, Rabaud C, Gil H, Khayat N, Estavoyer JM, et al. Low serum procalcitonin level accurately predicts the absence of bacteremia in adult patients with acute fever. Clin Infect Dis [Internet]. 2002 Jul 15 [cited 2021 Jan 13];35(2):156–61. Available from: https://pubmed.ncbi.nlm.nih.gov/12087521/

Part 2 / Chapter 7
Supplementary materials

Publication list

online only

CONTENT

Part 3

Sepsis and the Bigger picture

Chapter 8

Health related quality of life in sepsis survivors from the Prehospital Antibiotics Against Sepsis (PHANTASi) trial

PLoS One. 2019 Oct 1;14(10)

R. S. Nannan Panday

T. C. Minderhoud

D. S. Chantalou

N. Alam

P. W. B. Nanayakkara

ABSTRACT

Background

Due to the rise in incidence, the long term effect of sepsis are becoming more evident. There is increasing evidence that sepsis may result in an impaired health related quality of life. The aim of this study was to investigate whether health related quality of life is impaired in sepsis survivors and which clinical parameters are associated with the affected health related quality of life.

Methods

We analyzed 880 Short Form 36 (SF-36) questionnaires that were sent to sepsis survivors who participated in the Prehospital Antibiotics Against Sepsis (PHANTASi) trial. These questionnaires were sent by email, 28 days after discharge. Data entry and statistical analyses were performed in SPSS. The data from the general Dutch population, was obtained from the Netherlands Cancer Institute (NKI-AVL) and served as a control group. Subsequently, 567 sepsis survivors were matched to 567 controls. Non-parametric Wilcoxon signed-rank test was performed to compare these two groups. Within the group, we sought to explain the diminished health related quality of life by factor analysis.

Results

We found that sepsis survivors have a worse health related quality of life compared to the general Dutch population. This negative effect was more evident for the physical component than the mental component of health related quality of life. We found that health related quality of life was significantly altered by advancing age and female sex. We also found that the total length of stay (in the hospital) and (previous) comorbidity negatively affect the physical component of health related quality of life.

Conclusion

In our study we found that health related quality of life in sepsis survivors, 28 days after discharge, is severely diminished in comparison with the general Dutch population. The physical domain is severely affected, whereas the mental domain is less influenced. The length of stay, comorbidity, advancing age and female sex all have a negative effect on the Physical Component Scale of the health related quality of life.

INTRODUCTION

Sepsis is a syndrome of physiologic, pathologic and biochemical abnormalities induced by infection[1]. The worldwide incidence of sepsis is rising. This is caused by several factors, such as: the ageing population, antibiotic resistance, increased use of chemo-and immunotherapy and improved recognition[2],[1]. A meta-analysis of 27 international studies reported a global sepsis incidence of 437 per 100,000 person-years for the last decade [3]. Conversely, in-hospital mortality is decreasing. Previous studies found that in-hospital septic shock mortality decreased from 54.9% to 50.7% from 2005 to 2014 [4]. This decrease in case fatality can be attributed to improved recognition, clinical advances including early goal-directed therapy and mortality reduction campaigns[5].

Previous research has shown that HRQoL (Health-Related Quality of Life) is impaired in sepsis survivors [6,7,8]. A recent systematic review [2] found that 81.3% of ICU-sepsis survivors reported an impaired quality of life which lasted for years after the syndrome was treated.

However, research conducted on the HRQoL of sepsis survivors has been focused only on ICU patients. No studies have been conducted on the HRQoL in either a population of sepsis survivors with varying severities of sepsis or in sepsis patients who were transported to the ED by EMS.

Studies that compared quality of life of sepsis ICU survivors to the general population in several countries[7-10] found decreased HRQoL of the sepsis ICU survivors when their HRQoL was compared to that of the general population of these countries. However, it is not yet known if this effect is caused by the ICU admission or by the sepsis episode itself.

This prospective study will focus on the quality of sepsis survivors that were transported by EMS to the ED. A study in the Netherlands found that nearly half of all patients presenting to the ED were transported by Emergency Medical Services (EMS), and those transported by EMS were sicker[11]. Therefore this population is a representative sample of severely ill hospi-talized sepsis patients.

The patients in this study are those that were included in the Prehospital Antibiotics Against Sepsis (PHANTASi) trial recently published in The Lancet Respiratory Medicine[12,13].

The primary aim of this study is to evaluate the HRQoL of sepsis survivors. Subsequently the HRQoL of sepsis survivors was compared to the general Dutch population in order to determine to what extent the HRQoL of sepsis survivors differs from the general Dutch population. The secondary aims were to analyze which patient characteristics and clinical parame-ters were associated with the decreased HRQoL in sepsis survivors.

METHODS

Design and setting

This prospective study was part of the PHANTASi trial[12,13]. In brief, the PHANTASi trial was the first prospective randomized controlled trial in septic patients which investigated whether improved recognition and administration of antibiotics in the ambulance led to increased survival when compared to usual care. Sepsis severity was categorized in to three groups as defined by the 2001 SSCM/ ESCIM/ ACCP/ ATS/ SIS International Sepsis definitions Conference guidelines: (uncomplicated) sepsis, severe sepsis and septic shock[14]. The sepsis diagnosis was cross-checked at the ED by the attending physician. Patients characteristics and clinical parameters such as Charlson Comorbidity Index, clinical values, laboratory values, sepsis severity, organ dysfunction, hospital and ICU length of stay and readmission among others, were derived from patients medical charts.

Methodology

All patients in the study that survived their hospital admission were sent a SF-36 questionnaire [15] by mail within one month after discharge in order to measure their HRQoL. The SF-36 is a widely used [16], standardized questionnaire for measuring HRQoL. This questionnaire has been validated for the measurement of HRQoL in sepsis survivors [9] and is also validated in the Dutch Language [17]. The SF-36 measures HRQoL by addressing eight domains: physical functioning, role limitation due to physical problems, role limitation due to emotional problems, social functioning, bodily pain, mental health, vitality and general health. These eight domains are clustered into two summary scores: a physical component score (PCS) and a mental component score (MCS)[18]. The PCS consists of the domains: Physical Functioning, Role Functioning Physical, Bodily Pain and General Health. The MCS consist of the domains: Vitality, Social Functioning, Role Functioning Emotional and Mental Health.

Outcomes

The primary outcome is the HRQoL in sepsis survivors in this Dutch population compared to the HRQoL of the General Dutch population. To increase our understanding of the potential changes in HRQoL we we examined which clinical parameters were associated with the HRQoL. For this, we studied parameters that were associated with HRQOL in previous studies [17-19]. These factors consisted of demographic characteristics (age/sex), overall comorbidity (charlson comorbidity index), and specific comorbidity such as chronic pulmonary disease, heart failure, diabetes and cancer), sepsis severity, organ dysfunction and length of hospital stay.

Statistical analysis

Data are expressed as means and standard deviation (SD ±) if the data is normally distributed or median and interquartile range (IQR) if the data exhibited a non-normal distribution.

Raw scores from each of the 36 items were entered into SPSS (IBM version 22.0). These raw scores were transformed to scores that ranged from 0 to 100 as per the guidelines of the RAND corporation[19]. Higher scores of any domain correspond with a better HRQoL. Differences between SF-36 domains per age category were analysed by non-parametric Kruskal-Wallis one-way-ANOVA. A p-value < 0.05 was considered statistically significant.

Matching In order to compare the SF-36 data of our sepsis survivors population to the general Dutch population we retrieved data from the Netherlands Cancer Institute (NKI-AVL)[17]. This general Dutch population did not necessarily contain patients with malignant dis-eases, as it was a representative sample of the general Dutch population, which was used in a previous study. The data in that study was retrieved by sending out questionnaires to a randomly selected population in the Netherlands[17]. The age and gender matched data was com-pared by using a non-parametric Wilcoxon signed-rank test. Patients were matched by gender with a maximum age difference of 5 years between those in the study group and those in the control group. No information on the comorbidities of the patients in the control group was known and therefore matching for this aspect was not possible.

To analyse clinical parameters associated with HRQoL, non-parametric Mann-Whitney U test and Kruskal-wallis-1-way-ANOVA were used for prespecified factors. Clinical parameters with a $P < 0.05$ were included in a multiple linear regression analysis by blockwise entry.

Ethics

The study protocol of the PHANTASi trial was approved by the medical ethical committee of the Amsterdam University Medical Center, Location VU University Medical Center, the coordinating center and all ethical bodies of each participating hospital. Due to the complexity of the PHANTASi trial, the ethics committees granted approval to obtain deferred consent when necessary. Informed consent before study enrollment or deferred consent was obtained from all patients or their legal representatives or surrogates. All effort was made by EMS personnel to obtain informed consent before study inclusion provided the acuity of the situation allowed it.

RESULTS

Patient characteristics and demographics

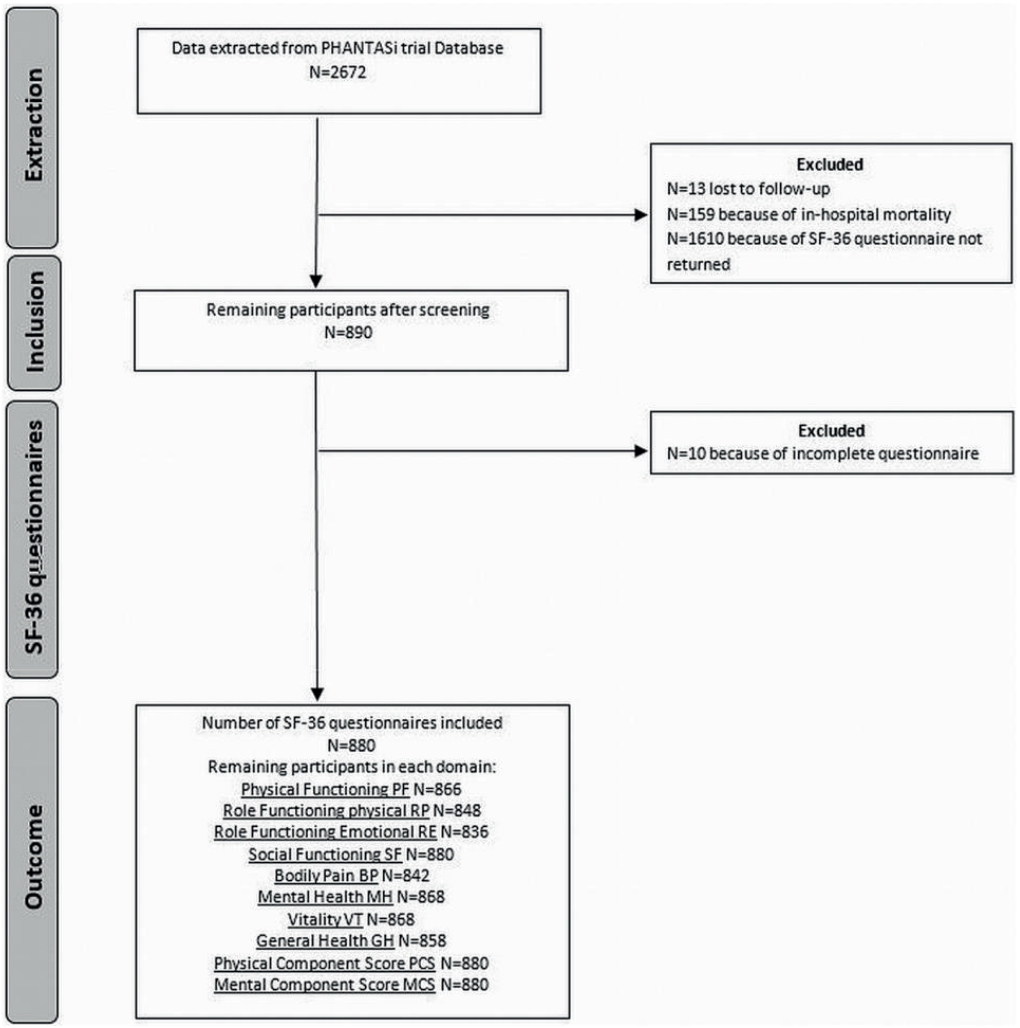
The PHANTASi included patients who had sepsis according to the SEPSIS-2 criteria which were used in the study. The SEPSIS-2 criteria are more sensitive but less specific in diagnosing sepsis[20]. 2672 patients were included in the PHANTASi trial. 13 patients were lost to follow-up and 159 patients did not survive their hospital admission. 1610 patients did not return the questionnaire or were excluded due to incomplete questionnaire. A total of 880 questionnaires remained for analysis (Fig 8.1).

The 880 patients who returned the questionnaire had a mean age of 72.8 ± 12.6 and 58.9% were male (Table 8.1). There were a number of important differences between patients who did and did not return the questionnaire. The patients who did not return the questionnaire had a significantly higher Charlson Comorbidity Index ($p = .016$), and had significantly more cere-brovascular diseases and dementia ($p < .001$). Patients who

did not return the questionnaire had a higher 90 day mortality rate of 5.2% versus 1.0% (p < .001).Data are presented as N (%), mean (SD/±) or median (IQR = Inter Quartile Range).

ED = Emergency Department. ICU = Intensive Care Unit.

Figure 8.1 Overview of patients included in the study



Data are presented as N (%), mean (SD/±) or median (IQR = Inter Quartile Range). ED = Emergency Department. ICU = Intensive Care Unit. ^independent samples T-test
χChi square test
 # Mann whitney U test

Table 8.1 Demographic and characteristics of the PHANTASi trial patients (Total N = 2649)

Characteristics	SF-36 questionnaire returned (N = 880)	SF-36 questionnaire not returned (in hospital mortality excluded)(N = 1610)	P
Age-years	72.9 ±12.51	72.0 ± 12.51	.132^
Male sex–no (%)	516(58.9)	869 (56.4)	.230χ
Charlson Comorbidity Index (median/IQR)	1 (0–2)	1 (1–3)	.016 #
Underlying chronic conditions (%)			
Chronic pulmonary disease	262 (29.9)	460 (29.9)	.976χ
Diabetes	205 (23.4)	372 (24.1)	.682χ
Malignancy <5yrs	113 (12.9)	198 (12.8)	.971χ
Congestive heart failure	75 (8.6)	166 (10.8)	.081χ
Dementia	19 (2.2)	95 (6.2)	< .001χ
Severity of sepsis (%)			
Sepsis	360 (41.6)	608 (40.2)	.508χ
Severe sepsis	477 (55.1)	863 (57.0)	.354χ
Septic shock	26 (3.0)	41 (2.7)	.678χ
Admission (%)			
Hospital	849 (96.9)	1458(94.6)	.009
ICU	78 (8.9%)	123(8.0)	.77χ
Readmission	56 (6.4%)	125 (8.1%)	.123
Length of stay (median/IQR)			
Hospital	5 (4–8)	6(4–10)	.961#
ICU/ MCU	0 (0–0)	0 (0–0)	.296#
Mortality (%)			
90 days	8 (1.0)	80 (5.2)	< .001χ

Data are presented as N (%), mean (SD/±) or median (IQR = Inter Quartile Range). ED = Emergency Department. ICU = Intensive Care Unit. ^independent samples T-test
χChi square test
 # Mann whitney U test

Primary outcome

Sepsis survivors compared to the general Dutch population

Sepsis survivors, 4 weeks after discharge, had a statistically significant lower median Physical Component Score of 33.2 (IQR 26-43) and Mental Component Score of 45.4 (IQR35-53) compared to the matched general Dutch population which had a median 48.3 (IQR 38–54) and median 54 (IQR 47–58) respectively. This difference was significant for both PCS and MCS (p< 0.001).

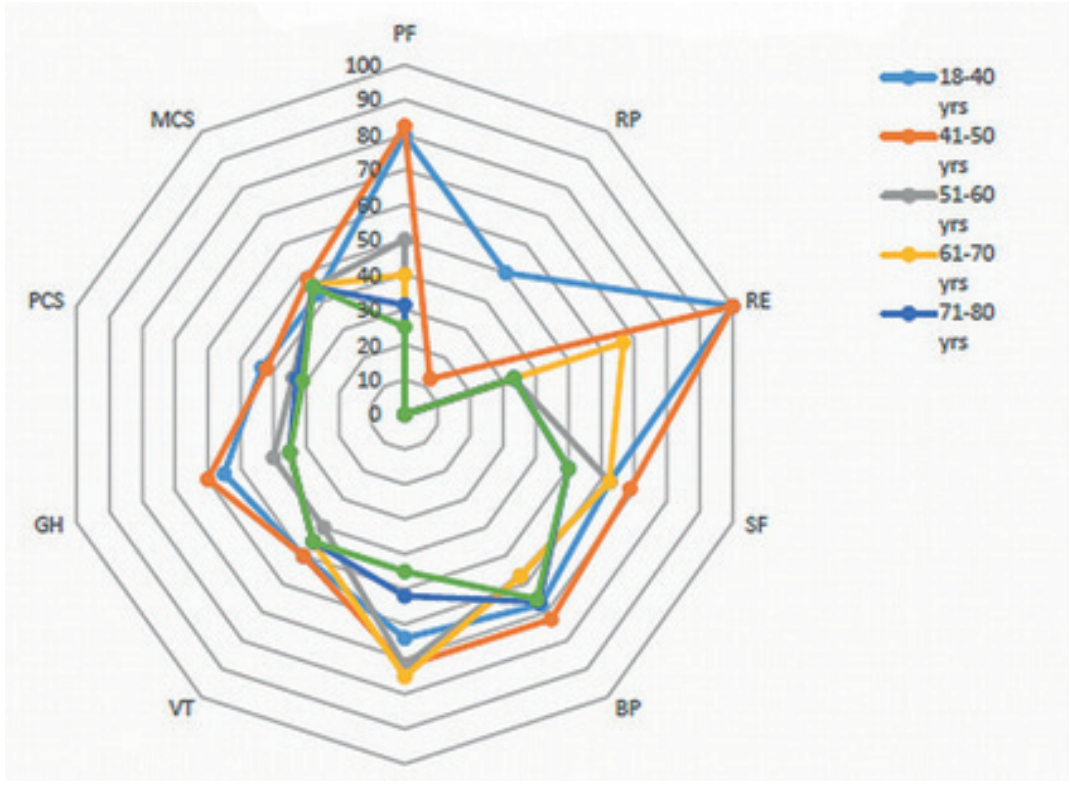
SF-36 HRQoL subdomains

The different subdomains also exhibit a statistically signifi-cant lower SF-36 score compared to the general Dutch population (Table 8.2).

Table 8.2 SF-36 scores of Sepsis Survivors compared to general Dutch population

	Median SF-36 scores		
SF-36 Domains	**Sepsis Survivors N = 567**	**Matched General Dutch Population N = 567**	***P***
Physical functioning (PF)N = 553	38.9 (15–75)	80 (55–90)	< .001
Role functioning Physical (RF) N = 520	0 (0–50)	100 (25–100)	< .001
Role functioning Emotional (RE) N = 509	66.7 (0–100)	100 (66.7–100)	< .001
Social functioning N = 563	62.5(38–75)	87.5	< .001
		(63–100)	
Bodily Pain N = 543	63.3 (IQR 35–90)	74 (IQR 51–100)	< .001
Mental Health N = 546	72 (IQR 52–84)	80 (IQR 64–88)	< .001
Vitality N = 546	45 (IQR 30–60)	70 (IQR 55–80)	< .001
General Health N = 532	40 (IQR 20–55)	67 (IQR 50–77)	< .001
Physical component score N = 514	33.2 (26–43)	48.3 (IQR 38–54)	< .001
Mental component score N = 514	45.4 (IQR 35–53)	54(IQR 47–58)	< .001

Figure 8.2 Spider diagram, HR-QOL by age category



Health related quality of life in sepsis survivors

Sepsis survivors had a median Physical Component Score (PCS) of 32.9 (IQR 26–41) and a median Mental Component Score (MCS) of 45.1 (IQR 35–53) (Table 8.2). When PCS is divided per age category, older respondents have significant worse scores compared to younger respon-dents (p<0.001). MCS did not differ throughout the different age groups (p = 0.970). Sepsis sur-vivors scored lowest for the Role Functioning Physical domain (median: 0.0, IQR 0–50) and the highest for the Mental Health domain (median = 72, IQR = 55–84). Physical Functioning, Role Functioning Physical, Role Functioning Emotional and General Health exhibit an overall decline with increasing age: (p < 0.001, p = 0.025, p < 0.001 and p = 0.022 respectively) (Fig 8.2).

Secondary outcome: What are contributing factors for decreased HRQoL

Age & sex

Bivariate analysis showed a significant, negative effect of age on the PCS (p = .002). The MCS however was not significantly affected by age (p = .765). Female sex was also associated with a small but significantly worse score on the PCS (p = .024), but likewise not in the MCS.

Length of stay and ICU versus hospital admission

ICU admission was not significantly associated with MCS or PCS scores (p = .063 for the PCS and p = 0.349 for the MCS). The total length of stay (including ICU days) was significantly associated with both MCS and PCS: an increase in length of stay resulted in a significant decrease in the PCS and MCS with p < .001.

Comorbidities and the effect on quality of life

A Charlson Comorbidity Index (CCI) > 3 was significantly associated with a lower PCS score (p < .001), patients with a CCI 0-3 had a median score of 33.4 (IQR 26-43) compared to a median score of 30.2 (IQR 23-35) in patients with a CCI > 3. MCS was also significantly associated with CCI (p = .039). Patients with CCI > 3 had a median score of 41.9(IQR 32-51-51) compared to patients with a CCI score of 0-3 with a median score of 45.7 (IQR 35-54). Chronic (Obstructive) pulmonary disease, heart failure and diabetes all had a significant nega-tive effect on the PCS compared to patients without comorbidity. The effect on MCS was not significant, except for patients with diabetes, who had a significantly worse median MCS of 41.7 versus 45.7 in patients without diabetes. Strikingly, malignancy (in the past 5 years) showed no significant effect on either PCS and MCS in the mann-whitney u test (p = .18 and p = .52 respectively).

Organ dysfunction and sepsis severity

Patients with Central Nervous System (CNS) organ dysfunction had a median MCS of 42.6 (IQR 31-52) versus 45.6 (IQR 36-54). CNS organ dysfunction was negatively associated with a lower MCS (p < .001). Other differences in organ dysfunction such as renal or pulmonary dys-function did not show significant effects on either the PCS or MCS. Sepsis severity studied by performing a Mann Whitney U test comparing patients with sepsis alone (SEP 2 definition) to the more severe forms (severe sepsis and septic shock). For the PCS, the group with sepsis alone had a median PCS of 34.0, whereas patients

with more severe forms of sepsis had a median PCS of 31.8. This was a significant difference with a p-value of .044.
The median MCS is in the group with sepsis was 44.3, compared to 45.6 in the group with more severe sepsis, which was not significantly different (p-value = .65)

Regression model
As stated before, all independent factors with a significant association were then studied in a linear regression model. For the PCS the following parameters were entered in the multiple regression model: age, sex, heart failure, diabetes, chronic pulmonary disease, sepsis severity and total length of stay. In the regression model, sepsis severity did not show a significant association with PCS with a p-value of .79 and was therefore excluded in the final model. Dia-betes also did not show a significant association with a p-value of 058 and was excluded. The final model is displayed in Table 8.3. Age is negatively associated with PCS. For every year increase in age, the PCS decreased -0,14 (95%CI -.20 to-0.08). Female sex showed a negative association with PCS, with a decrease of 1.93 (95%CI -3.42 to– 0.44) point on the PCS for females compared to males. A history of heart failure resulted in a decrease of 3.99 (95%CI -6.62 to -1.37) points on the PCS. Chronic pulmonary disease had the greatest effect on the PCS, with a decrease of 4,38 (95%CI -5.99 to -2.77).For MCS, sex, length of stay, diabetes and CNS organ dysfunction were entered in the model. Sex did not show significant association in the regression model and was excluded in the final model. The final model is displayed in Table 8.4. The presence of CNS organ dysfunction during the admission was associated with a decrease of the MCS of 2,56 points (95%CI -4.68 to -.45). Having diabetes also decreased MCS, with 2.59 points (95%CI -4.46 to -.73) Length of stay was also negatively correlated, for every day in the hospital the MCS decreased with .28 points (95%CI -4.46 to -.73)

Table 8.3 Linear regression: Factors associated with the physical component score

		Unstandardized regression coefficient	95% Confidence interval	Standardized regression coefficient	P-value
Final model[a]	**Age**	-0.14	-.20 to -.08	-.15	.000
	Seks	-1.93	-3.42 to -.44	-.09	.011
	Heart failure	-3.99	-6.62 to -1.37	-.10	.003
	Chronic Pulmonary disease	-4.38	-5.99 to -2.77	-.18	<001
	Length of stay (total)	-.023	-.33 to -.12	-.14	<.001

Variables entered in the equation stepwise were age, sex, heart failure, diabetes, chronic pulmonary disease, sepsis severity and total length of stay. Variables with a p-value < .05 were included in the final model.

Table 8.4 Linear regression: Factors associated with the mental component score

		Unstandardized regression coefficient	95% Confidence interval	Standardized regression coefficient	P-value
Final model[a]	**CNS organ dysfunction**	-2.56	-4.68 to -.45	-.08	.018
	Length of stay (total)	-.28	-,40 to -,17	-.17	<,001
	Diabetes	-2.59	-4.46 to -,73	-.09	.006

Variables entered in the equation were sex, length of stay, diabetes and CNS organ dysfunction. Variables with a p-value < 0,05 were included in the final model.

DISCUSSION

In this study on Health Related Quality of Life in sepsis survivors we investigated HRQoL by analyzing SF-36 questionnaires returned 28 days after discharge by sepsis survivors who were transported by ambulance to the ED. We found that sepsis survivors, 28 days after discharge had a significantly worse HRQoL compared to the Dutch general population. This negative effect was more evident for the physical partof the HRQoL compared to the mental part. In line with previously published data[21], we found a decrease of the physical component score (PCS) with increasing age. Regarding the clinical and vital parameters associated with HRQoL in sepsis survivors, our findings showed that the total length of stay (hospital and ICU length of stay combined), comorbidity such as heart failure and chronic pulmonary disease have a sig-nificant negative association with the physical component of the HRQoL.
The mental component score in our cohort was significantly associated with length of stay, but not age or sex. Of other factors described in literature, only diabetes and CNS organ dys-function showed a significant, negative association with the mental component of HRQoL.

Our findings regarding the decreased HRQoL in our cohort of sepsis survivors, is similar to those of studies conducted in ICU sepsis survivors [2, 6–10]. Our results indicate that this neg-ative effect tends to be greater for the domains regarding the physical component compared to the domains regarding the mental component.

Previous studies which compared quality of life of ICU sepsis survivors to the General French population [10], Scottish population [8], U.S. population [9] and Dutch population[7], show similar results. The fact that we found similar results in our study population, suggests that these effects can not only be attributed to the ICU admission, since only a small propor-tion of our study population was admitted to the ICU.

These studies also consistently show that not all of the domains are always negatively affected. Heyland and colleagues [9] compared the quality of life of 30 ICU sepsis survivors to the general US population, at discharge and two weeks after discharge. They found that the components related to physical and social function were lower than the

US population, but Bodily Pain, Role Emotional, Mental Health and Mental Component score were not signifi-cantly lower. Another explanation why the MCS seems less affected by sepsis, may lie in the so called response shift. Response shift is defined as a change in a person's perception of his own quality of life due to recalibration, reprioritization or redefinition of a person's value of a 'good' quality of life which can occur after an experience of hardship[22]. Thus this may explain why subjective measures seem to be minimally affected by sepsis. This might also explain the effect that we found septic shock survivors to be 2.5 times more likely to have a good Role Functional Emotional score compared to (uncomplicated) sepsis survivors. Granja and colleagues [23] assessed HRQoL by using the EuroQol five dimensions (EQ-5D) question-naire and found that sepsis survivors (septic shock and severe sepsis combined) have fewer problems on the depression and anxiety dimension compared to survivors of other critical illness.

Factors influencing PCS

For our secondary outcomes we observed that hospital length of stay and comorbidity such as heart failure, chronic pulmonary disease negatively affect the Physical Component Score. An association between pulmonary dysfunction and decreased quality of life are provided by sev-eral studies[10]. In regards to the negative association of hospital stay and HRQoL, length of stay has been described to be a risk factor for impaired Physical Functioning [24]. Further-more, there is evidence that bed rest may lead to muscle wasting[25].

Female sepsis survivors were found to have a lower physical component score of the HRQoL in our study, even after correcting for other factor such as age. We found no explanation for this difference, however studies on other subjects also found women to have lower general HRQoL than men, without an explanation for this phenomenon. Although some studies suggest that this difference might be caused by sociodemographic differences, no definitive explanation for this effect is known[26-29].

Factors influencing MCS

We found that central nervous system dysfunction and diabetes t both have a negative effect on the Mental Component score. Previous literature about the association between delirium an HRQoL are not consistent. Prior mental status has a negative effect on impaired quality of life as seen in the study conducted by Davydow and colleagues[30]. However, Boogard and colleagues[31] suggest that other factors associated with delirium instead of delirium itself may explain overall lower SF-36 scores as they found that patients with delirium exhibit no significant difference in HRQoL compared to patients without delirium. We found that diabetes mellitus had a negative effect on the Mental Component Score. There is evidence that the negative effect of diabetes on HRQoL is not due to the duration and type of diabetes itself, but that secondary complications, demographical and psychosocial factors have a negative effect [32, 33, 34].

Our findings regarding the impaired Health Related Quality of Life after sepsis and the associated clinical parameters, give us leads for possible interventions to ensure a better HRQoL. Interventions are important since HRQoL has a big impact on patients life for example in personal relationships [30] or might also be a financial burden[35]. One study protocol regarding a double-blinded randomized controlled trial analyzing the effect of a multidisciplinary intervention in sepsis survivors has been published. Data has yet to be published [36].

Strengths and weaknesses

Our study contains several strengths. Firstly, the overall number of returned SF-36 questionnaires are relatively high compared to other studies [7, 8, 10, 37]. To the best of our knowledge we have the largest sample of sepsis survivors who were timely recognized and treated by trained medical personnel. Moreover we are also the first to study HRQoL of sepsis survivors transported to the ED by ambulance, thus not restricted only to patients admitted to the ICU. Several other studies have compared quality of life to the general population and to our knowl-edge only one study used the adult general Dutch population[7]. What differentiates our study compared to Hofhuis and colleagues is that we used a large sample of the norm popula-tion which we matched by age and sex. Additionally, the patients included in this study come from different cities from the Netherlands including both urban and rural areas. Therefore a possible bias in quality of life due to a urban or rural setting was avoided[38]. Lastly, the SF-36 questionnaires were all self-administered in contrast to some studies which administered the questionnaire via telephone. Patients in phone interviews have a tendency to give socially desirable answers in contrast to patients who fill out the questionnaires themselves, possibly due to perceived anonymity [39, 40.] Thus self-administered questionnaires may portray more accurate answers.

Our study also holds several limitations. Firstly, we did not have a baseline measure of HRQoL, thus firm conclusions regarding impaired HRQoL solely influenced by sepsis cannot be made as several factors such as patient characteristics and premorbid quality of life may affect HRQoL[41]. Although the data of sepsis survivors was matched for age and sex, we were unable to match for comorbidities in the general Dutch population as this information was not available. However, previous studies have consistently shown that survivors of other acute illnesses also have a lower HRQoL when compared to the general population[42, 43].

Second, another limitation of our study is the lack of follow-up. Measuring the HRQoL one month after discharge gives the patient little time to recuperate. We chose this short time span, contrary to other HRQoL studies as other sepsis studies focused exclusively on ICU patients[2]. In our study population, 9,4% of patients were admitted to the ICU[12]. Therefore our study population was less severely ill which made us choose shorter follow-up period as they might require less time to recover. This is supported by the median duration of hospitalization of 5 days. Our study shows that even in patients not admitted to the ICU, the effect on PCS is still large, albeit relatively short after admission.

Third, at the time of inclusion of patients in this study, the SEPSIS-2 criteria were still the gold standard, which have nowadays been replaced by the SEPSIS-3 criteria, which would mean that our sample size would have been less if patients were included according to the new criteria as these are less sensitive than the SEPSIS-2 criteria[20]. However, the fact that patients who were included by using the more sensitive SEPSIS-2 criteria still

suffer from decreased HRQoL, underlines the fact that this diminished HRQoL might be even worse if patients were to be included by using the SEPSIS-3 criteria.

Our sample was not completely representative of all patients in our study, as patients that returned the questionnaire had a significantly lower CCI compared to patients that did not return the questionnaire. However, earlier studies (8–10) have already focused on the group of patients most severely ill (those in ICU), and show similar conclusions.

CONCLUSION

In a cohort of sepsis-survivors, 28 days after discharge, we found that HRQoL is considerably lower than the general Dutch population. This effect is most profound on the Physical Component Score. Length of stay and comorbidity, especially heart failure and chronic pulmonary disease, are significantly associated with a lower HRQoL in the physical domain. HRQoL.

REFERENCES

1 Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016; 315(8):801–10. doi: 10.1001/jama.2016.0287. PMID: 26903338; PubMed Central PMCID: PMC4968574.

2 Alam N, Nannan Panday RS, Heijnen JR, van Galen LS, Kramer MHH, Nanayakkara PWB. Long-term health related quality of life in patients with sepsis after intensive care stay: A systematic review. Acute Med. 2017; 16(4):164–9. PMID: 29300794.

3 Fleischmann C, Scherag A, Adhikari NK, Hartog CS, Tsaganos T, Schlattmann P, et al. Assessment of Global Incidence and Mortality of Hospital-treated Sepsis. Current Estimates and Limitations. Am J Respir Crit Care Med. 2016; 193(3):259–72. doi: 10.1164/rccm.201504-0781OC. PMID: 26414292.

4 Kadri SS, Rhee C, Strich JR, Morales MK, Hohmann S, Menchaca J, et al. Estimating Ten-Year Trends in Septic Shock Incidence and Mortality in United States Academic Medical Centers Using Clinical Data. Chest. 2017; 151(2):278–85. doi: 10.1016/j.chest.2016.07.010. PMID: 27452768; PubMed Central PMCID: PMC5310115.

5 Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, et al. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. Intensive Care Med. 2017; 43(3):304–77. doi: 10.1007/s00134-017-4683-6. PMID: 28101605.

6 Winters BD, Eberlein M, Leung J, Needham DM, Pronovost PJ, Sevransky JE. Long-term mortality and quality of life in sepsis: a systematic review. Crit Care Med. 2010; 38(5):1276–83. doi: 10.1097/CCM. 0b013e3181d8cc1d. PMID: 20308885.

7 Hofhuis JG, Spronk PE, van Stel HF, Schrijvers AJ, Rommes JH, Bakker J. The impact of severe sepsis on health-related quality of life: a long-term follow-up study. Anesth Analg. 2008; 107(6):1957–64. doi: 10.1213/ane.0b013e318187bbd8. PMID: 19020144.

8 Cuthbertson BH, Elders A, Hall S, Taylor J, MacLennan G, Mackirdy F, et al. Mortality and quality of life in the five years after severe sepsis. Crit Care. 2013; 17(2):R70. doi: 10.1186/cc12616. PMID: 23587132; PubMed Central PMCID: PMC4057306.

9 Heyland DK, Hopman W, Coo H, Tranmer J, McColl MA. Long-term health-related quality of life in survivors of sepsis. Short Form 36: a valid and reliable measure of health-related quality of life. Crit Care Med. 2000; 28(11):3599–605. https://doi.org/10.1097/00003246-200011000-00006 PMID: 11098960.

10 Nesseler N, Defontaine A, Launey Y, Morcet J, Malledant Y, Seguin P. Long-term mortality and quality of life after septic shock: a follow-up observational study. Intensive Care Med. 2013; 39(5):881–8. doi: 10.1007/s00134-013-2815-1. PMID: 23358541.

11 Groenewoudt M, Roest AA, Leijten FM, Stassen PM. Septic patients arriving with emergency medical services: a seriously ill population. Eur J Emerg Med. 2014; 21(5):330–5. doi: 10.1097/MEJ. 0000000000000091. PMID: 24185258.

12 Alam N, Oskam E, Stassen PM, Exter PV, van de Ven PM, Haak HR, et al. Prehospital antibiotics in the ambulance for sepsis: a multicentre, open label, randomised trial. Lancet Respir Med. 2018; 6(1):40–50. doi: 10.1016/S2213-2600(17)30469-1. PMID: 29196046.

13 Alam N, de Ven PM, Oskam E, Stassen P, Kramer MH, Exter PV, et al. Study protocol for a Multi-centre, Investigator-initiated, Randomized Controlled Trial to Compare the Effects of Prehospital Antibiotic Treatment for Sepsis Patients with Usual Care after Training Emergency Medical Services (EMS) Personnel in Early Recognition (- The Prehospital ANTibiotics Against Sepsis (PHANTASi) trial. Acute Med. 2016; 15(4):176–84. PMID: 28112286.

14 Levy MM, Fink MP, Marshall JC, Abraham E, Angus D, Cook D, et al. 2001 SCCM/ESICM/ACCP/ATS/SIS International Sepsis Definitions Conference. Crit Care Med. 2003; 31(4):1250–6. doi: 10.1097/01.CCM.0000050454.01978.3B. PMID: 12682500.

15 Ware JE Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992; 30(6):473–83. PMID: 1593914.

16 Ware JE Jr. SF-36 health survey update. Spine (Phila Pa 1976). 2000; 25(24):3130–9. https://doi.org/10.1097/00007632-200012150-00008 PMID: 11124729.

17 Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chro nic disease populations. J Clin Epidemiol. 1998; 51(11):1055–68. https://doi.org/10.1016/s0895-4356(98) 00097-3 PMID: 9817123.

18 Ware JE, New England Medical Center Hospital. Health I. SF-36 physical and mental health summary scales: a user's manual. Boston :: Health Institute, New England Medical Center; 1994.

19 Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. Health Econ. 1993; 2(3):217–27. PMID: 8275167.

20 Goulden R, Hoyle MC, Monis J, Railton D, Riley V, Martin P, et al. qSOFA, SIRS and NEWS for predicting inhospital mortality and ICU admission in emergency admissions treated as sepsis. Emerg Med J. 2018; 35(6):345–9. doi: 10.1136/emermed-2017-207120. PMID: 29467173.

21 Walters SJ, Munro JF, Brazier JE. Using the SF-36 with older adults: a cross-sectional community-based survey. Age Ageing. 2001; 30(4):337–43. https://doi.org/10.1093/ageing/30.4.337 PMID: 11509313.
22 Schwartz CE, Andresen EM, Nosek MA, Krahn GL, Measurement REPoHS. Response shift theory: important implications for measuring quality of life in people with disability. Arch Phys Med Rehabil. 2007; 88(4):529–36. doi: 10.1016/j.apmr.2006.12.032. PMID: 17398257.
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(2):R91–8. doi:10.1186/cc2818. PMID: 15025783; PubMed Central PMCID: PMC420036
24 Needham DM, Wozniak AW, Hough CL, Morris PE, Dinglas VD, Jackson JC, et al. Risk factors for physical impairment after acute lung injury in a national, multicenter study. Am J Respir Crit Care Med. 2014; 189(10):1214–24. doi: 10.1164/rccm.201401-0158OC. PMID: 24716641; PubMed Central PMCID: PMC4061900.
25 Brower RG. Consequences of bed rest. Crit Care Med. 2009; 37(10 Suppl):S422–8. doi: 10.1097/CCM.0b013e3181b6e30a. PMID: 20046130.
26 Charafeddine R, Demarest S, Cleemput I, Van Oyen H, Devleesschauwer B. Gender and educational differences in the association between smoking and health-related quality of life in Belgium. Prev Med. 2017; 105:280–6. doi: 10.1016/j.ypmed.2017.09.016. PMID: 28964851.
27 Choo J, Jeon S, Lee J. Gender differences in health-related quality of life associated with abdominal obesity in a Korean population. BMJ Open. 2014; 4(1):e003954. doi: 10.1136/bjopen-2013-003954. PMID: 24464522; PubMed Central PMCID: PMC3902435.
28 Gijsberts CM, Agostoni P, Hoefer IE, Asselbergs FW, Pasterkamp G, Nathoe H, et al. Gender differences in health-related quality of life in patients undergoing coronary angiography. Open Heart. 2015; 2 (1):e000231. doi: 10.1136/openhrt-2014-000231. PMID: 26339493; PubMed Central PMCID: PMC 4555073.
29 Hajian-Tilaki K, Heidari B, Hajian-Tilaki A. Are Gender Differences in Health-related Quality of Life Attributable to Sociodemographic Characteristics and Chronic Disease Conditions in Elderly People? Int J Prev Med. 2017; 8:95. doi: 10.4103/ijpvm.IJPVM_197_16. PMID: 29184646; PubMed Central PMCID: PMC5686916.
30 Davydow DS, Hough CL, Langa KM, Iwashyna TJ. Symptoms of depression in survivors of severe sepsis: a prospective cohort study of older Americans. Am J Geriatr Psychiatry. 2013; 21(9):887–97. https://doi.org/10.1016/j.jagp.2013.01.017 10.1097/JGP.0b013e31825c0aed. PMID: 23567391; PubMed Central PMCID: PMC3462893.
31 van den Boogaard M, Schoonhoven L, Evers AW, van der Hoeven JG, van Achterberg T, Pickkers P. Delirium in critically ill patients: impact on long-term health-related quality of life and cognitive functioning. Crit Care Med. 2012; 40(1):112–8. doi: 10.1097/CCM.0b013e31822e9fc9. PMID: 21926597.
32 Redekop WK, Koopmanschap MA, Stolk RP, Rutten GE, Wolffenbuttel BH, Niessen LW. Health-related quality of life and treatment satisfaction in Dutch patients with type 2 diabetes. Diabetes Care. 2002; 25(3):458–63. https://doi.org/10.2337/diacare.25.3.458 PMID: 11874930.
33 Glasgow RE, Ruggiero L, Eakin EG, Dryfoos J, Chobanian L. Quality of life and associated characteristics in a large national sample of adults with diabetes. Diabetes Care. 1997; 20(4):562–7. https://doi.org/10.2337/diacare.20.4.562 PMID: 9096981.
34 Goldney RD, Phillips PJ, Fisher LJ, Wilson DH. Diabetes, depression, and quality of life: a population study. Diabetes Care. 2004; 27(5):1066–70. https://doi.org/10.2337/diacare.27.5.1066 PMID: 15111522.
35 Tiru B, DiNino EK, Orenstein A, Mailloux PT, Pesaturo A, Gupta A, et al. The Economic and Humanistic Burden of Severe Sepsis. Pharmacoeconomics. 2015; 33(9):925–37. doi: 10.1007/s40273-015-0282-y. PMID: 25935211.
36 Paratz JD, Kenardy J, Mitchell G, Comans T, Coyer F, Thomas P, et al. IMPOSE (IMProving Outcomes after Sepsis)-the effect of a multidisciplinary follow-up service on health-related quality of life in patients postsepsis syndromes-a double-blinded randomised controlled trial: protocol. BMJ Open. 2014; 4(5):e004966. doi: 10.1136/bmjopen-2014-004966. PMID: 24861549; PubMed Central PMCID: PMC4039866.
37 Heyland DK, Groll D, Caeser M. Survivors of acute respiratory distress syndrome: relationship between pulmonary dysfunction and long-term health-related quality of life. Crit Care Med. 2005; 33(7):1549–56. https://doi.org/10.1097/01.ccm.0000168609.98847.50 PMID: 16003061.
38 Dhingra SS, Strine TW, Holt JB, Berry JT, Mokdad AH. Rural-urban variations in psychological distress: findings from the Behavioral Risk Factor Surveillance System, 2007. Int J Public Health. 2009; 54 Suppl 1:16–22. doi: 10.1007/s00038-009-0002-5. PMID: 19363589.
39 Perkins JJ, Sanson-Fisher RW. An examination of self- and telephone-administered modes of administration for the Australian SF-36. J Clin Epidemiol. 1998; 51(11):969–73. https://doi.org/10.1016/s0895-4356(98)00088-2 PMID: 9817114.
40 Lyons RA, Wareham K, Lucas M, Price D, Williams J, Hutchings HA. SF-36 scores vary by method of administration: implications for study design. J Public Health Med. 1999; 21(1):41–5. PMID: 10321858.
41 Oeyen SG, Vandijck DM, Benoit DD, Annemans L, Decruyenaere JM. Quality of life after intensive care: a systematic review of the literature. Crit Care Med. 2010; 38(12):2386–400. doi: 10.1097/CCM.0b013e3181f3dec5. PMID: 20838335.
42 Myhren H, Ekeberg O, Stokland O. Health-related quality of life and return to work after critical illness in general intensive care unit patients: a 1-year follow-up study. Crit Care Med. 2010; 38(7):1554–61. Epub 2010/05/18. doi: 10.1097/CCM.0b013e3181e2c8b1. PMID: 20473149.
43 Gerth AMJ, Hatch RA, Young JD, Watkinson PJ. Changes in health-related quality of life after discharge from an intensive care unit: a systematic review. Anaesthesia. 2019; 74(1):100–8. Epub 2018/10/07. doi: 10.1111/anae.14444. PMID: 30291744; PubMed Central PMCID: PMC6586053.

8

Chapter 9

Sepsis, what's in a name. Correspondence on sepsis

Lancet. 2020 Dec 5;396(10265):1804

T. C. Minderhoud

K. Azijli

P.W.B. Nanayakkara

TO THE EDITOR

Global sepsis incidence and mortality, what's in a name?

On January 18th, the report from Rudd et al1. was published, reporting the global burden of sepsis incidence and mortality. This report was widely circulated and has led to a lot of media attention, demanding more effort to be put into sepsis reduction. However, in our opinion, some questions can be raised if the current message is really helping the patients that are suffering and dying from infectious disease.

First, as the report states, sepsis can only be an intermediate cause of health loss, meaning that a underlying cause will always be present. As is stated in the results, the most common underlying cause of (probable) sepsis was diarrhoeal disease, with 9.21 million cases in 2017. In addition, the most common underlying cause of sepsis-related death was lower respiratory infection with 1.9 million sepsis related deaths in 2017. Both these diseases are already known to cause very large morbidity and mortality seven years ago, in 2013, WHO and UNICEF already introduced the Integrated Global Action Plan for Pneumonia and Diarrhea (GAPPD)2 to end preventable Child Deaths from diarrhea and pneumonia. In this plan, evidence based interventions are promoted such as exclusive breastfeeding, vaccinations, ORS and implementation of simple standardized guidelines for identification and treatment of pneumonia and diarrhoea. Similar (WHO) campaigns already exist targeting newborn health, malaria, HIV and tuberculosis. Of course, some of the interventions for these diseases overlap, such as ameliorating access to health care, but others are rather specific, such as HIV medication, or vector control approaches for reduction of the transmission of malaria

Second of all, the authors state that sepsis requires only the suspicion of infection, referencing the definitions in 1992. However, since then, several studies have raised concern that about one third of the patients with suspected sepsis do not, in fact, suffer from an infectious disease [4,5]. Others have raised concern that the emphasis on the suspicion of sepsis leads to overuse of antibiotics [6,7]. In addition, the benefit of early administration of antibiotics has not been firmly established, because many studies reporting benefit are retrospective and at a high risk of bias, and there is a lack of prospective data[8] .
The only prospective randomized control trial of early antibiotic treatment did not find a beneficial effect on mortality, but its external validity is limited due to a small number of patients with septic shock[9].

Summarizing, some of the root causes of death by infectious disease, such as diarrhoeal disease and malaria are already targeted by evidence based-programs. It is unclear to us how adding the label sepsis to these already known problems will aid the patients suffering from these diseases.
The authors state in their discussion that all sepsis patients, regardless of underlying source have a shared need for acces to basic acute care services, such as timely and appropriate antibiotic administration, microbiology facilities and capacity for organ support. However, how does early appropriate antibiotic treatment benefit a patient with malaria, if we are not thinking about the cause of his or her disease? And since the evidence on optimal timing for antibiotics is conflicting, and concerns do exist about overuse, should we not focus on the underlying disease and establishing evidence based treatment for that specific disease? A disease does not change by changing the name. The way forward for better treatment in cancer medicine turned out to be precision (molecular) medicine, and not research targeting all cancer patients as a collective. As Shakespeare pointed out 'That which we call a rose, by any other name would smell as sweet'. Let us just call the disease by its name, and acknowledge the effort that already has been made on targeting diarrheal disease, pneumonia, malaria and HIV, and what still needs to be done for other infectious diseases.

REFERENCES

1 Rudd et al, Global, regional and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study.
2 World Health Organization, The United Nations Children's Fund (UNICEF) 2013, End preventable deaths, Global Action Plan for Prevention and Control of Pneumonia and Diarrhoea. Accessed through on 20.1.2020 https://www.who.int
3 Bone RC, Balk RA, Cerra FB,et al. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. Chest. 1992; 101: 1644-1655
4 Klein Klouwenberg PM, Cremer OL, van Vught LA, et al. Likelihood of infection in patients with presumed sepsis at the time of intensive care unit admission: a cohort study. Crit Care 2015;19:319.
5 Minderhoud TC, Spruyt C, Huisman S, et al. Microbiological outcomes and antibiotic overuse in emergency department patients with suspected sepsis. Neth J Med 2017;75:196-203.
6 Singer M. Antibiotics for sepsis: does each hour really count, or is it incestuous amplification? Am J Respir Crit Care Med 2017;196:800-2.
7 Klompas M, Calandra T, Singer M. Antibiotics for sepsis-finding the equilibrium. JAMA 2018;320:1433-4
8 Michiel Schinkel, Rishi S. Nannan Panday, W. Joost Wiersinga, Prabath W. B. Nanayakkara. Timeliness of antibiotics for patients with sepsis and septic shock. Journal of Thoracic disease, Vol 12, Supplement 1 (February 2020):
9 Alam N, Oskam E, Stassen PM, et al. Prehospital antibiotics in the ambulance for sepsis: a multicentre, open label, randomised trial. Lancet Respir Med 2018;6:40-50.

Summary and general discussion

Chapter

10

Summary and general discussion

SUMMARY

Who has sepsis? Of all patients in the emergency department, approximately 20% present with a serious infection, and approximately 3% present with sepsis.[1] Even if these numbers might be different between countries and settings, it immediately makes it clear that a physicians in acute care will see many more patients with an infection, without direct threat to life, than patients with sepsis. How do we find out which patients has sepsis? In this thesis, we started out by studying the components of sepsis. In part one, we studied the accuracy of risk stratification scores, using organ failure to predict threat to life. In part two we studied how accurately we identify infections in the ED. And in part three we focus on the bigger picture; what is the impact of severe infections on our patients' lives.

Part one: Identifying threat to life

Before starting the diagnostic trajectory in a patient with suspected infection in the ED, we gauge if the situation of the patient is serious or not. If we think the patient's life is in danger, it influences our diagnostic and therapeutic strategy. For example, if a patient has a very severe pneumonia, we give more extensive antibiotic coverage than if the patient has pneumonia but is up and about ('walking pneumonia'). But how do we know the patient's life is in danger? For this purpose, all kind of risk stratification scores have been designed. Some scores were designed for specific diseases. The best example is the CURB-65. It has been extensively studied in pneumonia and is incorporated in guidelines and daily practice. However, to avoid having to use several different disease specific scoring systems, generic early warning scores were designed. These scores mostly used aggregated data. Points are allocated in a weighted manner, based on the derangement of a predetermined set of patient vital signs from an agreed "normal" range. The sum of the allocated points is aggregated and used to indicate a patient's severity of illness. Starting in the zero's, several aggregated early warning scores were designed, refined and studied in different settings, and in 2015 over 36 EWS existed with variable success rates**. In chapter 2** we evaluated the performance of 24 risk stratification scores that are used in the Emergency Department. The outcomes that were most frequently used were 30 day mortality, ICU admission or a composite of both. We studied overall performance, and performance in patients with suspected infection or pneumonia. For the outcome of 30 day mortality in the overall ED population NEWS performed best for most outcomes, although in patients with infections, the MEDS performed better, with an AUROC varying from poor to good 0.674 – 0.82. Since qSOFA was only introduced in 2016, and this review was performed in 2017, limited evidence was available to assess its acuity in the review in 2017.

Therefore we designed a prospective study to compare the performance of qSOFA to NEWS, MEWS and SIRS in the emergency department, in patients with suspected infection. This study is described in **chapter 3**. The entry criteria for this study were broad, since we wanted to address the applicability in the whole ED population with suspected infection, not just the sickest subset. We compared the generic scores of NEWS (≥5) and MEWS (≥3) and qSOFA ≥2, and also included SIRS (≥2,) as a comparison, to predict 30 day mortality. The results, described in chapter 3 indicate that the NEWS score showed favourable results for predicting 30-day mortality with an AUROC of 0.740 (95% Confidence Interval 0.682-0.798). NEWS≥ 5 also showed the highest sensitivity of 75,8% for predicting mortality with a specificity of 67,4%. In comparison with NEWS, qSOFA had a lower AUROC of 0.689 with a very low sensitivity 17.7 (9.2-29.5) at the conventional cut-off ≥2.

However, the differences between EWS systems are small, and the ease of calculation and implementation should be taken into account when choosing the appropriate approach. Just like every scoring system, it is important to know the limitations of the system, when interpreting the values. For example, the MEWS, which is implemented in Amsterdam UMC, is known to be very sensitive for changes in respiratory rate, leading to frequent false alarms in patients with underlying respiratory disease such as COPD. Separate scoring systems have been developed for specific groups, like CREWS (Chronic Respiratory Early Warning Score) for lung patients and MOEWS (Modified obstetric early warning scoring system). However, in the ED it is much easier to implement 1 score, than to implement 10 scores for 10 specific patient groups. However, in a respiratory ward, the implementation of a score such as CREWS will be more suitable and will lead to less false-alarms. We have to balance between ease of use and acuity. Therefore it is important to keep in mind that the sensitive cut-offs that we propose in the article are suitable for an ED environment, but will lead to many false alarms.

Summarizing, since the overall acuity of scoring systems is only moderate, we will inevitably be wrong about risk stratification in quite a large number of patients. How many false alarms can we accept to find one critically ill patient? In the ED environment, we choose to err on the side of caution, and we therefore searched for a sensitive cut-off. The inherent downside of this approach is that many patients will classified as critically ill, when their situation is really less severe.

Part two: Diagnosing infections in the Emergency Department

At the surface, diagnosing infection seems easy. However, many conditions can mimic infections. For example pulmonary embolism and pneumonia can both present with dyspnea, elevated temperature and tachycardia, but require very different treatments. Pancreatitis often causes abdominal pain and fever, but is not usually caused by an infection. On the other side is that many patients have atypical presentation of infections, especially those with an altered immune system. With the implementation of the surviving sepsis campaign, the perceived impact of missing an infection changed. Since the Kumar study [2], the message was spread that every hour that antibiotics were delayed the mortality would increase with 7,6%. It became common practice to start broad spectrum antibiotics early if any infection was suspected, leading to concern about the increased use of antibiotics and the associated rise in antimicrobial resistance[3]. In 2012 we wanted

[1] Wang HE, Jones AR, Donnelly JP. Revised National Estimates of Emergency Department Visits for Sepsis in the United States. Crit Care Med. 2017;45(9):1443-1449. doi:10.1097/CCM.0000000000002538

[2] Kumar et al Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. Crit Care Med. 2006 Jun;34(6):1589-96. doi: 10.1097/01.CCM.0000217961.75225.E9. PMID: 16625125.

[3] Fitzpatrick F, Tarrant C, Hamilton V, et al Sepsis and antimicrobial stewardship: two sides of the same coinBMJ Quality & Safety 2019;28:758-761.

to evaluate the surviving sepsis campaign in the Albert Schweitzer Hospital. We wanted to know how many patients that we treated for suspected sepsis with antibiotics, actually had sepsis. Because we studied patients that were given antibiotics for (suspected) sepsis, we wanted to answer the underlying question; how many have objective evidence of a bacterial infection? In **chapter 4** we describe how we classified patients in to microbiologically proven infection, probable bacterial infection, and absence of bacterial infection. In about one third (36%) of the patients, bacterial infection was confirmed, and in another 35 %, bacterial infection was likely. In the group without bacterial infection, the most prevalent condition was exacerbation of COPD. Antibiotics were discontinued in 32% of patients without bacterial infection, mostly in patients with an alternative (viral) diagnosis. The uncertainty in diagnosing infections is corroborated by studies targeting diagnostic uncertainty in urinary tract infections and pneumonia. In pneumonia, studies have found evidence that of patients admitted with pneumonia, this diagnosis was maintained in only ~ 60 % of the patients [4,5.] For urinary tract infection, only 40% of patients treated in the ED for presumed UTI had clinical or microbiological evidence of this infection.[6] Even in intensive care unit, a study found that 13% of patients treated for suspected sepsis had a post-hoc- likelihood of infection of none, and another 30% had a classification of only probable [7.]
Another important finding in this chapter is that only 26 of the 269 treated patients had hypotension or shock. It is key to remember that the 7,6% of mortality found by Kumar was based on a patient series that included only patients with septic shock. Our evaluation of practice in a single-center indicate that despite this limitation, early antibiotic treatment was used extensively in patients with a much milder clinical picture. Although the findings in this chapter are in line with other studies, its validity is limited because the data are relatively old, and from a single center. Both might have considerable influence on antibiotic practice and duration.

In **chapter 4** diagnosing a viral disease was a reason to discontinue antibiotics in a significant subset of patients. We hypothesized that rapid confirmation of a viral infection can be a valuable addition in the diagnostic process. However, how good are we in recognizing viral infections, such as influenza in the ED? During the 2017-2018 epidemic of influenza, no clear guidance was available which patients needed testing before hospital admission. In Amsterdam UMC, location VUmc, a local protocol was implemented to guide testing in patients that need hospitalization. In **chapter 5** the protocol was evaluated. The protocol was based on the case-definition of influenza of the National institute for Public Health and Environment (RIVM- LCI richtlijn). We studied if clinical symptoms could be used to predict the outcome of influenza. The typical symptoms of influenza, such as myalgia were present in only a small number of patients with influenza. Complaints that are less specific, such as coughing and fever were found often in influenza patients, but are also frequently found in other conditions. In our patients, clinical symptoms could not reliably separate patients with influenza from patients without influenza. Even though we combined the symptoms in several ways, total accuracy was around ~65%. Part if this can be explained by selection bias, since patient that need hospitalization are often comorbid and older, which is associated with atypical presentations of diseases. However, since this is exactly the population that needs admission, it is important to test for influenza in all patients with respiratory complaints that need hospitalization, during an influenza epidemic. In patients with fever, without an obvious explanation, testing should also be considered. While working on this piece, the updated 2018 IDSA guidelines were published in December 2018. In this guideline it is recommended to test any patient with respiratory symptoms that is hospitalized. The recommendation is based on research with similar findings, and it emphasizes that in patients needing hospitalization, clinical signs are unreliable.
However, even if we confirm that a patient has a viral infection, a bacterial co-infection is often suspected. We decided to study bacteremia and procalcitonin in a population of patients with suspected infection, during a viral epidemic. In these patients, the treating physician ordered both a blood culture and a rapid viral test. We hypothesized that in these patients, the treating clinician was facing diagnostic uncertainty. In addition, we thought that if patient has a proven viral infection, this lowers the prior chance of also having a bacteremia, based on the one-disease principle. In **chapter 6** we describe the results of our observational study. We found a significant lower percentage of bacteremia in patients with a viral infection (3.8% bacteremia) versus patients without a viral infection (11.7 % bacteremia). In the patients with a viral infection, all that had bacteremia also had a procalcitonin ≥ 0,5 µg/L. Therefore, in patients with a positive viral test, a procalcitonin of < 0,5 µg/L ruled out bacteremia. However, in patients without a viral infection, the cut-off ≥ 0,5 µg/L was unacceptable due to limited sensitivity. With a lower cut-off <0,25 µg/L, the sensitivity improved. All patients with a false negative procalcitonin of <0,25 µg/L had a short duration of symptoms. These results help us understand how we can use procalcitonin in Emergency Department. If we integrate the result with our other test results, we can adapt the cut-off and avoid overuse of antibiotics in patients with an alternative explanation of their symptoms. However, in patients with a short duration of symptoms, and in patients in whom bacterial infection is the most likely diagnosis, procalcitonin should be interpreted with caution.

To confirm bacterial infection, blood cultures are a vital diagnostic tool. The surviving sepsis guidelines, next to mandating early antibiotic treatment and fluid resuscitation, also suggested taking blood cultures in any patient with suspected sepsis. However, only around 10-12 percent of blood cultures performed in the Emergency Department are positive. In **chapter 7** we explored how machine learning could be used to design an algorithm that would estimate the probability of bacteremia (true positive cultures).

[4,5] Sikka R, Tommaso LH, Kaucky C, Kulstad EB. Diagnosis of pneumonia in the ED has poor accuracy despite diagnostic uncertainty. Am J Emerg Med. 2012 Jul;30(6):881-5. doi: 10.1016/j.ajem.2011.06.006. Epub 2011 Aug 19. PMID: 21855251.
[6] M. Kanwar, R. Brar, R. Khatib, M.G. Fakih Misdiagnosis of community acquired pneumonia and inappropriate utilization of antibiotics: side effects of the 4-hour antibiotic administration rule CHEST, 131 (2007), pp. 1865-1869
[7] Shallcross, L.J., Rockenschaub, P., McNulty, D. et al. Diagnostic uncertainty and urinary tract infection in the emergency department: a cohort study from a UK hospital. BMC Emerg Med 20, 40 (2020).
Klein Klouwenberg PM, Cremer OL, van Vught LA, Ong DS, Frencken JF, Schultz MJ, Bonten MJ, van der Poll T. Likelihood of infection in patients with presumed sepsis at the time of intensive care unit admission: a cohort study. Crit Care. 2015 Sep 7;19(1):319.

We found two machine learning models using information that would typically be available at the end of an ED visit. Both models show comparably good performance in predicting BC results with AUROCs of 0.77 (95%-CI = 0.73-0.82) in the test sets. The models can identify patients in the ED with low risk for bacteraemia. Since we have few tools available in the ED to help us estimate the chance of bacteremia, this machine learning model might be a helpful addition. However, further refinement, testing and incorporation in clinical practice is needed.

In summary, in **part 2** we learned that diagnosing infection in the emergency department is hard and has a high degree of diagnostic uncertainty. Rapid viral tests and procalcitonin can be used to reduce diagnostic uncertainty. We need to be incorporate them in our diagnostic reasoning, in order to reduce antibiotic overuse. Machine learning algorithms can help us estimate the probability of bacteremia, which might be a useful adjunct, but further studies are needed for clinical translation.

Part three: Sepsis and the bigger picture

For patients, infections can have serious and lasting impact on their functioning and the impact on the quality of life can be enormous. The patients that were included in the Phantasi trial, were sent a survey of the quality of life (SF36), 1 month after discharge from the hospital. The results, discussed in **chapter 8**, show decreased scores of quality of life in sepsis survivors compared to matched controls. The absolute decrease was larger on the physical component score, than on the mental component score. Although this study was interesting, the analysis had quite a few limitations. First of all, the comparison cohort was matched only on age and sex, not on comorbidity. Secondly, it might have been more interesting to compare our patients to other patients that were recently hospitalized, or ideally, to follow our patients over time. Unfortunately, these data were not available. The long-term sequelae of sepsis outside the ICU have been poorly studied. A study in older people found that severe sepsis was independently associated with substantial cognitive impairment and functional disability among survivors[8]. A review on recovery after sepsis[9] pointed out we need to provide better follow-up, to detect functional deterioration and to avoid readmissions. However, we still have much to learn in this area, since long-term complaints after sepsis are poorly understood.

In **chapter 9**, the global burden of sepsis is discussed. According to the report in Lancet an estimated 48·9 million (95% confidence interval 38·9–62·9) incident cases of sepsis were recorded worldwide in 2017[10]. About 1 in 5 deaths globally, was ascribed to sepsis. In a letter to the editor, we discuss if we really help our patients if we attach the word sepsis to these deaths. Sepsis is always an intermediate cause of death, since the root cause is an infection. The sepsis guidelines focus around early recognition of sepsis, early antibiotic treatment and fluids, and measurements. Most of the mortality caused by infection in the article is caused by pneumonia and diarrhoea in low-middle income countries. The solution to these problems does not lie with implementation of the SSC. Rather, it lies in combatting malnutrition, assuring access to clean drinking water, better access to preventative measures (malaria nets, vaccinations) and better access to healthcare. Do we need a new name to properly address these problems? The solution to most cases of sepsis still relies on accurately diagnosing and treating the underlying infection. By changing the name of the disease, the nature of the disease is not changed, nor is the treatment.

Overall, as we have seen in the chapters before, recognizing sepsis is difficult because the items that sepsis consists of, are difficult to measure. Since diagnosing infection has quite a poor accuracy and diagnosing organ dysfunction also has limited accuracy, ad ding them up can only lead to an imprecise result. As long as we cannot reliably measure our input variables (infection and organ dysfunction leading to threat to life), what does the outcome of the equation (sepsis) really mean to us? This question is the starting point for the general discussion.

10

[8] Iwashyna TJ, Ely EW, Smith DM, Langa KM. Long-term cognitive impairment and functional disability among survivors of severe sepsis. JAMA. 2010 Oct 27;304(16):1787-94. doi: 10.1001/jama.2010.1553. PMID: 20978258; PMCID: PMC3345288.

[9] Prescott HC, Angus DC. Enhancing Recovery From Sepsis: A Review. JAMA. 2018 Jan 2;319(1):62-75. doi: 10.1001/jama.2017.17687. PMID: 29297082; PMCID: PMC5839473.

[10] Rudd, Kristina E et al.Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study The Lancet, Volume 395, Issue 10219, 200 - 211

GENERAL DISCUSSION

Sepsis and its place in the Emergency Department

From the summary it has become clear that measuring sepsis and its components in the Emergency setting is difficult and imprecise. If we cannot reliably diagnose sepsis in the emergency department, then how should we approach the syndrome of sepsis? Just because the syndrome of sepsis is difficult to recognize in the ED, it does not mean that the concept is not important. The interplay between infection and sepsis can be envisioned as a large mountain, with perpetual snow on top. The snow on top represents the sepsis cases. As we ascend the mountain, the threat to life increases. Some parts of the mountain lie in the shadows which is metaphorical for the (dys)regulated host response. On the slopes, it is important to figure out which path (source of infection) the patient was on before she ended up on the slope. Because these patients generally have little organ dysfunction, the organ dysfunction is not the first priority. Thus, in the emergency department, we spend way most of our time finding the path across the mountain slopes than and we spend time in the snow. However, ICU physicians spend most of their time in the snow, and do not always have the time to see which path led the patient up the mountain. As we enter the snow, the priority shifts. In the ICU, the organ dysfunction is usually so severe that the first priority lies with stabilizing the patient, instead of finding the exact source of infection.

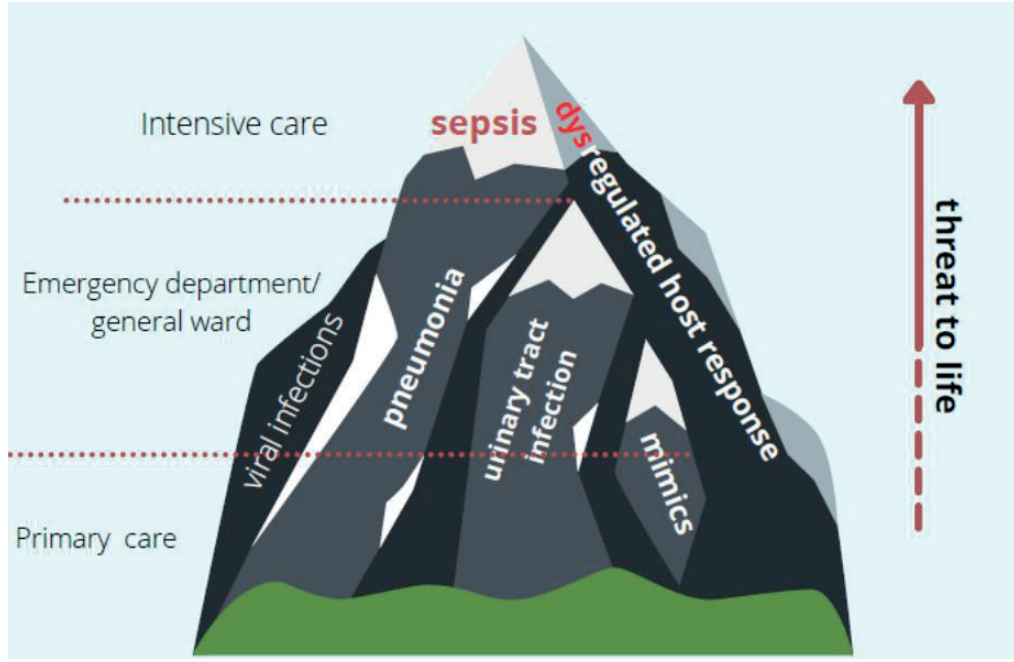
How should the concept of sepsis be approached in the Emergency Department? Should we focus on the snow, or on the mountain underneath? In the next section, we will propose how the concept of sepsis can be approached from the research, clinical and patient perspective.

Sepsis and research in the Emergency Department

As a researcher, the problem of the concept of sepsis is that diagnosis and severity are intertwined in the sepsis definition. For research purposes, a pragmatic approach is to separate diagnosis from severity. In many condition that we treat in the Emergency Department, diagnosis and severity are separated. For example, a patients has a pulmonary embolism (diagnosis usually by CT-PA) , and the severity is then classified by a validated severity score (PESI or sPESI). Both diagnosis and severity can be related to prognosis, but to create cohorts for treatment, we need to know if the patient has the diagnosis of interest, before stratifying for severity.

Given the degree of diagnostic uncertainty that was described in this thesis and other literature, our first priority should be to improve the diagnosis of infection in the Emergency Department. To reduce diagnostic uncertainty, we need to study new, fast microbiological techniques and translate them from bench to bedside. We have to optimalize the use of biomarkers by integrating the result in clinical reasoning. We might be able to use machine-learning to support our decision making.
In other words, we need to study suspected infection in the emergency department, before we can study sepsis. We need innovative study designs that combine diagnostic markers with clinical reasoning to improve the performance and clinical applicability of diagnostic tests.

Figure 10.1, concept of sepsis visualized as a mountain



On severity: selectivity versus generalizability

Most of the patients with infections that need admission will be admitted to general wards. In both chapter 2 and 4, around ~ 90% of the patients that needed admission were admitted to general wards. If we select only the sickest patients for sepsis studies, we lose a lot of information. It can result in poor generalizability of study results to other Emergency Department patients. And severity in the Emergency Department is not a static observation. Some patients might show great improvement in the ED, following some fluids and/or antibiotics, whereas others might deteriorate despite the same treatment. Studies on severity in infected patients need to include patients from the whole spectrum of severity, so we can study the differences between mild infection and severe infections with organ dysfunction.
In conclusion, we need to separate diagnostic studies from prognostic (severity) studies. Improvement of the diagnostic process is much needed to reduce diagnostic uncertainty in patients with suspected infection the Emergency Department.

THE CLINICAL PERSPECTIVE: SUSPECTED INFECTION AND EARLY ANTIBIOTIC TREATMENT

A specific diagnosis trumps the sepsis syndrome

As a clinician, we need to decide if the patients clinical picture is compatible with sepsis and what the appropriate treatment is. But how is sepsis different from other frequently encountered infectious diseases in the Emergency Department? If a patient comes in with sepsis, caused by pneumonia, which of the two is most important for the clinician? Pneumonia causes 30-50 % of the cases of sepsis and the second most prevalent cause

is urinary tract infection. That means that any study with sepsis patients, has considerable overlap with the patient population of pneumonia studies.
Consider the following situation: A patient comes in with shortness of breath, fever, and an infiltrate on x-ray. The blood pressure is low and the respiratory rate is high. The patient fulfills the criteria for severe pneumonia (CAP III), but also for septic shock (SEPIII). You have 2 guidelines; one on pneumonia, and one on sepsis. What is the best way to proceed? A brief comparison between the recommendations in the guidelines can be seen in box 10.1.
Scientifically, we can base our choice on the quality of evidence in the guidelines, and how much the research included in the guideline is applicable to our patient. One could also approach the problem in a more philosophical way. If pneumonia is one of the underlying causes of sepsis, why would we classify the patient in a syndromic term of sepsis if we can use the more specific term? We might also approach the problem on an interpersonal level. Does communicating the diagnosis of sepsis to the patient, family or team, confer benefits compared to discussing the diagnosis of pneumonia?

Box 10.1 Comparison of guideline recommendations

topic	SSC 2016	IDSA Pneumonia Guideline 2018
Risk stratification	-	Pneumonia Severity Index/ CURB 65
Microbiology	Blood + other appropriate sites	Blood cultures, sputum cultures (severe cases)
Antibody testing	No guidance	Urinary Pneum AG testing recommended (severe cases)
Influenza testing	No guidance	Recommended if high influenza activity
Fluids	30 ml/kg	No guidance
Antibiotic timing	< 1 hr of sepsis recognition	No guidance (2007: in the ED)
Antibiotic duration	7-10 days	5 days

In pneumonia, the choice seems easy. The pneumonia guidelines recommendations are widely accepted and the patients included in pneumonia trials are representative of the patient that we just described. If a disease specific guideline exists, this seems preferable to using a guideline based on a syndrome. Especially since the SSC guidelines are the topic of much discussion. And on the interpersonal level, pneumonia is a well-known entity that is probably better known than the concept of sepsis, since 4 out of 5 people in the Netherlands are not familiar with the term of sepsis.

The sepsis concept is beneficial as a red flag
However, this situation is different if a patient comes with a fever, low blood pressure, and does not have an obvious source of infection. In these patients, after recognizing the syndrome of sepsis, we can apply the sepsis guidelines recommendations. It provides the team with a shared framework about the nature of the condition, the severity of the condition and the urgency of treatment, even though the exact causes is unknown at that time. And, for the patient, and those around the patient, it can create clarity, given that they are familiar with the term of sepsis.
Sepsis, therefore, is a useful term to serve as a red flag in the Emergency department for patients with diagnostic uncertainty in a time-critical situation. However, it should be clear that sepsis in itself is not a diagnosis, but a broad term to indicate a serious threat to health, possibly caused by infection. After clarification of the cause, the disease episode should ideally be classified by the underlying cause, although in a minority, the underlying cause will not be found. These patients are a specifically interesting group for diagnostic research.

Early antibiotics and antibiotic overuse
If a clinician is faced with suspected sepsis, the first treatment that comes to mind is early antibiotic treatment. The amount of time that we take to assess the diagnosis and the severity has been influenced by the golden hour and sepsis guideline recommendations. However, it is important to realize that the evidence on early antibiotic treatment is strongest for patients with hypotension[11]. The perceived risk of missing an infection seems to be conflated, whereas the risks of overtreatment are poorly recognized[12]. For patients without hypotension, we have to weigh the benefit and risk of early antibiotic treatment. Estimated rates of unnecessary antibiotic use at the emergency department are between 30-60%[13] and it has been described as the most preventable cause of antibiotic resistance. Often, to avoid risk, it is decided to start antibiotics in the ED. It is assumed that later in the admission, the indication will be reviewed and de-escalation will be considered.

[11] IDSA Sepsis Task Force. Infectious Diseases Society of America (IDSA) POSITION STATEMENT: Why IDSA Did Not Endorse the Surviving Sepsis Campaign Guidelines. Clin Infect Dis. 2018 May 2;66(10):1631-1635. doi: 10.1093/cid/cix997. PMID: 29182749; PMCID: PMC6927848.
[12] Singer M, Inada-Kim M, Shankar-Hari M. Sepsis hysteria: excess hype and unrealistic expectations. Lancet. 2019 Oct 26;394(10208):1513-1514. doi: 10.1016/S0140-6736(19)32483-3. PMID: 31657730.
[13] May L, Cosgrove S, L'Archeveque M, Talan DA, Payne P, et al. 2013. A call to action for antimicrobial stewardship in the emergency department: approaches and strategies. Annals of emergency medicine 62:69-77.e2

However, the best approach of de-escalation of antibiotics in patient on wards is currently unclear. Even if cultures come back negative, antibiotics are often not stopped. An anthropological study found that teams are reluctant to overturn decisions on antibiotic treatment, and will continue antibiotics once they have been started[14] .

In conclusion, for clinicians, confirming a specific diagnosis like pneumonia gives us more information and more specific guidelines than the diagnosis of sepsis. The sepsis guideline could delineate more clearly for which patient groups other guidelines already exist. However, the sepsis concept has advantages in patients without a clear diagnosis. It can serve as a red flag and it creates a shared framework about the nature and urgency of the condition. However, if we can re-classify the patient to a more specific diagnosis, we should. If we do not focus on improving our diagnostic accuracy and antibiotic de-escalation, we have little in hands to battle the overuse of antibiotics.

Box 10.2 Sheila's story, continued (present)

Sheila comes to the emergency department with generalized weakness. In the ED, Sheila has a low blood pressure of 80/40mmHg and she is not quite alert. Her MEWS is 4 and her qSOFA score is 2. The sepsis team is alerted and she is rapidly assessed. Her blood tests show elevated creatinin (135ug/L) levels indicating an acute kidney injury, and mildly elevated inflammatory markers (CRP 62mg/L).

She is treated with fluids and empirical antibiotics in the Emergency Department. Her blood pressure improves after 1L of NaCL 0,9% and she is admitted to the ward for further treatment. Her medication (NSAID & perindopril) is stopped. Although stool cultures and viral PCR were ordered, they were never turned in. Twentyfour hours after admission, Sheila is feeling much better and her kidney function has improved.

She asks if she can go home, to her husband. However, she is still on IV ceftriaxone, and it is decided to keep her in the hospital for another 24 hours, awaiting the blood culture results. She is discharged after 3 days, on oral ciprofloxacin for a total duration of 7 days. She is advised to speak to her primary care physician about her medication.

One month later, Sheila develops a urinary tract infection, caused by a ciprofloxacin resistant E.coli. She develops pyelonefritis and needs to be admitted for 7 days for IV antibiotic treatment, because there are no suitable oral options.

[14] E Charani, R Ahmad, T M Rawson, E Castro-Sanchèz, C Tarrant, A H Holmes, The Differences in Antibiotic Decision-making Between Acute Surgical and Acute Medical Teams: An Ethnographic Study of Culture and Team Dynamics, Clinical Infectious Diseases, Volume 69, Issue 1, 1 July 2019, Pages 12–20

The patient perspective

To understand the patient perspective, we can only look in, from the outside. Patient experiences are available on the www.sepsisnet.nl and through conversations with patients it becomes clear that patients are often taken by surprise by the quick deterioration in sepsis. Awareness of sepsis is limited since many people have only encountered mild or self-limiting infections. For communication purposes, an overarching term such as sepsis can be used to create awareness and avoid patient/doctor delays. It can also ensure that patients have a common language when voicing their concerns with health professionals. In addition to this, public knowledge on the topic of sepsis might be helpful to draw attention for the long-term complaints that patients experience after sepsis. However, we must make clear that sepsis is not a diagnosis, but rather a syndrome consisting out of multiple diseases. As a patient, you want to get the treatment you need, on time. But for sepsis, it is both unclear how to diagnose it, and it is unclear how to treat it. Sheila's story continues in box 10.2.

CONCLUSION AND FUTURE IMPLICATIONS

For research in the Emergency Department, we need to focus on improving the diagnostic accuracy of suspected infection, before we can accurately identify sepsis. Structurally searching for viral diseases might improve diagnostic accuracy. Molecular testing and biomarker need to be embedded in clinical reasoning.

For clinicians, sepsis is a useful construct, to recognize patients with a time-critical condition. It also has benefits in communication. If our patients know that sepsis is a severe condition that needs timely diagnosis and treatment, we can avoid delays in treatment. But we have to decide the hierarchy between other infectious diseases and sepsis. Does a specific diagnosis trump the sepsis diagnosis?

We can parallel the use of the term sepsis to the term cancer. If a patient comes in, with suspected malignant disease, we try very hard to figure out precisely what type cancer the patient suffers from. If a precise diagnosis can be made, we should seek it. Therefore it is not uncommon to delay (chemo)therapy to await diagnostic testing. Precise categorisation based on histopathological characteristics and cytogenetics, has proved to be a very useful way to guide therapy and estimate prognosis. In some patients, despite the effort that is made, we cannot categorize their illness, and these patients are often grouped together, for example ACUP (adenocarcinoma of Unknown Primary). Despite the common characteristics that cancer types have in common, the focus of therapy has not been to find a cure for all, but rather to find a precise treatment that targets the specific events in specific tumor types. We should learn from this approach in sepsis, and rather than continue to find a treatment for all, we should focus on meaningful subcategories. These should not only be based on convention (eg location) but also on phenotypes, biomarkers and -omics data.

Patients need timely treatment, but it should also be the right treatment, for the right condition. Overuse of antibiotics may cause harm on a personal level, and has greater repercussion on microbiological resistance. To achieve this, we need:

- A shared research agenda and strategy to improve the diagnostic approach of suspected infections in the Emergency Department setting.

- A multidisciplinary approach, with representation of (acute) internists, ED physicians, infectious disease specialists, microbiologists, and critical care physicians.

- A collaborative, nationwide database of patients with suspected infections in the Emergency department. Many (academic) hospitals already have sepsis databases and biobanks of patients with suspected sepsis in the Emergency Department. If we collaborate, we can improve diagnostic strategies of both common and rare infections.

- Agreement on the hierarchy of sepsis and the underlying infections. If we can diagnose a specific condition, patient should be included in specific studies, and patients with unclear diagnosis might be grouped together.

Maybe in the future, Sheila's journey will look like this:

Sheila comes to the emergency department with generalized weakness. In the emergency department, Sheila has a low blood pressure, 80/40 mmHg and a high pulse of 110 beats per minute. She is not quite alert. Her MEWS is 4 and her qSOFA score is 2. The sepsis team is alerted and she is rapidly assessed. Her blood tests show evidence of an acute kidney injury, as well as signs of mild inflammation (normal leucocyte count with a moderately elevated CRP).

A rapid biomarker panel showed a low procalcitonin value but elevated biomarkers consistent with acute kidney injury. Point-of-care ultrasound was compatible with hypovolemia, with good cardiac function. During her ED stay, Sheila is closely monitored. After 1 hour, her response to fluids is re-assessed and the results of the rapid testing are discussed. After 1 L of fluids, Sheila is alert and her blood pressure has increased to 100/60 mmHg, which means that her qSOFA score is now 1 and MEWS is 3.

The sepsis team decides to withhold antibiotics, but to repeat the procalcitonin measurement in 6 hours if her MEWS is still ≥3. Fortunately , 24 hours later, Sheila feels much better. She says she would like to go home, to her husband and son. Her MEWS in the morning is 0. Her rapid bacterial PCR was negative and her kidney function is improving. Her final diagnosis is dehydration and acute kidney injury due to gastroenteritis and medication. The diagnosis is discussed with Sheila and her intravenous fluids are stopped. She will be allowed to go home if she is able to drink enough and if she mobilizes safely. A follow-up appointment is arranged for 1 week after discharge to discuss her recovery and her medication. At the follow-up appointment, the admission is discussed with Sheila, and the effect of her medication on her kidney function is explained to her. Further follow-up on the blood pressure is arranged through the primary care physician.

Appendices

About the Author

Tannetje Cornelia(Tanca) Minderhoud started medical school in 2001, at the University of Utrecht. During her studies, her interest in infectious diseases was raised during internships in the Harbour hospital in Rotterdam and in the hospital for tropical infectious diseases (Fundação de Medicina Tropical) in Manáus. In her last year, she performed research on a rat-model for vitamin K absorption in bile-deficient conditions. After her graduation, she continued to work as a research-assistant to complete some of the studies she had worked on. Although she enjoyed the research and was interested in a PhD trajectory at the time, she decided that she wanted to pursue clinical training first. Her idea at the time was that as a specialist, you can shape the PhD to address the knowledge gaps you have encountered as a clinician.

After working as a resident in the Harbour Hospital and in the Haga hospital (The Hague), she started her training for internal medicine in May 2010 in the Albert Schweitzer Hospital (ASZ) in Dordrecht. Having encountered patients with severe infections during her earlier work experience, she was very interested in the Surviving Sepsis Campaign which was actively promoted in the Albert Schweitzer Hospital. In the fall of 2011, she started a research project (supervised by dr. Mark-David Levin) to evaluate the outcomes of the sepsis screening that was implemented in the ASZ. In 2014 she started subspecialty training for acute medicine in the Erasmus MC and the Sint Franciscus Gasthuis in the Rotterdam region. She continued to work on the sepsis-project and finished the manuscript in 2016, which was finally published in 2017.

In January 2017 she continued her education in Amsterdam, in the VU medical centre (VUmc). Under supervision of prof. dr. Prabath Nanayakkara she completed her specialty training in June 2017. Now a newly minted specialist, she quickly adjusted and found her place in the research group. Aside from her research, she acquired skills as a teacher (Basis Kwalificatie Onderwijs), implemented point-of-care ultrasound in the internal medicine department and played a key role in the sub-specialty training for acute fellows in location VUmc.

Tanca is married to Jan and has 4 children, Jonne (2010), Krijn (2012), Doris, (2016) and Idris (2020). She enjoys reading, DIY and sewing. She is also a Twitter enthusiast and is never shy to share her (feminist) opinion on the current state of affairs.

Publications in this thesis

Azijli K & **Minderhoud TC**, C. de Gans, A. Lieveld, Nanayakkara PWB Optimal Use of Procalcitonin to Rule Out Bacteremia in Patients with Possible Viral Infections, accepted for publication JACEP Open.

Boerman AW, Schinkel M, Meijerink L, van den Ende ES, Pladet LC, Scholtemeijer MG, Zeeuw J, van der Zaag AY, **Minderhoud TC**, Elbers PWG, Wiersinga WJ, de Jonge R, Kramer MH, Nanayakkara PWB. Using machine learning to predict blood culture outcomes in the emergency department: a single-centre, retrospective, observational study. BMJ Open. 2022 Jan 4;12(1):e053332. doi: 10.1136/bmjopen-2021-053332. PMID: 34983764; PMCID: PMC8728456.

Azijli K, **Minderhoud T**, Mohammadi P, Dekker R, Brown V, Attaye T, Huisman SJ, Hettinga-Roest AA, Nanayakkara P. A prospective, observational study of the performance of MEWS, NEWS, SIRS and qSOFA for early risk stratification for adverse outcomes in patients with suspected infections at the emergency department. Acute Med. 2021;20(2):116-124. PMID: 34190738.

Minderhoud TC, Azijli K, Nanayakkara PWB. Sepsis. Lancet. 2020 Dec 5;396(10265):1804. doi: 10.1016/S0140-6736(20)32401-6. PMID: 33278929.

Nannan Panday RS, **Minderhoud TC**, Chantalou DS, Alam N, Nanayakkara PWB. Health related quality of life in sepsis survivors from the Prehospital Antibiotics Against Sepsis (PHANTASi) trial. PLoS One. 2019 Oct 1;14(10):e0222450. doi: 10.1371/journal.pone.0222450. PMID: 31574094; PMCID: PMC6772145.

Nannan Panday RS, **Minderhoud TC**, Alam N, Nanayakkara PWB. Prognostic value of early warning scores in the emergency department (ED) and acute medical unit (AMU): A narrative review. Eur J Intern Med. 2017 Nov;45:20-31. doi: 10.1016/j.ejim.2017.09.027. Epub 2017 Oct 7. PMID: 28993097

Minderhoud TC, Spruyt C, Huisman S, Oskam E, Schuit SCE, Levin MD. Microbiological outcomes and antibiotic overuse in Emergency Department patients with suspected sepsis. Neth J Med. 2017 Jun;75(5):196-203. PMID: 28653945.

Other publications

Douiyeb S, de la Court JR, Tuinte B, Sombogaard F, Schade RP, Kuijvenhoven **M, Minderhoud T**, Sigaloff KCE. Risk factors for readmission among patients receiving outpatient parenteral antimicrobial therapy: a retrospective cohort study. Int J Clin Pharm. 2022 Apr;44(2):557-563. doi: 10.1007/s11096-022-01379-7. Epub 2022 Feb 14. Erratum in: Int J Clin Pharm. 2022 Apr 21;: PMID: 35157228; PMCID: PMC9007809.

Rooimans T, M**inderhoud T**, Leal N, Vromans H, van Nostrum C, van Hasselt P; Lipophilic Drug Delivery Despite Cholestasis Study Group. Novel Orally Formulated Mixed Micelles Optimize Vitamin K Absorption Under Bile-Deficient Conditions. Gastroenterology. 2021 Sep;161(3):1056-1059.e5. doi: 10.1053/j.gastro.2021.05.056. Epub 2021 Jun 4. PMID: 34090886.

T.C. (Tanca) Minderhoud, Maurits P.A. van Meer, Robert J. van Thiel, Caroline M. den Hoed, Paul L.A. van Daele en Carolina A.M. Schurink, Infecties bij gebruik van glucocorticoïden, Ned Tijdschr Geneeskd. 2018;162D2215

Boonstra G, Cahn W, Schnack HG, Hulshoff Pol HE, **Minderhoud TC**, Kahn RS, van Haren NE. Duration of untreated illness in schizophrenia is not associated with 5-year brain volume change. Schizophr Res. 2011 Oct;132(1):84-90. Epub 2011 Aug 10. PMID: 21835595

van Hasselt PM, Janssens GE, Slot TK, van der Ham M, **Minderhoud TC**, Talelli M, Akkermans LM, Rijcken CJ, van Nostrum CF. The influence of bile acids on the oral bioavailability of vitamin K encapsulated in polymeric micelles. J Control Release. 2009 Jan 19;133(2):161-8. doi: 10.1016/j.jconrel.2008.09.089. Epub 2008 Oct 12. PMID: 18955093.

Dankwoord

In 2016 begon het einde van mijn opleiding tot internist in zicht te komen. Ik had veel tijd aan het ontwikkelen van klinische vaardigheden besteed maar nu wilde ik mij meer bekwamen in wetenschappelijke vaardigheden. Al sinds 2011 was ik bezig met een evaluatie van een sepsis screening en ik wilde gaag meer doen op dit gebied. Daarvoor ging ik in gesprek met dr. Prabath Nanayakkara in het VUmc, wat leidde tot tal van ideeën. Uiteindelijk was het resultaat dat ik in januari 2017 als fellow de overstap maakte naar VUmc, in juni 2017 internist werd, en ondertussen aan een promotieproject begon. Nu dit traject afgerond is, wil ik graag terugkijken en iedereen bedanken die hieraan bijgedragen heeft.

Mijn promoteren wil ik als eerste bedanken: Prabath, jij geeft iedereen met ambitie en drive op wetenschappelijk gebied een kans en daarbij laat je je niet weerhouden door gebruikelijke kaders of verwachtingen. Met jouw - out of the box- manier van denken heb je een hele diverse onderzoeksgroep bij elkaar gebracht. Dit leidt tot spannende samenwerking met allerlei verschillende invalshoeken. Als dat even leidt tot chaos dan houd je je hoofd koel en zeg je; slowly, slowly catch a monkey. Je gaf ook mij een kans, en daar ben ik heel dankbaar voor. Michiel, jij was het rustige baken in mijn promotietraject. Hoewel we niet erg frequent overleg hadden heb je me altijd gesteund en me op moeilijke momenten het vertrouwen gegeven om dit project af te maken.
Alle leden van de commissie, Jan Prins, Stephanie Klein-Nagelvoort Schuit, Louella Vaughan, Elske Sieswerda en Peter Pickkers wil ik graag bedanken voor het lezen en beoordelen van mijn proefschrift.

Dank aan alle patiënten die hebben meegewerkt en hun data met ons hebben gedeeld. Uiteindelijk doen we het onderzoek om jullie zorg beter te maken, ik hoop dat ik daaraan een steentje mag bijdragen.

Dank aan de VU en alle collega´s van de afdeling interne geneeskunde. In de afgelopen jaren heb ik veel collegialiteit mogen ervaren. In het bijzonder wil ik mijn kamergenoten van 4a46: Kim, Marije en Edgar. Jullie hebben me opgebeurd als mijn artikelen werden afgewezen en mee gevierd als er iets goed ging. Kim, wij begonnen hier samen als ‘jonge klare’ in 2017. Sindsdien hebben we heel wat haastbiertjes gedronken met Marije op vrijdagmiddag. Marije, dankzij jouw inzet en plan heeft griep een heel hoofdstuk gekregen in dit boek. Dank voor alle gezelligheid en steun! Nu de alliantie steeds verder vorm krijgt wil ik naast de collega’s op de Boelelaan, natuurlijk ook de collega’s van Meibergdreef bedanken voor hun collegialiteit en steun.

Mijn paranimfen, Claartje en Kim, ik ben ontzettend blij dat jullie me hebben bijgestaan in dit traject. Claartje, tijdens mijn allereerste wetenschaps-stage in 2005 heb je mij de beginselen van SPSS en statistiek bijgebracht, waar jij als socioloog natuurlijk veel beter in was. Ik weet nog precies het etentje in Utrecht waar je mij het zetje gaf wat ik nodig had om mijn boekje af te ronden Ontzettend veel dank voor alle gezellige avonden, hulp en leuke gesprekken. Ook als paranimf heb je alles weer perfect geregeld!

Dankwoord

Kim, we hebben sinds 2017 lief en leed gedeeld: variërend van onderzoekstress, opvoedleed tot feministische bingo; we kunnen samen lachen en samen huilen. Bedankt voor alles!

Alle collega-promovendi waar ik mee samen heb gewerkt wil ik graag bedanken. Kaoutar, mijn maatje, vanaf het begin toen Prabath ons aan elkaar voorstelde is er een mooie samenwerking ontstaan. Dit boekje was er niet geweest zonder jouw hulp en bijdrage. Bij een nieuw project had jij altijd binnen korte tijd vele studenten geworven en alles geregeld om het project in goede banen te leiden! Ik blijf onder de indruk van jouw kracht en je rustige volharding. Ik hoop dat ik jou ook kan ondersteunen bij de laatste loodjes van jouw proefschrift. Je bent een voorbeeld voor velen!

Rishi, de review die we samen schreven terwijl ik op vakantie was in Italië is nog steeds mijn meest geciteerde publicatie. Ik heb genoten van onze discussies over van alles (groente, feminisme) maar heb ook veel respect voor je productiviteit. Eva, Roos, Michiel, Bo, Hanneke, Christel, Karlinde, Siham, Jara, David, Sabrine, bedankt voor de leuke researchbesprekingen, de klaagmomentjes en de hoogtepunten die we gedeeld hebben. Roos en Michiel, jullie ambitieuze samenwerking op het gebied van data-analyse, artificial intelligence en bloedkweken heeft onze onderzoeksgroep echt een nieuwe invalshoek en inspiratie geboden (en mij een extra hoofdstuk). Ik hoop dat we de komende jaren nog veel mogen samenwerken op dit gebied. Rashudy, het was altijd heerlijk om met je bij te praten en vooral om elkaar te vinden in onze directe communicatietijl. Je bent nu in opleiding tot ziekenhuisapotheker en ik weet zeker dat je een mooie tijd tegemoet gaat.

Zonder hulp van wetenschappelijke studenten had ik dit proefschrift nooit kunnen voltooien. Parisa, Rachelle, Karlijn, Amber, Shiragani, Caspar en nog vele anderen! Jullie inzichten, nieuwsgierigheid en (moeilijke) vragen hebben me geholpen om het onderzoek te verbeteren. Bedankt voor jullie hulp met includeren, coderen en analyseren en alle andere dingen die jullie gedaan hebben.

Dank aan collega's op de SEH in het VUmc die hebben geholpen om patiënten te includeren voor onze studies. In het bijzonder wil ik alle verpleegkundigen bedanken die ons hebben geholpen met de inclusies voor DISC, ook van het SFG-Vlietland en Reinier de Graaf Ziekenhuis Delft. Ook Vanessa Brown, Tamana Attaye en Bas Huisman wil ik bedanken voor hun hulp en coördinatie daarbij.

Alle collega's op de acute as en AOA (o.a. Joost, Amy, Bart); bedankt voor jullie peptalks en altijd fijn als jullie me even inseinden als 'de professor' weer gesignaleerd was op de gang op 1 D (we gaan hem binnenkort echt uitrusten met een GPS chip).
Beste Natalina, Annelies, Ilse, Elise en Karin, bedankt voor jullie ondersteuning en hulp bij planning en praktische zaken. Naast jullie hulp zijn jullie ook vaak een luisterend oor en het is heel fijn met jullie samenwerken!

Ik wil graag alle aios, fellows en inmiddels 'jonge klare' collega's bedanken voor jullie gezelligheid, steun en hulp. Jullie hebben geholpen met inclusies (vooral de DISC) en eindeloze discussies over procalcitonine (bij COVID) aangehoord bij de overdracht, en soms daarna nóg een keer bij supervisie of de acute bespreking. Ik vind het inspirerend om met jullie samen te werken en een bijdrage aan jullie opleiding te mogen leveren.

Dank aan diegenen die tijdens mijn opleidingstraject me hebben gevormd en geïnspireerd om dit onderzoek te doen. Peter van Hasselt, jij liet me zien hoe mooi het is om onderzoek te doen naar een onderwerp wat echt je hart raakt. Pieter Wismans, in het Havenziekenhuis werd de basis voor mijn opleiding en uiteindelijk ook voor dit onderzoek. De diversiteit van infecties, en hun beloop fascineert me nog steeds. Mark-David Levin, met jou zette ik de eerste stappen om me te verdiepen in onderzoek naar sepsis. Je legt de lat hoog en juist daardoor werd ik gestimuleerd om de sprong te wagen. Je hebt mij geduldig begeleid, ook al was ik onervaren en behoorlijk eigenwijs. Veel dank daarvoor.
Mijn opleidingstijd in het ErasmusMC was een ontzettende leuke tijd waarin ik veel heb gehad aan het Rotterdamse adagium 'niet lullen maar poetsen'. Binnen de acute waren we gezamenlijk aan het pionieren om de acute zorg nóg beter te maken. Ik wil graag alle opleiders en collega's bedanken! Stefanie, jij hebt me overtuigd om het aandachtsgebied acute te doen in plaats van infectieziekten; een keuze waar ik nooit spijt van heb gehad. Je bent een bron van inspiratie en ik ben ontzettend blij dat je deel kon uitmaken van mijn commissie.

Dank aan mijn vrienden en familie. Lobke, al sinds onze kleutertijd kennen we elkaar en een vriendschap die zo ver teruggaat is van onschatbare waarde. Zonder onze vriendschap was ik een ander persoon geweest. En dankzij jou heb ik de muziek van Bram Vermeulen leren kennen die een mooie plek heeft gekregen in dit boek. De GNK-meiden (Feikje, Jolien, Masha, Elize), wij kennen elkaar sinds we in 2001 samen in een kennismakingsgroepje terecht kwamen, bedankt voor jullie vriendschap. Feikje en Jolien, jullie zijn er steeds voor mij geweest. Jullie hebben meegelezen, meegedacht en meegeleefd en dat heeft me door moeilijke tijden heen geholpen. Masha, je zat de afgelopen jaren in Boston op afstand, mede door COVID. Maar je weet er altijd het beste van te maken, ik ben geïnspireerd door jouw kracht en incasseringsvermogen.

Dank aan alle (schoon) familie, iedereen die af en toe heeft geïnformeerd hoe het toch ook weer ging met dat proefschrift. Hanna, ontzettend bedankt voor de schitterende illustratie die je gemaakt hebt om de zoektocht van dit proefschrift te symboliseren. Wico & Ilkay en Froukje, ook jullie hebben me gesteund en geholpen. Bedankt voor alle keren dat jullie een van de kinderen meenamen voor een gezellig uitje of logeerpartij.

Mijn zussen Mirjam, Marieke en Hanne-Maria. We zijn allemaal heel anders maar naarmate de jaren verstrijken is dat een verrijking en blijft over wat ons bindt. Ik ben blij en dankbaar dat jullie mijn zussen zijn.

Frank, sinds 2016 kennen we elkaar en ben je onze steun en toeverlaat. Met veel humor en plezier ben jij er voor de kinderen, ik ben ontzettend blij met al jouw hulp.

Dank aan mijn ouders. Papa, al van kleins af aan heb je me laten zien hoe je mensen kan betrekken bij een klus door ze een verantwoordelijkheid te geven, en hun bijdrage, ook als die klein is, te waarderen. Jouw rustige, niet-oordelende kijk op zaken heeft me altijd geïnspireerd. Mama, van jou leerde ik om te improviseren, om een probleem van meerdere kanten te bekijken, en om je niet te laten weerhouden door de verwachtingen van mensen om je heen. En zoals je altijd zegt: er zit een hoop graniet onder het fluweel.

Deze alinea is eigenlijk veel te kort om mijn gezin te bedanken. Jonne, Krijn, Doris en Idris, jullie energie, humor en eigenheid geven zoveel vreugde. Ik kijk ernaar uit om deze zomer heerlijk lang met elkaar op vakantie te gaan, en om een beetje in te halen wat ik soms aan familietijd gemist heb. Jan, jij bent mijn anker. Jij houdt me veilig op een woelige zee. Je hebt ook aan dit proefschrift ontzettend veel bijgedragen en ik denk dat je bij elk stuk wel hebt meegelezen of hebt meegedacht. Samen zijn wij meer dan de som der delen. Jij bent mijn grote liefde!

Dank aan de Eeuwige. Waar liefde is en vriendschap, daar is God.

Infections leading to severe illness and ultimately death have been described for more than 2000 years, and have been described by the word sepsis since Homer and Hippocrates.

With the discovery of antimicrobial therapy and intravenous fluid therapy, it was recognized that the prognosis could be improved, especially if the treatment was started early. After studies confirmed this finding, screening programs stimulating early antibiotic treatment were widely implemented.

However, severe infections and the ensuing organ failure can present in very diverse ways which makes accurate screening difficult. In addition, early antimicrobial treatment focused mostly on antibiotics, even though virus infections (and other organisms) can also cause severe illness. In this thesis, several articles on the topic of infections in the ED have been brought together. In part 1, we focus on how we identify severely ill patients in the Emergency Department. Part 2 focuses on microbiological evidence and recognition of viral or bacterial infection in the ED. Part 3 is about the bigger picture. In the discussion, we discuss how we the concept of sepsis in the ED can be applied in the ED and what the challenges for the future are. After all, do we really know what we are looking for?