PUBLICATIONS OF
THE UNIVERSITY OF EASTERN FINLAND

Dissertations in Health Sciences

ANSSI PÖLKKI

ORGAN FAILURE ASSESSMENT

INSIGHTS INTO MEASURING ORGAN DYSFUNCTIONS
AND INTENSIVECARE UNIT BENCHMARKING

ORGAN FAILURE ASSESSMENT

INSIGHTS INTO MEASURING ORGAN DYSFUNCTIONS AND INTENSIVE CARE UNIT BENCHMARKING

Anssi Pölkki

ORGAN FAILURE ASSESSMENT

INSIGHTS INTO MEASURING ORGAN DYSFUNCTIONS AND INTENSIVE CARE UNIT BENCHMARKING

To be presented by permission of the Faculty of Health Sciences, University of Eastern Finland for public examination in Kuopio University Hospital Auditorium 1, Kuopio on September 20th, 2024, at 12 o'clock noon

Publications of the University of Eastern Finland
Dissertations in Health Sciences
No 841

Department of Anaesthesiology and Intensive Care / School of Medicine
University of Eastern Finland, Kuopio
2024

Series Editors
Research Director Jari Halonen, MD, PhD.
Institute of Clinical Medicine, Surgery
Faculty of Health Sciences

Professor Joonas Sirola, MD, PhD.
Institute of Clinical Medicine, Surgery
Faculty of Health Sciences

Adjunct professor Toni Rikkonen, MD, PhD.
Institute of Clinical Medicine, School of Medicine
Faculty of Health Sciences

Professor Tarja Palm, PhD.
A.I. Virtanen Institute for Molecular Sciences
Faculty of Health Sciences

Lecturer Veli-Pekka Ranta, PhD.
School of Pharmacy
Faculty of Health Sciences

Lecturer Tarja Välimäki, PhD.
Department of Nursing Science
Faculty of Health Science

Printing office
Punamusta
Joensuu, 2024
Distributor: University of Eastern Finland
Kuopio Campus Library

ISBN: 978-952-61-5288-2 (Print)
ISSNL: 1798-5706
ISSN: 1798-5706
ISBN: 978-952-61-5289-9 (PDF)
ISSN: 1798-5714 (PDF)

Author's address: School of Medicine, Faculty of Health Sciences,
University of Eastern Finland
Department of Anaesthesiology and Intensive Care,
Kuopio University Hospital,
KUOPIO
FINLAND

Doctoral programme: Doctoral Programme in Clinical Research

Supervisors: Professor Matti Reinikainen, MD, PhD.
School of Medicine, Faculty of Health Sciences,
University of Eastern Finland
Department of Anaesthesiology and Intensive Care,
Kuopio University Hospital
KUOPIO
FINLAND

Adjunct Professor Pirkka T. Pekkarinen, MD, PhD.
Division of Intensive Care Medicine, Department of Anaesthesiology and Intensive Care, University of Helsinki and Helsinki University Hospital
HELSINKI
FINLAND

Adjunct Professor Rahul Raj, MD, PhD.
Department of Neurosurgery, University of Helsinki and Helsinki University Hospital
HELSINKI
FINLAND

Reviewers:

Adjunct Professor Panu Uusalo, MD, PhD.
Department of Anaesthesiology and Intensive Care, University of Turku, and Turku University Hospital,
TURKU
FINLAND

Adjunct Professor Kati Järvelä, MD, PhD.
Heart Hospital, Tampere University Hospital, and Tampere University
TAMPERE
FINLAND

Opponent:

Adjunct Professor Anne Kuitunen, MD, PhD.
Tampere University Hospital, Department of Intensive Care and Tampere University, Faculty of Medicine and Health Technology
TAMPERE
FINLAND

Pölkki, Anssi
Organ Failure Assessment
Insights into measuring organ dysfunctions and intensive care unit benchmarking
Kuopio: University of Eastern Finland
Publications of the University of Eastern Finland
Dissertations in Health Sciences 841. 2024, 140 p.
ISBN: 978-952-61-5288-2 (Print)
ISSNL: 1798-5706
ISSN: 1798-5706
ISBN: 978-952-61-5289-9 (PDF)
ISSN: 1798-5714 (PDF)

ABSTRACT

The Sequential Organ Failure Assessment (SOFA) score is a universal and practical tool to quantify the severity of organ dysfunctions in intensive care. It concludes six components: respiratory, coagulation, hepatic, cardiovascular, renal, and neurologic subscores. Patients receive 0–4 points for each component depending on the severity of organ failure. Introduced in 1996, the SOFA score has remained unrevised for nearly 30 years. There have been significant changes in clinical practices during that period, and it is probable that the score's criteria no longer accurately reflect the severity of organ failure in contemporary intensive care.

Within the context of intensive care unit (ICU) benchmarking, calculation of standardised mortality ratios (SMR) is a useful method for comparing the performance of ICUs. In SMR, the observed mortality is compared to the predicted mortality, which is calculated using a specific mortality prediction model. Patients with an untreatable brain injury that are admitted to ICUs purely for the purpose of potential organ donation have an almost 100% mortality. However, mortality prediction models may not identify the extremely high probability of death among these patients. The confounding effect of potential organ donors (PODs) on SMRs is unknown.

Aims

The aim of this study was to assess whether the SOFA score aligns with current practices and whether the SOFA subscores derived from different organ systems have an equal association with mortality. Subsequently, the aim was to investigate whether the accuracy of the SOFA score could be improved by replacing the original cardiovascular score by a Vasoactive Inotropic Score (VIS)-based score. Another goal was to study whether simply the required dose of the predominant vasopressor in intensive care, which is noradrenaline (norepinephrine), could be used to determine the severity of cardiovascular dysfunction. Moreover, we investigated which would be the optimal cutoffs for noradrenaline doses to determine low, intermediate, and high dose. Furthermore, the aim was to explore to what extent PODs influence the SMRs and ICU benchmarking.

Materials and Methods

The study comprised four cohort observations utilising single-centre data or larger ICU register databases. In Study I, we used the Finnish Intensive Care Consortium (FICC) database. Study II was conducted as a single-centre study, and the study population included patients admitted to the Kuopio University Hospital (KUH) ICU. In Study III, we used the KUH database in the cutoff development and a large eICU open ICU database for external validation. In Study IV, we utilised the FICC database, which also included units from Estonia and Switzerland. We used mortality (either hospital mortality or 30-day mortality) as the primary outcome.

Main Results

In Study I, we found that increase in the cardiovascular subscores are neither linearly nor consistently associated with mortality. In addition, high cardiovascular subscores did not reflect as high a risk of death as high subscores derived from other organ systems. In study II, the VIS-based criteria outperformed the current cardiovascular score (area under the receiver operating characteristic curve 0.816 for the current SOFA score vs. 0.822 for the maximum VIS-based SOFA score in predicting 30-day mortality). In Study III, utilising a pragmatic statistical approach, we found that cutoffs of 0.2 μg/kg/min and 0.4 μg/kg/min were optimal to determine low, intermediate, and high doses of noradrenaline in quantifying the

severity of circulatory failure. The cutoffs were proven in the external validation. In Study IV, we demonstrated that the benchmark ranking alters in 70% of the units when PODs were excluded. Although this group represents a small proportion of all ICU admissions (0.9%), it has a significant influence on ICU benchmarking. The predicted mortality of PODs was 37%, while the observed mortality was 93%, and PODs accounted for 7% of all deaths.

Conclusion
The SOFA score, particularly its cardiovascular component score, is outdated and requires revision. The replacement of the cardiovascular SOFA component with VIS-based criteria improves the accuracy of the SOFA score. The cutoffs of 0.2 μg/kg/min and 0.4 μg/kg/min are suitable for low, intermediate, and high noradrenaline dose determination in the description of severity of circulatory failure. Patients admitted to the ICU for the sole purpose of organ donation cause bias in SMR calculations and they should be excluded from ICU benchmarking programmes.

Keywords: SOFA score, Sequential Organ Failure Assessment score, Multiorgan-failure, Multiple Organ Dysfunction Syndrome, Mortality prediction, Standardised Mortality Ratio, SMR, Intensive care benchmarking, Intensive care quality

Pölkki, Anssi
Elinvaurioiden vakavuusasteen arviointi
Elinvaurioiden vakavuusasteen mittaamisesta ja tehohoidon laadun vertaisarvioinnista
Kuopio: Itä-Suomen yliopisto
Publications of the University of Eastern Finland. Dissertations in Health Sciences 841. 2024, 140p.
ISBN: 978-952-61-5288-2 (Print)
ISSNL: 1798-5706
ISSN: 1798-5706
ISBN: 978-952-61-5289-9 (PDF)
ISSN: 1798-5714 (PDF)

TIIVISTELMÄ

Sequental Organ Failure Assessment (SOFA) -pisteytys on yleisesti käytetty tehohoitopotilaan monielinvaurion vaikeusasteen mittari. SOFA-pisteytyksellä arvioidaan hengityselimistön, veren hyytymisjärjestelmän, maksan, verenkierron, munuaisten ja keskushermoston elinhäiriöiden vaikeusastetta asteikolla nollasta neljään. SOFA-pisteytys otettiin käyttöön vuonna 1996 eikä sitä ole päivitetty sen jälkeen, vaikka tehohoidon käytännöt ovat muuttuneet huomattavasti. Onkin mahdollista, että pisteytys on vanhentunut eivätkä elinjärjestelmien häiriöistä saadut pisteet enää kuvaa luotettavasti elinvaurion vakavuusastetta.

Vakioitu kuolleisuussuhde, Standardised Mortality Ratio (SMR) on vakiintunut mittari tehohoidon laadulle. Sitä hyödynnetään teho-osastojen vertaisarvioinnissa. SMR kuvaa toteutuneen kuolleisuuden ja ennustemallien perusteella lasketun ennakoidun kuolleisuuden suhdetta. Ennustemallit eivät kuitenkaan todennäköisesti havaitse niiden potilaiden, joita hoidetaan teho-osastolla mahdollisina elinluovuttajina lähes varmaa menehtymistä sairaalahoidon aikana. Mahdollisten elinluovuttajien vertaisarviointiin aiheuttaman harhan laajuutta ei tiedetä.

Tavoitteet

Tutkimuksen tavoitteena oli tarkastella vastaavatko SOFA-pisteytyksen kriteerit nykyaikaista tehohoitoa ja ovatko sen elinjärjestelmäpisteet yhteismitallisia kuoleman ennustamisessa. Tutkimme, voiko SOFA-pisteytyksen ennustuskykyä parantaa vaihtamalla verenkierron elinvauriota kuvaava komponentti Vasoactive Inotrope Scoreen (VIS) perustuvilla kriteereillä. Lisäksi laadimme optimaaliset kynnysarvot noradrenaliini (norepinefriini) -infuusiolle (pieni, keskisuuri ja suuri annos). Tarkastelimme myös, kuinka suuren harhan mahdolliset elinluovuttajat aiheuttavat teho-osastojen SMR-lukuihin ja tehohoidon vertaisarviointiin.

Aineisto ja menetelmät

Väitöskirjan osatutkimukset olivat taaksepäin katsovia. Ensimmäisen osatutkimuksen aineistona käytettiin suomalaisen tehohoidon vertaisarvioinnin laatutietokantaa. Toisessa osatutkimuksessa aineistona olivat Kuopion yliopistollisen sairaalan (KYS) teho-osastolla 2013–2019 hoidetut potilaat. Kolmannessa osatyössä noradrenaliinin kynnysarvot laadittiin ensin KYS:in tehohoitopotilasaineiston avulla vuosilta 2013–2019, jonka jälkeen raja-arvot validoitiin laajassa yhdysvaltalaisessa tehohoitorekisterissä, eICU:ssa, vuosilta 2014–2015. Neljännessä osatyössä hyödynnettiin suomalaisen tehohoidon vertaisarvioinnin tietokantaa, johon oli otettu mukaan myös teho-osastot Virosta ja Sveitsistä.

Tärkeimmät tulokset

Ensimmäisessä osatutkimuksessa havaitsimme, että verenkiertovajauksen vaikeutta kuvaavat SOFA-pisteet poikkeavat ennustearvoltaan muista SOFA:n elinjärjestelmiä kuvaavista osapisteistä. Suuri SOFA:n verenkiertokomponentista saatu pistemäärä ei ollut yhteydessä erityisen suureen kuolleisuuteen verrattuna muihin elinjärjestelmiin. SOFA:n ennustekyky 30:n vuorokauden kuolleisuudelle parani, kun verenkiertokomponentti vaihdettiin VIS-pisteisiin perustuvaksi. SOFA:n Area Under the Receiver Operating Characteristic (AUROC) -luku nousi 0.816:sta 0.822:een. Kolmannessa osatutkimuksessa pienen, keskisuuren ja suuren noradrenaliiniannoksen optimaalisiksi kynnysarvoiksi määritettiin 0.2 ja 0.4 µg/kg/min. Kynnysarvot validoitiin ulkoisesti laajassa rekisteriaineistossa. Neljännessä osatutkimuksessa

havaittiin, että mahdollisina elinluovuttajina tehohoitoon kirjattujen potilaiden todellinen kuolleisuus oli 93 %. Heidän ennustettu kuolleisuutensa ennustusmallin perusteella oli 37 %. Vertaisarvioinnin sijoitus muuttui 70 %:lla teho-osastoista, kun mahdolliset elinluovuttajat poistettiin ennustemallista

Yhteenveto
SOFA-pisteytys, etenkin sen verenkiertohäiriötä kuvaava komponentti, on vanhentunut, ja se tulisi päivittää. SOFA-pisteytyksen erottelukyky parantui, kun verenkiertokomponentti korvattiin VIS-pisteisiin perustuvilla kriteereillä. 0.2µg/kg/min ja 0.4µg/kg/min ovat sopivat kynnysarvot määrittämään pientä, keskisuurta ja suurta noradrenaliiniannosta. Mahdolliset elinluovuttajat aiheuttavat teho-osastojen SMR-lukuihin harhan, ja ne sotkevat tehohoidon vertaisarviointia. Nämä potilaat tulisi jättää pois SMR-laskennoista.

Avainsanat: SOFA, SOFA -pisteytys, Sequential Organ Failure Score -pisteytys, Monielinvaurio, Kuoleman ennustemalli, Vertaisarviointi, Tehohoidon laatu

ACKNOWLEDGEMENTS

My interest in discovering new things has been present since my youth. However, I began my dissertation project after I had already completed my MD degree and was somewhat advanced in my specialisation in anaesthesiology and intensive care. Throughout this journey, there have been many people who initially supported my interest in science and, ultimately, helped me to create something concrete in the field of medical research.

Without my excellent supervisors, this project would never have even been initiated. The first I should mention is Professor Matti Reinikainen, who was the first to introduce me to the topic. He guided me through the entire project. Matti, your dedication to research is exceptional. This, combined with your thorough knowledge of medicine, teaching skills, interpersonal skills, and patience in guiding your students, makes you a perfect supervisor. Thank you for always finding the time to guide me, review my work, and enthusiastically contribute new ideas.

Adjunct Professor Pirkka T. Pekkarinen, thank you for contributing brilliant ideas throughout the project, and for your support and encouragement. Your talents in research are remarkable. I appreciate our conversations, which gave me the confidence to carry on. Adjunct Professor Rahul Raj, thank you for providing your knowledge of benchmarking. Your expertise has improved the quality of not only intensive care but also this dissertation.

I would like to thank Adjunct Professors Panu Uusalo and Kati Järvelä for their contributions to the reviewing process and for spending their valuable time.

A special thanks is due to Jukka Takala for providing insights based on exceptional experience, as well as to Tuomas Selander and André Moser for their invaluable support in statistical methodology. Moreover, I want to thank all the co-authors for their contributions to the publications and the effort they invested and to everyone who have contributed to the Finnish Intensive Care Consortium.

Without time and funding, this dissertation would never have been completed. The shared doctoral researcher position system of the

University of Eastern Finland provided me the opportunity to spend 50% of my working time on this project for four years, along with clinical work. I also received grants from the Finnish Society of Anaesthesiology, the Finnish Society of Intensive Care, and the Kuopio University Hospital Research Foundation. ACTA Foundation provided me with a high-quality research methodology course in Copenhagen. The Department of Anaesthesiology and Intensive Care at Kuopio University Hospital made a major financial effort to enable me to attend national and international conferences and present the study results. Thanks are also due to Stepani Bendel, Minna Niskanen and Juha Rutanen, who enabled me to take time off from clinical work to promote this study.

I am astonished by the clinical skills of my colleagues in the Kuopio University Hospital Intensive Care Unit. In addition to Matti and Stepani, I want to thank Ari U, Maarit L, Timo N, Ville K, Lauri W, Christiane H, Martti S, Ville I and Reetta S. A special thanks is also due to all the intensive care nurses who I have had the privilege to work with. You all, doctors and nurses, do amazing work saving people's lives.

There are a huge number of friends and relatives who have supported me along the way and occasionally provided me with other things to think about. Among these, I want to thank Ville L, Pauliina L, Miika V, Antti K, Mikko Ö, Outi Ö, Anna J, Laura M, Sami C, Pertti S, Niina K, Antti M, Ville M, Hanna M, Erkka K, Miika K, Pertti N, Pekka S, Teemu O, Sanna K, Alisa O, Satu L, Tuomo L, Kirsi M, Marko M, Miia K, Jussi K, Emmi K, Jarkko L, Arno P and many others who are not mentioned here.

There are no words to express my gratitude to my father, Sakari, and my mother, Anneli. You have a wealth of wisdom and have always seen education as the path forwards in life. Your humane perspective on the world is something everyone can learn from. I am proud to be your son. I would also like to extend my heartfelt thanks to my sisters Hannaleena and Minna, including her family, Juha, Niilo, Toivo, and Hilma. Your support has been priceless. Thanks are also due to Sarianna S and Juha S for their invaluable support.

Finally, Tytti, my beloved partner, you have stood by me through the difficult times and during periods when I have been absorbed in work. There are no words to express my gratitude for your unwavering support.

Thank you. I look forward to spending more time together in the coming years.

Kuopio, August 2024
Anssi Pölkki

LIST OF ORIGINAL PUBLICATIONS

This dissertation is based on the following original publications:

I Pölkki A, Pekkarinen PT, Takala J, Selander T, Reinikainen M. Association of Sequential Organ Failure Assessment (SOFA) components with mortality. Acta Anaesthesiol Scand. 2022;66:731-741.

II Pölkki A, Pekkarinen PT, Lahtinen P, Koponen T, Reinikainen M. Vasoactive Inotropic Score compared to the Sequential Organ Failure Assessment cardiovascular score in intensive care. Acta Anaesthesiol Scand. 2023;67:1219-1228.

III Pölkki A, Pekkarinen PT, Hess B, Reintam Blaser A, Bachmann KF, Lakbar I, Hollenberg SM, Lobo SM, Rezende E, Selander T, Reinikainen M. Noradrenaline dose cutoffs to characterise the severity of cardiovascular failure: data-based development and external validation. Acta Anaesthesiol Scand. In Press

IV Pölkki A & Moser A, Raj R, Takala J, Bendel S, Reinikainen M & Jakob S. The Influence of Potential Organ Donors on Standardized Mortality Ratios and ICU Benchmarking. Crit Care Med. 2024;52:387-395.

The publications were adapted with the permission of the copyright owners. For study IV, only the reprint of the final version of the manuscript was permitted in the printed version of the dissertation.

CONTENTS

ABBREVIATIONS

ADQI Acute Disease Quality Initiative

AIDS Acquired Immune Deficiency Syndrome

AKI Acute Kidney Injury

ACLF Acute-on-Chronic Liver Failure

ALF Acute Liver Failure

APACHE Acute Physiology and Chronic Health Evaluation

APS Acute Physiology Score

AUROC Area Under the Receiver Operating Characteristic

BUN Blood Urea Nitrogen

cvSOFA Cardiovascular SOFA score

DIC Disseminated Intravascular Coagulation

ECMO Extracorporeal Membrane Oxygenation

ESICM European Society of Intensive Care Medicine

FICC The Finnish Intensive Care Consortium

FiO_2 Fraction of inspired Oxygen

GCS Glasgow Coma Scale

FOUR Full Outline of UnResponsiveness

GDF-15 Growth Differentiation Factor 15

HFNC High Flow Nasal Cannula

HUS Hemolytic Uremic Syndrome

HR Hazard Ratio

ICU Intensive Care Unit

IGFBP7 Insulin-Like Growth Factor-Binding Protein 7

IQR Inter-Quartile Range

KDIGO Kidney Disease: Improving Global Outcomes

KIM-1 Kidney Injury Molecule 1

LODS Logistic Organ Dysfunction Score

MAP Mean Arterial Pressure

MODS Multiple Organ Dysfunction Score

MOF Multiorgan Failure

MPM Mortality Probability Model

NEE Noradrenaline Equivalent

NEQ Noradrenaline Equivalent

NGAL Neutrophil Gelatinase-Associated Lipocalin

OSF Organ-System Failure

PaO_2 Partial pressure of arterial oxygen

$PaCO_2$ Partial pressure of arterial Carbon dioxide

$PaFiO_2$ Fractional inspired oxygen

POD Potential Organ Donor

PAR Pressure-Adjusted heart Rate

PEEP Positive End-Expiratory Pressure

SAPS Simplified Acute Physiology Score

SA-AKI Sepsis-Associated Acute Kidney Injury

SI-AKI Sepsis-Induced Acute Kidney Injury

SMR Standardised Mortality Ratio

SOFA Sequential Organ Failure Assessment (Previously Sepsis-related Organ Failure Assessment)

TRIPOD Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis

TTP Thrombotic Thrombocytopenic Purpura

VIS Vasoactive Inotropic Score

WBC White Blood Cell

1 INTRODUCTION

In August 1952, a devastating polio epidemic swept through Denmark. Several young patients were admitted to Blegdam Hospital, Copenhagen, with bulbospinal poliomyelitis, which affected the nerves responsible for breathing. In the most severe cases, the result was total paralysis. The chances of saving these patients from suffocation were limited. Pioneer anaesthesiologist Bjørn Ibsen proposed providing positive pressure ventilation to polio patients, enabling them to overcome the infection until their spontaneous breathing power was restored.[1] The patients were tracheostomised and manually ventilated, many of them for several weeks, by medical and dental students from Copenhagen University. It is estimated that around 120 individuals were saved. The specialty of intensive care medicine was born (Figure 1).[2–4]

The development of techniques supporting severely deteriorating vital organs led to the establishment of specialised wards with advanced resources to treat the most critically ill patients, namely, intensive care units (ICUs).[5] The World Federation of Societies of Intensive and Critical Care Medicine task force defined an intensive care medicine as a 'multidisciplinary and interprofessional specialty dedicated to the comprehensive management of patients having, or at risk of developing, acute, life-threatening organ dysfunction'.[6] This definition is an accurate portrayal of intensive care today.

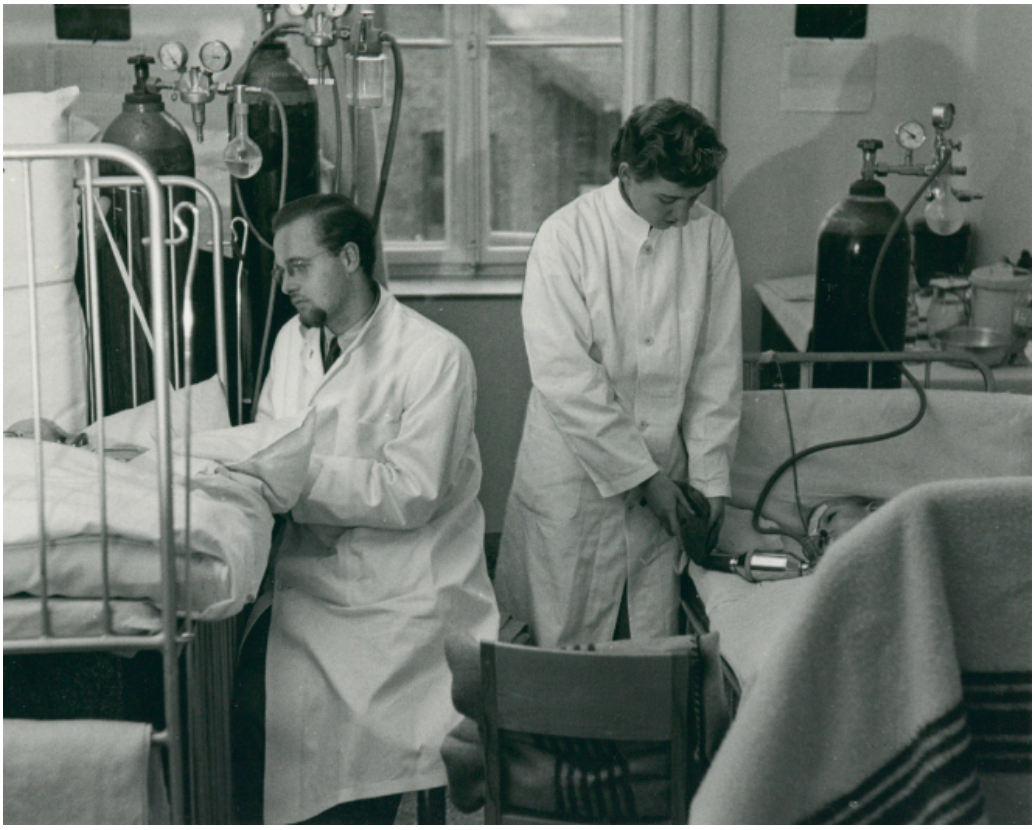

Figure 1. Children affected by a paralysing form of polio were ventilated manually. Reprinted with the permission of Medicinisk Museion, University of Copenhagen, Copenhagen, Denmark.

The most severely ill ICU patients experience life-threatening organ dysfunction involving not just one but several deteriorating organ systems. This dysfunction can occur in conditions such as septic shock or severe trauma and is referred to as multiorgan failure (MOF).[7] Organ dysfunction scores are a practical way to describe the severity and number of organ failures. Several such scores have been published, but Sequential Organ Failure Assessment (SOFA) score is the most widely adopted. The SOFA score quantifies the degree of failure of six organ systems: respiratory, coagulation, hepatic, cardiovascular, neurologic, and renal.[8]

Over the years, the results of intensive care have improved, and the accelerating development of knowledge and medical technology has changed what is considered life-threatening. For example, the 28-day mortality for severe sepsis in randomised controlled trials has decreased

from 47% to 29% over the course of two decades.[9] Such developments should prompt updates to organ failure scores, including the SOFA score. The SOFA score was introduced in 1996 and has remained unchanged despite the rapid pace of development in intensive care.

In a broader assessment of ICU patients' probability of surviving critical illnesses, a more comprehensive array of parameters, in addition to the severity of the acute disease, must be considered. These include the patient's age, admission cause, and physiological measures in greater detail. Various mortality prediction models tailored for this purpose are available. The most used of these are the Acute Physiology and Chronic Health Evaluation (APACHE), Simplified Acute Physiology Score (SAPS), and Mortality Prediction Model (MPM).[10–12] Estimating the probability of survival is essential for two reasons. First, at the individual patient level, the health care professionals and the patient's family members need to gain a basic understanding of the prognosis through the patient's background and the severity of the acute disease. Second, at the institutional administrative level, it is imperative to understand the capability of the ICU to save lives. Gaining this understanding involves comparing the number of patients admitted to the unit who have died to the expected number of deaths. This enables the comparison of the unit´s performance to its peers. This practise is referred to as ICU benchmarking, and it has achieved a well-established status within the intensive care field.[13,14] However, the benchmarking results in intensive care are not the ultimate goal themselves; rather, they are a tool to standardise intensive care quality and to ensure equal treatment for all patients.

The data collection for the prediction models is often automated, which also facilitates the establishment of extensive databases of ICU data that can be used also for research purposes.[15]

2 REVIEW OF THE LITERATURE

2.1 INTENSIVE CARE INSIGHTS INTO MULTIPLE ORGAN FAILURE

2.1.1 The conceptualisation and incidence of organ failures

The first known description of MOF in critically ill patients was authored by Dr Nicholas L. Tilney along with his colleagues from Harvard Medical School's Department of Surgery in 1973. They published a case series presenting 18 patients who had undergone surgery for ruptured aortic aneurysms and suffered major blood loss. The patients developed failures of the pancreatic, pulmonary, central nervous, hepatic, gastrointestinal tract, and cardiac systems. Ultimately, only one of these patients survived, highlighting the challenges in managing MOF patients in the early 1970s.[16] The first publications to follow those of Tinley et al. on MOF mainly focus on severely injured or other surgical patients with significant blood loss.[17–19] Intensive care has developed at an astonishing pace since the 1970s. However, even today, the development and progression of organ dysfunction across multiple organs is part of the final stages of the pathway leading to the death of severely ill individuals.[20,21] Therefore, it is essential to establish an accurate and generally recognised concept of how the failure of an organ system is determined that aligns with current intensive care practices and prognoses.

The assessment of the trend in the incidence of MOF in ICU patients was challenging during the first decades of intensive care as there was no standardised determination of MOF. The first appropriate organ failure scores were introduced in the 1990s´to overcome this problem. Currently, the most widely adopted definition of organ failure relies on the SOFA score. Achieving organ-specific SOFA subscores of 3 or 4 is often referred as organ failure, whereas more limited malfunction, indicated by organ specific SOFA subscores of 1 or 2 can be referred to as organ dysfunction.[22,23]

Previous studies have estimated the incidence of organ failure in ICU patients to be in the range of 50–83%.[22,24–26] The Sepsis Occurrence in Acutely Ill Patients (SOAP) study, conducted in nearly 200 ICUs in 24

European countries including over 3000 ICU patients with sepsis, found that during the ICU period, the respiratory system (43%) experienced the highest rate of failure of all organ systems, followed by the renal (38%), cardiovascular (34%), neurologic (26%), coagulation (10%), and hepatic (6%) systems. Cardiovascular failure was the most frequent organ failure at ICU admission (24%), whereas respiratory and renal failure appeared to develop later during ICU stays.[25] In another study observing 872 patients in 10 Scottish ICUs, the findings revealed similar frequencies of organ failures during ICU stays. Respiratory failure was notably predominant, occurring in 83% of cases, followed by cardiovascular failure at 45%, renal failure at 23%, coagulation failure at 15%, and hepatic failure at 8%. Data on the frequency of neurologic failure were not collected.[26]

2.1.2 MOF in specific ICU patient groups

Patients particularly susceptible to MOF during intensive care include those admitted with trauma, sepsis, and acute pancreatitis. Among ICU patients, these groups have likely been the most extensively studied in relation to MOF and warrant more in-depth consideration. However, MOF can impact patients admitted to the ICU for various reasons, such as severe drug intoxication or massive myocardial infarct.[27,28]

2.1.2.1 Trauma

The incidence of MOF in trauma patients admitted to the ICU appears to have increased in comparison to the first observatory studies of MOF in the ICU patients in the 1970s. However, the mortality of these patients has recently decreased due to improved treatments.[29] In addition to the SOFA score and other organ dysfunction scores, trauma-specific scoring systems have been widely used in research involving trauma patients. These specific scores include the Injury Severity Score (ISS) and Denver Score.[30,31]

The incidence of MOF in trauma patients during the ICU period has been reported to range from 11% to as high as 78%, depending on the MOF definition and population studied.[31–35] In a large observational study involving 440 severely injured trauma patients admitted to 29 ICUs in the United Kingdom, 55% of the patients developed MOF during their ICU stays. The in-hospital mortality for patients with MOF was 22%, compared

to only 0.5% in patients without MOF. Most trauma patients developed MOF on the day of their arrival in the ICU.[32] In another observational study investigating patterns of prolonged MOF in a population of 595 American trauma ICU patients, MOF appeared to be prolonged when hepatic and renal organ systems were involved. Prolonged MOF was strongly associated with a higher rate of infection.[34] In addition to infection, the other known risk factors for developing severe MOF are increased age, the presence of immunosuppression, and ongoing catabolic state.[36,37]

Trauma-related MOF appears to exhibit a bimodal pattern. The first subgroup presents MOF promptly after admission, and the instability of these patients is most typically characterised by cardiovascular shock. The most severely injured patients in this group tend to die soon after admission. The other subgroup consists of patients who develop MOF later during their care. These patients often experience hepatic and renal function failure, and trauma-related or nosocomial infections are strongly involved. Mortality does not significantly differ between these two groups.[34,38,39] Although the prognosis of trauma patients with MOF appears to be gradually improving, prognosis remains poor for the most severely injured patients.[40]

2.1.2.2 Sepsis

Besides trauma patients, another group strongly associated with MOF consists of sepsis patients. As early as 1975, Iain Ledingham and colleagues demonstrated that MOF is a significant predictor of mortality in patients with septic shock.[41] In the SOAP study, the prevalence of organ failure involving at least two organ systems was approximately 40%. One third of these patients died during their ICU stays.[25] In an observational study of over 2,000 American ICU patients, an even higher mortality of 54% was observed among those patients who developed MOF during their ICU stays.[42]

Advancements in treatments are leading to improved outcomes also for sepsis patients.[43] A Catalonian observational study by Cárdenas et al. investigated the trend of sepsis patients with MOF between 2005 and 2019. This study, which considered 296,554 patients with sepsis, found that mortality decreased most in sepsis patients with cardiovascular organ failure during the study period (47% in 2005 vs. 31% in 2019). Twenty-six

percent of the patients had MOF, and 36% did not survive to hospital discharge. However, in this study, the definition of organ failure was determined by the International Classification of Diseases 9 (ICD-9) instead of the SOFA score.[44]

The development of new organ failures is very closely linked to the concept of sepsis. In the current diagnosis criterion for sepsis, SEPSIS-3, the presence of organ failure defines the sepsis diagnosis,[45] whereas it was previously defined by the concept of systemic inflammatory response syndrome (SIRS).[46] Studies have shown that SIRS criteria are too sensitive and nonspecific for diagnosing sepsis. On the other hand, in Australian and New Zealand populations, studies have found that one patient in eight with severe sepsis was missed when SIRS criteria was used.[47–49] SEPSIS-3 defines organ failure as an increase of at least two points in the SOFA score.[45]

2.1.2.3 Acute pancreatitis

The occurrence of organ failures is the key determinant for the prognosis in severe acute pancreatitis. According to population-based studies, 8-20% of patients with acute pancreatitis get affected by organ failure, whereas the incidence of organ failure may exceed 40% in patients admitted to tertiary hospitals.[50–55] High age, alcohol aetiology of pancreatitis, high serum triglyceride levels and the extent of fat necrosis are associated with increased risk of organ failure.[56–59] The organ failures in acute pancreatitis are divided into early onset, which is usually sterile and resemble SIRS, and late onset, which is typically caused by infected pancreatic necrosis. Particularly non-transient organ failure is associated with poor prognosis. The mortality of acute pancreatitis with persistent organ failure lasting for at least 48 hours is very high, exceeding 40%, according to the recent observational trials.[50,51,60] Early-stage organ failure is associated with a slightly poorer prognosis compared to late-onset organ failure, and an increasing number of failing organ systems substantially increase the risk of death.[55,61] According to Finnish study, the respiratory failure is the most common and hepatic failure the most fatal organ failure in acute pancreatitis.[62]

The Atlanta classification determines the severity level of acute pancreatitis based on the occurrence of organ failures: mild (no organ

failure), moderate (transient organ failure), and severe (persistent organ failure lasting over 48 hours).[63]

2.1.3 Organ dysfunction scores

The first assessments of MOF severity were limited to reporting the number of organ failures. Of these preliminary scores, Knaus et al. demonstrated a strong association between the number of failing organs and mortality in their paper introducing the Organ System Failure (OSF) score (Table 1).[64] In a study population including over 5,000 admissions to surgical and mixed ICUs at 13 hospitals in the United States, mortality reached 98% for patients experiencing three or more organ failures persisting by day three (only two out of 99 patients survived). In addition to OSF, similar reports of an association between an increasing number of failing organ systems and worse outcomes were published in the 1980s prior to the publication of the present organ dysfunction scores.[65] The early scores aimed to present organ failure as a binary condition, that is, either present or absent. In contrast, the later organ dysfunction scores characterise organ dysfunction as a continuum that ranges from relatively mild to a complete breakdown of the organ system.

Table 1. The criteria of the OSF score as introduced by Knaus et al. (1985). OSF existed if at least one of the criteria was fulfilled during a 24-hour period.[64]

Cardiovascular Failure	HR ≤ 54/min
	MAP ≤ 49 mm Hg
	Occurrence of ventricular tachycardia and/or ventricular fibrillation
	Serum pH ≤ 7.24 with a $PaCO_2$ of ≤ 49 mm Hg
Respiratory Failure	Respiratory rate ≤ 5/min or ≥ 49/min
	$PaCO_2$ ≥ 50mm Hg
	$AaDO_2$ ≥ 350 mm Hg ($AaDO_2$ = 713 FiO_2 – $PaCO_2$ – PaO_2)
	Dependent on ventilator on the fourth day of OSF, e.g. not applicable for the initial 72 hours of OSF
Renal Failure*	Urine output ≤ 479ml/24 hours or ≤ 159ml/8 hours
	Serum BUN ≥ 100mg/100ml
	Serum creatinine ≥ 3.5 mg/100 ml
Haematologic Failure	WBC ≤ 1000 mm^3
	Platelets ≤ 20,000 mm
	Haematocrit ≤ 20%
Neurologic Failure	Glasgow Coma Score ≤ 6 (in absence of sedation at any one point in the day)

*Excluding patients on chronic dialysis before hospital admission

Abbreviations: AaDO2, Alveolar-arterial oxygen gradient; BUN, Blood Urea Nitrogen; FiO_2*, Fraction of inspired oxygen delivered, HR, Heart rate; MAP, Mean Arterial Pressure; OSF, Organ System Failure;* PaO_2*, Partial pressure of Oxygen in arterial blood,* $PaCO_2$*, Partial pressure of Carbon dioxide in arterial blood; WBC White Blood Cells*

During the mid-1990s, four parallel working groups undertook almost simultaneously the task of formulating more precise organ dysfunction scores. These scores had considerable similarities: all encompassed six

organ systems, used a four- or five-step scale, and were based mainly on the same variables, with slightly differing cutoffs.

Marshall et al. were the first to publish results. Their score was created through a two-stage process. In the initial phase, they conducted a systematic literature review to examine the previous measures used for MOF severity evaluation. Thereafter, based on the results of the review, they developed the Multiple Organ Dysfunction Score (MODS) (Table 2). [66,67]

Table 2. Multiorgan Dysfunction Score (MODS).[67]

Organ component	**1**	**2**	**3**	**4**
Respiratory, PaO_2/FiO_2, mm Hg	226-300	151-225	75-150	≤ 75
Renal, Serum creatinine (μmol/L)	101-200	201-350	351-500	> 500
Hepatic, Serum bilirubin (μmol/L)	21-60	61-120	121-240	> 240
Cardiovascular, PAR*	10.1-15.0	15.1-20.0	20.1-30.0	> 30.0
Haematologic, Platelet count ($x10^3/mm^3$)	81-120	51-80	21-50	≤ 20
Neurologic, GCS	13-14	10-12	7-9	≤ 6

*Pressure-adjusted heart rate: (Heart rate (HR) x (Central venous pressure / Mean Arterial Pressure)
Abbreviations: GCS, Glasgow Coma Scale; PaO_2/FiO_2, The ratio of partial Pressure of arterial Oxygen to Fraction of inspired Oxygen

During the same year, Le Gall et al. introduced the Logistic Organ Dysfunction Score (LODS) (Table 3). Among the almost simultaneously published organ dysfunction scores, LODS was the only one developed through a multiple logistic regression analysis undertaken in a study that

included 14,734 admissions to 137 ICUs in 12 countries.[68] Originally, it was designed to measure the severity of MOF on the day of arrival in the ICU, but it has also been shown to effectively quantify the severity of MOF during subsequent days.[69]

The so-called 'Brussels score' was introduced merely as an abstract without following validating studies.[70]

Table 3. Logistic Organ Dysfunction Score (LODS). [71]

Organ component	**5**	**3**	**1**	**0**	**1**	**3**	**5**
Neurologic, GCS	3-5	6-8	9-13	14-15			
Cardiovascular, HR Beats/min	< 30			30-139	> 140		
SBP, mmHg	< 40	40-69	70-89	90-239	240-269	≥ 270	
Renal, Urea nitrogen (mmol/L)				< 6 and	6-9.9 or	10-19.9	≥ 20
Creatinine (μmol/L)				< 106 and	106-140	≥ 140	
Urine Output, (Litres)	< 0.5	0.5-0.74		0.75-0.99		≥ 10	
Pulmonary, PaO_2/FiO_2, mm Hg, kPa		< 150* < 19.9*	≥ 150* ≥ 19.9*	No IPAP/ CPAP/MV			
Haematologic, Leukocytes ($x10^9$/L)		< 1.0	1.0-2.4	2.5-49.9	≥ 50.0		
Platelets ($x10^9$/L)			< 50	≥ 50			
Hepatologic, Bilirubin μmol/L				< 34.2	≥ 34.2		
PTT secs (% above standard)			(< 25%)	< 3s, (≥ 25%)	≥ 3s		

*IPAP/CPAP/MV required
Abbreviations: CPAP, Continuous Positive Airway Pressure; GCS, Glasgow Coma Scale; HR, Heart Rate; IPAP, Inspiratory Positive Airway Pressure; MV, Mechanical Ventilation; PaO_2/FiO_2, Ratio of arterial oxygen partial pressure to fraction of inspired oxygen; PTT, Partial Thromboplastin time; SBP, Systolic Blood Pressure

2.1.4 Purpose and definition of the SOFA score

The SOFA score originated from the consensus meeting of a working group organised by the European Society of Intensive Care Medicine (ESICM) in Paris in October 1994.[8] The SOFA score was not primarily designed as a mortality prediction model. Instead, its primary objective was to quantify the severity of MOF and identify the number of failing organ systems, particularly in sepsis patients. Consequently, the initial title of the scoring system was Sepsis-related Organ Failure Assessment. Later, the score's title was changed to the Sequential Organ Failure Assessment score. According to the working group, the ideal scoring system for this purpose should exhibit the following characteristics:

- Objectivity: Inter-observer variability should be minimal.
- Simplicity: The scoring system should be easy to understand and implementable across different healthcare settings.
- The variables (physical parameters or blood samples) should be routinely obtained in most healthcare institutions.
- Organ-specificity: The organ-specific scoring should evaluate the function of the organ system.
- The variables should be continuous.
- Generalisability: The scoring system should be applicable across diverse patient groups (for example, different demographics and clinical characteristics).
- Independence from therapeutic interventions.

The working group integrated six organ systems (respiratory, coagulation, hepatic, cardiovascular, neurologic, and renal) into the scoring system (Table 4) and recognised the omission of the gastrointestinal

system as a limitation . Devising a simple and straightforward method for quantifying gut failure turned out to be impossible.

Following the SOFA score publication, the working group conducted a prospective validation study to assess its utility in 40 ICUs in 16 countries.[22,72] Here, the working group defined organ dysfunction as 1–2 organ-specific subscores, while organ failure was defined as obtaining 3–4 organ-specific points. In addition to the increase in number of failing organ systems, an increase in the maximum total SOFA score during the ICU stay was associated with increased mortality. In patients with a maximum score of 10 or more SOFA points, mortality was over 40%. As a conclusion, the working group further suggested that, in addition to measuring the severity of MOF, the change in the SOFA score could be used as an outcome in clinical trials.[22]

The SOFA score was intended to be dynamic, that is, the results would evolve during treatment and the worst value for each organ system would be recorded daily. Several derivations of the SOFA score have been suggested, including admission SOFA, daily maximum, study period maximum, mean SOFA of the study period, and change in SOFA during the study period (ΔSOFA). In a large meta-analysis investigating the relationship between SOFA and mortality in 87 RCTs, ΔSOFA was found to be more associated with mortality than the fixed-day SOFA score.[73] According to a prospective single-centre study by Ferreira et al., a positive ΔSOFA (increasing SOFA points) during the first 48 hours predicted at least 50% mortality.[74] However, an earlier multinational observation by Moreno et al. showed that the admission SOFA score is at least as prognostic as ΔSOFA. This finding was confirmed by a recent observation of more than 20,000 ICU patients.[23,75]

Table 4. SOFA score as introduced by Vincent et al.[8]

Subscore	1	2	3	4
Respiration PaO_2/FiO_2, mm Hg	< 400	< 300	< 200*	< 100*
Coagulation Platelets $x10^3/mm^3$	< 150	< 100	< 50	< 20
Liver Bilirubin, mg/dl (µmol/l)	1.2-1.9 (20-32)	2.0-5.9 (33-101)	6.0-11.9 (102-204)	> 12.0 (> 204)
Cardiovascular Hypotension or vasopressor doses	MAP < 70mm Hg	Dopamine ≤ 5 or dobutamine at any dose†	Dopamine > 5 or adrenaline < 0.1 or noradrenaline < 0.1†	Dopamine > 15 or adrenaline > 0.1 or noradrenaline > 0.1†
Neurologic GCS	13-14	10-12	6-9	< 6
Renal Creatinine, mg/dl (µmol/l) or urine output	1.2–1.9 (110–170)	2.0–3.4 (171–299)	3.5–4.9 (300–440) or < 500 ml/day	> 5.0 (> 440) or < 200 ml/day

*With respiratory support
†Adrenergic agents administered for at least 1h (doses given are in µg/kg/min)
Abbreviations: GCS, Glasgow Coma Scale; PaO_2/FiO_2, Ratio of arterial oxygen partial pressure to fractional inspired oxygen

2.1.5 Organ dysfunction score comparisons

The SOFA, MODS, LODS, and Brussels score use almost the same criteria to describe the failure of five organ systems. The respiratory component is based on the PaO_2/FiO_2 ratio, the haematologic component on platelet count, the hepatic component on bilirubin concentration, the neurologic component on Glasgow Coma Scale (GCS), and the renal component on

creatinine concentration (or urine output). However, some of the criteria used differ across scores, particularly in cardiovascular scores. In MODS, the cardiovascular criterion is based on a calculation that the authors call the pressure-adjusted heart rate (PAR). The PAR involves multiplying the heart rate by the ratio of the central venous pressure to the mean arterial pressure. Cardiovascular SOFA score involves blood pressure and dosages of vasopressors, cardiovascular LODS heart rate and blood pressure, and cardiovascular Brussels score blood pressure, fluid responsiveness, and pH.[72]

The results of observational studies comparing MODS, LODS, and the SOFA score in prognostication are conflicting. In a Belgian prospective single-centre study, the predictive ability of the first 24-hour SOFA score was slightly better than that of the MODS (Area under the receiver operating characteristic curve [AUROC] for SOFA 0.872 vs. AUROC for MODS 0.856) [76]. The superiority of the SOFA score to the MODS has also been shown in patients with traumatic brain injury.[77] In an examination of German trauma patients admitted to ICU, MODS was superior in predicting mortality, while the SOFA score demonstrated better predictive ability for the length of stay (LOS) in ICU.[78] In a French multicentre study comparing the SOFA and LODS, both scores showed good consistency without significant difference in performance.[79]

2.1.6 Components of the SOFA score

2.1.6.1 Respiratory system

The classification of severity of respiratory failure is based on the ratio of the partial pressure of arterial blood oxygen to the fraction of oxygen in inhaled gas (PaO_2/FiO_2). The respiratory failure classification of the SOFA score is almost identical to the classification of the severity of Acute Respiratory Distress Syndrome (ARDS), as determined in the Berlin definition.[80] In the Berlin definition, $PaO_2/FiO_2 \leq 300$ mm Hg is considered mild, $PaO_2/FiO_2 \leq 200$ mm Hg moderate, and $PaO_2/FiO_2 \leq 100$ mm Hg severe ARDS. The definition requires positive end-expiratory pressure (PEEP) or continuous positive airway pressure (CPAP) of greater than or equal to 5 cm H_2O.[80] The SOFA score requires mechanical ventilation in respiratory subscore categories of 3 and 4.

The severity of respiratory failure in the SOFA score focuses on hypoxemic gas exchange dysfunction (type 1 respiratory failure), disregarding the type 2 failure (failure to remove CO2), and level of breathing exhaustion. Moreover, the respiratory component of the SOFA score does not take into account the use of high-flow nasal cannula (HFNC) in less severe cases and veno-veno extracorporeal membrane oxygenation (VV-ECMO) in the most extreme cases of respiratory failure. The utilisation of both treatments has significantly increased, especially during the Covid-19 pandemic.[81–84]

2.1.6.2 Coagulation system

The coagulation pathway is a cascade of events responsible for preventing haemorrhages and, conversely, inhibiting thromboses and clot formation [85]. In the SOFA score, the functionality of the coagulation system is measured by thrombocyte levels. In critically ill patients, the prevalence of thrombocytopenia (platelet count < $150x10^9$) ranges from 35% to 44%.[86–88]

In sepsis, the decrease of platelet product is combined with increased platelet consumption.[89] These together lead often to thrombocytopenia. Some MOF-triggering conditions requiring intensive care affect platelets directly. These include thrombotic microangiopathy: thrombotic thrombocytopenic purpura (TTP) and haemolytic uremic syndrome (HUS).[90] Iatrogenic causes of thrombocytopenia may also occur during intensive care. Heparin-induced thrombocytopenia (HIT) occurs in 0.5–1% of patients receiving unfractionated heparin treatment. Previous observational studies have reported HIT-associated mortality up to 30%.[91–93] Furthermore, thrombocytopenia is often associated with sepsis, especially gram-negative pathogens.[94]

Disseminated intravascular coagulation (DIC) is commonly associated with severe sepsis, but it can also be linked to various conditions, including malignancies, obstetrical complications, and trauma, particularly head trauma.[95] No single diagnostic definition of DIC exists, but the diagnosis is typically made based on suspicion and the use of various diagnostic scoring systems.[96,97] DIC is associated with significant morbidity, and the mortality is high, up to 50%.[98]

Focusing solely on platelet levels may oversimplify the concept of coagulation, but its advantage lies in its wide availability, at least in high- and middle-income countries. A more comprehensive assessment of the coagulation system would require the measurement of both the intrinsic and extrinsic pathways as well as access to more advanced laboratory facilities.[99]

Viscoelastic measurements of whole blood using rotational thromboelastography (TEG) and thromboelastometry (ROTEM) have shown promise as methods for measuring coagulation status. In a German cohort study involving 98 sepsis patients admitted to ICUs, coagulopathy measured using thromboelastography at admission even outperformed the Simplified Acute Physiologic Score (SAPS) II and SOFA scores in predicting mortality.[100] However, it is important to note that these tests currently lack proper standardisation, and many studies examining their efficacy have methodological flaws.[101]

2.1.6.3 Hepatic system

Hepatic failure is the least frequent of the organ failures included in the SOFA score.[22,102] One possible reason for this is the ability of hepatocytes to regenerate to some extent in non-fulminant sepsis, which provides a certain level of tolerance to biological insults. [26,103,104] Measuring hepatic dysfunction or failure, in the SOFA score, is based on serum bilirubin concentration.

Hepatic failure can be classified by its development window: acute liver failure (ALF), chronic liver failure (CLF), and acute-on-chronic-failure (ACLF). ALF refers to acute onset development of jaundice (hyperbilirubinemia), impairment in synthetic function (international normalised ratio [INR] >1.5), and hepatic encephalopathy in patients without a previous history of liver disease. According to O'Grady's definition, acute onset of hepatic failure is characterized by a period of 4 weeks from jaundice to encephalopathy, while subacute onset is defined as occurring within 5-12 weeks. [105,106] ALF is a rare cause of ICU admission. The incidence of ALF likely varies depending on the geographic location. In epidemiological studies, the incidence of ALF has ranged from less than 10 cases per million inhabitants per year in the United States and Europe to 69 cases per million

inhabitants in Thailand.[107–109] Moreover, the aetiology of ALF differs. Drug-related causes are more common in high-income countries, while viral aetiology is more frequent in low- and middle-income countries.[109]

The concept of ACLF was introduced and defined by the European Association for the Study of the Liver-Chronic Liver Failure (EASL-CLIF) consortium in 2013. ACLF is defined as the sudden hepatic decompensation observed in patients with pre-existing chronic hepatic disease. In contrast to acute hepatic decompensation, ACLF is characterised by the new onset of organ failures determined by a marginally modified SOFA score, known as CLIF-SOFA. The grading of ACLF depends on the number of failing organ systems: grade 1 (one failing organ system), grade 2 (two failing organ systems), grade 3 (three or more failing organ systems) in addition to the hepatic system. In a validation study of ACLF in European patients, grade 1 ACLF was associated with a 28-day mortality of 22%, grade 2 with 52%, and grade 3 with 76%.[110]

Hepatic function plays a pivotal role in sepsis patients. The mortality in sepsis patients with hepatic failure is high, ranging from 54% to 68%.[111] In the French EPISEPSIS study, incidence of hepatic dysfunction and hepatic failure was documented in 47% and 6% of sepsis patients, respectively.[112]

2.1.6.4 Cardiovascular system

Circulatory shock is common in patients admitted to ICUs. It is characterised by 1) hypotension (in adults, mean arterial pressure below 70 mm Hg or a systolic pressure below 90 mm Hg); 2) clinical signs of organ hypoperfusion (skin mottling, decreased urine output, and deteriorated level of consciousness); and 3) hyperlactatemia as a sign of disturbed cellular metabolism. The principal causes of hypotension are hypovolemia, distributive shock (for example in sepsis), anaphylaxis, obstructive shock, and cardiac dysfunction. In many cases, the aetiology is a combination of these.[113]

In the SOFA score, the cardiovascular component is measured by assessing hypotension (MAP above or below 70 mm Hg) and the need for vasopressors or inotropes to maintain normotension. There are some concerns regarding this determination.

It is uncertain whether the MAP target of > 70 mm Hg should be pursued. In a British trial involving ICU patients above the age of sixty-five

and a Danish trial in post-cardiac arrest patients, the pursuit of MAP > 65 mm Hg with the use of vasopressors did not lead to a reduction in mortality.[114,115]

The recommended MAP target in guidelines may vary significantly from 70 mm Hg, depending on the underlying disease. For instance, in the acute phase of traumatic spinal cord injury, the current approach is to maintain a MAP > 85 mm Hg to enhance spinal cord perfusion.[116]

The approach to managing hypotension has evolved since the introduction of the SOFA score. In the past, it was more common to administer significant amounts of fluids to patients before initiating vasoactive infusions. The shift towards more restrictive fluid resuscitation has resulted in a more liberal use of vasoactive medications.[117]

It is possible that changes in clinical practice have influenced the functionality of cardiovascular component. Dopamine, dobutamine, noradrenaline (norepinephrine), and adrenaline (epinephrine) are the agents considered in the cardiovascular SOFA score, but their clinical use has changed remarkably. Moreover, several new vasopressors have entered the market, and several old vasopressors have been reintroduced into practice that are not considered by the cardiovascular SOFA score.[118,119]

According to the SOFA score definition, the occurrence organ failure is assumed to be independent of therapeutic interventions. However, many sedative agents, including propofol, are associated with significant hypotension. Additionally, several inotropic agents, for example milrinone, dobutamine, and levosimendan, have vasodilating effects. Use of these agents in low cardiac output event, increases the likelihood of initiating vasopressor agents, such as noradrenaline.[120]

2.1.6.5 Neurologic system

The assessment of the central nervous system in the SOFA score is directly derived from the GCS. The GCS was proposed by Jennett and Teasdale in 1974 to evaluate the severity of impairment of level of consciousness in patients with acute brain injury.[121] Remarkably, for half a century, the GCS has withstood the test of time, serving as the primary method to assess the level of consciousness and as a communication tool

among professionals [121]. Alongside the SOFA score, LODS, and MODS, it is a major component of ICU mortality prediction models.[122,123]

GCS assesses patients' behavioural responses to various stimuli. It is calculated as the sum of three sections: motor response, verbal response, and eye opening (Table 5). In sedated patients, GCS cannot be assessed. In the SAPS II severity-of-illness score, the guidance is to evaluate the level of consciousness before the initiation of sedation and to record the corresponding GCS score.[122]

Table 5. The Glasgow Coma Scale (GCS) as introduced by Jennett and Teasdale.[121]

Score	Eye response	Motor response	Verbal response
6		Obeys commands	
5		Localises pain	Oriented
4	Eyes open spontaneously	Withdraws from pain	Confused
3	Eye opening to verbal command	Flexion response to pain	Inappropriate words
2	Eye opening to pain	Extension response to pain	Incomprehensible sounds
1	No eye opening	No motor response	No verbal response

It has been suggested that in patients with head trauma, the motor component alone can predict patient survival as well as or even better than the entire GCS. Hence, for simplicity, it has been proposed that the eye and verbal components could be omitted, as is done, for example, in the International Mission for Prognosis and Analysis of Clinical Trials (IMPACT) score in the determination of prognoses of patients with traumatic brain injury.[124,125]

Of the SOFA components, the neurological component is the least objective[126] primarily due to the relatively high inter-rater variability of evaluators.[127] In addition, evaluation of GCS may be challenging in sedated patients, particularly those who are orally intubated.

2.1.6.6 Renal system

Acute kidney injury (AKI) affects approximately 30–50% of the patients admitted to ICUs.[128,129] The SOFA score criteria for renal failure are based on serum creatinine and daily urine output (dUO). These criteria differ partially from the more recent definition of Kidney Disease: Improving Global Outcomes (KDIGO) guidelines, which are widely adopted for defining AKI.[130] For example, the initiation of renal replacement therapy (RRT) is not considered in the SOFA score definition.[131] It should be noted that in AKI (or, in the context of the SOFA score, renal organ dysfunction), the two determinants urine output and creatinine increase are not fully comparable. Renal failure is most frequently determined by the criterion of low urine output alone, but low urine output is less associated with the need for RRT and mortality than the fulfilment of the creatinine level criterion alone or combined with the urine output criterion. The decrease in urine output may sometimes indicate a physiological response to acute illness rather than signalling organ failure.[132]

Besides creatinine and blood urea nitrogen (BUN), which is also widely used in measuring kidney function, there are several novel biomarkers that are very sensitive and detectable in the earlier stages of AKI. The consensus statement of the Acute Disease Quality Initiative (ADQI), a network of highly recognised experts in various acute care specialisations, encourages the implementation of novel biomarkers or their combinations in the prevention and management of AKI.[133] These novel biomarkers include Neutrophil gelatinase-associated lipocalin (NGAL), Urinary Tissue Inhibitor of Metalloproteinase-2 (TIMP-2), Insulin-Like Growth Factor-Binding Protein 7 (IGFBP7), cystateine C, Kidney Injury Molecule-1 (KIM-1), growth differentiation factor-15 (GDF-15), and calcoprotectin.[134–137]

The mechanisms of AKI development are miscellaneous and complex but not irrelevant in terms of recovery. Recently, ADQI has proposed a novel concept based on AKI pathophysiology, namely, the Sepsis-Associated Acute Kidney Injury (SA-AKI) and Sepsis-Induced AKI (SI-AKI).[138] [139]. Moreover, ADQI has proposed another novel concept of drug-induced kidney injury to move towards a more pathophysiology-based categorisation of AKI.[132]

2.1.7 Vasoactive Inotropic Score

The cardiovascular system, given its complexity, poses challenges to adopting a straightforward approach to quantifying the level of dysfunction. One commonly adopted method involves assessing the level of pharmacological support necessary to achieve an appropriate blood pressure level.

Wernovsky and his colleagues introduced the idea of measuring the amount of inotropic support as a prognostic tool for neonatal cardiac surgery patients.[140] Wernovsky et al. incorporated dopamine, dobutamine, and adrenaline in μg/kg/min (with adrenaline weighted by a factor of 100) in the score formulation. Gaies et al. added a range of the most used vasopressors and inotropes: noradrenaline with a factor of 100, vasopressin with a factor of 10,000, and milrinone with a factor of 10 in the formulation.[141] The new score was entitled the Vasoactive Inotropic Score (VIS).

Gaies et al. introduced VIS in a retrospective study that observed 174 infants undergoing heart surgery with cardiopulmonary bypass. The maximum and mean values during the 48 hours following surgery were associated with the composite outcome of in-hospital or 30-day mortality, cardiac arrest, mechanical circulatory support, the need for renal replacement therapy, and neurologic injury. Compared to patients with low maximum VIS, patients with high maximum VIS had an eight-fold increased risk of a poor outcome.[141] Subsequently, Gaies et al. validated the VIS in a larger prospective multi-centre cohort study comprising 391 cardiac surgery infants and found a similar association with poor outcomes: high-VIS patients had a six-fold increased risk of poor outcomes compared to low-VIS patients.[142]

VIS is associated with various adverse outcomes beyond increased short-term mortality. These include prolonged mechanical ventilation, extended ICU stays, prolonged hospital stays, AKI, sepsis, and increased healthcare costs. Most of the observations stem from relatively small, single-centre studies.[143–147] Since the VIS introduction study by Gaies et al., several other vasoactive drugs have been included in modified versions of the score, including levosimendan, phenylephrine, enoximone, and olprinone.[148–150] Outside cardiac surgery, VIS has been associated with mortality and other adverse outcomes in paediatric and adult sepsis populations in an increasing number of observational studies.[151–156]

In adult cardiac surgery patients, the most well-established validation study to date was conducted by Koponen et al. in 2019. A retrospective single-centre study investigated the relationship between the VIS and a composite outcome that included 30-day mortality, mediastinitis, stroke, AKI, and myocardial infarction in adult postoperative cardiac surgery patients.[157] The results demonstrated the association of higher VIS with increased mortality as well as prolonged ICU stay and need for mechanical ventilation in patients undergoing coronary artery bypass surgery. A smaller study by Yamazaki et al. found comparable results.[158]

There are several ways to determine the VIS. The most common method has been to determine the maximum VIS during the 24–48 hours following admission to the hospital, ICU, or operation theatre.[159] Some further variations of VIS have also been proposed. Bangalore et al. suggested a score that considers the duration of time the patients have been exposed to the drug. However, no further validation studies have been conducted on this proposal.[160]

2.2 BENCHMARKING THE ICU PERFORMANCE

2.2.1 Quality in healthcare

It has been estimated that up to 9 million patients die annually due to poor-quality healthcare, constituting 15% of the total deaths in low- and middle-income countries.[161] In high-income countries, the issue of healthcare quality is also significant. According to systematic review by De Vries et al., one out of ten patients admitted to care has been harmed during hospitalisation. Up to half of these incidents were preventable.[162] Medical errors and shortcomings in quality of care and patient safety lead to significant increases in healthcare costs. The World Bank has reported that 15% of hospitalisation expenditures result from preventable harm caused to patients. Investing in the quality of care could not only prevent such incidents but also save money.[163]

Although only around 3% of the incidents recorded by De Vries et al. occurred in the ICU (compared to 41% in the operating theatre and 25% in patient rooms on the ward), ICUs cannot be considered entirely safe in terms of patient safety.[162]

2.2.2 ICU performance

Benchmarking is a concept used in the industrial field that has been adapted to the healthcare environment. A well-known example of this is the 'Leadership through Quality' programme run by David T. Kearns, Chief Executive Officer of Xerox, in the 1970s. Benchmarking is a strategic management tool used to evaluate and improve performance by comparing it to best practices or standards[164], and it can be divided into two broad categories. Internal benchmarking takes place within a single institution and can be performed, for example, to follow the sequence of specific administrative interventions. However, the more common method is external benchmarking, in which the results of performance metrics are compared to those of peer institutions. Trends in the quality metrics and distinct outlier units should raise suspicions on the quality deviation.[15]

The most important mission of intensive care is to prevent patients with acute life-threatening diseases from dying. Thus, mortality is the most important quality indicator of the ICU. However, comparing crude

mortalities in different units does not make sense if there are significant differences in populations (often referred to as case-mix differences).[165] The mortality at an ICU that treats the most severely ill patients is assumed to be higher than that at an ICU admitting less severely ill patients. Therefore, mortality results must be standardised. Standardised mortality ratios (SMRs) are calculated by dividing the observed mortality by the expected mortality, which is determined by a specific mortality prediction model. Hence, the lower the SMR of the unit, the better is the unit´s performance compared to its peers.[166–168]

The ESICM has regarded external benchmarking, which involves comparing SMRs and auditing ICUs, as the primary tool to improve the quality of intensive care.[169] Numerous ICU benchmarking programmes around the world use SMRs as the primary method to compare performances. National or regional intensive care registries are mainly used as the primary sources of the data.[170–173]

Standardised mortality is not the only metric used to determine the quality of intensive care. Depending on the ICU stakeholder, a variety of estimates of ICU performance are of interest. From both political and healthcare funding perspectives, cost of treatment and cost-effectiveness are of high importance. Meanwhile, the relatives of patients appreciate effective and understandable communication skills from caregivers. Further examples of additional benchmarking variables are the effective utilisation of resources, efficient staff utilisation, patient satisfaction, complications during the ICU stay, adherence to best practices, and the length of ICU stays.[15] There are some metrics that the ESICM has raised specifically as a measure of quality in intensive care. Those include, in addition to SMRs, incidences of readmission rate within 48 hours of ICU discharge, catheter-related blood-stream infections, and unintended extubations.[169] In a review by Flaatten, 63 different ICU quality indicators were found to be used around the world, with 26 in use in more than one country.[174]

2.2.3 Principles of mortality prediction

Mortality prediction models are crucial parts of SMR calculations and ICU benchmarking. It is important to recognise that prediction models are

intended for use at the population level rather than the individual level.[175] Forecasting the outcome of a heterogeneous population, such as patients admitted to the ICU, is complicated. Therefore, the variables included in a good mortality prediction model must be selected carefully.

Firstly, age and the presence of concurrent frailty-inducing chronic diseases have a substantial impact on the probability of survival in critical care.[176] Increasing age increases the risk of both short-term and long-term mortality. Apart from age, frailty decreases but female gender increases the probability of surviving critical care.[177–179]

Secondly, the circumstances and causes of admission have an impact on survival probability. Various admission diagnoses are associated with entirely different prognoses. Furthermore, medical admissions compared to surgical ones, and emergency admissions compared to elective ones, are associated with an increased risk of mortality.[180–182]

Thirdly, the severity of the acute disease matters. Numerous physiological measures serve as surrogates for abnormal organ functions, for example, lactate levels, blood pressure, tissue oxygenation measured through various methods, and level of consciousness. These factors should be carefully assessed when creating the prediction model.

From a statistical perspective, a prediction model must meet a couple of essential requirements to perform well. Firstly, the model should have good discrimination ability: it should be able to distinguish patients with a poor prognosis from those with a good prognosis. Discrimination is commonly reported as the AUROC (or concordance statistics [c-statistics], which is equal to the AUROC).

Secondly, the predictive ability of the model may not be equally good throughout the stratum. For example, the model may accurately estimate low risk but over- or underestimate high risk. This is referring to poor calibration of the model. Despite its necessity in terms of the reliability of the model, the calibration is often overlooked in the model developments.[183,184]

A working group of recognised statisticians, epidemiologists, methodologists, clinicians, and medical journal editors has published the 'Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis' (TRIPOD) statement, a 22-item guideline for prognostic model development and reporting.[185,186] Predictive models

become outdated relatively fast. Hence, the models require frequent recalibration.[187]

2.2.4 Prediction models in intensive care

In the latter half of the 1970s, intensive care physicians William A. Knaus and Jack E. Zimmerman were concerned about the quality of their ICU in George Washington University Hospital in Washington, United States. More importantly, they were concerned on how its quality compared to peer ICUs. For standardised method of ICU performance reporting, the first version of APACHE was published in 1981 by Knaus et al. The APACHE evaluation was validated in a study that considered 805 admissions to two ICUs .[10,188]

There were two distinct components in the original APACHE score: the factors that existed before the admission indicating the patient's premorbid capability of recovering from the acute disease (pre-admission health status with a four-step evaluation by reviewing the patient's health history) and the severity of acute illness indicated by the Acute Physiology Score (APS), which is a sum of 32 physiological variables weighted with a scale of 0 to 4 and measured during the 24 hours after ICU admission. Following findings from single studies indicating that neurological abnormalities were underweighted in the APACHE score, this prototype score was updated after only four years to yield a refined and remarkably simplified APACHE II score.[189]

Since the introduction of APACHE II, neurological deficits have gained major (usually the most significant) weight in prediction models.[190,191] In APACHE II, the number of physiological variables was reduced from 32 to 12. APACHE II also introduced categorisation by cause of ICU admission (APACHE II diagnosis) in addition to the type of elective/emergency admission and medical/surgical admission. The APACHE II categorisation has been widely adopted, particularly in study population descriptions. Hence, the 1985 paper introducing APACHE II has become the most cited study in the critical care field.[192] APACHE has been further customised multiple times as APACHE III (1991), APACHE III-I (1998). APACHE IV (2006). APACHE III has been frequently recalibrated in Australia and New Zealand.[172]

'The European counterpart' of APACHE followed some years later, when the SAPS was presented and validated with 679 admissions to eight French ICUs. As its name suggests, SAPS is a simplified version of the APS from the original APACHE and comprises 13 APS values in addition to age.[11]

The second version, SAPS II, was introduced in 1993 through an international development and validation study conducted in 137 ICUs across 12 countries, including Finland. SAPS II included a total of 17 variables: 12 physiological variables, and admission type and the presence of three chronic diseases (acquired immunodeficiency syndrome [AIDS], metastatic cancer, and haematologic malignancy) (Table 6).[193]

The 'second-generation' prognostic scores – APACHE II, APACHE III, and SAPS II – were still claimed to exhibit inaccuracies and calibration flaws, underestimating the risk of mortality in low-risk patients, and overestimating it in high-risk patients.[194–196]

SAPS 3 was an aim to improve the prognostication with completely new model, which would also take into account the circumstances preceding ICU admission and that would be less influenced by treatment effects.

In SAPS 3 validation, efforts were made to increase geographical heterogeneity by extending to Europe, North America, Central America, South America, and Australasia. The physiological variables were aimed to be recorded within one hour (before or after) of ICU admission to mitigate the impact of treatments. In SAPS II and APACHE II, the physiological values were determined as the most abnormal ones during the first 24 hours after ICU admission. SAPS 3 provided a prediction model that considered estimates from three different areas. The admission-preceding variables were the presence of chronic diseases (such as cancer therapy, advanced chronic heart failure, cirrhosis, AIDS, and metastatic cancer), length of stay at the hospital before ICU admission, the ICU dispatching location, and the use of major therapeutic medications before ICU admission (vasopressors). The admission-related variables were admission type (elective/emergency), the type of possible preceding surgery (emergency/elective, anatomical site), and the type of possible infection at admission (nosocomial, respiratory). The physiological variables followed those in APS.[197,198] However, despite these refinements, observational trials have not proven the performance of SAPS 3 to be significantly superior to SAPS II.[199,200]

The MPM was published in 1985, the same year as APACHE II. The model comprised seven variables related to admission and seven related to treatments and physiology.[12] The MPM branched into two derivatives: MPM_0 assesses the prognostic variables at the ICU admission, whereas MPM_{24} assesses those 24 hours after admission. The variable included in MPM_0 and MPM_{24} slightly differ. Seven variables are considered at both time points. To assess the risk development beyond 24 hours, prognostic evaluations at 48 and 72 hours (MPM_{48}, and MPM_{72}) after the admission were also involved in the refined versions of MPM.[201] Two main updates of MPM have been released (MPM II in 1993 and MPM III in 2007) (Figure 2).[202,203]

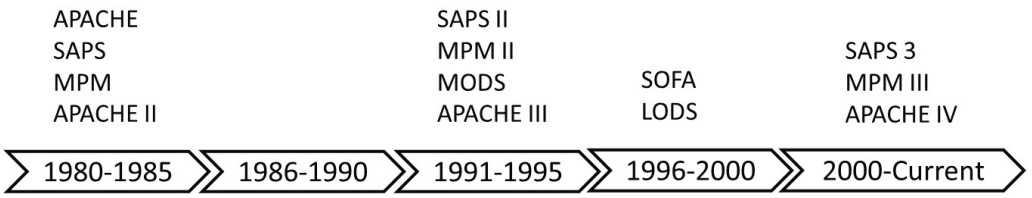

Figure 2. Generations of common severity-of-illness scores and mortality prediction scores.

One of the most significant differences between the models is the time point at which the measurement variables are recorded. MPM_0-III and SAPS 3 consider the factors one hour prior to ICU admission (emergency department, operation ward, etc.). These variables may provide a considerable amount of prognostic information about the circumstances under which the admission was made. On the other hand, the need for variables prior to admission may result in a significant amount of missing data, as data collection is likely to be less structured. However, all these mortality prediction models were created in a 'static' manner, meaning the models did not have the dynamic dimension. This is notably different from, for example, the SOFA score, which can be calculated daily, and the result evolves during the ICU stay.

Several studies have been conducted to compare the discrimination ability, and calibration of the prediction models. Kuzniewicz et al. studied 11,300 patients admitted to 35 Californian ICUs and noted that there were substantial differences between the units in SMRs across the ICUs studied.

These differences in SMRs persisted and were minimally affected by whether the mortality prediction was conducted with APACHE IV, SAPS II, or MPM_0 III.[204] In other studies, most of which compare different versions of APACHE scores to a variety of other prediction scores, the accuracy of the scores has been evaluated in different populations. Ultimately, none of the studies has been comprehensively able to prove the superiority of one mortality prediction score over the others that are commonly used.[205–208] Thus, regional customisations of the prediction scores are imperative for their reliability.[209] The SMR results are also markedly dependable on the prediction model used. The agreement of SMRs calculated by APACHE IVa and ICU Outcomes Mortality Model was studied in observation of 47 ICUs in the USA. There was concordance of the SMRs within confidence intervals in only 45% of the ICUs depending on which prediction model was used.[208]

Table 6. Simplified Acute Physiology Score (SAPS) II.[122]

	Variable	Value (Score)			
Acute Phyisologic Score	Heart Rate (BPM)	<40 (+11)	40-69 (+2)	120-159 (+4)	≥160 (+7)
	SAP (mm Hg)	<70 (+13)	70-99 (+5)	≥200 (+2)	
	Temperature (°C)	≥39 (+3)			
	PaO_2/FiO_2 (mm Hg)	<100 (+11)	100-199 (+9)	≥200 (+6)	
	Serum Urea	10-29.6 (+6)	>30 (+10)		
	Urine Output (ml/d)	<500 (+11)	500-999 (+4)		
	WBC ($x10^3/mm^3$)	<1.0 (+12)	≥20.0 (+3)		
	Potassium (mmol/L)	<3.0 (+3)	≥5.0 (+3)		
	Sodium (mmol/L)	<125 (+5)	≥145 (+1)		
	HCO_3 (mmol/L)	<15 (+6)	15-19 (+3)		
	Bilirubin (µmol/L)	<68.4-102.5 (+4)	≥102.6 (+9)		
	GCS	<6 (+26)	6-8 (+13)	9-10 (+7)	11-13 (+5)
Age (Score)	<40 (+0)	**Chronic Disease** (Score)		**Admission type** (Score)	
	40-59 (+7)	Metastatic cancer (+9)		Medical (+6)	
	60-69 (+12)	Haematologic malignancy (+10)		Unscheduled surgery (+8)	
	70-74 (+15)	AIDS (+17)		Scheduled surgery (+0)	
	75-79 (+16)				
	≥80 (+18)				

Abbreviations: AIDS, Acquired Immune Deficiency Syndrome; BPM, Beats Per Minute; GCS, Glasgow Coma Scale; PaO_2/FiO_2, Ratio of arterial oxygen to fraction of inspired oxygen; SAP, Systolic Arterial Pressure; WBC, White Blood Cell

2.2.5 Sources of bias in ICU benchmarking

It is essential that the prediction model score is optimally calibrated for the case-mix of the ICUs it is measuring. Otherwise, it may overestimate the risk for some populations while underestimating it for others, or these discrepancies can occur for the entire population. Some examples of this can be seen in observations from a study of Californian ICUs, in which all common prognostic scores (APACHE IV, SAPS II, and MPM_0 III) underestimated the risk for pulmonary patients, as well as in two studies from the UK in which APACHE II, APACHE III, and SAPS II underestimated the mortality of the entire population.[204,210,211] The calibration of a model may become poor if the model is applied to predict mortality in a different population than the one for which it was initially calibrated because of differences in case-mixes.[208,212] A model calibrated for a large teaching hospital may not be usable in a small regional hospital. On the other hand, the model can lose its calibration even when used within the same group of hospitals if there are changes in admission policies (resulting in the disappearance of some patient groups) or significant advancements in the treatment of certain diseases, leading to improved prognoses.[213–216]

Patient transfers between ICUs have the potential to cause bias in the SMRs. When patients are transferred from smaller ICUs to larger ones due to the wider availability of resources, they are most likely among the most severely ill, and the risk of short-term post-discharge mortality is high. This transfer bias is likely to cause unfairly poor performance results for the larger ICUs and improve the results of units from which the patients are transferred from.[217,218] A Monte Carlo simulation study conducted on 131,618 patients admitted to 104 American ICUs demonstrated that an increase in patient transfers by 2% and 6% resulted in SMR decreases of 0.10 and 0.14, respectively, in the overall population, proving the transfer bias significant.[219] In general, if available, fixed-day observation times (for example, 30-day mortality) should be used instead of in-hospital mortality.[220–222]

Treatment limitations have recently been proven to be major confounding factors for SMRs.[223] An admission of dying patient for the sole purpose of potential organ donation is an extreme treatment limitation

and is likely to influence the benchmark results. The extent of this bias has not been evaluated to date.

Overall, there is need for caution when interpreting the ICU benchmark results measured by SMRs. In addition to the many sources of bias, random effects play a major role in ICU benchmarking.[224] The greater the heterogeneity among the ICUs, the more caution should be exercised in interpretation.[225] Even when the prediction model is optimally calibrated, a major imbalance between high-risk and low-risk patients distributed among the ICUs may lead to misleading SMR results. This statistical confounder, known as Simpson's paradox, refers to a phenomenon in which findings, such as SMR results, show a particular tendency within multiple subgroups, but this tendency disappears or even reverses when the data from the subgroups are aggregated.[226]

2.2.6 Brain death and organ donation at the ICU

In 1968, a committee from Harvard Medical School published a paper aiming to define the concept of brain death. This publication has since been recognised as the Harvard Report. The Harvard Report suggested that the irreversible cessation of brain function could be used as a determination of death, similar to the end of circulation.[227,228] This broadening of the concept of death significantly aided the activity of organ transplantation, which had commenced a few years earlier. Finland was the first country in the world to adopt this determination of death in its legislation.[229]

Currently, patients with devastating and untreatable brain injury can be admitted to the ICU with the intention of treating them as potential organ donors (PODs). The prognosis for these patients is considered hopeless. There is an anticipation of imminent cessation of all brain functioning and brain death. Therefore, the only purpose of the ICU admission for these patients is to maintain the organ viability for potential transplantation.[230,231]

With an ageing population and increasing rates of end-stage renal and liver diseases, the need for transplantable organs is consistently increasing.[230,232] In Europe, the rate of organ donation varies significantly between countries, ranging from three donations per million inhabitants in Romania to 48 in Spain.[233] The proportion of PODs among all ICU patients

is unknown. In Canada, it has been estimated that approximately 2% of the ICU population is treated with the intention of potential organ donation.[234] However, it is possible that this rate is increasing. In 2021 alone, there was over 10% increase in the organ transplantations according to the global database on donation and transplantation.[235] This trend poses challenges for ICU benchmarking, as the objectives for PODs differ substantially from those of ICU patients whose lives are still deemed salvageable.

3 AIMS OF THE STUDY

The purpose of this study was to find out whether the component scores of the original SOFA score are still valid measures of organ dysfunction, whether severity of cardiovascular dysfunction can be better measured based on doses of vasopressor medications, and whether potential organ donors affect the SMRs of the ICUs.

The specific research questions were:

1. Do the organ-specific subscores of the SOFA score equally prognosticate mortality?

2. Is VIS a valid prognostic measure of cardiovascular dysfunction in the general ICU population, and does replacing the cardiovascular component with a VIS-based score enhance the accuracy of the SOFA score?

3. Is it reliable to determine the severity of cardiovascular dysfunction by using only the level of noradrenaline dose needed? Furthermore, what would be the optimal cutoffs for determining low, intermediate, and high dose noradrenaline dose?

4. Do patients admitted to ICUs for the sole purpose of possible organ donation after brain death affect the SMRs and ICU benchmarking?

4 MATERIALS AND METHODS

4.1 STUDY DESIGNS AND STUDY POPULATIONS

4.1.1 Study designs and permissions

The study consisted of four registry-based observational studies. No interventions were conducted on the study patients. According to Finnish legislation, there was no need for individual consent from the patients. Studies I, III, and IV were multicentre studies, and, of which, Studies III and IV had a multinational context. Study II was conducted as a single-centre study.

The protocol for Study I was approved by the Research Ethics Committee of the Northern Savo Hospital District (225/13.02.00/2016). Permission to conduct the study was obtained from the National Institute for Health and Welfare (THL/1585/5.05.00/2015). Regarding Studies II and III, the research ethics committee of the Northern Savo Hospital District reviewed and approved the study protocol (Reference number: 478/2021). For Study IV, the Finnish Institute for Health and Welfare approved the data management plan, database contents, and study process (THL/1524/5.05.00/2017; THL/1173/05/00/2018; THL/3795/14.06.00/2021).

4.1.2 Finnish Intensive Care Consortium

The Finnish Intensive Care Consortium (FICC) is a quality benchmarking programme for Finnish ICUs. FICC's activity began in 1994 with nine participating ICUs. Over the following years, the number of member units grew. As of May 2021, 26 ICUs were participating in FICC, providing comprehensive coverage of Finnish intensive care.[173] Furthermore, ICUs at Inselspital in Bern, Switzerland, and Tartu University Hospital in Tartu, Estonia, have also joined FICC. The healthcare districts are the data controllers for the data from their respective units in the FICC database. Since 2023, the FICC database, along with several other national quality registries, has been under the record-keeping responsibility of the Institute of Health and Welfare. This change was implemented following a decree issued by the Finnish Ministry of Social Affairs and Health in 2022.

4.1.3 eICU database

The eICU programme is a multi-centre telehealth system developed by Philips Healthcare. The eICU-Collaborative Research Database (eICU-CRT) is an ICU database containing data, for example, on physiological measurements, vital care plan documentation, severity-of-illness measures, diagnosis information, and outcome data. It encompasses data from over 200,859 admissions to ICUs at 208 hospitals across the United States. The database is open to registered researchers.[236]

4.1.4 Study populations and exclusions

In Study I, the study population consisted of all adult patients admitted to 26 Finnish ICUs between January 1st, 2013, and December 31st, 2015. No exclusions were made based on admission cause. The study data was extracted from FICC database.

In Study II, the study population consisted of adult patients admitted to Kuopio University Hospital ICU, Finland, between January 1st, 2013 and December 31st, 2019. All scheduled (non-emergency) surgical admissions and all cardiac surgery (both emergency and non-emergency) admissions were excluded.

In Study III, for the determination of noradrenaline cutoffs, we used the same study population as in Study II, namely, adult patients admitted to the ICU at Kuopio University Hospital between 2013 and 2019. For external validation of the cutoffs, we used the eICU database. To ensure that our study predominantly focuses on intensive care patients, we applied a threshold of excluding units where less than 5% of patients received noradrenaline.

Study IV was a secondary analysis of the study population primarily collected in an observational study by Takala et al.[237] The study population was gathered from 20 ICUs in Finland, Estonia, and Switzerland. Data on patients admitted to ICUs from 2015 to 2017 were considered, as this period coincided with the availability of information on potential organ donation as a cause of ICU admissions.

In all studies, only the first admission to the ICU during the hospitalisation was included, and subsequent admissions were excluded.

4.1.5 Handling of missing data

In Studies I and II, the most frequently missing data were the hepatic points of the SOFA score due to missing data on bilirubin concentration at ICU admission (39% in Study I and 21% in Study II). When bilirubin data were missing, we gained patients zero hepatic scores, assuming normality of bilirubin concentrations and hepatic function. However, for other SOFA components with missing data, we did not make any assumptions regarding normality, and patients with missing data on SOFA components other than the hepatic score were excluded.

In Study III, to avoid selection bias, we included only units that had recorded near complete data of noradrenaline dose in µg /kg/min. Thus, we excluded all patients who were admitted to ICUs where data on noradrenaline infusion rate, drug concentration used, or body weight were missing for more than 5% of patients.

In all studies, we excluded patients with missing data on the primary outcome (hospital mortality in Studies I, III, and IV and 30-day mortality in Study II).

4.1.6 Determination of organ failure

We determined organ dysfunction to exist when the patient received 1 or 2 organ-specific SOFA points. Organ failure was determined when the patient reached 3 or 4 organ-specific SOFA scores. This description was in line with that of the working group on sepsis-related problems of the ESICM, which introduced the SOFA score in 1996.[22]

4.1.7 VIS coefficients and cutoffs

We defined two separate VIS-based derivatives: VIS_{max} (maximum value) and VIS_{mean} (mean value). These variables were calculated using data from the vasopressor and inotrope infusions administered during the first 24 hours after ICU admission. To ensure comparability between doses of different vasopressor and inotropic medications, we multiplied the doses by coefficients specific to each medication, as suggested by Gaies et al., and, in the case of levosimendan, Favia et al.[141,150] In most previous VIS studies, the coefficients are presented in dopamine equivalents. However,

for clarity and to emphasise the more modern use of vasopressors, we presented the coefficients in noradrenaline equivalents (Table 7).

Table 7. Coefficients for each vasopressor/inotrope.

Agent	Coefficient
Noradrenaline	1
Adrenaline	1
Dopamine	0.01
Dobutamine	0.01
Levosimendan	0.1
Milrinone	0.5
Vasopressin	100

The infusion rates are reported in μg /kg/min except for vasopressin units/kg/min.

Regarding the categories of VIS derivatives, our goal was to determine cutoff values that fulfil three specific criteria. First, they should be easy to remember and practical for use in clinical practice. Second, they should establish a relatively linear association between the increasing VIS category and increasing mortality. Third, they should yield categories with a patient distribution comparable to other SOFA component scores. For the optimal cutoff selection, we also employed cubic spline analysis to identify non-linear patterns in the data. However, upon analysing the results, we did not find clear non-linearities that could have produced suitable cutoff points. In other words, the analysis did not reveal distinct thresholds or breakpoints that could be used as optimal cutoff values for the VIS derivatives of interest. The cutoff values utilised in Study II are presented in Table 8.

Table 8. The categorical cutoffs for VIS derivatives (VIS_{max} and VIS_{mean}).

Category	VIS_{max}	VIS_{mean}
0	0	0
1	>0	>0
2	>0.15	>0.05
3	>0.3	>0.1
4	>0.45	>0.15

To calculate VIS_{mean}, the first step involved calculating the total amount of each medication administered during the first 24 hours after ICU admission. This total was adjusted using the specific coefficients for each medication, as mentioned earlier. Next, we divided the sum of these coefficient-adjusted infusions by 1440 to convert the mean dose from micrograms per kilogram per day into micrograms per kilogram per minute (μg/kg/min). Unlike other medications, for vasopressin, the unit of infusion rate used was U/kg/min. It is noteworthy that the VIS_{mean} value for patients with LOS of less than 24 hours was calculated based on their actual ICU stay duration rather than the full 24-hour period.

4.1.8 Requirements and determination of noradrenaline cutoffs

We established pre-set requirements for optimal cutoffs for noradrenaline doses to determine low-dose, intermediate-dose, and high-dose groups. The requirements for the cutoffs were that they should be statistically rational, practical to use (rounded to the first decimal), and that they should results in increasing mortality in groups with increasing severity.

Optimal cutoffs were determined in the development cohort that was a single-centre cohort (patients admitted to ICU at KUH between 2013 and 2019). To ensure the generalisability of the cutoffs, we performed an external validation in eICU register database with varying ICU typologies (validation cohort). Moreover, we tested the cutoffs in patients admitted for septic circulatory failure and other patients, patients needing

mechanical ventilation and those who did not, and in patients of different age quartiles.

4.1.9 Determination of PODs

To identify PODs, we used the information recorded at the ICU admission: 'admission because of possible organ donation'. Patients who were initially admitted to ICUs without treatment restrictions but whose active treatment was later withdrawn and who became organ donors during the ICU period were not considered as PODs in the study.

4.2 STATISTICAL METHODS

4.2.1 Data processing

We assessed differences between survivors and non-survivors using the non-parametric Mann–Whitney test for continuous parameters and Fisher's exact test or Pearson's chi-square test, depending on appropriateness, for categorical variables. The statistical analyses were performed using R, Version 4.1.1, developed in Vienna, Austria, and IBM SPSS Statistics for Windows, Version 26.0, developed by IBM Corp. in Armonk, NY. P values of less than 0.05 were considered statistically significant.

4.2.2 Study I

The increased in-hospital mortality associated with organ failures was the primary endpoint in this study. ICU and 12-month mortalities were the secondary endpoints. We conducted a multivariable logistic regression analysis to calculate the increasing risk of mortality for each step increase in the organ-specific SOFA score. The analysis was adjusted for age and sex. Moreover, we studied whether certain organ failures tend to occur concurrently. We used a standardised occurrence ratio (SOR) to determine if some combinations of organ failures occur more frequently than anticipated. The formulation for SOR was $\mathrm{N}(o) \div [\mathrm{N} \times \mathrm{P}(a) \times \mathrm{P}(b)]$. Here, N(*o*) represents the observed number of patients with organ failure of both *a* and *b*. N is the number of all admissions. P(*a*) and P(*b*) are the proportions of patients with organ failures a and b, respectively.

4.2.3 Study II

The primary statistical outcome measured in this study was the AUROC of VIS-derivatives and the change in AUROC when replacing the cardiovascular SOFA score with the VIS-based score.

We assessed the discrimination ability of scoring methods by utilising the predicted probabilities for each patient. These predicted probabilities were derived from the cardiovascular SOFA score, VIS_{max}, and VIS_{mean} categories by using binary logistic regression analysis adjusted for age and sex. Likewise, we calculated the risk of death adjusted for age and sex for each patient using the total SOFA score and alternative SOFA scores in which the cardiovascular subscore was replaced by categorised VIS_{max} or VIS_{mean} values. The difference in discrimination between these scoring methods was assessed using the DeLong test.[238]

Regarding the alternative VIS-based cardiovascular SOFA scores, patients in the VIS category 0 were assigned 0 points, while patients in other categories received points corresponding to their respective categories. The maximum score assigned was 4 points for category 4.

4.2.4 Study III

We determined the cutoffs through a log-rank statistic test. The most significant split among low-dose, intermediate-dose, and high-dose noradrenaline groups, in their association with mortality, was identified. Noradrenaline dose was treated as a continuous variable, and the Contal and O'Quigley method was employed to pinpoint these noradrenaline dose thresholds.[239] Hospital mortality was used as the primary outcome. The highest dose of noradrenaline during the first 24 hours after ICU admission was used. We estimated the association of the noradrenaline dose groups with mortality with the Kaplan–Meier survival plot, and the results were interpreted with the log-rank test (Mantel–Cox). Moreover, we evaluated the relative risk of death across the noradrenaline dose groups by conducting a Cox regression analysis, which was adjusted for the unit to which the patients were admitted, age, and need of mechanical ventilation. The patient group not administered noradrenaline was used as the reference.

4.2.5 Study IV

To investigate the effect of PODs on SMRs, we compared the SMR of the cohort without PODs to the one with PODs. The ratio of the SMRs of the cohorts was calculated using a Poisson regression model adjusted for calendar year. The analysis was carried out for the overall cohorts and separately for each ICU.

We also evaluated the influence of different case-mixes (different ICU typologies: university hospital, small non-university hospital, and large university hospital) by additive analysis. Here, the two cohorts were the main predictors, and hospital typology was set as a confounder.

5 RESULTS

5.1 STUDY I – ASSOCIATION OF SOFA SCORE COMPONENTS WITH MORTALITY

5.1.1 Study population

In total, 63,756 ICU patients were included in the analysis. ICU mortality was 5.3%, in-hospital mortality was 10.7%, and 12-month mortality was 21.6%. A majority of the patients were male (63%). The total SOFA score of the admission day was strongly associated with mortality. In-hospital mortality was 69% in patients with 15 or more SOFA points (Figure 3).

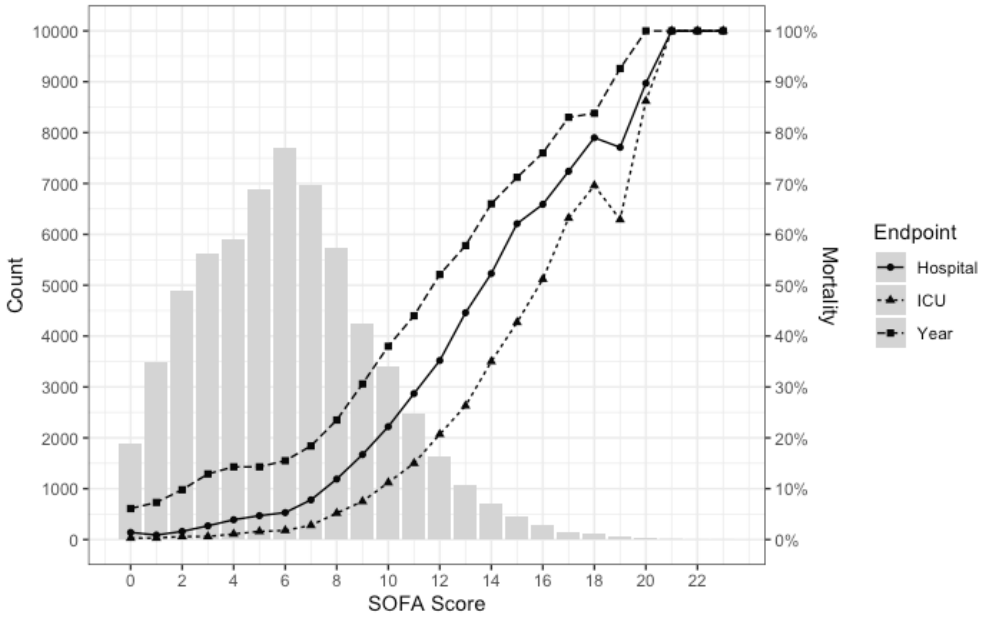

Figure 3. The association of the SOFA score with ICU mortality (dotted line with triangles), in-hospital mortality (solid line with filled circles), and 12-month mortality (dashed line with squares) (right-hand y-axis). The grey columns represent the numbers of patients (left-hand y-axis) in groups according to the SOFA score (x-axis).

5.1.2 The SOFA score components and in-hospital mortality

The most frequent organ failure was cardiovascular failure, affecting 52.7% of the patients, followed by respiratory failure (22.8%), neurologic failure (17.8%), renal failure (8.3%), coagulation failure (2.7%), and hepatic failure (1.0%) (Figure 4). The failing organ systems were not equally associated with increased mortality. The most lethal organ failure was hepatic failure, with in-hospital mortality of 40.1%, followed by coagulation failure (34.8%), renal (34.6%), neurologic (26.9%), and respiratory failure (22.5%). For patients with cardiovascular failure, the in-hospital mortality was 14.9%, markedly lower than the mortalities associated with failures of other organ systems (Figure 5).

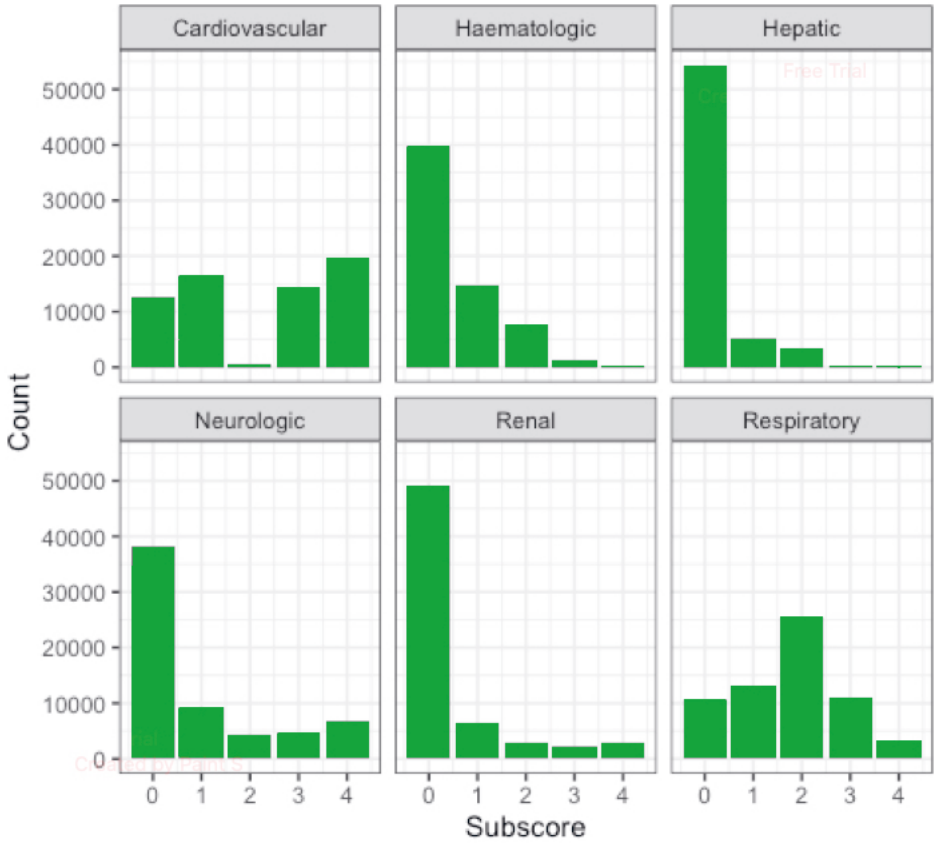

Figure 4. The distribution of patients according to organ-specific SOFA subscores.

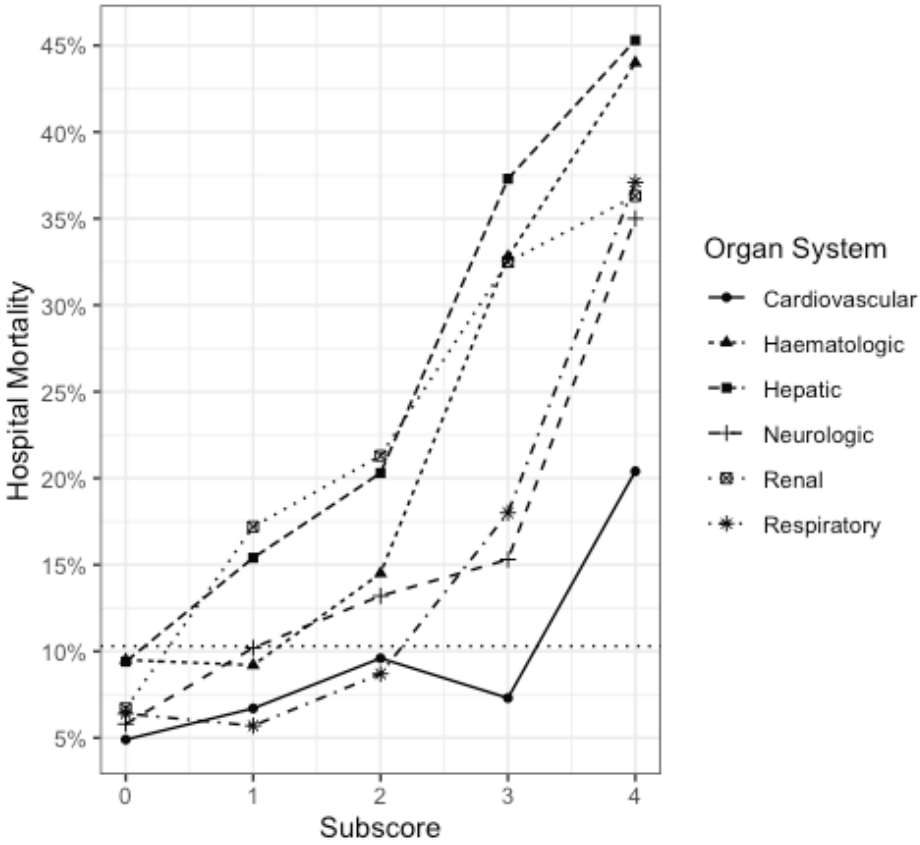

Figure 5. The hospital mortality according to organ-specific SOFA subscores. The dotted horisontal line represents the hospital mortality in the overall population.

In the age- and sex-adjusted logistic regression model, cardiovascular failure was the least associated with mortality. Renal failure was the most associated with ICU mortality, neurologic failure with hospital mortality, and hepatic failure with 12-month mortality (Table 9).

Table 9. The association of organ failures with mortality – respiratory, coagulation, hepatic, cardiovascular, neurological, and renal. For each organ failure, patients, without that particular organ failure were used as the reference category. Organ failure was determined as 3 or 4 organ-specific SOFA subscores.

		OR	95% CI		OR	95% CI
ICU Mortality	Respiratory	2.92	2.70–3.16	Cardiovascular	2.15	1.95–2.36
	Coagulation	4.18	3.63–4.82	Neurologic	4.63	4.28–5.01
	Hepatic	2.27	1.78–2.89	Renal	5.99	5.48–6.55
Hospital Mortality	Respiratory	2.41	2.27–2.56	Cardiovascular	1.57	1.47–1.67
	Coagulation	4.04	3.57–4.57	Neurologic	5.00	4.71–5.30
	Hepatic	4.24	3.47–5.17	Renal	4.93	4.58–5.32
12-month Mortality	Respiratory	1.71	1.63–1.79	Cardiovascular	1.05	1.01–1.10
	Coagulation	3.24	2.89–3.63	Neurologic	4.13	3.93–4.34
	Hepatic	4.27	3.53–5.17	Renal	3.81	3.56–4.07

5.1.3 Organ failure combinations

Mortality was higher in patients with multiple failing organ systems. 47.4% of the patients had at least two concurrent organ failures, 12.7% had at least three, and 2.0% had at least four. Hospital mortality was 35.8% for patients with at least two organ failures, 54.1% for those with at least three organ failures, and 71.8% for those with at least four organ failures.

Of the organ failure combinations, coagulation and hepatic failures tended to appear concurrently the most frequently (SOR 8.21), followed by hepatic and renal (SOR 4.53), and coagulation and renal failures (SOR 3.08). The in-hospital mortalities of patients with these combinations were 57.0%, 55.4%, and 57.6%, respectively. Of the combinations with three concurrent

organ failures, the most frequently appearing, compared to the expected occurrence, were coagulation, hepatic, and renal (SOR 43.97); coagulation, cardiovascular, and hepatic (SOR 11.76); and coagulation, neurologic, and hepatic failures (SOR 11.07). It was noted that 66.1%, 62.0%, and 68.6%, respectively, of patients with these organ failure combinations died.

5.2 STUDY II – VIS COMPARED TO THE CARDIOVASCULAR SOFA SCORE

5.2.1 Study population

We analysed 8,079 admissions to the ICU. A majority of the patients were male (62%), and the median age of the patients was 61 years. Seven percent of the patients died during the ICU stay, 11.4% during the hospitalisation, and 13.7% during the 30 days following the ICU admission.

5.2.2 Association between VIS and mortality

One third (34.7%) of the patients were administered vasopressors or inotropes during the first 24 hours in the ICU. Noradrenaline was by a considerable margin the most administered drug (99.1%) to patients receiving any vasopressor or inotrope, followed by dobutamine (10.4%), adrenaline (4.5%), levosimendan (3.2%), milrinone (1.9%), vasopressin (1.8%), and dopamine (0.4%).

The increasing values of both VIS_{max} and VIS_{mean} were consistently associated with increasing ICU mortality, hospital mortality, and 30-day mortality (Figure 6).

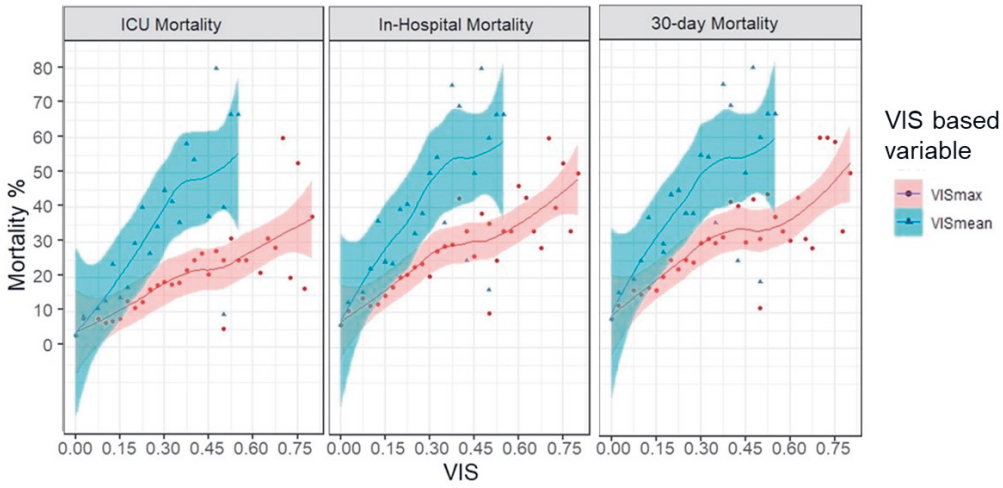

Figure 6. Association of the first 24-hour VIS_{max} and VIS_{mean} values with ICU mortality, hospital mortality, and 30-day mortality. VIS values are represented on the x-axis and mortality (in percentages) on the y-axis. The red line (and filled circles) represents VIS_{max}, whereas the turquoise line (with filled circles) represents VIS_{mean}. The coloured areas represent the standard error.

When categorised to five groups, the number of patients decreased consistently with increasing VIS category (except for categories 3–4 in VIS_{mean}). This differed markedly from the numbers of patients in the cardiovascular SOFA score categories. Only eight patients received a cardiovascular SOFA score of 2 (Figure 7).

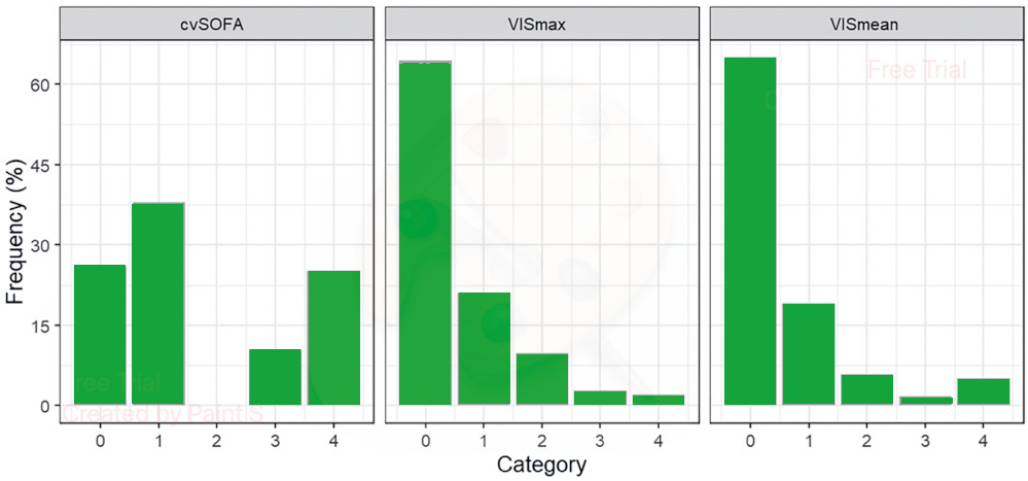

Figure 7. The distribution of study patients according to the cardiovascular SOFA score and categorised VIS_{max} and VIS_{mean} scores.

The 30-day mortality rates associated with increasing VIS_{max} categories were 8.0%, 16.4%, 23.4%, 33.0%, and 44.1%, while for increasing VIS_{mean} categories, the rates were 8.5%, 17.6%, 23.7%, 33.4%, and 51.0%, respectively. The mortality rates were dissimilar to those in the increasing cardiovascular SOFA score, at 6.6%, 9.1%, 22.2%, 14.9%, and 26.1%. The risk of mortality with increasing VIS derivative categories was moderately linear, while in the cardiovascular SOFA score, the association was less consistent (Figure 8).

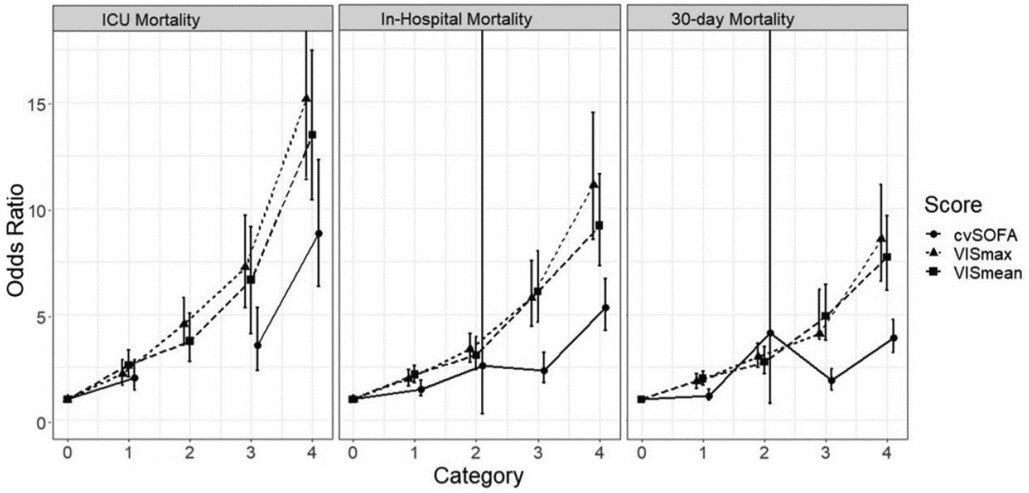

Figure 8. The odds ratios for mortality associated with each score of the cardiovascular SOFA score and categorised VIS_{max} and VIS_{mean} values. The leftmost panel represents ICU mortality, the middle panel represents hospital mortality, and the rightmost panel represents 30-day mortality.

The predictive value in 30-day mortality prediction was acceptable for the cardiovascular SOFA subscore, for categorised VIS_{max}, and for categorised VIS_{mean}, with AUROC values of 0.737 (95% CI 0.722–0.752) for cardiovascular SOFA subscore, 0.750 (95% CI 0.735–0.765) for VIS_{max}, and 0.746 (95% CI 0.731–0.760) for VIS_{mean} (Table 10). The VIS derivatives had significantly better discriminative ability compared to cardiovascular SOFA ($p<0.001$) and VIS_{max} outperformed VIS_{mean} ($p=0.02$).

The conventional total SOFA score performed rather well in predicting 30-day mortality, with an AUROC of 0.813 (95% CI 0.800–0.825). However, the discrimination ability significantly improved when the cardiovascular SOFA component was replaced with a VIS_{max}-based component, resulting in an AUROC of 0.822 (95% CI 0.809–0.834; a difference of 0.009 compared to the conventional SOFA score, $p<0.001$), or with a VIS_{mean}-based component, resulting in an AUROC of 0.816 (95% CI 0.803–0.828; a difference of 0.003 compared to the conventional SOFA score, $p=0.004$) (Table 10).

Table 10. The results of the AUROC analysis for categorised VIS derivatives and cardiovascular SOFA score (cvSOFA). Additionally, the table represents the AUROC values with 95% confidence intervals (CI) for the total SOFA score and alternative VIS-based total SOFA scores (based on VIS_{max} and VIS_{mean}). The results are presented at various time intervals (ICU discharge, hospital discharge, and 30 days after ICU admission).

		AUROC	95% CI
ICU Mortality	VIS_{max}	0.751	0.735–0.768
	VIS_{mean}	0.746	0.729–0.762
	cvSOFA	0.738	0.722–0.754
	VIS_{max}-based SOFA score	0.853	0.834–0.868
	VIS_{max}-based SOFA score	0.846	0.830–0.861
	Conventional SOFA score	0.842	0.827–0.858
In-Hospital Mortality	VIS_{max}	0.775	0.754–0.795
	VIS_{mean}	0.767	0.747–0.787
	cvSOFA	0.759	0.739–0.778
	VIS_{max}-based SOFA score	0.826	0.812–0.840
	VIS_{max}-based SOFA score	0.820	0.807–0.834
	Conventional SOFA score	0.818	0.804–0.832
30-day Mortality	VIS_{max}	0.750	0.735–0.765
	VIS_{mean}	0.746	0.731–0.760
	cvSOFA	0.737	0.722–0.752
	VIS_{max}-based SOFA score	0.822	0.809–0.834
	VIS_{max}-based SOFA score	0.816	0.803–0.828
	Conventional SOFA score	0.813	0.800–0.825

5.3 STUDY III – NORADRENALINE DOSE CUTOFFS TO DETERMINE LOW, INTERMEDIATE, AND HIGH DOSE

5.3.1 Study population

The study population of the development cohort is previously described in the results of Study II in Section 5.3.1.

The final population in the validation cohort consisted of 39,007 admissions to 29 ICUs. The majority of the patients (70.3%) were admitted to medical or medical–surgical ICUs; 23.3% were admissions to cardiac or cardiothoracic ICUs and 4.6% to neuro ICUs. The majority (54.0%) of the patients were male, and the median age was 65 years.

5.3.2 The development of the cutoffs

The absolute values for the 20 highest values of the log rank statistic estimate are presented in Table 11. The two most prominent peaks for the log-rank statistic were identified at noradrenaline doses of 0.20 µg/kg/min and 0.44 µg/kg/min (Figure 9, Table 11). Using these cutoffs would result in the most significant separation of the groups in the log-rank test, which examines survival disparities among the groups. Therefore, in line with the pre-set requirement for rounding to the first decimal, we determined the cutoffs for noradrenaline doses at 0.2 and 0.4 µg/kg/min. This resulted in three groups of patients receiving different doses of noradrenaline: low dose (<0.2 µg/kg/min), intermediate dose (0.2–0.4 µg/kg/min), and high dose (>0.4 µg/kg/min). There was a significant separation in survival between the groups receiving noradrenaline as well as between those groups and the group not receiving any noradrenaline ($p<0.001$) (Figure 10). Furthermore, the third predefined criterion for the cutoffs was met: mortality associated consistently with increasing noradrenaline dose groups (Figure 11).

Nearly all patients (99%) receiving any vasoactive medication were administered noradrenaline. The hazard ratio (HR), compared to patients receiving no noradrenaline, was 1.7 (95% CI 1.4-2.1) for the low-dose group, 3.1 (95% CI 2.5-3.9) for the intermediate-dose group, and 4.2 (95% CI 3.3-5.4) for the high-dose group ($p<0.001$). The AUROC for predicting

mortality based on noradrenaline dose categories was 0.685 (95% CI 0.667-0.703).

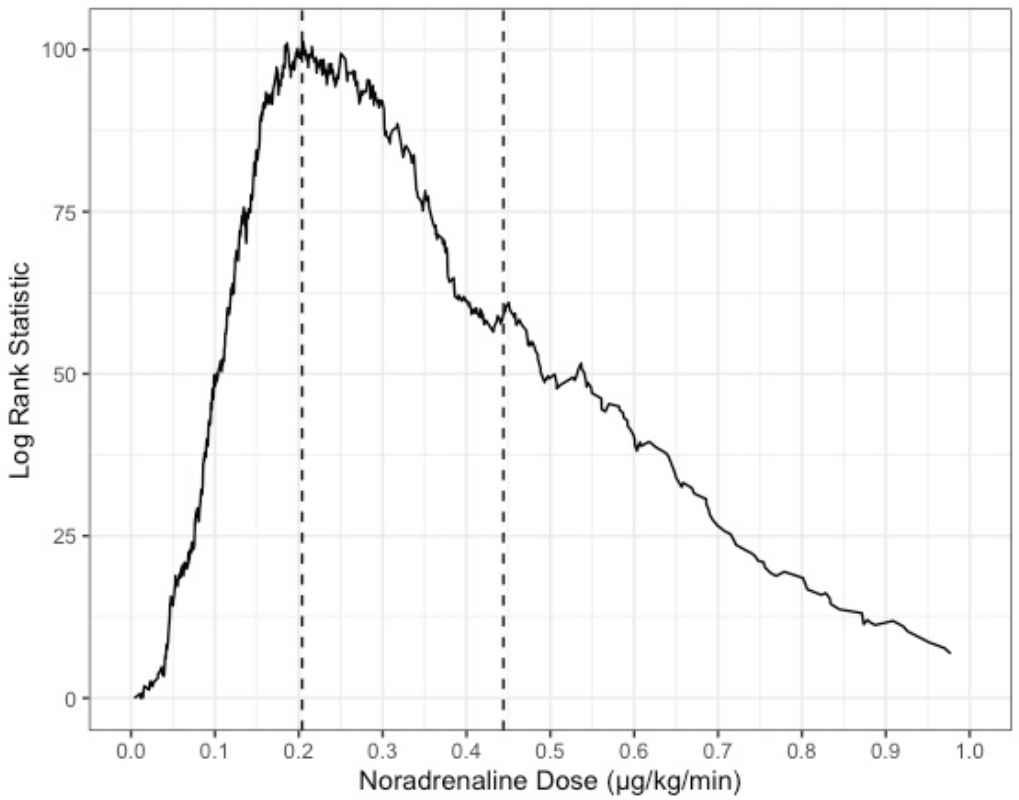

Figure 9. The results of the log-rank statistic analysis by the Contal and O'Quigley method.[239] The prominent peaks were identified at noradrenaline doses of 0.20 and 0.44 µg/kg/min.

Table 11. Absolute estimates of noradrenaline cutoffs derived from log-rank statistical analysis within the Contal and O'Quigley method. The 20 highest values for log-rank statistic estimates.

Lower cutoff	Upper cutoff	Log rank statistic
0.20	0.44	106.01
0.18	0.44	103.59
0.16	0.42	103.09
0.20	0.60	102.73
0.22	0.44	102.10
0.16	0.32	101.74
0.27	0.60	100.87
0.24	0.44	100.87
0.20	0.32	100.63
0.18	0.32	100.33
0.27	0.44	100.13
0.24	0.60	99.98
0.22	0.60	99.82
0.18	0.60	98.61
0.32	0.60	98.57
0.15	0.32	98.50
0.15	0.44	97.52
0.13	0.32	96.57
0.16	0.60	96.45
0.32	0.44	96.39

5.3.3 The validation cohort

As in the development cohort, noradrenaline was the predominant vasopressor, and it was used in 95% of the patients receiving any vasoactive agent. It was combined in in 92% of the patients receiving vasopressin, 63% receiving dobutamine, 65% receiving milrinone, 79% receiving adrenaline, 40% receiving dopamine, and 48% receiving phenylnephrine.

There was significant separation between the three noradrenaline groups and the fourth group, which did not receive any noradrenaline in their associations with mortality ($p<0.001$) (Figure 10). The cutoffs resulted in a consistently decreasing number of patients and increased mortality with groups of increasing noradrenaline dose (Figure 11).

Compared to patients who did not receive noradrenaline, HR for in-hospital death was 1.4 (95% CI 1.2-1.7) for the low-dose group, 4.0 (95% CI 3.5-4.7) for the intermediate dose group, and 7.5 (95% CI 6.5-8.5) for the high-dose group ($p<0.001$). The AUROC for noradrenaline group cutoffs in the validation group was 0.659 (95% CI 0.651–0.667).

The validity of the cutoffs remained consistent across patients admitted for septic circulatory failure and other patients, those receiving and not receiving mechanical ventilation, and across different age quartiles.

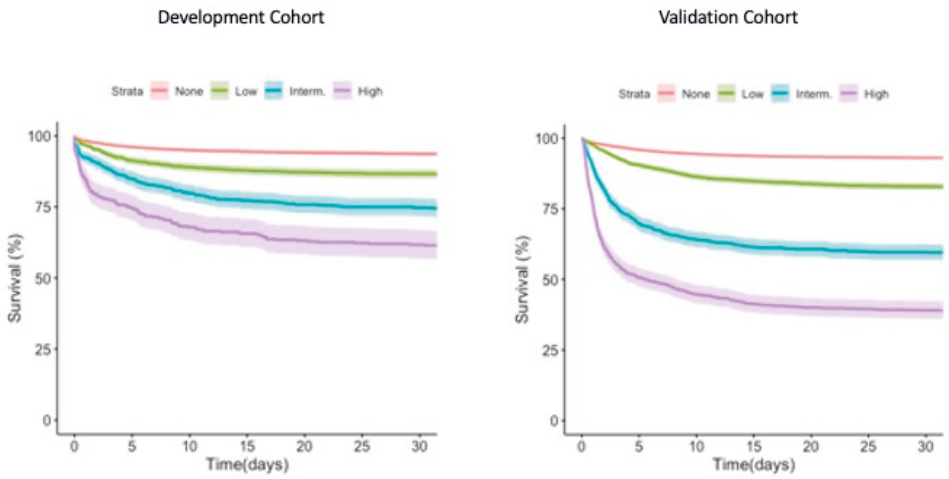

Figure 10. The Kaplan–Meier plots, showing survival over time, demonstrates the separation of the noradrenaline dose groups in addition to the group not receiving any noradrenaline. The development cohort is in the left panel and the validation cohort in the right panel. There was significant separation between all groups ($p<0.001$).

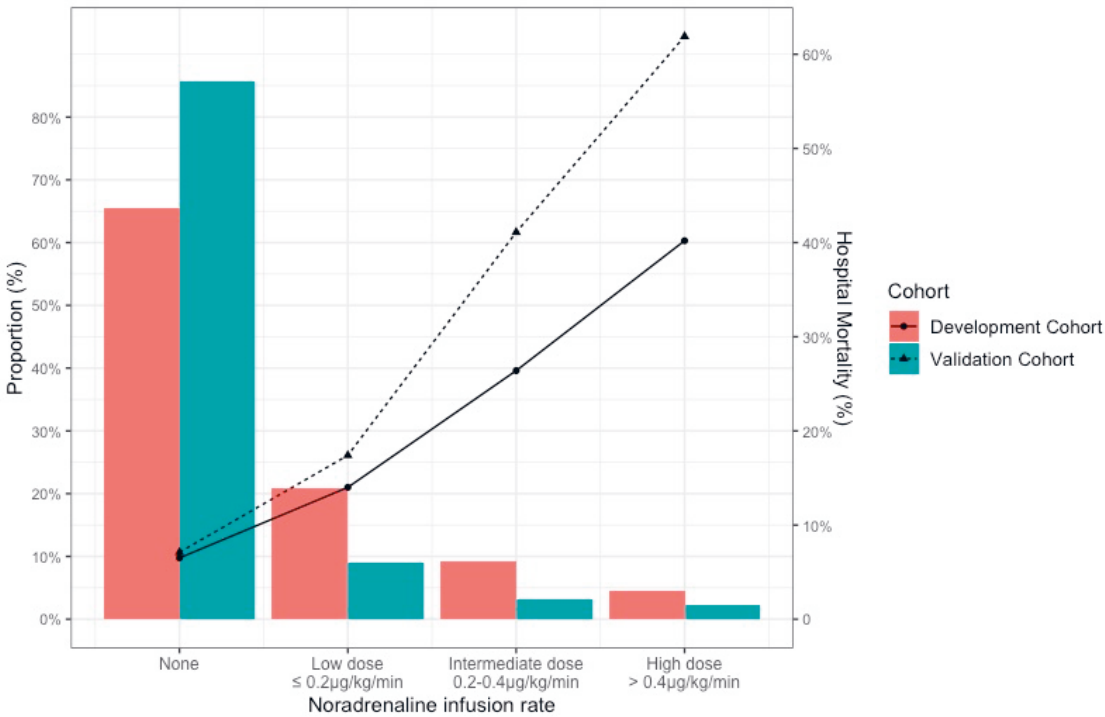

Figure 11. The distribution of patients to groups according to noradrenaline dose (x-axis). The red bars represent the patients (in percentages) in the development cohort, and the turquoise bars represent the patients in the validation cohort (left y-axis). The solid line with filled circles represents the hospital mortality of patients in the development cohort, and triangles with a dotted line represent the mortality in the validation cohort (right y-axis).

5.4 STUDY IV – THE EFFECT OF PODS ON SMRS AND ICU BENCHMARKING

5.4.1 Study population

We studied 60,047 admissions to 20 ICUs. The in-hospital mortality for the overall population was 11.2%. Of the patients admitted to ICUs, 514 (0.9%) were admitted as PODs. The frequency of POD admissions ranged between 0.5 and 18.3 per 1000 admissions depending on the ICU. The predicted in-hospital mortality for these patients was 37%, while their observed mortality was 93%.

The most common cause of brain damage leading to POD admission was intracerebral haemorrhage, accounting for 44%. This was followed by trauma at 22%, subarachnoid haemorrhage at 15%, hypoxemic brain injury at 5%, ischaemic stroke at 4%, and miscellaneous aetiologies at 10%.

5.4.2 The effect of PODs on the SMRs

The SMR for the cohort including PODs was 1.04 (95% CI 1.01–1.06), whereas in the cohort excluding PODs, it was 0.99 (95% CI 0.97–1.02). The ratio of the SMRs between the cohort without PODs and that with PODs was 0.96 (95% CI 0.93–0.99). There was no interaction effect between the two cohorts due to different hospital sizes (p=0.89). On an individual ICU level, there was extensive overlap between the SMRs of the two cohorts (Figure 12).

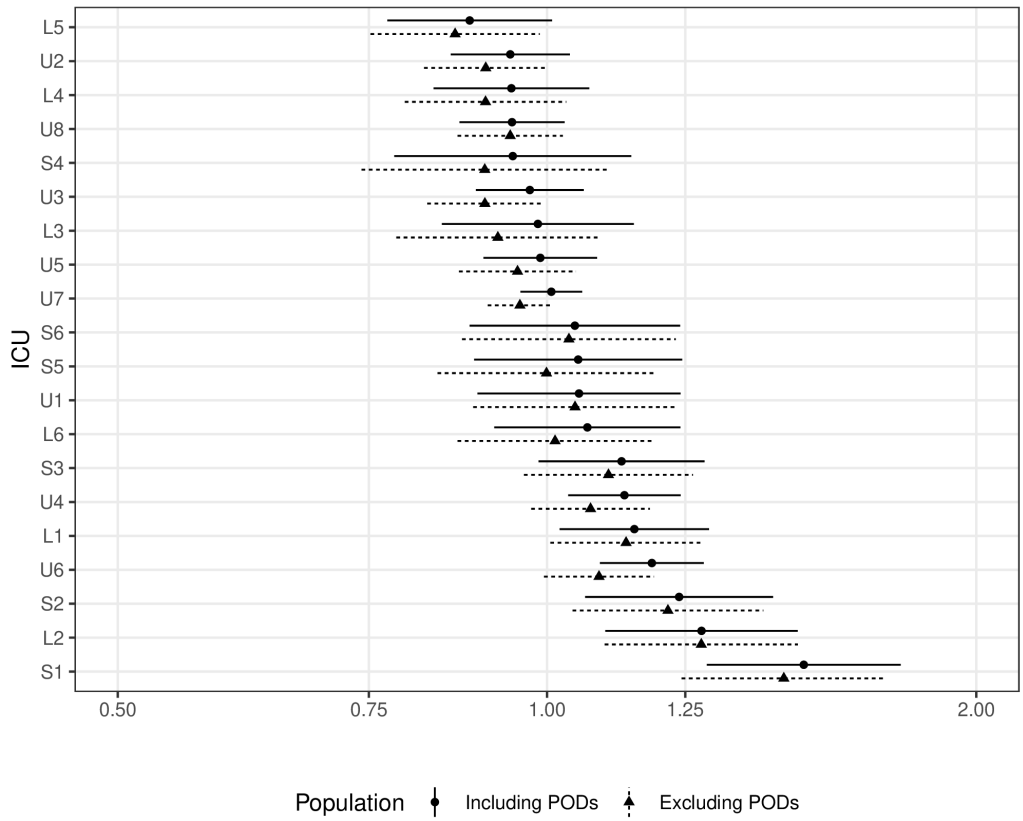

Figure 12. The figure depicts the SMRs at the individual ICU level. The ICUs are represented on the y-axis. U1–U8 represent the university hospitals, L1–L6 represent the large non-university hospitals, and S1–S6 represent the small non-university hospitals. The filled circles represent the corresponding SMRs including PODs, and the triangles represent SMRs excluding PODs. The lines represent the 95% CIs (solid line with PODs and dotted line without PODs). The figure is reprinted with permission from Lippincott Williams & Wilkins.

5.4.3 The influence of PODs on the ICU benchmark ranking

The inclusion of PODs altered the benchmark ranking position in 70% of the units (Figure 13). A higher number of PODs was associated with a positive change in the ranking position after excluding the PODs.

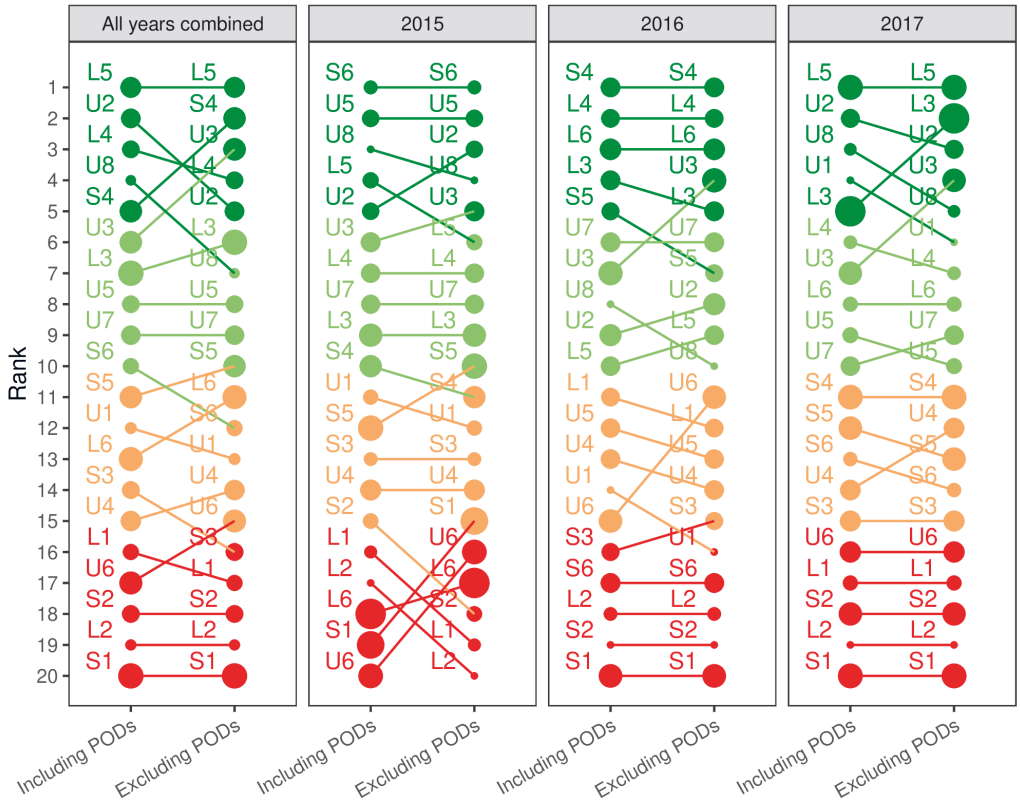

Figure 13. The y-axis represents the ranking position. In each panel, the left circles represent the cohort including PODs, and the right circles represent the cohort excluding PODs. The left panel represents the benchmark results across the study period, and the subsequent panels represent the results for each year. U1–U8 represent the university hospitals, L1–L6 represent the large non-university hospitals, and S1–S6 represent the small non-university hospitals. The size of the circle corresponds to the proportion of PODs. The figure is reprinted with permission from Lippincott Williams & Wilkins.

6 DISCUSSION

6.1 ORGAN DYSFUNCTION SCORES

6.1.1 Summary of the findings of Studies I, II, and III

In Study I, we compared the weights of the organ-specific component scores of the SOFA score in terms of their association with mortality. The cardiovascular component was less strongly associated with mortality than the other organ-system components. In Study II, our aim was to propose a solution to overcome this discrepancy by replacing the cardiovascular component with a VIS-based score. The discrimination ability of the total SOFA score in predicting mortality improved with this replacement. Moreover, in this study, we conducted the first validation of VIS in a general ICU population. The increasing VIS was consistently associated with increased mortality in this population. However, the results of Study II showed that practically all patients receiving any vasopressor or inotrope received noradrenaline and that the VIS sum was mainly composed of noradrenaline.

Subsequently, in Study III, we examined whether the dose of noradrenaline alone could characterise the severity of cardiovascular failure by assessing the association with increased the risk of mortality. We identified the optimal noradrenaline dose cutoffs for this approach, and we resulted in cutoffs of 0.2 μg/kg/min and 0.4 μg/kg/min, which met the predefined criteria.

6.1.2 Evolution of vasopressor and inotrope use

The use of vasopressors and inotropes has changed markedly over the past 30 years. According to observational studies from the late 1990s, the cardiovascular SOFA had the strongest association with mortality among the organ specific subscores.[23] Strikingly, some recent studies report the complete opposite, indicating that treatment practices and vasopressor use have changed.[119,240] In 1996, when the SOFA score was introduced, fluid resuscitation was the primary treatment for circulatory shock. Vasopressors, on the other hand, were used with caution.[241] However,

studies conducted since have proven that there is an association between persistent positive fluid balance and mortality.[242,243] This has led to more liberal trend in vasopressor use. Moderate administration of vasopressors in the early stages of intensive care is nowadays less indicative of a poor prognosis. In fact, the results of recent studies have suggested that earlier initiation of vasopressors may decrease mortality.[244–246] Therefore, the cardiovascular SOFA score component has experienced an inflation in prognostic ability.[22,23,247]

The change in vasopressor use practice has affected the distribution of the cardiovascular SOFA score. The criterion for obtaining 2 cardiovascular points is dopamine infusion of < 5 μg/kg/min or dobutamine at any dose (without adrenaline or noradrenaline). There has been a significant decline in the use of dopamine in clinical practice for treating cardiovascular shock in the ICU. In many centres, dopamine is no longer used.[248–251] Moreover, dobutamine is rarely used as a single agent. As a result, few patients meet the criteria for obtaining 2 cardiovascular SOFA points.

In SOAP II randomised trial, noradrenaline and dopamine were compared as a first-line treatment for shock. Patients in the dopamine group experienced more arrhythmias, and those patients with cardiogenic shock had increased mortality when treated with dopamine.[249] Otherwise, to date, there has been no clear evidence that any vasopressor or inotrope is superior to any other. For example, the Surviving Sepsis Campaign guidelines for septic shock recommend noradrenaline as first-line therapy and vasopressin as a second-line vasopressor in refractory septic shock.[252] In a questionnaire study conducted by the Acute Cardiovascular Care Association Research, a community under the European Society of Cardiology (ESC-ACVC) that spans 60 countries, 89% of clinicians preferred noradrenaline as the first-line vasoactive agent in cardiovascular shock. This was followed by dopamine (9%) and other agents (1%).[253] In a survey study conducted among ESICM member physicians, 97% of the responders, the majority of whom were intensive care practitioners, chose noradrenaline as the first-line vasopressor for septic shock.[254] In our Study II, practically all patients receiving any vasopressor or inotrope (99.1%) also received noradrenaline, and most VIS points were derived from noradrenaline. Therefore, it might be justified to consider the

noradrenaline infusion rate as the main measure of the need for pharmacological circulatory support.

The guidelines for shock treatment and common practices for vasoactive agents are likely to continue evolving. The utilisation of novel (or previously underutilised) agents, such as levosimendan, angiotensin II, methylene blue, and phosphodiesterase inhibitors, has increased, and some medications are making their way into clinical practice.[255–258] These agents are not included in the current SOFA score criteria. The results of the VANISH trial suggest that adding vasopressin decreases the demand for noradrenaline[259] and the ATHOS-3 trial showed similar effects with angiotensin II.[260] The results of these studies suggest that in the future, the range of vasopressors to choose from might be broader, and the selection of a vasopressor will depend on the underlying cause of shock.[261] Therefore, severity-of-illness scores should be updated much more frequently than every 30 years, or the scores should be capable of adapting to these changes in practice. VIS might offer an adaptable method of measuring cardiovascular failure. Several papers have recently demonstrated good prognostication of noradrenaline equivalent dose (NEE or NEQ) and its derivatives, such as MAP/NEQ-ratio.[262–264] In practice, the only difference between the VIS and NEE is the name. However, the equivalence between vasopressors and inotropes should be accurate. Kotani et al. recently published a proposal for such equivalences based on the available evidence.[265] These factors are comparable to the factors that we used, apart from vasopressin.

6.1.3 Comparison with the previous literature

Our finding, indicating that the cardiovascular SOFA score is less prognostic of death compared to other organ component scores of SOFA, is in accordance with results of a recent a post hoc analysis on gastrointestinal dysfunction in a population of 1,031 ICU patients, with a 90-day mortality of 20%, Bachmann et al. conclusively revealed a lack of linearity of the cardiovascular SOFA score component relationship with mortality and few patients scoring two cardiovascular points.[119,266]

In the old ICU population (aged ≥ 80 years old), the validity of the current SOFA score was studied by Pollock et al.[247] They performed a post hoc analysis of the VIP2 study, observing frailty, cognition, activity of daily

life, and comorbidities in elderly ICU patients.[267] As in our results, the association with mortality was not equally strong for all organ failures.

Problems with the cardiovascular SOFA score have been acknowledged, and many proposals for replacing it have been presented with numerous thresholds to determine refractory shock. Some proposals for refractory shock include the need for vasopressors equivalent to a noradrenaline dose of up to 2 μg/kg/min, which is considered somewhat excessive based on clinical experience.[268–274] Lee et al. compared the modified cardiovascular SOFA score with the current component retrospectively in 1,015 Korean ICU patients with sepsis.[263,275] In the modified cardiovascular SOFA score, the criteria for obtaining 2-4 cardiovascular SOFA scores included noradrenaline equivalent cutoffs of 0.2 μg/kg/min and 0.5 μg/kg/min, while the criterion for distinguishing between patients receiving 0 or 1 points remained the presence of hypotension (MAP 70 mm Hg). The proposed cutoffs closely resembled the cutoffs suggested in our Study III. Lee et al. also proposed adding one point to the score for lactatemia (lactate ≥2 mmol/L). The AUROC for 28-day mortality of the total SOFA score (AUROC 0.730 [95% CI 0.698–0.763]) increased substantially when the cardiovascular SOFA score was replaced by the modified cardiovascular SOFA score (AUROC 0.747 [95% CI 0.715–0.779]). Although there was a substantial improvement in the discrimination ability, this comes at the cost of simplicity. Adding the presence of lactatemia to the criteria increases the criteria derived from three categories: physiological (presence of hypotension), laboratory value (serum lactate), and the amount of drug infusion needed. When the discrimination ability of Lee et al.'s cutoff proposal was tested without the lactate criterion, the AUROC showed a slightly poorer performance (AUROC 0.625–0.640 depending on the observed cohort) compared to the AUROC of 0.659–0.685 in our study.

Several proposals for assessing circulatory shock have been made in scores that have been studied in cardiologic or cardiothoracic surgery populations. Examples are the Society for Cardiovascular Angiography and Intervention (SCAI) definition, Mayo Clinic definition, Hamburg definition, and IABP-SHOCK II for indicating the severity of cardiogenic shock.[276–280] These definitions take into account, through varying methods, the number of vasopressors used, serum lactate concentration, and other signs of hypoperfusion. However, these scores are mainly validated in patients

suffering from acute myocardial infarction, those who have undergone cardiothoracic surgery, or those admitted to the cardiac ICU for circulatory shock. These scores have not been validated in the general ICU population as we did for VIS, and these scores may not be generalisable for use in the general definition of MOF. Moreover, there is increased complexity related to these scores, whereas the goal of the SOFA criteria is to keep them as simple as possible.

6.1.4 The noradrenaline equivalent formulation

The variation in presentations of noradrenaline formulations may cause challenges for the noradrenaline-based estimation of the severity of cardiovascular failure. The pure form of noradrenaline (trihydroxy-substituted phenethylamine) is called noradrenaline base. However, it is poorly soluble in water, alcohol, or ether. Therefore, noradrenaline is usually prepared in the form of noradrenaline salt (for example, bitartrate or tartrate). Noradrenaline tartrate is the most common formulation. When reporting doses, it is essential, however, to specify whether the dose is described as the amount of noradrenaline tartrate or noradrenaline base, as 2 mg of noradrenaline tartrate correspond to 1 mg of noradrenaline base. However, there may be variation in this, which can challenge the comparability between the reported doses.[281–285] A position paper from the Society of Critical Care Medicine and the ESICM joint task force recommended that the infusion rate should always be presented as the amount of noradrenaline base administered per patient weight per minute.[286] This is well-reasoned guidance and should be followed.

6.1.5 Pharmacologic cardiovascular support comparison

In VIS calculations, it is important that VIS including agents have reliable equivalent coefficients. This matters also because decatecholaminisation, which involves attenuating the catecholamine load in selected patients using alternative vasopressors, is an increasingly debated strategy to avoid the adverse effects associated with catecholamines.[287,288] Moreover, there has been a temporary shortage of noradrenaline in recent years, highlighting the importance of the appropriate equivalent dose determinations.[289]

Numerous studies have compared the effects of various vasopressors to noradrenaline, including dopamine[249], phenylephrine[290], vasopressin[259,291], angiotensin II[292], terlipressin[293], methylene blue[294], metaraminol[295], hydroxycobalamin[296], and midodrine[297]. Gauging the noradrenaline equivalent doses is quite feasible with agents that primarily act via vasoconstriction.

However, VIS calculation encounters a challenge in drug comparability when inotropes without vasoconstricting effect are included in the score. In addition to increased inotropy, phosphodiesterase 3 inhibitors (milrinone, olprinone) and calcium sensitiser levosimendan have vasodilating effects leading to hypotension.[120] Furthermore, dobutamine exerts a beta-2-agonist effect causing vasodilation, which potentially leads to significant hypotension upon administration.[298] The hypotension-inducing effect of these inotropes may increase the demand for noradrenaline, potentially increasing the VIS as a result of the treatment. Therefore, it is important to be cautious when vasoconstricting and inotropic agents are combined in the same equation.

6.2 SMR AND ICU PERFORMANCE ASSESSMENT

6.2.1 Summary of Study IV

In Study IV, our hypothesis was that including patients with catastrophic brain damage admitted to the ICU solely for the purpose of possible organ donation might result in a significant bias in the SMRs and, ultimately, in ICU benchmarking. We found that, despite being a relatively small number (less than 1% of all admissions), these patients made up a significant proportion of all patients who died in the ICU (7%). The ranking position in the benchmarking league table changed for 70% of the ICUs when PODs were excluded. This change suggests that PODs cause a major bias in SMR calculations. Therefore, excluding PODs in ICU benchmarking is advisable.

6.2.2 Interpretation and comparison with previous studies

Benchmarking the quality of ICUs is based on the difference between observed mortality and predicted mortality. This does not apply to PODs. If PODs are carefully identified, their probability of death should approach 100%. However, prediction scores do not capture this extremely high likelihood of death. For PODs, it is common that the only failing organ system is the central nervous system, whereas other organ systems function properly. Therefore, PODs do not score high severity-of-illness scores and the predicted probability of death remains erroneously low. This was clearly illustrated in Study IV, as the predicted mortality according to the prediction model was only 37%, whereas the observed mortality was 93%. This occurred despite using a recently developed mortality prediction model that had previously showed good calibration and discrimination.[299] The goal of treatment for PODs differs markedly from that of other ICU admissions. PODs are considered non-salvageable, and the primary objective of intensive care is not the patient's survival. Study IV was the first study, to our knowledge, to quantify the influence of PODs on ICU benchmarking. The results indicate that admitting more PODs negatively affects the SMR.

The only ICU benchmark programme that publicly announces the exclusion of PODs in the British Intensive Care National Audit & Research Centre (ICNARC) model.[171.300]. In other models, information regarding the inclusion or exclusion of PODs is generally not published, so it is assumed that, currently, the exclusion of PODs is not common practise across the benchmarking programmes.

In an observation conducted by Friele et al. of 80 Dutch hospitals and 868 organ donors in 1998–2002, it was noticed that 81% of the donors came from one quarter of hospitals. These hospitals were large and had a neurosurgery department.[301] An earlier study by Sheehy et al. of over 18,000 PODs in the United States and a smaller Danish cohort study confirmed these results: larger hospitals tend to have higher proportions of organ donors compared to overall admissions. Both Friele et al. and Sheehy et al. speculated that this is probably due to higher resources for identifying PODs.[302,303] However, the results of previous studies do not completely align with those of Study IV. In the FICC, there were POD

admissions in all ICUs. Nevertheless, the proportion of PODs varied significantly, namely, from one to 18 out of 1000 admissions during the study period, but there was no association between hospital size and the POD admission rate.

6.3 LIMITATIONS

In Studies I and II, observing SOFA scores, the validated data of SOFA scores were limited to the first 24 h after ICU admission. Therefore, we were not able to assess the comparability of the subscores on the following days and the differences in ΔSOFA scores. Moreover, in both Studies I and II, approximately one third of the hepatic SOFA scores were missing, and we were forced to use the normality assumption, thereby potentially losing data from patients who might have met the criteria for hepatic dysfunction. However, such patients are most likely few in number.

For Study II, the main limitation was the single-centre study design. The generalisability of the VIS results is therefore limited. In general, creating a score using only the development cohort without internal or external validation is not recommended.[304] However, in this study, we examined a score (VIS), which has been previously validated in several studies, and confirmed that it is justified to expand its use in the general ICU population. Moreover, we were not able to track the duration of the maximum infusion rate of each agent.

In Study III, the data on noradrenaline infusion rates were missing for a substantial number of admissions in the validation cohort. Therefore, we had to exclude a significant number of admissions. In contrast to the development cohort cardiac surgery and elective admissions were included in the validation cohort. Including these data improved the comprehensiveness of the cohort but may limit comparability to the development cohort. Furthermore, there were some other significant differences in patient characteristics between the development cohort and validation cohort. For instance, noradrenaline was administered to 35% of patients in the development cohort compared to only 14% in the validation cohort.

For Study IV, the number of PODs was relatively small (514 admissions). The impact of PODs might have been even greater if they had constituted a

larger proportion of the patients admitted. The units included were limited to Finland three units from Estonia and one from Switzerland. Thus, the results may not be generalisable to other geographical regions.

6.4 FUTURE PERSPECTIVES

6.4.1 SOFA II

Over the past few years, an increasing number of experts have recognised the outdatedness of the SOFA score and the need to update it.[119,247,305] Thus, nearly 30 years after its original publication, a consensus working group has initiated the update of the SOFA score.[306,307]

Of all the subscores that make up the SOFA score, the cardiovascular one is in most urgent need of revision. For the SOFA II cardiovascular subscore, there are several important points to consider. The vasoactive agent cutoffs must be revised so that the score truly reflects the severity of circulatory failure. The scoring for additional vasoactive agents should be applicable either by the VIS or NEE. This allows individualised vasoactive treatment and deviation from noradrenaline to other vasopressors in limited-resource settings.[244,308–310] The data-based noradrenaline dose cutoffs and the data on VIS categories found in this study might be useful for the SOFA score update process. According to the results of study III, cutoffs of 0.2 and 0.4 µg/kg/min might be justified for determination of noradrenaline dose. Moreover, mechanical support devices should be included in the score. At the time the SOFA score was under development, the use of such devices was rare; indeed, the use of VA-ECMO increased 23-fold in the United States between 2002 and 2022.[311]

In respiratory component of the SOFA score, the PaO_2/FiO_2 cutoffs may also need reassessment. The respiratory SOFA calculation requires blood gas analysis, which may not always be available in resource-limited settings. There have been suggestions of using peripheral oxygen saturation (SpO_2) measurements as an alternative for original respiratory SOFA score.[312,313] Moreover, the use of ECMO refers to severe respiratory failure regardless of the pO_2 levels.

Neurologic points (GCS) are the least reliable of all SOFA subscores.[126] This concerns particuarly sedated and intubated patients. To overcome the

reliability problem related to GCS, the Full Outline of UnResponsiveness (FOUR) score has been developed for more accurate neurologic evaluation sedated and intubated patients.[314] In the FOUR score, motor and eye responses, respiratory pattern and brainstem reflexes are evaluated. FOUR score has been as valid as the GCS – or even superior to it – in predicting outcomes and it might be a possible complementation to neurologic organ dysfunction evaluation.[314–317] However, this still doesn't fully eliminate the significant confounding factor of sedation when evaluating the level of neurological status.

Furthermore, the current SOFA score doesn't take into account the dysfunction of the gastrointestinal and immunologic systems. It is relevant to consider including these organ systems in the SOFA II score. Recently, it has been shown that there are slight sex-specific differences in the current SOFA score components. However, these differences are small, and it is questionable whether sex-specific SOFA score criteria should be incorporated into the next SOFA score version.[318]

The SOFA score update should be done carefully. An optimal compromise between simplicity and prognostic accuracy should be attained. Prior to its implementation, it is advisable to conduct appropriate validation on the SOFA II score.

6.4.1.1 Future of ICU performance measures

The performance benchmarking of healthcare systems is a growing field worldwide. For decision-makers, it is imperative to understand how limited resources are utilised. Benchmarking programmes should provide expectations on how the invested resources lead to benefits in terms of improved outcomes.[237,319,320] However, there is insufficient concrete evidence to determine whether benchmarking improves the performance of healthcare units.[321,322] More research is warranted to answer the question of whether and how benchmarking efforts result in better quality and, ultimately, improved health for patients. SMRs, which are considered more appropriate measures compared to crude mortality rates, are currently the most used measure in the ICU benchmarking.

SMRs are prone to several confounding factors. For POD admissions, the goal of treatment deviates significantly from the other ICU admissions.

This causes major bias to SMRs and ICU benchmarking. When calculating SMRs for ICU benchmarking purpose, it is imperative to exclude the PODs as well as other admission groups with markedly deviating goals of treatment. An example of such admissions, where the treatment goal is not primarily to preserve the patient's life, includes Palliative ICU admissions. Palliative ICU admissions constitute a small but slowly growing proportion of all admissions.[323] The effect of these admissions on SMRs is not studied. Moreover, ICU transfers cause substantial bias in SMRs.[187, 217].

To overcome the transfer bias, fixed-day outcomes, such as 30-day mortality, are likely a better option for assessing outcomes. Furthermore, medium-term, and long-term outcomes, such as 3-month or 12-month survival, might be advisable.[324] From the patient´s perspective it is also important to consider the quality of life after the critical illness. The use of health-related quality of life has been increasingly examined as an outcome measure and it may also serve as an ICU quality benchmarking measure.[325,326] From a funder's perspective, the frequency of readmissions and, in addition to severity-adjusted outcomes, cost-adjusted outcomes are of great interest in the future, with a focus on resource allocation.[237]

The current mortality prediction models are based on relatively simple data, such as diagnoses, a few physiological parameters, age, and a couple of variables describing chronic diseases. However, structured patient record systems, monitoring devices, drug infusion devices, ventilators, and renal replacement devices continuously generate an enormous amount of data.[327] This enables development of machine learning-based prediction models. In recent years, an increasing number of papers has reported the superiority of machine learning models over the traditional mortality prediction models.[328–330] Thereby, it is possible that machine learning and artificial intelligence driven models will have a significant role in benchmarking programmes in the future.

7 CONCLUSION

1. The organ-specific components of the SOFA score do not have equal weights as predictors of mortality. The cardiovascular component suffers from inflation and high scores do not necessarily mean severe cardiovascular failure. The score needs to be updated.
2. An increase in VIS is associated with an increased risk of mortality in the general ICU population. Replacing the cardiovascular SOFA score with a VIS-based score improves the performance of the total SOFA score.
3. The noradrenaline dose alone might be a viable way to determine the level of cardiovascular dysfunction. Cutoffs of 0.2 µg/kg/min and 0.4 µg/kg/min are suitable for categorising low-dose, intermediate-dose, and high-dose ranges of noradrenaline.
4. Despite their small numbers, patients admitted to the ICU for the sole purpose of potential organ donation cause bias in SMR calculations and ICU benchmarking. PODs should be excluded from SMR calculations.

REFERENCES

1. Zorab J. The resuscitation greats. Bjørn Ibsen. *Resuscitation*. 2003;57(1):3-9. doi:doi: 10.1016/s0022-3476(54)80220-4.
2. Wunsch H. The outbreak that invented intensive care. *Nature*. 2020. doi:10.1038/d41586-020-01019-y
3. Barsky B. Tracheotomy in an eleven-week-old infant with bulbospinal poliomyelitis: case report and brief summary of the literature. *J Pediatr*. 1954;44(4):445-447. doi:10.1016/s0022-3476(54)80220-4
4. Harries J, Lawes W. Intermittent positive-pressure respiration in bulbo-spinal poliomyelitis; use of the Radcliffe respiration pump. *Br Med J*. 1955;1(4911):448-454. doi:10.1136/bmj.1.4911.448
5. Berthelsen PG, Cronqvist M. The first intensive care unit in the world: Copenhagen 1953. *Acta Anaesthesiol Scand*. 2003;47(10):1190-1195. doi:10.1046/j.1399-6576.2003.00256.x
6. Marshall JC, Bosco L, Adhikari NK, et al. What is an intensive care unit? A report of the task force of the World Federation of Societies of Intensive and Critical Care Medicine. *J Crit Care*. 2017;37:270-276. doi:10.1016/j.jcrc.2016.07.015
7. Baue AE. Multiple Organ Failure, Multiple Organ Dysfunction Syndrome, and Systemic Inflammatory Response Syndrome: Why No Magic Bullets? *Archives of Surgery*. 1997;132(7):703-707. doi:10.1001/archsurg.1997.01430310017002
8. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med*. 1996;22(7):707-710. doi:10.1007/BF01709751
9. Stevenson EK, Rubenstein AR, Radin GT, Wiener RS, Walkey AJ. Two decades of mortality trends among patients with severe sepsis: a comparative meta-analysis*. *Crit Care Med*. 2014;42(3):625-631. doi:10.1097/CCM.0000000000000026
10. Knaus WA, Zimmerman JE, Wagner DP, Draper EA, Lawrence DE. APACHE-acute physiology and chronic health evaluation: a physiologically based classification system. *Crit Care Med*. 1981;9(8):591-597. doi:10.1097/00003246-198108000-00008

11. Gall JRLE, Loirat P, Alperovitch A, et al. A simplified acute physiology score for ICU patients. *Crit Care Med*. 1984;12(11). doi:10.1097/00003246-198411000-00012
12. Lemeshow S, Teres D, Pastides H, Avrunin JS, Steingup J. A method for predicting survival and mortality of ICU patients using objectively derived weights. *Crit Care Med*. 1985;13(7). doi:10.1097/00003246-198507000-00001
13. Watson SR, Scales DC. Improving Intensive Care Unit Quality Using Collaborative Networks. *Crit Care Clin*. 2013;29(1):77-89. doi:10.1016/j.ccc.2012.10.008
14. Quintairos A, Pilcher D, Salluh JIF. ICU scoring systems. *Intensive Care Med*. 2023;49(2):223-225. doi:10.1007/s00134-022-06914-8
15. Salluh JIF, Soares M, Keegan MT. Understanding intensive care unit benchmarking. *Intensive Care Med*. 2017;43(11):1703-1707. doi:10.1007/s00134-017-4760-x
16. Tilney N, Bailey G, Morgan A. Sequential system failure after rupture of abdominal aortic aneurysms: an unsolved problem in postoperative care. *Ann Surg* . 1973;178(2):117-122. doi:10.1097/00000658-197308000-00001
17. Shoemaker W.C. Editorial: Multiple injuries and multiple organ failure. *Crit Care Med*. 1973;1(3):157. doi:10.1097/00003246-197305000-00010
18. Bauel AE. Multiple, Progressive, or Sequential Systems Failure: A Syndrome of the 1970s. *Archives of Surgery*. 1975;110(7):779-781. doi:10.1001/archsurg.1975.01360130011001
19. Grenvik A. Respiratory care in multiple organ failure. *The Japanese journal of anesthesiology*. 1973;22(10):973-978.
20. Vincent JL, Sakr Y, Sprung CL, et al. Sepsis in European intensive care units: Results of the SOAP study. *Crit Care Med*. 2006;34(2). doi:10.1097/01.ccm.0000194725.48928.3a
21. Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: Analysis of incidence, outcome, and associated costs of care. *Crit Care Med*. 2001;29(7). doi:10.1097/00003246-200107000-00002
22. Vincent JL, de Mendonca A, Cantraine F, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care

units: Results of a multicenter, prospective study. *Crit Care Med*. 1998;26(11). doi:10.1097/00003246-199811000-00016
23. Moreno R, Vincent JL, Matos R, et al. The use of maximum SOFA score to quantify organ dysfunction/failure in intensive care. Results of a prospective, multicentre study. *Intensive Care Med*. 1999;25(7):686-696. doi:10.1007/s001340050931
24. Pedersen PB, Hrobjartsson A, Nielsen DL, Henriksen DP, Brabrand M, Lassen AT. Prevalence and prognosis of acutely ill patients with organ failure at arrival to hospital: A systematic review. *PLoS One*. 2018;13(11):e0206610-e0206610. doi:10.1371/journal.pone.0206610
25. Sakr Y, Lobo SM, Moreno RP, et al. Patterns and early evolution of organ failure in the intensive care unit and their relation to outcome. *Crit Care*. 2012;16(6):R222. doi:10.1186/cc11868
26. Lone NI, Walsh TS. Impact of Intensive Care Unit Organ Failures on Mortality during the Five Years after a Critical Illness. *Am J Respir Crit Care Med*. 2012;186(7):640-647. doi:10.1164/rccm.201201-0059OC
27. Zhao Y, Li X, Xue Q, Gao W, Wang S. Impact factors of multiple organ dysfunction syndromes complicating acute myocardial infarction in the elderly: multivariate logistic regression analysis. *Crit Care*. 2008;12(Suppl 2). doi:10.1186/cc6465
28. Feng G, Luo Q, Guo E, et al. Multiple organ dysfunction syndrome, an unusual complication of heroin intoxication: A case report and review of literature. *Int J Clin Exp Pathol*. 2015;8(9).
29. Nast-Kolb D, Aufmkolk M, Rucholtz S, Obertacke U, Waydhas C. Multiple Organ Failure Still a Major Cause of Morbidity but Not Mortality in Blunt Multiple Trauma. *Journal of Trauma and Acute Care Surgery*. 2001;51(5). doi:10.1097/00005373-200111000-00003
30. Baker S, O'Neill B, Haddon W, Long W. The injury severity score: A method for describing patients with multiple injuries and evaluating emergency care. *Journal of Trauma and Acute Care Surgery*. 1974;14(3).
31. Hutchings L, Watkinson P, Young JD, Willett K. Defining multiple organ failure after major trauma: A comparison of the Denver, Sequential Organ Failure Assessment, and Marshall scoring systems. doi:10.1097/TA.0000000000001328

32. Cole E, Gillespie S, Vulliamy P, et al. Multiple organ dysfunction after trauma. *British Journal of Surgery*. 2020;107(4):402-412. doi:10.1002/bjs.11361
33. Regel G, Grotz M, Weltner T, Sturm JA, Tscherne H. Pattern of Organ Failure following Severe Trauma. *World J Surg*. 1996;20(4):422-429. doi:10.1007/s002689900067
34. Shepherd JM, Cole E, Brohi K. Contemporary Patterns of Multiple Organ Dysfunction in Trauma. *Shock*. 2017;47(4). doi:10.1097/SHK.0000000000000779
35. Fröhlich M, Lefering R, Probst C, et al. Epidemiology and risk factors of multiple-organ failure after multiple trauma: An analysis of 31,154 patients from the TraumaRegister DGU. *Journal of Trauma and Acute Care Surgery*. 2014;76(4). doi:10.1097/TA.0000000000000199
36. Namas RA, Almahmoud K, Mi Q, et al. Individual-specific principal component analysis of circulating inflammatory mediators predicts early organ dysfunction in trauma patients. *J Crit Care*. 2016;36:146-153. doi:10.1016/j.jcrc.2016.07.002
37. Vanzant EL, Lopez CM, Ozrazgat-Baslanti T, et al. Persistent inflammation, immunosuppression, and catabolism syndrome after severe blunt trauma. *Journal of Trauma and Acute Care Surgery*. 2014;76(1). doi:10.1097/TA.0b013e3182ab1ab5
38. Moore FA, Sauaia A, Moore EE, Haenel JB, Burch JM, Lezotte DC. Postinjury Multiple Organ Failure: A Bimodal Phenomenon. *Journal of Trauma and Acute Care Surgery*. 1996;40(4). doi:10.1097/00005373-199604000-00001
39. Eriksson J, Nelson D, Holst A, Hellgren E, Friman O, Oldner A. Temporal patterns of organ dysfunction after severe trauma. *Crit Care*. 2021;25(1):165. doi:10.1186/s13054-021-03586-6
40. Durham RM, Moran JJ, Mazuski JE, Shapiro MJ, Baue AE, Flint LM. Multiple Organ Failure in Trauma Patients. *Journal of Trauma and Acute Care Surgery*. 2003;55(4).
41. Ledingham IM. Septic shock. *British Journal of Surgery*. 1975;62(10):777-780. doi:10.1002/bjs.1800621008
42. Barie P, Hydo L. Epidemiology of Multiple Organ Dysfunction Syndrome in Critical Surgical Illness. *Surg Infect (Larchmt)*. 2000;1(3):173-186. doi:10.1089/109629600750018105

43. Kaukonen KM, Bailey M, Suzuki S, Pilcher D, Bellomo R. Mortality Related to Severe Sepsis and Septic Shock Among Critically Ill Patients in Australia and New Zealand, 2000-2012. *JAMA*. 2014;311(13):1308-1316. doi:10.1001/jama.2014.2637
44. Lorencio Cárdenas C, Yébenes JC, Vela E, et al. Trends in mortality in septic patients according to the different organ failure during 15 years. *Crit Care*. 2022;26(1):302. doi:10.1186/s13054-022-04176-w
45. 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
46. 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(6):1644-1655. doi:10.1378/chest.101.6.1644
47. Vincent JL. Dear SIRS, I'm sorry to say that I don't like you. *Crit Care Med*. 1997;25(2). doi:10.1097/00003246-199702000-00029
48. Drewry AM, Hotchkiss RS. Revising definitions of sepsis. *Nat Rev Nephrol*. 2015;11(6):326-328. doi:10.1038/nrneph.2015.66
49. Kaukonen KM, Bailey M, Pilcher D, Cooper DJ, Bellomo R. Systemic Inflammatory Response Syndrome Criteria in Defining Severe Sepsis. *New England Journal of Medicine*. 2015;372(17):1629-1638. doi:10.1056/NEJMoa1415236
50. Padhan RK, Jain S, Agarwal S, et al. Primary and Secondary Organ Failures Cause Mortality Differentially in Acute Pancreatitis and Should be Distinguished. *Pancreas*. 2018;47(3). doi:10.1097/MPA.0000000000000998
51. Schepers NJ, Bakker OJ, Besselink MG, et al. Impact of characteristics of organ failure and infected necrosis on mortality in necrotising pancreatitis. *Gut*. 2019;68(6). doi:10.1136/gutjnl-2017-314657
52. Kamal A, Sinha A, Hutfless SM, et al. Hospital admission volume does not impact the in-hospital mortality of acute pancreatitis. *HPB*. 2017;19(1). doi:10.1016/j.hpb.2016.10.013
53. Hamada S, Masamune A, Kikuta K, Hirota M, Tsuji I, Shimosegawa T. Nationwide epidemiological survey of acute pancreatitis in Japan. *Pancreas*. 2014;43(8). doi:10.1097/MPA.0000000000000200
54. Shen HN, Lu CL. Incidence, resource use, and outcome of acute pancreatitis with/without intensive care: A nationwide population-

based study in Taiwan. *Pancreas*. 2011;40(1). doi:10.1097/MPA.0b013e3181f7e750
55. Garg PK, Singh VP. Organ Failure Due to Systemic Injury in Acute Pancreatitis. *Gastroenterology*. 2019;156(7). doi:10.1053/j.gastro.2018.12.041
56. Isenmann R, Rau B, Beger HG. Early severe acute pancreatitis: Characteristics of a new subgroup. *Pancreas*. 2001;22(3). doi:10.1097/00006676-200104000-00008
57. Noel P, Patel K, Durgampudi C, et al. Peripancreatic fat necrosis worsens acute pancreatitis independent of pancreatic necrosis via unsaturated fatty acids increased in human pancreatic necrosis collections. *Gut*. 2016;65(1). doi:10.1136/gutjnl-2014-308043
58. Nawaz H, Koutroumpakis E, Easler J, et al. Elevated serum triglycerides are independently associated with persistent organ failure in acute pancreatitis. *American Journal of Gastroenterology*. 2015;110(10). doi:10.1038/ajg.2015.261
59. Amma Preethi T, Nizarudeen A. Prevalence of multi-organ dysfunction syndrome amongst patients with acute pancreatitis. *Kerala Surgical Journal*. 2022;28(1). doi:10.4103/ksj.ksj_19_22
60. Sternby H, Bolado F, Canaval-Zuleta HJ, et al. Determinants of Severity in Acute Pancreatitis: A Nation-wide Multicenter Prospective Cohort Study. *Ann Surg*. 2019;270(2). doi:10.1097/SLA.0000000000002766
61. Machicado JD, Gougol A, Tan X, et al. Mortality in acute pancreatitis with persistent organ failure is determined by the number, type, and sequence of organ systems affected. *United European Gastroenterol J*. 2021;9(2). doi:10.1002/ueg2.12057
62. Halonen KI, Pettilä V, Leppäniemi AK, Kemppainen EA, Puolakkainen PA, Haapiainen RK. Multiple organ dysfunction associated with severe acute pancreatitis. *Crit Care Med*. 2002;30(6). doi:10.1097/00003246-200206000-00019
63. Banks PA, Bollen TL, Dervenis C, et al. Classification of acute pancreatitis - 2012: Revision of the Atlanta classification and definitions by international consensus. *Gut*. 2013;62(1). doi:10.1136/gutjnl-2012-302779

64. Knaus W.A. Draper EA, Wagner DP, Zimmerman JE. Prognosis in acute organ-system failure. *Ann Surg*. 1985;202(6):685-693. doi:10.1097/00000658-198512000-00004
65. Goris RJA, te Boekhorst TPA, Nuytinck JKS, Gimbrère JSF. Multiple-Organ Failure: Generalized Autodestructive Inflammation? *Archives of Surgery*. 1985;120(10):1109-1115. doi:10.1001/archsurg.1985.01390340007001
66. Marshall JC, Cook DJ, Christinou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple Organ Dysfunction Score: A reliable descriptor of a complex clinical outcome. *Crit Care Med*. 1995;23(10):1638-1652.
67. Marshall JC. The Multiple Organ Dysfunction (MOD) Score. *Sepsis (Boston)*. 1994;1:49-52.
68. Le Gall JR, Klar J, Lemeshow S, et al. The Logistic Organ Dysfunction System: A New Way to Assess Organ Dysfunction in the Intensive Care Unit. *JAMA*. 1996;276(10):802-810. doi:10.1001/jama.1996.03540100046027
69. Metnitz PG Lang T, Valentin A, Steltzer H, Krenn KC, Le Gall JR. Evaluation of the logistic organ dysfunction system for the assessment of organ dysfunction and mortality in critically ill patients. *Intensive Care Med*. 2001;27(6):992-998. doi:10.1007/s001340100888
70. Bernard GR, Doig BG, Hudson G. Quantification of organ failure for clinical trials and clinical practice. *Am J Respir Crit Care Med*. 1995;151:A323.
71. Le Gall JR, Klar J, Lemeshow S, et al. The Logistic Organ Dysfunction System: A New Way to Assess Organ Dysfunction in the Intensive Care Unit. *JAMA*. 1996;276(10):802-810. doi:10.1001/jama.1996.03540100046027
72. Bernard GR. Quantification of organ dysfunction: Seeking standardization. *Crit Care Med*. 1998;26(11). doi:10.1097/00003246-199811000-00001
73. Lambden S, Laterre PF, Levy MM, Francois B. The SOFA score—development, utility and challenges of accurate assessment in clinical trials. *Crit Care*. 2019;23(1):374. doi:10.1186/s13054-019-2663-7

74. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial Evaluation of the SOFA Score to Predict Outcome in Critically Ill Patients. *JAMA*. 2001;286(14):1754-1758. doi:10.1001/jama.286.14.1754
75. Soo A, Zuege DJ, Fick GH, et al. Describing organ dysfunction in the intensive care unit: a cohort study of 20,000 patients. *Crit Care*. 2019;23(1):186. doi:10.1186/s13054-019-2459-9
76. Peres Bota D, Melot C, Lopes Ferreira F, Nguyen Ba V, Vincent JL. The Multiple Organ Dysfunction Score (MODS) versus the Sequential Organ Failure Assessment (SOFA) score in outcome prediction. *Intensive Care Med*. 2002;28(11):1619-1624. doi:10.1007/s00134-002-1491-3
77. Zygun D, Berthiaume L, Laupland K, Kortbeek J, Doig C. SOFA is superior to MOD score for the determination of non-neurologic organ dysfunction in patients with severe traumatic brain injury: a cohort study. *Crit Care*. 2006;10(4):R115-R115. doi:10.1186/cc5007
78. Fröhlich M, Wafaisade A, Mansuri A, et al. Which score should be used for posttraumatic multiple organ failure? - Comparison of the MODS, Denver- and SOFA- Scores. *Scand J Trauma Resusc Emerg Med*. 2016;24(1):130. doi:10.1186/s13049-016-0321-5
79. Timsit JF, Fosse JP, Troché G, et al. Calibration and discrimination by daily Logistic Organ Dysfunction scoring comparatively with daily Sequential Organ Failure Assessment scoring for predicting hospital mortality in critically ill patients. *Crit Care Med*. 2002;30(9). doi:10.1097/00003246-200209000-00009
80. Ranieri V. Marco, Rubenfield Gordon, Thompson B. Taylor, et al. Acute respiratory distress syndrome: The Berlin definition. *JAMA*. 2012;307(23):2526-2533. doi:10.1001/jama.2012.5669
81. Matthay MA, Thompson BT, Ware LB. The Berlin definition of acute respiratory distress syndrome: should patients receiving high-flow nasal oxygen be included? *Lancet Respir Med*. 2021;9(8):933-936. doi:10.1016/S2213-2600(21)00105-3
82. Murugappan K, Walsh D, Mittel A, Sontag D, Shaefi S. Veno-venous extracorporeal membrane oxygenation allocation in the COVID-19 pandemic. *J Crit Care*. 2021;61:221-226. doi:10.1016/j.jcrc.2020.11.004

83. Tomii K, Terada J, Chin K. Increased usage of the high flow nasal cannula in COVID-19 cases in Japan -from the online questionnaire survey by the Japanese Respiratory Society. *Respir Investig*. 2021;59(5):666-669. doi:10.1016/j.resinv.2021.05.009
84. Tonna JE, Abrams D, Brodie D, et al. Management of Adult Patients Supported with Venovenous Extracorporeal Membrane Oxygenation (VV ECMO): Guideline from the Extracorporeal Life Support Organization (ELSO). *ASAIO J*. 2021;67(6):601-610. doi:10.1097/MAT.0000000000001432
85. Sang Y, Roest M, de Laat B, de Groot PG, Huskens D. Interplay between platelets and coagulation. *Blood Rev*. 2021;46:100733. doi:10.1016/j.blre.2020.100733
86. Baughman RR, Lower EE, Flessa HC, Tollerud DJ. Thrombocytopenia in the Intensive Care Unit. *Chest*. 1993;104(4):1243-1247. doi:10.1378/chest.104.4.1243
87. Strauss R, Wehler M, Mehler K, Kreutzer D, Koebnick C, Hahn EG. Thrombocytopenia in patients in the medical intensive care unit: Bleeding prevalence, transfusion requirements, and outcome. *Crit Care Med*. 2002;30(8). doi:10.1097/00003246-200208000-00015
88. Vanderschueren S, De Weerdt A, Malbrain M, et al. Thrombocytopenia and prognosis in intensive care. *Crit Care Med*. 2000;28(6). doi:10.1097/00003246-200006000-00031
89. Gonzalez DA, Kumar R, Asif S, Bali A, Dang AK. Sepsis and Thrombocytopenia: A Nowadays Problem. *Cureus*. Published online 2022. doi:10.7759/cureus.25421
90. Forzley BR, Clark WF. TTP/HUS and prognosis: the syndrome and the disease(s). *Kidney Int*. 2009;75:S59-S61. doi:https://doi.org/10.1038/ki.2008.623
91. Arepally GM, Padmanabhan A. Heparin-Induced Thrombocytopenia. *Arterioscler Thromb Vasc Biol*. 2021;41(1):141-152. doi:10.1161/ATVBAHA.120.315445
92. Gallo T, Curry SC, Padilla-Jones A, et al. A computerized scoring system to improve assessment of heparin-induced thrombocytopenia risk. *Journal of Thrombosis and Haemostasis*. 2019;17(2):383-388. doi:https://doi.org/10.1111/jth.14359

93. Stoll F, Gödde M, Leo A, Katus HA, Müller OJ. Characterization of hospitalized cardiovascular patients with suspected heparin-induced thrombocytopenia. *Clin Cardiol*. 2018;41(12):1521-1526. doi:10.1002/clc.23061
94. Thrombocytopenia in gram-negative bacteria. *JAMA*. 1966;195(5):680-681.
95. Iba T, Helms J, Neal MD, Levy JH. Mechanisms and management of the coagulopathy of trauma and sepsis: trauma-induced coagulopathy, sepsis-induced coagulopathy, and disseminated intravascular coagulation. *Journal of Thrombosis and Haemostasis*. 2023;21(12). doi:10.1016/j.jtha.2023.05.028
96. Bakhtiari K, Meijers JCM, de Jonge E, Levi M. Prospective validation of the International Society of Thrombosis and Haemostasis scoring system for disseminated intravascular coagulation. *Crit Care Med*. 2004;32(12). doi:10.1097/01.ccm.0000147769.07699.e3
97. Gando S, Saitoh D, Ogura H, et al. Natural history of disseminated intravascular coagulation diagnosed based on the newly established diagnostic criteria for critically ill patients: Results of a multicenter, prospective survey. *Crit Care Med*. 2008;36(1). doi:10.1097/01.CCM.0000295317.97245.2D
98. Roy AM, Konda M, Goel A, Sasapu A. Epidemiology and Factors Affecting Mortality of Hospitalized Patients with Disseminated Intravascular Coagulation in the United States. *Blood*. 2019;134(Supplement_1):2404. doi:10.1182/blood-2019-132071
99. Palta S, Saroa R, Palta A. Overview of the coagulation system. *Indian J Anaesth*. 2014;58(5):515-523. doi:10.4103/0019-5049.144643
100. Adamzik M, Langemeier T, Frey UH, et al. Comparison of Thrombelastometry with Simplified Acute Physiology Score II and Sequential Organ Failure Assessment Scores for the Prediction of 30-Day Survival: A Cohort Study. *Shock*. 2011;35(4). doi:10.1097/SHK.0b013e318204bff6
101. Müller MC, Meijers JCM, Vroom MB, Juffermans NP. Utility of thromboelastography and/or thromboelastometry in adults with sepsis: a systematic review. *Crit Care*. 2014;18(1):R30. doi:10.1186/cc13721

102. Group TES. EPISEPSIS: a reappraisal of the epidemiology and outcome of severe sepsis in French intensive care units. *Intensive Care Med*. 2004;30(4):580-588. doi:10.1007/s00134-003-2121-4
103. Abcejo A, Andrejko KM, Ochroch EA, Raj NR, Deutschman CS. Impaired Hepatocellular Regeneration in Murine Sepsis Is Dependent on Regulatory Protein Levels. *Shock*. 2011;36(5). doi:10.1097/SHK.0b013e31822d60ff
104. Weiss YG, Bellin L, Kim PK, et al. Compensatory hepatic regeneration after mild, but not fulminant, intraperitoneal sepsis in rats. *American Journal of Physiology-Gastrointestinal and Liver Physiology*. 2001;280(5):G968-G973. doi:10.1152/ajpgi.2001.280.5.G968
105. Wendon J, Cordoba J, Dhawan A, et al. EASL Clinical Practical Guidelines on the management of acute (fulminant) liver failure. *J Hepatol*. 2017;66(5):1047-1081. doi:10.1016/j.jhep.2016.12.003
106. Williams R, Schalm SW, O'Grady JG. Acute liver failure: redefining the syndromes. *The Lancet*. 1993;342(8866). doi:10.1016/0140-6736(93)91818-7
107. Bretherick AD, Craig DGN, Masterton G, et al. Acute liver failure in scotland between 1992 and 2009; incidence, aetiology and outcome. *QJM: An International Journal of Medicine*. 2011;104(11). doi:10.1093/qjmed/hcr098
108. Bower WA, Johns M, Margolis HS, Williams IT, Bell BP. Population-based surveillance for acute liver failure. *American Journal of Gastroenterology*. 2007;102(11). doi:10.1111/j.1572-0241.2007.01388.x
109. Thanapirom K, Treeprasertsuk S, Soonthornworasiri N, et al. The incidence, etiologies, outcomes, and predictors of mortality of acute liver failure in Thailand: A population-base study. *BMC Gastroenterol*. 2019;19(1). doi:10.1186/s12876-019-0935-y
110. Moreau R, Jalan R, Gines P, et al. Acute-on-Chronic Liver Failure Is a Distinct Syndrome That Develops in Patients With Acute Decompensation of Cirrhosis. *Gastroenterology*. 2013;144(7):1426-1437.e9. doi:10.1053/j.gastro.2013.02.042
111. Yan J, Li S, Li S. The role of the liver in sepsis. *Int Rev Immunol*. 2014;33(6):498-510. doi:10.3109/08830185.2014.889129

112. The EPISEPSIS Study Group. EPISEPSIS: a reappraisal of the epidemiology and outcome of severe sepsis in French intensive care units. *Intensive Care Med*. 2004;30(4):580-588. doi:10.1007/s00134-003-2121-4
113. Vincent JL, De Backer D. Circulatory Shock. *New England Journal of Medicine*. 2013;369(18):1726-1734. doi:10.1056/NEJMra1208943
114. Kjaergaard J, Møller JE, Schmidt H, et al. Blood-Pressure Targets in Comatose Survivors of Cardiac Arrest. *New England Journal of Medicine*. 2022;387(16):1456-1466. doi:10.1056/NEJMoa2208687
115. Lamontagne F, Richards-Belle A, Thomas K, et al. Effect of Reduced Exposure to Vasopressors on 90-Day Mortality in Older Critically Ill Patients With Vasodilatory Hypotension: A Randomized Clinical Trial. *JAMA*. 2020;323(10):938-949. doi:10.1001/jama.2020.0930
116. Menacho ST, Floyd C. Current practices and goals for mean arterial pressure and spinal cord perfusion pressure in acute traumatic spinal cord injury: Defining the gaps in knowledge. *J Spinal Cord Med*. 2021;44(3):350-356. doi:10.1080/10790268.2019.1660840
117. Thongprayoon C, Cheungpasitporn W, Harrison AM, et al. Temporal trends in the utilization of vasopressors in intensive care units: an epidemiologic study. *BMC Pharmacol Toxicol*. 2016;17(1):19. doi:10.1186/s40360-016-0063-z
118. Burger A, Bellomo R, Chapman M, Finfer S, Hickling K, Myburgh J. Low-dose dopamine in patients with early renal dysfunction: A placebo-controlled randomised trial. *Lancet*. 2000;356(9248). doi:10.1016/S0140-6736(00)03495-4
119. Bachmann KF, Arabi YM, Regli A, Starkopf J, Reintam Blaser A. Cardiovascular SOFA score may not reflect current practice. *Intensive Care Med*. 2022;48(1):119-120. doi:10.1007/s00134-021-06536-6
120. Bloom JE, Chan W, Kaye DM, Stub D. State of Shock: Contemporary Vasopressor and Inotrope Use in Cardiogenic Shock. *J Am Heart Assoc*. 2023;12(15). doi:10.1161/JAHA.123.029787
121. Teasdale G, Jennett B. Assessment of coma and impaired consciousness: A Practical Scale. *The Lancet*. 1974;304(7872):81-84. doi:10.1016/S0140-6736(74)91639-0
122. Le Gall JR, Lemeshow S, Saulnier F. A New Simplified Acute Physiology Score (SAPS II) Based on a European/North American

Multicenter Study. *JAMA*. 1993;270(24):2957-2963. doi:10.1001/jama.1993.03510240069035
123. Wagner DP, Knaus WA, Draper EA. *Statistical Validation of a Severity of Illness Measure*.; 1983;73(8):878-884. doi: 10.2105/ajph.73.8.878
124. Murray GD, Butcher I, McHugh GS, et al. Multivariable Prognostic Analysis in Traumatic Brain Injury: Results from The IMPACT Study. *J Neurotrauma*. 2007;24(2):329-337. doi:10.1089/neu.2006.0035
125. Healey C, Osler TM, Rogers FB, et al. Improving the Glasgow Coma Scale Score: Motor Score Alone Is a Better Predictor. *Journal of Trauma and Acute Care Surgery*. 2003;54(4). doi:10.1097/01.TA.0000058130.30490.5D
126. Tallgren M, Bäcklund M, Hynninen M. Accuracy of Sequential Organ Failure Assessment (SOFA) scoring in clinical practice. *Acta Anaesthesiol Scand*. 2009;53(1):39-45. doi:10.1111/j.1399-6576.2008.01825.x
127. Reith FCM, Synnot A, van den Brande R, Gruen RL, Maas AIR. Factors Influencing the Reliability of the Glasgow Coma Scale: A Systematic Review. *Neurosurgery*. 2017;80(6). doi:10.1093/neuros/nyw178
128. Faubel S. Acute kidney injury and multiple organ dysfunction syndrome. *Minerva Urol Nefrol*. 2009;61(3):171-188.
129. Ronco C, Bellomo R, Kellum JA. Acute kidney injury. *The Lancet*. 2019;394(10212):1949-1964. doi:10.1016/S0140-6736(19)32563-2
130. Kellum JA, Lameire N, Aspelin P, et al. Kidney disease: Improving global outcomes (KDIGO) acute kidney injury work group. KDIGO clinical practice guideline for acute kidney injury. *Kidney Int Suppl (2011)*. 2012;2(1). doi:10.1038/kisup.2012.1
131. Summary of Recommendation Statements. *Kidney Int Suppl (2011)*. 2012;2(1):8-12. doi:10.1038/kisup.2012.7
132. Karimzadeh I, Barreto EF, Kellum JA, et al. Moving toward a contemporary classification of drug-induced kidney disease. *Crit Care*. 2023;27(1):435. doi:10.1186/s13054-023-04720-2
133. Ostermann M, Zarbock A, Goldstein S, et al. Recommendations on Acute Kidney Injury Biomarkers From the Acute Disease Quality Initiative Consensus Conference: A Consensus Statement. *JAMA Netw Open*. 2020;3(10):e2019209-e2019209. doi:10.1001/jamanetworkopen.2020.19209

134. Lim JH, Jeon Y, Ahn JS, et al. GDF-15 Predicts In-Hospital Mortality of Critically Ill Patients with Acute Kidney Injury Requiring Continuous Renal Replacement Therapy: A Multicenter Prospective Study. *J Clin Med*. 2021;10(16):3660. doi:10.3390/jcm10163660
135. Dessing MC, Tammaro A, Pulskens WP, et al. The calcium-binding protein complex S100A8/A9 has a crucial role in controlling macrophage-mediated renal repair following ischemia/reperfusion. *Kidney Int*. 2015;87(1):85-94. doi:10.1038/ki.2014.216
136. Leem AY, Park MS, Park BH, et al. Value of Serum Cystatin C Measurement in the Diagnosis of Sepsis-Induced Kidney Injury and Prediction of Renal Function Recovery. *Yonsei Med J*. 2017;58(3):604-612. doi:10.3349/ymj.2017.58.3.604
137. Han WK, Bailly V, Abichandani R, Thadhani R, Bonventre J V. Kidney Injury Molecule-1 (KIM-1): A novel biomarker for human renal proximal tubule injury. *Kidney Int*. 2002;62(1):237-244. doi:10.1046/j.1523-1755.2002.00433.x
138. White KC, Serpa-Neto A, Hurford R, et al. Sepsis-associated acute kidney injury in the intensive care unit: incidence, patient characteristics, timing, trajectory, treatment, and associated outcomes. A multicenter, observational study. *Intensive Care Med*. 2023;49(9). doi:10.1007/s00134-023-07138-0
139. Zarbock A, Nadim MK, Pickkers P, et al. Sepsis-associated acute kidney injury: consensus report of the 28th Acute Disease Quality Initiative workgroup. *Nat Rev Nephrol*. 2023;19(6). doi:10.1038/s41581-023-00683-3
140. Wernovsky G, Wypij D, Jonas RA, et al. Postoperative course and hemodynamic profile after the arterial switch operation in neonates and infants: A comparison of low-flow cardiopulmonary bypass and circulatory arrest. *Circulation*. 1995;92(8). doi:10.1161/01.CIR.92.8.2226
141. Gaies MG, Gurney JG, Yen AH, et al. Vasoactive–inotropic score as a predictor of morbidity and mortality in infants after cardiopulmonary bypass. *Pediatric Critical Care Medicine*. 2010;11(2). doi:10.1097/PCC.0b013e3181b806fc
142. Gaies MG, Jeffries HE, Niebler RA, et al. Vasoactive-inotropic score is associated with outcome after infant cardiac surgery: an analysis

from the Pediatric Cardiac Critical Care Consortium and Virtual PICU System Registries. *Pediatr Crit Care Med*. 2014;15(6):529-537. doi:10.1097/PCC.0000000000000153

143. Barge-Caballero E, Segovia-Cubero J, González-Vilchez F, et al. Evaluation of the preoperative vasoactive-inotropic score as a predictor of postoperative outcomes in patients undergoing heart transplantation. *Int J Cardiol*. 2015;185:192-194. doi:10.1016/j.ijcard.2015.03.098
144. Butts RJ, Scheurer MA, Atz AM, et al. Comparison of Maximum Vasoactive Inotropic Score and Low Cardiac Output Syndrome As Markers of Early Postoperative Outcomes After Neonatal Cardiac Surgery. *Pediatr Cardiol*. 2012;33(4):633-638. doi:10.1007/s00246-012-0193-z
145. Davidson J, Tong S, Hancock H, Hauck A, da Cruz E, Kaufman J. Prospective validation of the vasoactive-inotropic score and correlation to short-term outcomes in neonates and infants after cardiothoracic surgery. *Intensive Care Med*. 2012;38(7):1184-1190. doi:10.1007/s00134-012-2544-x
146. Kumar M, Sharma R, Sethi SK, et al. Vasoactive Inotrope Score as a tool for clinical care in children post cardiac surgery. *Indian J Crit Care Med*. 2014;18(10):653-658. doi:10.4103/0972-5229.142174
147. Sanil Y, Aggarwal S. Vasoactive–inotropic score after pediatric heart transplant: A marker of adverse outcome. *Pediatr Transplant*. 2013;17(6):567-572. doi:10.1111/petr.12112
148. Landoni G, Lomivorotov V V., Alvaro G, et al. Levosimendan for Hemodynamic Support after Cardiac Surgery. *New England Journal of Medicine*. 2017;376(21). doi:10.1056/nejmoa1616325
149. Nguyen H V., Havalad V, Aponte-Patel L, et al. Temporary biventricular pacing decreases the vasoactive-inotropic score after cardiac surgery: A substudy of a randomized clinical trial. *Journal of Thoracic and Cardiovascular Surgery*. 2013;146(2). doi:10.1016/j.jtcvs.2012.07.020
150. Favia I, Vitale V, Ricci Z. The Vasoactive-Inotropic Score and Levosimendan: Time for LVIS? *J Cardiothorac Vasc Anesth*. 2013;27(2):e15-e16. doi:10.1053/j.jvca.2012.11.009

151. Haque A, Siddiqui NR, Munir O, Saleem S, Mian A. Association between vasoactive-inotropic score and mortality in pediatric septic shock. *Indian Pediatr*. 2015;52(4). doi:10.1007/s13312-015-0630-1
152. Kara İ, Sargin M, Bayraktar YŞ, Eyiol H, Duman İ, Çelik JB. The use of vasoactive-inotropic score in adult patients with septic shock in intensive care. *Journal of Medical and Surgical Intensive Care Medicine*. 2019;10(1). doi:10.33381/dcbybd.2019.2057
153. Song J, Cho H, Park DW, et al. Vasoactive–inotropic score as an early predictor of mortality in adult patients with sepsis. *J Clin Med*. 2021;10(3). doi:10.3390/jcm10030495
154. Demirhan S, Topcuoglu S, Karadag N, Ozalkaya E, Karatekin G. Vasoactive inotropic score as a predictor of mortality in neonatal septic shock. *J Trop Pediatr*. 2022;68(6). doi:10.1093/tropej/fmac100
155. Kallekkattu D, Rameshkumar R, Chidambaram M, Krishnamurthy K, Selvan T, Mahadevan S. Threshold of Inotropic Score and Vasoactive–Inotropic Score for Predicting Mortality in Pediatric Septic Shock. *Indian J Pediatr*. 2022;89(5). doi:10.1007/s12098-021-03846-x
156. McIntosh AM, Tong S, Deakyne SJ, Davidson JA, Scott HF. Validation of the Vasoactive-Inotropic Score in Pediatric Sepsis. *Pediatr Crit Care Med*. 2017;18(8):750-757. doi:10.1097/PCC.0000000000001191
157. Koponen T, Karttunen J, Musialowicz T, Pietiläinen L, Uusaro A, Lahtinen P. Vasoactive-inotropic score and the prediction of morbidity and mortality after cardiac surgery. *Br J Anaesth*. 2019;122(4):428-436. doi:10.1016/j.bja.2018.12.019
158. Yamazaki Y, Oba K, Matsui Y, Morimoto Y. Vasoactive-inotropic score as a predictor of morbidity and mortality in adults after cardiac surgery with cardiopulmonary bypass. *J Anesth*. 2018;32(2):167-173. doi:10.1007/s00540-018-2447-2
159. Belletti A, Lerose CC, Zangrillo A, Landoni G. Vasoactive-Inotropic Score: Evolution, Clinical Utility, and Pitfalls. *J Cardiothorac Vasc Anesth*. 2021;35(10):3067-3077. doi:10.1053/j.jvca.2020.09.117
160. Bangalore H, Gaies M, Ocampo EC, et al. The Total Inotrope Exposure Score: An extension of the Vasoactive Inotrope Score as a predictor of adverse outcomes after paediatric cardiac surgery. *Cardiol Young*. 2017;27(6). doi:10.1017/S1047951116002602

161. Kruk ME, Gage AD, Joseph NT, Danaei G, García-Saisó S, Salomon JA. Mortality due to low-quality health systems in the universal health coverage era: a systematic analysis of amenable deaths in 137 countries. *The Lancet*. 2018;392(10160). doi:10.1016/S0140-6736(18)31668-4
162. De Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: A systematic review. *Qual Saf Health Care*. 2008;17(3). doi:10.1136/qshc.2007.023622
163. Kieny M, Evans T, Scarpetta S, et al. *Delivering Quality Health Services: A Global Imperative for Universal Health Coverage (English)*. 2018. Available online: https://www.who.int/publications/i/item/9789241513906
164. Kearns DT. Leadership through quality. *Academy of Management Perspectives*. 1990;4(2):86-89. doi:10.5465/ame.1990.4274800
165. Manktelow BN, Evans TA, Draper ES. Differences in case-mix can influence the comparison of standardised mortality ratios even with optimal risk adjustment: An analysis of data from paediatric intensive care. *BMJ Qual Saf*. 2014;23(9). doi:10.1136/bmjqs-2013-002608
166. Liddell FD. Simple exact analysis of the standardised mortality ratio. *J Epidemiol Community Health (1978)*. 1984;38(1):85. doi:10.1136/jech.38.1.85
167. Manaseki-Holland S, Lilford RJ, Te AP, et al. Ranking Hospitals Based on Preventable Hospital Death Rates: A Systematic Review With Implications for Both Direct Measurement and Indirect Measurement Through Standardized Mortality Rates. *Milbank Q*. 2019;97(1):228-284. doi:10.1111/1468-0009.12375
168. Pavlou M, Ambler G, Omar RZ, et al. Outlier identification and monitoring of institutional or clinician performance: an overview of statistical methods and application to national audit data. *BMC Health Serv Res*. 2023;23(1):23. doi:10.1186/s12913-022-08995-z
169. Rhodes A, Moreno RP, Azoulay E, et al. Prospectively defined indicators to improve the safety and quality of care for critically ill patients: a report from the Task Force on Safety and Quality of the

European Society of Intensive Care Medicine (ESICM). *Intensive Care Med*. 2012;38(4):598-605. doi:10.1007/s00134-011-2462-3
170. Engerström L, Kramer AA, Nolin T, et al. Comparing Time-Fixed Mortality Prediction Models and Their Effect on ICU Performance Metrics Using the Simplified Acute Physiology Score 3. *Crit Care Med*. 2016;44(11). doi:10.1097/CCM.0000000000001877
171. Harrison DA, Parry GJ, Carpenter JR, Short A, Rowan K. A new risk prediction model for critical care: The Intensive Care National Audit & Research Centre (ICNARC) model. *Crit Care Med*. 2007;35(4). doi:10.1097/01.CCM.0000259468.24532.44
172. Paul E, Bailey M, Kasza J, Pilcher D. The ANZROD model: better benchmarking of ICU outcomes and detection of outliers. *Crit Care Resusc*. 2016;18:25-36.
173. Reinikainen M, Mussalo P, Hovilehto S, et al. Association of automated data collection and data completeness with outcomes of intensive care. A new customised model for outcome prediction. *Acta Anaesthesiol Scand*. 2012;56(9):1114-1122. doi:10.1111/j.1399-6576.2012.02669.x
174. Flaatten H. The present use of quality indicators in the intensive care unit. *Acta Anaesthesiol Scand*. 2012;56(9):1078-1083. doi:10.1111/j.1399-6576.2012.02656.x
175. Vincent JL, Moreno R. Clinical review: Scoring systems in the critically ill. *Crit Care*. 2010;14(2):207. doi:10.1186/cc8204
176. Vallet H, Guidet B, Boumendil A, et al. The impact of age-related syndromes on ICU process and outcomes in very old patients. *Ann Intensive Care*. 2023;13(1):68. doi:10.1186/s13613-023-01160-7
177. Atramont A, Lindecker-Cournil V, Rudant J, et al. Association of Age With Short-term and Long-term Mortality Among Patients Discharged From Intensive Care Units in France. *JAMA Netw Open*. 2019;2(5):e193215-e193215. doi:10.1001/jamanetworkopen.2019.3215
178. Flaatten H, De Lange DW, Morandi A, et al. The impact of frailty on ICU and 30-day mortality and the level of care in very elderly patients (≥ 80 years). *Intensive Care Med*. 2017;43(12):1820-1828. doi:10.1007/s00134-017-4940-8

179. Reinikainen M, Niskanen M, Uusaro A, Ruokonen E. Impact of gender on treatment and outcome of ICU patients. *Acta Anaesthesiol Scand*. 2005;49(7):984-990. doi:10.1111/j.1399-6576.2005.00759.x
180. Timmers T, Verhofstad M, Moons K, Leenen L. Long-term survival after surgical intensive care admission: 50% die within 10 years. *Crit Care*. 2009;13(1):P515. doi:10.1186/cc7679
181. Namikata Y, Matsuoka Y, Ito J, et al. Association between ICU admission during off-hours and in-hospital mortality: a multicenter registry in Japan. *J Intensive Care*. 2022;10(1):41. doi:10.1186/s40560-022-00634-3
182. Ingraham NE, Vakayil V, Pendleton KM, et al. Recent Trends in Admission Diagnosis and Related Mortality in the Medically Critically Ill. *J Intensive Care Med*. 2020;37(2):185-194. doi:10.1177/0885066620982905
183. Van Calster B, McLernon DJ, van Smeden M, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17(1):230. doi:10.1186/s12916-019-1466-7
184. Van Calster B, Vickers AJ. Calibration of Risk Prediction Models: Impact on Decision-Analytic Performance. *Medical Decision Making*. 2014;35(2):162-169. doi:10.1177/0272989X14547233
185. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med*. 2015;13(1):1. doi:10.1186/s12916-014-0241-z
186. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Ann Intern Med*. 2015;162(1):W1-W73. doi:10.7326/M14-0698
187. Moreno RP. Outcome prediction in intensive care: why we need to reinvent the wheel. *Curr Opin Crit Care*. 2008;14(5). doi:10.1097/MCC.0b013e328310dc7d
188. Knaus WA. APACHE 1978-2001: The Development of a Quality Assurance System Based on Prognosis: Milestones and Personal Reflections. *Archives of Surgery*. 2002;137(1):37-41. doi:10.1001/archsurg.137.1.37

189. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med*. 1985;13(10):818-829.
190. Teres D, Brown RD, Lemeshow S. Predicting mortality of intensive care unit patients. The importance of coma. *Crit Care Med*. 1982;10(2):86-95. doi:10.1097/00003246-198202000-00004
191. Zimmerman JE, Kramer AA. A history of outcome prediction in the ICU. *Curr Opin Crit Care*. 2014;20(5). doi:10.1097/MCC.0000000000000138
192. Zhang Z, Van Poucke S, Goyal H, Rowley D, Zhong M, Liu N. The top 2,000 cited articles in critical care medicine: A bibliometric analysis. *J Thorac Dis*. 2018;10. doi:10.21037/jtd.2018.03.178
193. Le Gall JR, Lemeshow S, Saulnier F. A New Simplified Acute Physiology Score (SAPS II) Based on a European/North American Multicenter Study. *JAMA*. 1993;270(24):2957-2963. doi:10.1001/jama.1993.03510240069035
194. Moreno R, Miranda DR, Fidler V, Van Schilfgaarde R. Evaluation of two outcome prediction models on an independent database. *Crit Care Med*. 1998;26(1). doi:10.1097/00003246-199801000-00016
195. Zimmerman JE, Wagner DP, Draper EA, Wright L, Alzola C, Knaus WA. Evaluation of Acute Physiology and Chronic Health Evaluation III predictions of hospital mortality in an independent database. *Crit Care Med*. 1998;26(8). doi:10.1097/00003246-199808000-00012
196. Moreno R, Morais P. Outcome prediction in intensive care: results of a prospective, multicentre, Portuguese study. *Intensive Care Med*. 1997;23(2):177-186. doi:10.1007/s001340050313
197. Moreno RP, Metnitz PGH, Almeida E, et al. SAPS 3--From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. *Intensive Care Med*. 2005;31(10):1345-1355. doi:10.1007/s00134-005-2763-5
198. Metnitz PGH, Moreno RP, Almeida E, et al. SAPS 3—From evaluation of the patient to evaluation of the intensive care unit. Part 1: Objectives, methods and cohort description. *Intensive Care Med*. 2005;31(10):1336-1344. doi:10.1007/s00134-005-2762-6

199. Capuzzo M, Scaramuzza A, Vaccarini B, et al. Validation of SAPS 3 Admission Score and comparison with SAPS II. *Acta Anaesthesiol Scand*. 2009;53(5):589-594. doi:10.1111/j.1399-6576.2009.01929.x
200. Strand K, Søreide E, Aardal S, Flaaten H. A comparison of SAPS II and SAPS 3 in a Norwegian intensive care unit population. *Acta Anaesthesiol Scand*. 2009;53(5):595-600. doi:10.1111/j.1399-6576.2009.01948.x
201. Lemeshow S, Klar J, Teres D, et al. Mortality probability models for patients in the intensive care unit for 48 or 72 hours: A prospective, multicenter study. *Crit Care Med*. 1994;22(9). doi:10.1097/00003246-199409000-00003
202. Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J. Mortality Probability Models (MPM II) Based on an International Cohort of Intensive Care Unit Patients. *JAMA*. 1993;270(20):2478-2486. doi:10.1001/jama.1993.03510200084037
203. Higgins TL, Teres D, Copes WS, Nathanson BH, Stark M, Kramer AA. Assessing contemporary intensive care unit outcome: An updated Mortality Probability Admission Model (MPM0-III). *Crit Care Med*. 2007;35(3). doi:10.1097/01.CCM.0000257337.63529.9F
204. Kuzniewicz MW, Vasilevskis EE, Lane R, et al. Variation in ICU Risk-Adjusted Mortality: Impact of Methods of Assessment and Potential Confounders. *Chest*. 2008;133(6):1319-1327. doi:https://doi.org/10.1378/chest.07-3061
205. Sánchez-Casado M, Hostigüela-Martín VA, Raigal-Caño A, et al. Predictive scoring systems in multiorgan failure: A cohort study. *Medicina Intensiva (English Edition)*. 2016;40(3). doi:10.1016/j.medine.2015.03.009
206. Ko M, Shim M, Lee SM, Kim Y, Yoon S. Performance of APACHE IV in medical intensive care unit patients: Comparisons with APACHE II, SAPS 3, and MPM0 III. *Acute and Critical Care*. 2018;33(4). doi:10.4266/acc.2018.00178
207. Keegan MT, Gajic O, Afessa B. Comparison of APACHE III, APACHE IV, SAPS 3, and MPM0III and influence of resuscitation status on model performance. *Chest*. 2012;142(4). doi:10.1378/chest.11-2164
208. Kramer AA, Higgins TL, Zimmerman JE. Comparing Observed and Predicted Mortality Among ICUs Using Different Prognostic Systems:

Why Do Performance Assessments Differ? *Crit Care Med*. 2015;43(2). doi:10.1097/CCM.0000000000000694
209. Salluh JIF, Soares M. ICU severity of illness scores: APACHE, SAPS and MPM. *Curr Opin Crit Care*. 2014;20(5). doi:10.1097/MCC.0000000000000135
210. Livingston BM, MacKirdy FN, Howie JC, Jones R, Norrie JD. Assessment of the performance of five intensive care scoring models within a large Scottish database. *Crit Care Med*. 2000;28(6). doi:10.1097/00003246-200006000-00023
211. Beck DH, Taylor BL, Millar B, Smith GB. Prediction of outcome from intensive care: A prospective cohort study comparing Acute Physiology and Chronic Health Evaluation II and III prognostic systems in a United Kingdom intensive care unit. *Crit Care Med*. 1997;25(1). doi:10.1097/00003246-199701000-00006
212. Glance LG, Osler TM, Papadakos P. Effect of mortality rate on the performance of the Acute Physiology and Chronic Health Evaluation II: A simulation study. *Crit Care Med*. 2000;28(10). doi:10.1097/00003246-200010000-00008
213. Angus DC, Sirio CA, Clermont G, Bion J. International comparisons of critical care outcome and resource consumption. *Crit Care Clin*. 1997;13(2):389-408. doi:10.1016/S0749-0704(05)70317-1
214. Van Calster B, McLernon DJ, van Smeden M, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17(1):230. doi:10.1186/s12916-019-1466-7
215. Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54(8):774-781. doi:10.1016/S0895-4356(01)00341-9
216. Thai TN, Ebell MH. Prospective validation of the Good Outcome Following Attempted Resuscitation (GO-FAR) score for in-hospital cardiac arrest prognosis. *Resuscitation*. 2019;140:2-8. doi:10.1016/j.resuscitation.2019.05.002
217. Reineck LA, Pike F, Le TQ, Cicero BD, Iwashyna TJ, Kahn JM. Hospital factors associated with discharge bias in ICU performance measurement. *Crit Care Med*. 2014;42(5):1055-1064. doi:10.1097/CCM.0000000000000132

218. Drye EE, Normand SLT, Wang Y, et al. Comparison of Hospital Risk-Standardized Mortality Rates Calculated by Using In-Hospital and 30-Day Models: An Observational Study With Implications for Hospital Profiling. *Ann Intern Med*. 2012;156(1_Part_1):19-26. doi:10.7326/0003-4819-156-1-201201030-00004
219. Kahn JM, Kramer AA, Rubenfeld GD. Transferring critically ill patients out of hospital improves the standardized mortality ratio: A simulation study. *Chest*. 2007;131(1). doi:10.1378/chest.06-0741
220. Pouw ME, Peelen LM, Moons KGM, Kalkman CJ, Lingsma HF. Including post-discharge mortality in calculation of hospital standardised mortality ratios: retrospective analysis of hospital episode statistics. *BMJ*. 2013;347:f5913-f5913. doi:10.1136/bmj.f5913
221. Glance LG, Szalados JE. Benchmarking in Critical Care: The Road Ahead. *Chest*. 2002;121(2):326-328. doi:10.1378/chest.121.2.326
222. Flaatten H, Reinikainen M. Severity scoring, outcome prediction and mortality endpoints in intensive care. *Acta Anaesthesiol Scand*. 2015;59(7). doi:10.1111/aas.12553
223. Kaufmann M, Perren A, Cerutti B, Dysli C, Rothen HU, Medicine on behalf of the SS of IC. Severity-Adjusted ICU Mortality Only Tells Half the Truth—The Impact of Treatment Limitation in a Nationwide Database. *Crit Care Med*. 2020;48(12). doi:10.1097/CCM.0000000000004658
224. Richardson DB, Cole SR, Chu H. Random effects regression models for trends in standardised mortality ratios. *Occup Environ Med*. 2013;70(2):133. doi:10.1136/oemed-2012-100792
225. Roessler M, Schmitt J, Schoffer O. Can we trust the standardized mortality ratio? A formal analysis and evaluation based on axiomatic requirements. *PLoS One*. 2021;16(9):e0257003-e0257003. doi:10.1371/journal.pone.0257003
226. Marang-van De Mheen PJ, Shojania KG. Simpson's paradox: How performance measurement can fail even with perfect risk adjustment. *BMJ Qual Saf*. 2014;23(9). doi:10.1136/bmjqs-2014-003358
227. A Definition of Irreversible Coma: Report of the Ad Hoc Committee of the Harvard Medical School to Examine the Definition of Brain

Death. *JAMA*. 1968;205(6):337-340. doi:10.1001/jama.1968.03140320031009
228. Truog RD, Pope TM, Jones DS. The 50-Year Legacy of the Harvard Report on Brain Death. *JAMA*. 2018;320(4):335-336. doi:10.1001/jama.2018.6990
229. Randell TT. Medical and legal considerations of brain death. *Acta Anaesthesiol Scand*. 2004;48(2):139-144. doi:10.1111/j.0001-5172.2004.00304.x
230. Manara A, Procaccio F, Domínguez-Gil B. Expanding the pool of deceased organ donors: the ICU and beyond. *Intensive Care Med*. 2019;45(3):357-360. doi:10.1007/s00134-019-05546-9
231. Kotloff RM, Blosser S, Fulda GJ, et al. Management of the Potential Organ Donor in the ICU: Society of Critical Care Medicine/American College of Chest Physicians/Association of Organ Procurement Organizations Consensus Statement. *Crit Care Med*. 2015;43(6). doi:10.1097/CCM.0000000000000958
232. Pall V, Laisaar T, Ericzon BG. A Regional Report on the Evolution of Organ Transplantation in Estonia. *Transplantation*. 2019;103(10). doi:10.1097/TP.0000000000002730
233. Müller PC, Kabacam G, Vibert E, Germani G, Petrowsky H. Current status of liver transplantation in Europe. *International Journal of Surgery*. 2020;82:22-29. doi:10.1016/j.ijsu.2020.05.062
234. Sauthier N, Bouchakri R, Carrier FM, et al. Automated screening of potential organ donors using a temporal machine learning model. *Sci Rep*. 2023;13(1):8459. doi:10.1038/s41598-023-35270-w
235. Vergano M, Jung C, Metaxa V. Intensive care admission aiming at organ donation. Con. *Intensive Care Med*. 2024;50(3):440-442. doi:10.1007/s00134-024-07326-6
236. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU collaborative research database, a freely available multi-center database for critical care research. *Sci Data*. 2018;5. doi:10.1038/sdata.2018.178
237. Takala J, Moser A, Raj R, et al. Variation in severity-adjusted resource use and outcome in intensive care units. *Intensive Care Med*. 2022;48(1):67-77. doi:10.1007/s00134-021-06546-4

238. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-845.
239. Contal C, O'Quigley J. An application of changepoint methods in studying the effect of age on survival in breast cancer. *Comput Stat Data Anal*. 1999;30(3):253-270. doi:10.1016/S0167-9473(98)00096-6
240. Jee W, Jo S, Lee JB, et al. Mortality difference between early-identified sepsis and late-identified sepsis. *Clin Exp Emerg Med*. 2020;7(3). doi:10.15441/ceem.19.009
241. Carlos Sanchez E, Pinsky MR, Sinha S, Mishra RC, Lopa AJ, Chatterjee R. Fluids and Early Vasopressors in the Management of Septic Shock: Do We Have the Right Answers Yet? *J Crit Care Med*. 2023;9(3). doi:10.2478/jccm-2023-0022
242. Acheampong A, Vincent JL. A positive fluid balance is an independent prognostic factor in patients with sepsis. *Crit Care*. 2015;19(1). doi:10.1186/s13054-015-0970-1
243. Meyhoff TS, Hjortrup PB, Wetterslev J, et al. Restriction of Intravenous Fluid in ICU Patients with Septic Shock. *New England Journal of Medicine*. 2022;386(26):2459-2470. doi:10.1056/NEJMoa2202707
244. Ammar MA, Ammar AA, Wieruszewski PM, et al. Timing of vasoactive agents and corticosteroid initiation in septic shock. *Ann Intensive Care*. 2022;12(1):47. doi:10.1186/s13613-022-01021-9
245. Tong X, Xue X, Duan C, Liu A. Early administration of multiple vasopressors is associated with better survival in patients with sepsis: a propensity score-weighted study. *Eur J Med Res*. 2023;28(1):249. doi:10.1186/s40001-023-01229-w
246. Ospina-Tascón GA, Hernandez G, Alvarez I, et al. Effects of very early start of norepinephrine in patients with septic shock: a propensity score-based analysis. *Crit Care*. 2020;24(1):52. doi:10.1186/s13054-020-2756-3
247. Polok K, Fronczek J, Putowski Z, et al. Validity of the total SOFA score in patients ≥ 80 years old acutely admitted to intensive care units: a post-hoc analysis of the VIP2 prospective, international cohort study. *Ann Intensive Care*. 2023;13(1):98. doi:10.1186/s13613-023-01191-0

248. Suzuki R, Uchino S, Sasabuchi Y, Kawarai Lefor A, Sanui M. Dopamine use and its consequences in the intensive care unit: a cohort study utilizing the Japanese Intensive care PAtient Database. *Crit Care*. 2022;26(1):90. doi:10.1186/s13054-022-03960-y
249. De Backer D, Biston P, Devriendt J, et al. Comparison of Dopamine and Norepinephrine in the Treatment of Shock. *New England Journal of Medicine*. 2010;362(9):779-789. doi:10.1056/NEJMoa0907118
250. Annane D. Physicians no longer should consider dopamine for septic shock!. *Crit Care Med*. 2012;40(3). doi:10.1097/CCM.0b013e31823d779b
251. De Backer D, Foulon P. Minimizing catecholamines and optimizing perfusion. *Crit Care*. 2019;23(1):149. doi:10.1186/s13054-019-2433-6
252. Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Med*. 2021;47(11):1181-1247. doi:10.1007/s00134-021-06506-y
253. Tavazzi G, Rossello X, Grand J, et al. Epidemiology, monitoring, and treatment strategy in cardiogenic shock. A multinational cross-sectional survey of ESC-acute cardiovascular care association research section . *Eur Heart J Acute Cardiovasc Care*. 2022;11(9):706-711. doi:10.1093/ehjacc/zuac087
254. Scheeren TWL, Bakker J, De Backer D, et al. Current use of vasopressors in septic shock. *Ann Intensive Care*. 2019;9(1). doi:10.1186/s13613-019-0498-7
255. Nagao K, Kato T, Yaku H, et al. Current use of inotropes according to initial blood pressure and peripheral perfusion in the treatment of congestive heart failure: findings from a multicentre observational study. *BMJ Open*. 2022;12(1):e053254-e053254. doi:10.1136/bmjopen-2021-053254
256. Wakefield BJ, Sacha GL, Khanna AK. Vasodilatory shock in the ICU and the role of angiotensin II. *Curr Opin Crit Care*. 2018;24(4). doi:10.1097/MCC.0000000000000517
257. Kuzmiszyn AK, Selli AL, Smaglyukova N, et al. Treatment of Cardiovascular Dysfunction with PDE3-Inhibitors in Moderate and Severe Hypothermia—Effects on Cellular Elimination of Cyclic

Adenosine Monophosphate and Cyclic Guanosine Monophosphate. *Front Physiol*. 2022;13. doi:10.3389/fphys.2022.923091
258. Ibarra-Estrada M, Kattan E, Aguilera-González P, et al. Early adjunctive methylene blue in patients with septic shock: a randomized controlled trial. *Crit Care*. 2023;27(1):110. doi:10.1186/s13054-023-04397-7
259. Gordon AC, Mason AJ, Thirunavukkarasu N, et al. Effect of early vasopressin vs norepinephrine on kidney failure in patients with septic shock: The VANISH randomized clinical trial. *JAMA - Journal of the American Medical Association*. 2016;316(5). doi:10.1001/jama.2016.10485
260. Quinn C, Moukhachen H, Bagnola A, et al. 1280: Successful Use of Angiotensin II in a Case of Refractory Vasodilatory Cardiogenic Mixed Shock. *Crit Care Med*. 2021;49(1). doi:10.1097/01.ccm.0000731008.71855.c1
261. Wieruszewski PM, Khanna AK. Vasopressor Choice and Timing in Vasodilatory Shock. *Crit Care*. 2022;26(1). doi:10.1186/s13054-022-03911-7
262. Ammar MA, Limberg EC, Lam SW, et al. Optimal norepinephrine-equivalent dose to initiate epinephrine in patients with septic shock. *J Crit Care*. 2019;53. doi:10.1016/j.jcrc.2019.05.024
263. Lee HJ, Ko BS, Ryoo SM, et al. Modified cardiovascular SOFA score in sepsis: development and internal and external validation. *BMC Med*. 2022;20(1). doi:10.1186/s12916-022-02461-7
264. Wang L, Yang H, Cheng Y, et al. Mean arterial pressure/norepinephrine equivalent dose index as an early measure of initiation time for enteral nutrition in patients with shock: A prospective observational study. *Nutrition*. 2022;96. doi:10.1016/j.nut.2021.111586
265. Kotani Y, Di Gioia A, Landoni G, Belletti A, Khanna AK. An updated “norepinephrine equivalent” score in intensive care as a marker of shock severity. *Crit Care*. 2023;27(1). doi:10.1186/s13054-023-04322-y
266. Reintam Blaser A, Padar M, Mändul M, et al. Development of the Gastrointestinal Dysfunction Score (GIDS) for critically ill patients - A prospective multicenter observational study (iSOFA study). *Clinical Nutrition*. 2021;40(8):4932-4940. doi:10.1016/j.clnu.2021.07.015

267. Guidet B, de Lange DW, Boumendil A, et al. The contribution of frailty, cognition, activity of daily life and comorbidities on outcome in acutely admitted patients over 80 years in European ICUs: the VIP2 study. *Intensive Care Med*. 2020;46(1):57-69. doi:10.1007/s00134-019-05853-1
268. Auchet T, Regnier MA, Girerd N, Levy B. Outcome of patients with septic shock and high-dose vasopressor therapy. *Ann Intensive Care*. 2017;7(1). doi:10.1186/s13613-017-0261-x
269. Leone M, Albanèse J, Delmas A, Chaabane W, Garnier F, Martin C. Terlipressin in catecholamine-resistant septic shock patients. *Shock*. 2004;22(4). doi:10.1097/01.shk.0000136097.42048.bd
270. O'Brien A, Clapp L, Singer M. Terlipressin for norepinephrine-resistant septic shock. *Lancet*. 2002;359(9313). doi:10.1016/S0140-6736(02)08225-9
271. Katsaragakis S, Kapralou A, Theodorou D, et al. Refractory septic shock: Efficacy and safety of very high doses of norepinephrine. *Methods Find Exp Clin Pharmacol*. 2006;28(5). doi:10.1358/mf.2006.28.5.990203
272. Svoboda P, Scheer P, Kantorová I, et al. Terlipressin in the treatment of late phase catecholamine-resistant septic shock. *Hepatogastroenterology*. 2012;59(116). doi:10.5754/hge10550
273. Monti G, Terzi V, Calini A, et al. Rescue therapy with polymyxin B hemoperfusion in high-dose vasopressor therapy refractory septic shock. *Minerva Anestesiol*. 2015;81(5).
274. Martin C, Medam S, Antonini F, et al. Norepinephrine: Not too much, too long. *Shock*. 2015;44(4). doi:10.1097/SHK.0000000000000426
275. Ko BS, Ryoo SM, Han E, et al. Modified Cardiovascular Sequential Organ Failure Assessment Score in Sepsis: External Validation in Intensive Care Unit Patients. *J Korean Med Sci*. 2023;38(50). doi:10.3346/jkms.2023.38.e418
276. Thayer KL, Zweck E, Ayouty M, et al. Invasive Hemodynamic Assessment and Classification of In-Hospital Mortality Risk Among Patients With Cardiogenic Shock. *Circ Heart Fail*. 2020;13(9):E007099. doi:10.1161/CIRCHEARTFAILURE.120.007099

277. Jentzer JC. Understanding Cardiogenic Shock Severity and Mortality Risk Assessment. *Circ Heart Fail*. 2020;13(9). doi:10.1161/CIRCHEARTFAILURE.120.007568
278. Alviar CL, Li BK, Keller NM, et al. Prognostic Performance of the IABP-SHOCK II Risk Score Among Cardiogenic Shock Subtypes in the Critical Care Cardiology Trials Network Registry. *Am Heart J*. Published online January 6, 2024. doi:10.1016/J.AHJ.2023.12.018
279. Baran DA, Grines CL, Bailey S, et al. SCAI clinical expert consensus statement on the classification of cardiogenic shock. *Catheterization and Cardiovascular Interventions*. 2019;94(1):29-37. doi:10.1002/CCD.28329
280. Jentzer JC, van Diepen S, Barsness GW, et al. Cardiogenic Shock Classification to Predict Mortality in the Cardiac Intensive Care Unit. *J Am Coll Cardiol*. 2019;74(17):2117-2128. doi:10.1016/J.JACC.2019.07.077
281. Kotani Y, Landoni G, Belletti A, Khanna AK. Response to: norepinephrine formulation for equivalent vasopressive score. *Crit Care*. 2023;27(1). doi:10.1186/s13054-023-04404-x
282. Mongardon N, de Roux Q, Leone M, Guerci P. Norepinephrine formulation for equivalent vasopressive score. *Crit Care*. 2023;27(1). doi:10.1186/s13054-023-04354-4
283. Leone M, Goyer I, Levy B, Dünser MW, Asfar P, Jentzer JC. Dose of norepinephrine: the devil is in the details. *Intensive Care Med*. 2022;48(5). doi:10.1007/s00134-022-06652-x
284. Leone M, Goyer I, Levy B, Dünser MW, Asfar P, Jentzer JC. Norepinephrine dose and the grain of salt: a task force is needed. *Intensive Care Med*. 2023;49(6). doi:10.1007/s00134-023-07087-8
285. Kotani Y, Belletti A, D'Andria Ursoleo J, Salvati S, Landoni G. Norepinephrine Dose Should Be Reported as Base Equivalence in Clinical Research Manuscripts. *J Cardiothorac Vasc Anesth*. 2023;37(9). doi:10.1053/j.jvca.2023.05.013
286. Wieruszewski PM, Leone M, Kaas-Hansen BS, et al. Position Paper on the Reporting of Norepinephrine Formulations in Critical Care from the Society of Critical Care Medicine and European Society of Intensive Care Medicine Joint Task Force. *Crit Care Med*. 2024;;52(4):521-530. doi:10.1097/CCM.0000000000006176

287. Heavner MS, McCurdy MT, Mazzeffi MA, Galvagno SM, Tanaka KA, Chow JH. Angiotensin II and Vasopressin for Vasodilatory Shock: A Critical Appraisal of Catecholamine-Sparing Strategies. *J Intensive Care Med*. 2021;36(6). doi:10.1177/0885066620911601
288. Rudiger A, Singer M. Decatecholaminisation during sepsis. *Crit Care*. 2016;20(1). doi:10.1186/s13054-016-1488-x
289. Vail E, Gershengorn HB, Hua M, Walkey AJ, Rubenfeld G, Wunsch H. Association between US norepinephrine shortage and mortality among patients with septic shock. *JAMA - Journal of the American Medical Association*. 2017;317(14). doi:10.1001/jama.2017.2841
290. Ngan Kee WD, Lee SWY, Ng FF, Tan PE, Khaw KS. Randomized double-blinded comparison of norepinephrine and phenylephrine for maintenance of blood pressure during spinal anesthesia for cesarean delivery. *Anesthesiology*. 2015;122(4). doi:10.1097/ALN.0000000000000601
291. Russell JA, Walley KR, Singer J, et al. Vasopressin versus Norepinephrine Infusion in Patients with Septic Shock. *New England Journal of Medicine*. 2008;358(9). doi:10.1056/nejmoa067373
292. Khanna A, English SW, Wang XS, et al. Angiotensin II for the Treatment of Vasodilatory Shock. *New England Journal of Medicine*. 2017;377(5). doi:10.1056/nejmoa1704154
293. Liu ZM, Chen J, Kou Q, et al. Terlipressin versus norepinephrine as infusion in patients with septic shock: a multicentre, randomised, double-blinded trial. *Intensive Care Med*. 2018;44(11). doi:10.1007/s00134-018-5267-9
294. Kirov MY, Evgenov O V., Evgenov N V., et al. Infusion of methylene blue in human septic shock: A pilot, randomized, controlled study. *Crit Care Med*. 2001;29(10). doi:10.1097/00003246-200110000-00002
295. Natalini G, Schivalocchi V, Rosano A, Taranto M, Pletti C, Bernardini A. Norepinephrine and metaraminol in septic shock: A comparison of the hemodynamic effects. *Intensive Care Med*. 2005;31(5). doi:10.1007/s00134-005-2607-3
296. Papadopoulos G, Sintou E, Siminelakis S, Koletsis E, Baikoussis NG, Apostolakis E. Perioperative infusion of low- dose of vasopressin for prevention and management of vasodilatory vasoplegic syndrome in patients undergoing coronary artery bypass grafting-A double-blind

randomized study. *J Cardiothorac Surg*. 2010;5:17. doi:10.1186/1749-8090-5-17
297. Adly DH El, Bazan NS, El Borolossy RM, Anan IF, Fakher MA, El Wakeel LM. Midodrine improves clinical and economic outcomes in patients with septic shock: a randomized controlled clinical trial. *Ir J Med Sci*. 2022;191(6). doi:10.1007/s11845-021-02903-w
298. Vallet B, Dupuis B, Chopin C. Dobutamine: mechanisms of action and use in acute cardiovascular pathology. *Ann Cardiol Angeiol (Paris)*. 1991;40(6).
299. Moser A, Reinikainen M, Jakob SM, et al. Mortality prediction in intensive care units including premorbid functional status improved performance and internal validity. *J Clin Epidemiol*. 2022;142:230-241. doi:10.1016/j.jclinepi.2021.11.028
300. Ferrando-Vivas P, Jones A, Rowan KM, Harrison DA. Development and validation of the new ICNARC model for prediction of acute hospital mortality in adult critical care. *J Crit Care*. 2017;38. doi:10.1016/j.jcrc.2016.11.031
301. Friele RD, Coppen R, Marquet RL, Gevers JKM. Explaining differences between hospitals in number of organ donors. *American Journal of Transplantation*. 2006;6(3). doi:10.1111/j.1600-6143.2005.01200.x
302. Sheehy E, Conrad SL, Brigham LE, et al. Estimating the Number of Potential Organ Donors in the United States. *New England Journal of Medicine*. 2003;349(7). doi:10.1056/nejmsa021271
303. Christiansen CL, Gortmaker SL, Williams JM, et al. A method for estimating solid organ donor potential by organ procurement region. *Am J Public Health*. 1998;88(11). doi:10.2105/AJPH.88.11.1645
304. Labarère J, Bertrand R, Fine MJ. How to derive and validate clinical prediction models for use in intensive care medicine. *Intensive Care Med*. 2014;40(4). doi:10.1007/s00134-014-3227-6
305. Szczeklik W, Fronczek J. Should we revise the SOFA score? Yes. *Acta Anaesthesiol Scand*. 2022;66(6):658-659. doi:10.1111/aas.14060
306. Moreno R, Rhodes A, Piquilloud L, et al. The Sequential Organ Failure Assessment (SOFA) Score: has the time come for an update? *Crit Care*. 2023;27(1):15. doi:10.1186/s13054-022-04290-9
307. Moreno R, Rhodes A, Singer M, Vincent JL, authors. O behalf of the. Real-world inter-observer variability of the sequential organ failure

assessment (SOFA) score in intensive care medicine: the time has come for an update: authors' reply. *Crit Care*. 2023;27(1):182. doi:10.1186/s13054-023-04468-9
308. Meresse Z, Medam S, Mathieu C, Duclos G, Vincent JL, Leone M. Vasopressors to treat refractory septic shock. *Minerva Anestesiol*. 2020;86(5). doi:10.23736/S0375-9393.20.13826-4
309. Ma P, Liu J, Shen F, et al. Individualized resuscitation strategy for septic shock formalized by finite mixture modeling and dynamic treatment regimen. *Crit Care*. 2021;25(1). doi:10.1186/s13054-021-03682-7
310. Quan AT, Li F. Hyperinflation of Vasopressors (Vasopressin, Norepinephrine, Ephedrine, etc). *J Pharm Pract*. 2018;31(4). doi:10.1177/0897190018770062
311. Syed M, Khan MZ, Osman M, et al. Sixteen-Year National Trends in Use and Outcomes of VA-ECMO in Cardiogenic Shock. *Cardiovascular Revascularization Medicine*. 2022;44:1-7. doi:10.1016/j.carrev.2022.06.267
312. Schenck EJ, Hoffman KL, Oromendia C, et al. A comparative analysis of the respiratory subscore of the sequential organ failure assessment scoring system. *Ann Am Thorac Soc*. 2021;18(11). doi:10.1513/AnnalsATS.202004-399OC
313. Riviello ED, Kiviri W, Twagirumugabe T, et al. Hospital Incidence and Outcomes of the Acute Respiratory Distress Syndrome Using the Kigali Modification of the Berlin Definition. *Am J Respir Crit Care Med*. 2015;193(1):52-59. doi:10.1164/rccm.201503-0584OC
314. Wijdicks EFM, Bamlet WR, Maramattom B V, Manno EM, McClelland RL. Validation of a new coma scale: The FOUR score. *Ann Neurol*. 2005;58(4):585-593. doi:10.1002/ana.20611
315. Bruno MA, LeDoux D, Lambermont B, et al. Comparison of the Full Outline of UnResponsiveness and Glasgow Liege Scale/Glasgow Coma Scale in an Intensive Care Unit Population. *Neurocrit Care*. 2011;15(3):447. doi:10.1007/s12028-011-9547-2
316. McNett M, Amato S, Gianakis A, et al. The FOUR Score and GCS as Predictors of Outcome After Traumatic Brain Injury. *Neurocrit Care*. 2014;21(1):52-57. doi:10.1007/s12028-013-9947-6

317. Iyer VN, Mandrekar JN, Danielson RD, Zubkov AY, Elmer JL, Wijdicks EFM. Validity of the FOUR score coma scale in the medical intensive care unit. *Mayo Clin Proc*. 2009;84(8):694-701. doi:10.1016/S0025-6196(11)60519-3
318. Zimmermann T, Kaufmann P, Amacher SA, et al. Sex differences in the SOFA score of ICU patients with sepsis or septic shock: a nationwide analysis. *Crit Care*. 2024;28:209. doi:10.1186/s13054-024-04996-y
319. Lindemark F, Haaland ØA, Kvåle R, Flaatten H, Norheim OF, Johansson KA. Costs and expected gain in lifetime health from intensive care versus general ward care of 30,712 individual patients: a distribution-weighted cost-effectiveness analysis. *Crit Care*. 2017;21(1):220. doi:10.1186/s13054-017-1792-0
320. Rothen HU, Stricker K, Einfalt J, et al. Variability in outcome and resource use in intensive care units. *Intensive Care Med*. 2007;33(8):1329-1336. doi:10.1007/s00134-007-0690-3
321. Willmington C, Belardi P, Murante AM, Vainieri M. The contribution of benchmarking to quality improvement in healthcare. A systematic literature review. *BMC Health Serv Res*. 2022;22(1):139. doi:10.1186/s12913-022-07467-8
322. Woodhouse D, Berg M, Van Der Putten J, Houtepen J. Will benchmarking ICUs improve outcome? *Curr Opin Crit Care*. 2009;15(5). doi:10.1097/MCC.0b013e32833079fb
323. Melville A, Kolt G, Anderson D, Mitropoulos J, Pilcher D. Admission to intensive care for palliative care or potential organ donation: Demographics, circumstances, outcomes, and resource use. *Crit Care Med*. 2017;45(10). doi:10.1097/CCM.0000000000002655
324. Salluh JIF, Chiche JD, Reis CE, Soares M. New perspectives to improve critical care benchmarking. *Ann Intensive Care*. 2018;8(1). doi:10.1186/s13613-018-0363-0
325. Timmers TK. Intensive care performance: How should we monitor performance in the future? *World J Crit Care Med*. 2014;3(4). doi:10.5492/wjccm.v3.i4.74
326. Zheng S, He A, Yu Y, Jiang L, Liang J, Wang P. Research trends and hotspots of health-related quality of life: a bibliometric analysis from

2000 to 2019. *Health Qual Life Outcomes*. 2021;19(1). doi:10.1186/s12955-021-01767-z

327. Falini S, Angelotti G, Cecconi M. ICU management based on big data. *Current Opinion in Anesthesiology*. 2020;33(2). doi:10.1097/ACO.0000000000000834
328. Choi MH, Kim D, Choi EJ, et al. Mortality prediction of patients in intensive care units using machine learning algorithms based on electronic health records. *Sci Rep*. 2022;12(1):7180. doi:10.1038/s41598-022-11226-4
329. Delahanty RJ, Kaufman D, Jones SS. Development and Evaluation of an Automated Machine Learning Algorithm for In-Hospital Mortality Risk Adjustment Among Critical Care Patients. *Crit Care Med*. 2018;46(6). doi:10.1097/CCM.0000000000003011
330. Iwase S, Nakada T aki, Shimada T, et al. Prediction algorithm for ICU mortality and length of stay using machine learning. *Sci Rep*. 2022;12(1):12912. doi:10.1038/s41598-022-17091-5

ORIGINAL PUBLICATIONS (I-IV)

I

Association of Sequential Organ Failure Assessment (SOFA) components with mortality

Pölkki A, Pekkarinen PT, Takala J, Selander T, Reinikainen M

Acta Anaesthesiol Scand. 2022;66:731-741

Received: 25 October 2021 | Revised: 16 March 2022 | Accepted: 21 March 2022

DOI: 10.1111/aas.14067

acta Anaesthesiologica Scandinavica

RESEARCH ARTICLE

Association of Sequential Organ Failure Assessment (SOFA) components with mortality

Anssi Pölkki[1,2] | Pirkka T. Pekkarinen[3] | Jukka Takala[4] | Tuomas Selander[5] | Matti Reinikainen[1,2]

[1]Department of Anaesthesiology and Intensive Care, Kuopio University Hospital, Kuopio, Finland

[2]University of Eastern Finland, Kuopio, Finland

[3]Division of Intensive Care Medicine, Department of Anaesthesiology, Intensive Care and Pain Medicine, Helsinki University Hospital, University of Helsinki, Helsinki, Finland

[4]Department of Intensive Care Medicine, University Hospital Bern (Inselspital), University of Bern, Bern, Switzerland

[5]Science Service Center, Kuopio University Hospital, Kuopio, Finland

Correspondence
Anssi Pölkki, M.D, Department of Anaesthesiology and Intensive Care, Kuopio University Hospital, P.O. Box 100, FI 70029 KYS, Finland.
Email: anssi.polkki@kuh.fi

Funding information
Dr Pölkki was supported by institutional funding from Kuopio University Hospital and University of Eastern Finland. Otherwise the authors declare that they received no specific funding for this study.

Abstract

Background: Sequential Organ Failure Assessment (SOFA) is a practical method to describe and quantify the presence and severity of organ system dysfunctions and failures. Some proposals suggest that SOFA could be employed as an endpoint in trials. To justify this, all SOFA component scores should reflect organ dysfunctions of comparable severity. We aimed to investigate whether the associations of different SOFA components with in-hospital mortality are comparable.

Methods: We performed a study based on nationwide register data on adult patients admitted to 26 Finnish intensive care units (ICUs) during 2012–2015. We determined the SOFA score as the maximum score in the first 24 hours after ICU admission. We defined organ failure (OF) as an organ-specific SOFA score of three or higher. We evaluated the association of different SOFA component scores with mortality.

Results: Our study population comprised 63,756 ICU patients. Overall hospital mortality was 10.7%. In-hospital mortality was 22.5% for patients with respiratory failure, 34.8% for those with coagulation failure, 40.1% for those with hepatic failure, 14.9% for those with cardiovascular failure, 26.9% for those with neurologic failure and 34.6% for the patients with renal failure. Among patients with comparable total SOFA scores, the risk of death was lower in patients with cardiovascular OF compared with patients with other OFs.

Conclusions: All SOFA components are associated with mortality, but their weights are not comparable. High scores of other organ systems mean a higher risk of death than high cardiovascular scores. The scoring of cardiovascular dysfunction needs to be updated.

KEYWORDS
Multiorgan Failure, SOFA, SOFA score, SOFA score components, SOFA score weights, Surrogate endpoint

Editorial Comment

In this large study from the Finnish ICU registry, evidence is provided to show poor performance of the cardiovascular component of the SOFA score. The authors suggest that a revision of this sub-score relative weight might improve the predictive value of the overall score for mortality.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Acta Anaesthesiologica Scandinavica* published by John Wiley & Sons Ltd on behalf of Acta Anaesthesiologica Scandinavica Foundation.

1 | INTRODUCTION

The Sequential Organ Failure Assessment (SOFA), at first named Sepsis-related Organ Failure Assessment, was introduced by The Working Group on Sepsis-Related Problems in 1996.[1] The SOFA score describes and quantifies the severity of dysfunction or failure of six essential organ systems (Table S1). Primarily, the SOFA score was not meant for outcome prediction. Multiple studies have shown, however, that it can rather well predict mortality in groups of critically ill patients.[1-13] This has notably widened the employment of the SOFA score beyond its original purpose.

In randomised controlled trials, the gold standard has been to use all-cause mortality as an endpoint. However, interventional trials often fail to detect any difference between study arms in mortality.[14] Therefore, there is growing interest to use surrogate endpoints, for example SOFA scores.[15-17] Regulatory authorities, including the European Medicines Agency, can under certain limitations approve the use of surrogate endpoints instead of mortality as primary endpoints.[18]

The change in the SOFA score during critical illness has been proposed to reflect the benefit or harm of the intervention of interest. The SOFA score, which is a scalar variable, is presumably more sensitive in detecting the effects of an intervention than mortality, a binary variable. However, the total SOFA score cannot be an unbiased trial endpoint unless all its components have comparable weights as measures of organ dysfunction severity. Moreover, some organ failures (OFs) are more likely to occur concurrently.[3] It is unclear whether different combinations of OFs affect the predictive value of total SOFA score.

The aim of this study was to investigate whether different SOFA score components, recorded during the first 24 h of intensive care, carry comparable weights in terms of their association with mortality. In other words, do patients with comparable total SOFA scores have comparable probabilities to perish regardless of which OFs they suffer from? We evaluated how combinations of different organ system failures are associated with mortality. Furthermore, we assessed the association of increasing SOFA scores with mortality across different admission groups. Mortality at hospital discharge was the primary endpoint. Mortality at ICU discharge and mortality within 12 months were secondary endpoints.

2 | METHODS

2.1 | Study design and participants

The study protocol was approved by the Research Ethics Committee of the Northern Savo Hospital District Data (225/13.02.00/2016), and research authorisation was obtained from the National Institute for Health and Welfare (THL/1585/5.05.00/2015). Due to the retrospective nature of the study, the Research Ethics Committee waived the written informed consent in line with Finnish act of personal data.

We performed a retrospective cohort study of data collected prospectively in the Finnish ICU quality register, the Finnish Intensive Care Consortium (FICC) database. The FICC is a national programme for benchmarking intensive care in Finland.[19] FICC covers all 26 general ICUs of central and university-level hospitals in Finland.

We included all adult patients admitted to Finnish ICUs between January 1, 2012 and December 31, 2015. For patients with multiple ICU treatment periods during the same hospitalisation, we included only the first ICU admission. In line with the 1998 paper by the working group that created the SOFA system,[2] we defined OF as an organ-specific SOFA score of three or higher. OF could appear isolated or as part of multiorgan failure.

We performed subgroup analyses to observe whether the findings were consistent, regardless of the admission type—medical, elective surgery and emergency surgery.

2.2 | Extracted variables

We extracted following variables from the FICC database: the most severe values of SOFA score components within the first 24 h after admission to the ICU and the outcome variables: vital status at ICU discharge, at hospital discharge and 12 months after ICU admission. Moreover, we gathered baseline data on Acute Physiology, Age, Chronic Health Evaluation (APACHE) II,[20] The Simplified Acute Physiology Score (SAPS) II,[21] age and sex. We also retrieved data on length of stay in the ICU and length of stay in hospital.

2.3 | Data handling and statistical methods

In the neurologic component, the SOFA score is based on the Glasgow Coma Score (GCS). For anaesthetised or sedated patients, the GCS recorded to the FICC registry is the last reliable GCS preceding sedation, in line with the SAPS II score.[21]

The hepatic SOFA score is based on the plasma bilirubin concentration. Bilirubin is normally measured when there is a clinical reason to suspect hepatic problems. Therefore, we consider normality of bilirubin concentrations as likely in patients for whom the data on bilirubin were missing. In these patients, we assumed the hepatic SOFA score to be 0. We made no assumption of normality for other SOFA components in cases of missing data. Therefore, we excluded patients with missing SOFA data concerning all other components except for the hepatic component. In addition, we excluded patients with missing mortality data.

We compared the characteristics of survivors and non-survivors at hospital discharge employing the Mann–Whitney U-test for continuous data and chi square test for categorical data. Using age-adjusted multivariable logistic regression, we evaluated the association between SOFA score components and mortality. All components as well as age were included in the analysis simultaneously. P-value of less than 0.05 was considered as statistically significant in all tests. We calculated standardized occurrence ratio (SOR) for each

set of at least two, three, or four concurrently occurring failing organ systems. SOR is a tool to evaluate whether particular OFs occur concurrently more frequently than anticipated by merely observing the frequencies of OFs. SOR was calculated as $N(o) \div [N \times p(a) \times p(b)]$, where $N(o)$ is the number of patients with OF of a and b, N is the total number of admissions and $p(a)$ and $p(b)$ are the proportions of patients with failure of organ systems a and b, respectively. In the same way, we calculated the SOR for patients with three and four concurrent OFs. SOR >1 signals that the odds of concurrent occurrence of these particular failing organ systems are increased. Bonferroni correction was used for multiple comparisons regarding the SOR analysis.

We used IBM SPSS Statistics, Version 22 (IBM Corp., Amonk, NY, USA) and R statistical software version 4.0.4 for the statistical analyses.

3 | RESULTS

3.1 | Study population

There were totally 71,492 ICU admissions during the study period. We excluded 4289 (6%) readmissions. Data were missing most commonly for the hepatic component, for 26,435 (39.3%) admissions. For other components, data were missing for few admissions: 14 (0%) in respiratory, 2144 (3.2%) in coagulation, 14 (0%) in cardiovascular, 1318 (2%) in neurologic and 14 (0%) in the renal component. We excluded 104 (0.2%) cases with missing data on vital status at hospital discharge. The final study population included 63,756 patients (Figure 1). For ICU and 12-month mortality calculations, we excluded 16 (0%) cases with missing data on vital status at ICU discharge and 3,717 (5.5%) cases with missing data on vital status at 12 months, respectively.

The median age of the patients was 64 years (inter-quartile range 52–73), and the majority (63.7%) were male. During the ICU stay, 66.9% of the patients needed mechanical ventilation and 6.1% renal replacement therapy. Baseline data are presented in Table 1. The median score in the respiratory component was 2 and in the cardiovascular component 3. In the cardiovascular component, the scores were almost equally distributed among the patients except for score 2, which was documented for only 678 (1.1%) patients. In all other components (coagulation, hepatic, neurological and renal), the median score was 0, with the score 1 being second most common (Figure 2). Of OFs, defined as an organ-specific SOFA score ≥3, the most common OF was cardiovascular failure, in 53.6% of patients. The second most common OF was respiratory failure, in 22.5% of patients.

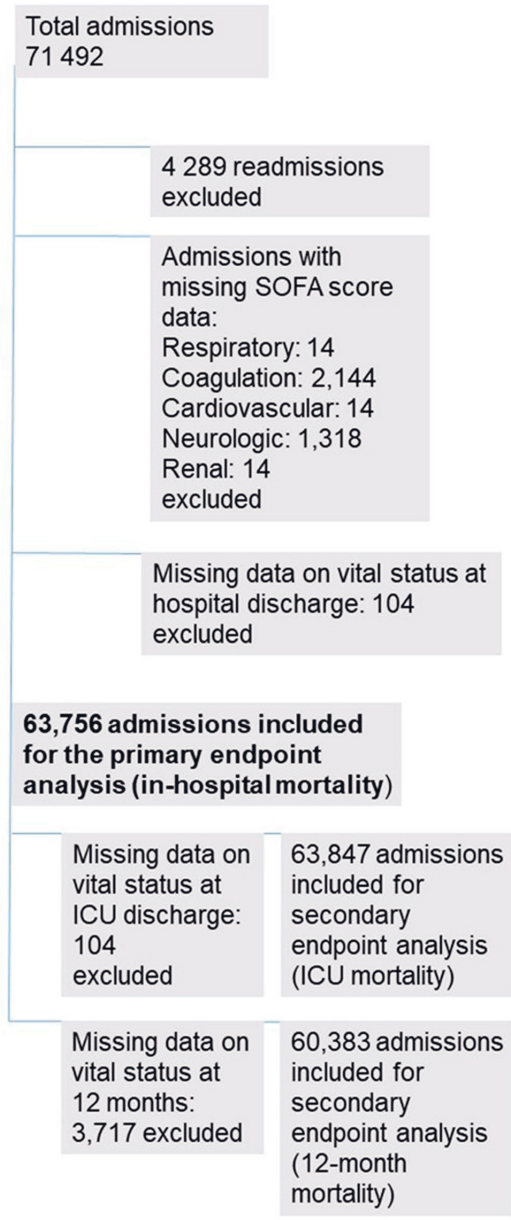

FIGURE 1 Flowchart

3.2 | ICU, hospital and 12-month mortality

Overall, 6,851 (10.7%) patients died in hospital. The first day total SOFA score was strongly associated with mortality (Figure 3). Mortality was 5.3% at ICU discharge and 21.6% in 12 months. Mortality increased with increasing SOFA scores (Figure 3). In-hospital mortality was 15.0% in those patients with LOS at the ICU more than 48 h. There were 642 (1%) patients with a SOFA score over 15. In these patients, ICU mortality was 60%, hospital mortality was 72%, and 12-month mortality was 80%.

TABLE 1 Demographics, baseline characteristics and lengths of stay in ICU and hospital

	Overall (n=63,756)	Survivors (n=56,905)	Non-survivors (n=6,851)	p-value
Age, median (IQR)	64 (52–73)	63 (51–73)	69 (61–77)	<0.001
Female, n (%)	23 121 (36.3%)	20 642 (36.3%)	2 479 (36.2%)	0.87
SOFA	6 (4–8)	6 (3–8)	10 (7–12)	<0.001
SAPS II	31 (23–44)	29 (22–40)	56 (44–69)	<0.001
APACHE II	18 (13–24)	17 (12–22)	29 (24–35)	<0.001
Metastatic cancer	1 550 (2.4%)	1 295 (2.3%)	255 (3.7%)	<0.001
Haematologic malignancy	886 (1.4%)	650 (1.1%)	236 (3.6%)	<0.001
AIDS	71 (0.1%)	60 (0.1%)	11 (0.2%)	0.02
Admission type				<0.001
Medical	34 987 (55.2%)	29 651 (52.5%)	5 336 (78.1%)	
Elective surgery	17 034 (26.8%)	16 774 (29.6%)	260 (3.8%)	
Emergency surgery	11 401 (17.9%)	10 165 (17.9%)	1236 (18.1%)	
Diagnostic category				<0.001
Cardiovascular surgery	15 130 (23.7%)	14 619 (25.7%)	517 (7.5%)	
Neurologic	10 753 (16.9%)	9 756 (17.1%)	997 (14.6%)	
Cardiovascular insufficiency	8 967 (14.1%)	6 651 (11.7%)	2 316 (33.8%)	
Metabolic or renal	6 862 (10.8%)	5 961 (10.5%)	901 (13.2%)	
Respiratory insufficiency	5 989 (9.4%)	4 987 (8.8%)	1 002 (14.6%)	
Gastrointestinal surgery	4 886 (7.7%)	4 288 (7.5%)	598 (8.7%)	
Trauma	4 316 (6.8%)	4 043 (7.1%)	273 (4.0%)	
Other postoperative cause	2 769 (4.3%)	2 664 (4.7%)	105 (1.5%)	
Intoxication	2 666 (4.2%)	2 615 (4.6%)	51 (0.7%)	
Miscellaneous	1 397 (2.2%)	1 307 (2.3%)	90 (1.3%)	
LOS ICU (days), median (IQR),	1.4 (0.9–3.1)	1.3 (0.9–2.9)	2.1 (0.9–5.1)	<0.001
LOS Hospital (days), median (IQR)	8 (5–14)	8 (5–14)	5 (2–13)	<0.001

Note: Data are presented as numbers with percentages or as medians (inter-quartile ranges). Characteristics of hospital survivors and non-survivors were compared with the Mann–Whitney U-test for continuous data and Chi-squared test for categorical data.

Abbreviations: AIDS, acquired immune deficiency syndrome; APACHE, acute physiology and health evaluation; ICU, intensive care unit; LOS, length of stay; QR, Inter-quartile range; SAPS, Simplified Acute Physiology Score; SD, standard deviationSOFA, Sequential Organ Failure Assessment.

Mortality mostly increased consistently with increasing SOFA component points (Figure 4). The cardiovascular component, however, was an exception. In this component, a clear increase in mortality occurred only in the group with the score 4. For the respiratory and coagulation components, mortality was similar for the scores of 0 and 1 points but increased consistently with increasing points thereafter. This pattern appeared rather similar regardless of whether the vital status was observed at ICU or hospital discharge or 12 months after ICU admission (Figures 4 and 5).

Mortality in patients with OFs (organ-specific SOFA score 3 or 4) increased with increasing total SOFA scores. However, within groups of patients with comparable total SOFA scores, mortality was lower in patients with cardiovascular OF compared with patients with other OFs in patients with a total SOFA score lower than 12. In fact, mortality in patients with cardiovascular OF did not exceed the mortality in patients with no first-day OF at all in patients with a total SOFA score lower than 9 (Figure 6).

Respiratory failure was observed for 16 277 (22.8%) patients, coagulation failure for 1 932 (2.7%), hepatic failure for 704 (1.0%), cardiovascular failure for 37 672 (52.7%), neurologic failure for 12 714 (17.8%), and renal failure for 5 958 (8.3%) patients (Figure 2). Hospital mortality was 22.5% for patients with respiratory failure, 34.8% for those with coagulation failure, 40.1% for those with hepatic failure, 14.9% for those with cardiovascular failure, 26.9% for those with neurologic failure and 34.6% for the patients with renal failure. Concerning patients with LOS more than 48 h, the in-hospital mortality was 20.4% for patients with respiratory failure, 30.1% for those with coagulation failure, 36.2% for those with hepatic failure, 17.1% for those with cardiovascular failure, 22.3% for those with neurologic failure, and 24.6% for the patients with renal failure.

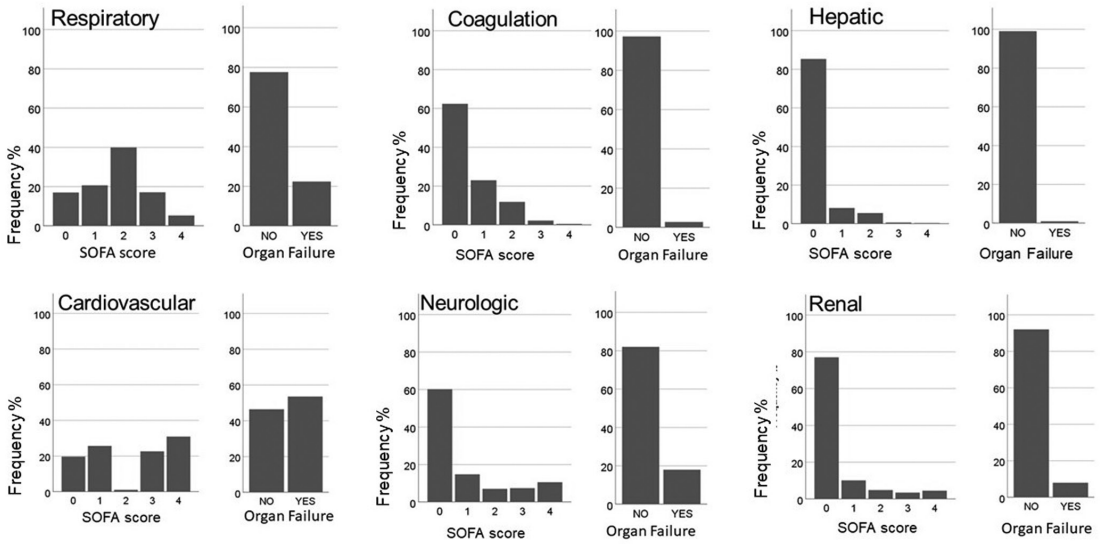

FIGURE 2 The distribution of SOFA component scores and frequency of organ failures

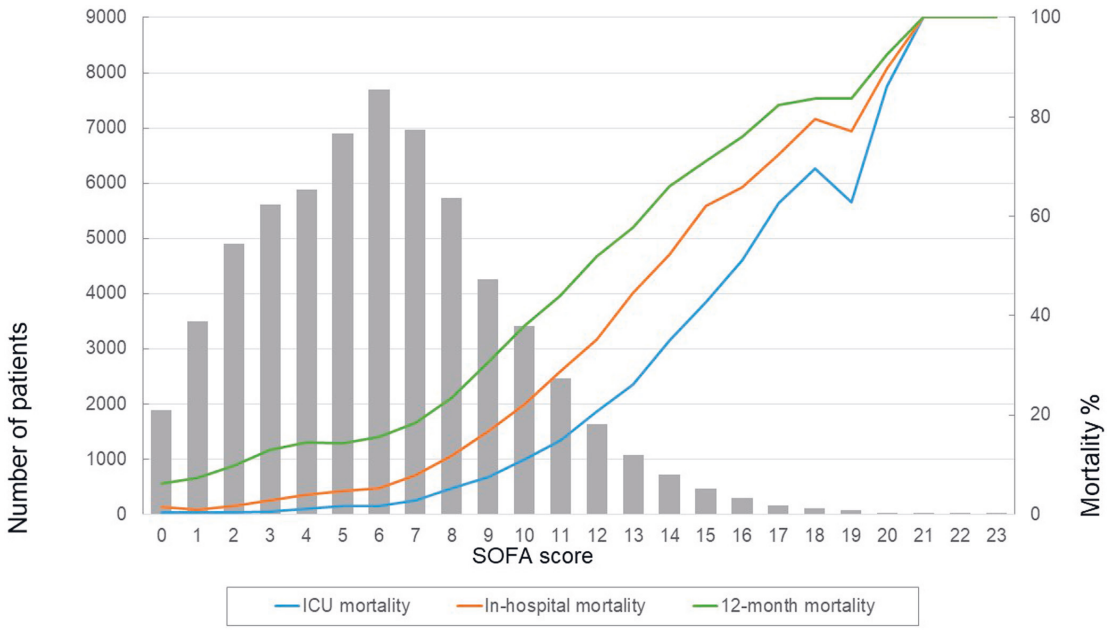

FIGURE 3 The number of patients and mortality according to first-day total SOFA score. ICU mortality (blue line), in-hospital mortality (red line) and 12-month mortality (green line) increased with increasing total SOFA score. The bars present the number of patients within each total SOFA score group

The results of age-adjusted multivariable logistic regression analysis are presented in Table 2. The odds of in-hospital death were highest for patients with neurologic failure, whereas the odds of death were lowest in patients with cardiovascular failure. Especially for 12-month outcome, cardiovascular OF had little influence on the risk of death.

3.3 | Combinations of organ system failures and mortality

Mortality increased with increasing numbers of concurrent OFs (Figure 7). Of all patients, 47.4% had at least two, 12.7% had at least

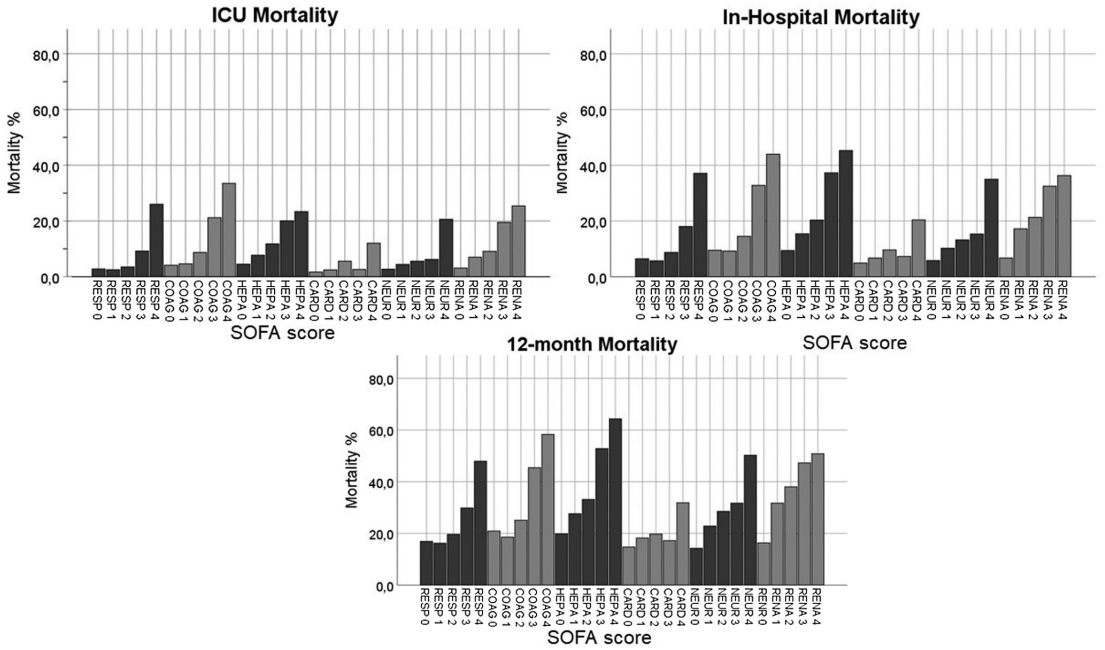

FIGURE 4 ICU mortality, in-hospital mortality and 12-month mortality according to SOFA component scores. ICU mortality, in-hospital mortality and 12-month mortality are presented in separate panels. The bars present the mortality in each SOFA score category recorded in the first 24 hours after ICU admission

three and 2.0% had at least four concurrent OFs. In-hospital mortalities in these groups were 35.8%, 54.1% and 71.8%, respectively. SOR was >1 in 48 (94.1%) out of all 51 OF combinations (Table S2), suggesting that OFs are likely to occur concurrently. In-hospital mortality ranged between 25.7%–65.2% in patients with two, 41.4%–82.4% in those with three, and 52.9%–85.7% in those with four failing organ systems, depending on which organ systems were failing. The variation in mortality according to the different sets of OFs decreased towards 12-month mortality observation.

4 | DISCUSSION

We evaluated Finnish ICU patients' SOFA scores during the first 24 h in the ICU and assessed the prevalence of different OFs, defined as the organ system-specific SOFA score of 3 or 4, and their associations with mortality. Cardiovascular failure, observed in 53% of patients, was the most common, followed by respiratory failure (23%), neurologic (18%), renal failure (8%), coagulation failure (3%) and hepatic failure (1%).

Mortality increased with increasing SOFA scores. However, scores reflecting dysfunctions of different organ systems were not equivalent as metrics of risk. In particular, high cardiovascular SOFA scores did not imply as high a risk of death as high scores of other SOFA components. In addition, OF combinations including cardiovascular failure were associated with lower mortality than other OF combinations: hospital mortality was in the range 25%–45% for patients with cardiovascular failure together with another OF, whereas mortality exceeded 50% for all other OF combinations except the combination of neurologic and respiratory failure (40%). Moreover, within a group of patients with comparable total SOFA scores, the risk of death was lower in patients with cardiovascular OF compared with patients with other OFs in patients with a total SOFA score lower than 12.

The contributions of SOFA component scores to outcome has not been studied much previously. However, our findings contradict those of the 1999 study by the Working Group on Sepsis Related Problems, where cardiovascular SOFA scores contributed more strongly than scores of other components to poor outcomes.[3] On the contrary, our results are in accordance with the study by Gupta et al. on 2796 septic patients with in-hospital mortality of 10%. Coagulation dysfunction or failure predicted a higher and cardiovascular dysfunction or failure a lower risk increase compared with dysfunctions of other organ systems.[22]

Recently, Bachmann et al. found that there are few patients with 2 cardiovascular SOFA points, and the prognostic value of cardiovascular SOFA was poor in patients assessed for intra-abdominal hypertension and gastrointestinal dysfunction.[23-25] Our findings in a large sample of general ICU patients confirm this. The distribution of the cardiovascular SOFA score had two peaks, made up of categories 0–1 and 3–4. A score of 2 was uncommon, present for roughly 1% of the patients. Two cardiovascular SOFA points are

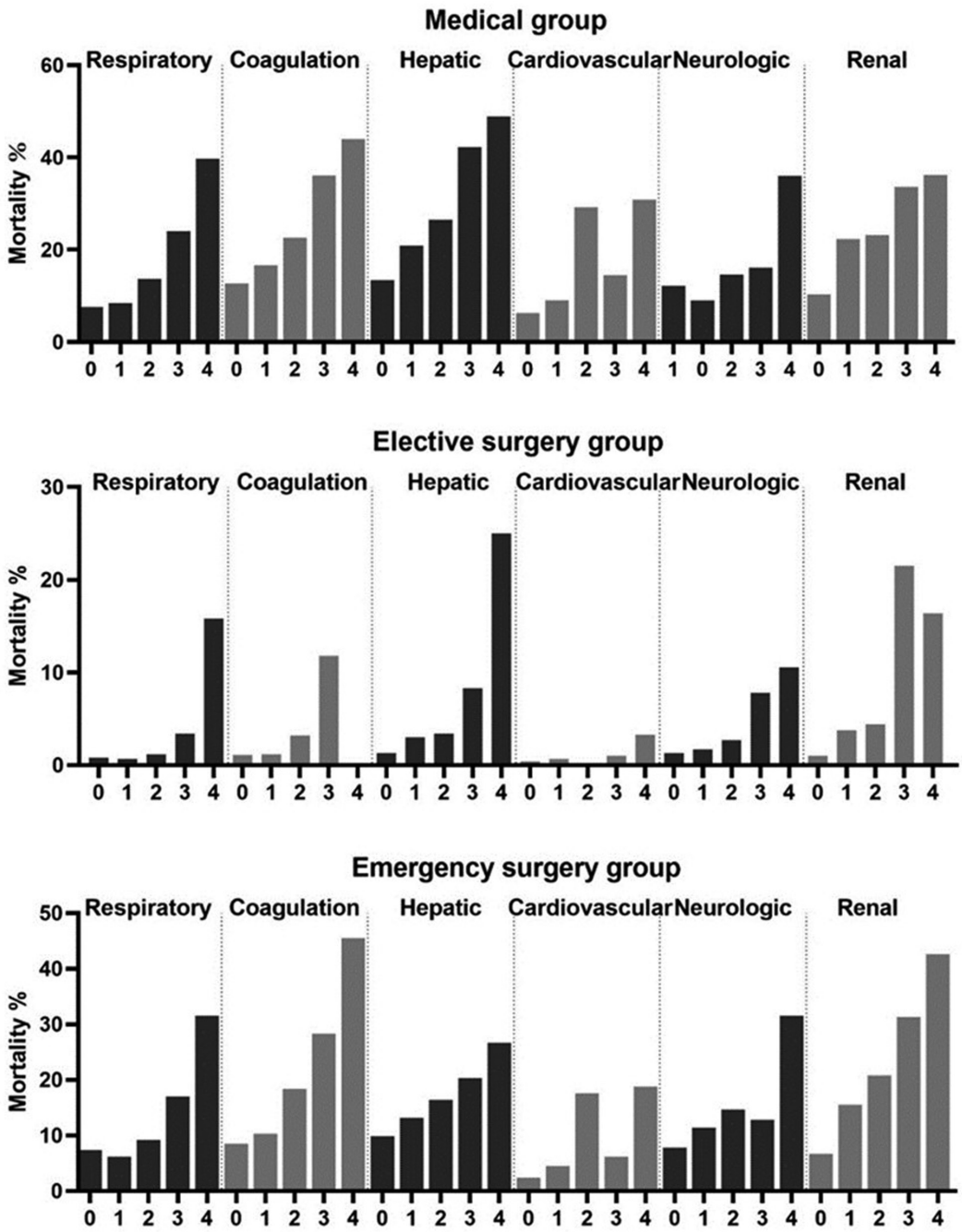

FIGURE 5 Hospital mortality according to SOFA component scores in different admission categories. The bars present the in-hospital mortality according to SOFA component scores in different admission categories (medical, elective surgical and emergency surgical). Hospital mortality increased with increasing SOFA component scores

scored to patients who are administered dopamine at a dose less than 5μg $kg^{-1} \cdot min^{-1}$ or dobutamine at any dose. Recent guidelines recommend against or advise specific caution for monotherapy use of these inotropes in circulatory shock.[26,27] However, administering dopamine to brain-dead organ donors with the intention to support renal function was relatively common in Finland during the study

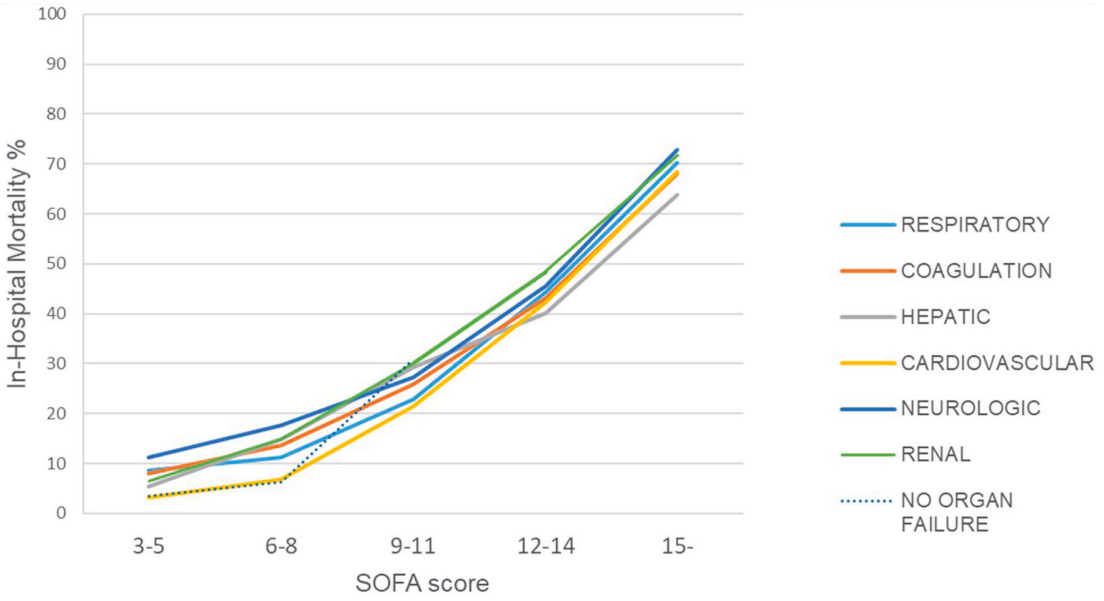

FIGURE 6 Mortality in patients with different organ failures according to total SOFA score. The lines represent in-hospital mortality in patients with respiratory (light blue), coagulation (orange), hepatic (grey), cardiovascular (yellow), neurologic (purple) and renal (green) failure. The organ failure was determined as organ-specific SOFA score 3 or 4. Mortality in patients without any first-day organ failure is shown with black dashed line

TABLE 2 The association of failures of different organ systems and age with ICU, hospital and 12-month mortality

	ICU mortality			Hospital mortality			12-month mortality		
	OR	95% CI		OR	95% CI		OR	95% CI	
Respiratory OF	2.92	2.70	3.16	2.41	2.27	2.56	1.71	1.63	1.79
Coagulation OF	4.18	3.63	4.82	4.04	3.57	4.57	3.24	2.891	3.64
Hepatic OF	2.27	1.78	2.89	4.24	3.47	5.17	4.27	3.53	5.17
Cardiovascular OF	2.15	1.95	2.36	1.57	1.47	1.67	1.05	1.01	1.10
Neurologic OF	4.63	4.28	5.01	5.00	4.71	5.30	4.13	3.93	4.34
Renal OF	5.99	5.48	6.55	4.93	4.58	5.32	3.81	3.56	4.07
Age (for each year)	1.01	1.01	1.02	1.03	1.03	1.04	1.04	1.04	1.04
Female sex	1.16	1.07	1.26	1.02	0.96	1.08	0.95	0.91	0.99

Note: Abbreviations: CI, confidence intervalOF, organ failure; OR, odds ratio.

period,[28] which may partly explain the high mortality in this SOFA category.

Outcomes of ICU patients have improved over the years. In 1998, Vincent et al. reported an ICU mortality of 90% in patients with a SOFA score above 15,[2] whereas in-hospital mortality for patients with first-day SOFA score above 15 was 72% in our study. The overall ICU and in-hospital mortality was lower in our study compared with that reported in the 1990s.[2,3] In addition to assumed improvements in prognosis of ICU patients, a plausible explanation for this is that we also included patients with preceding scheduled surgery. Our results show, however, that the cardiovascular SOFA score was associated with lower risk of mortality in the whole cohort, in both emergency and elective admissions, as well as those with at least 48 h length of ICU stay.

Although high SOFA scores often indicate a poor prognosis, cardiovascular scores seem to be an exception. This may reflect a change in clinical practices in recent years. The SOFA score was introduced in an era of more restricted use of vasopressors. During the last two decades, the use of norepinephrine has become more common.[29-31] Vasopressor treatment is initiated earlier without preceding large doses of resuscitation fluids.[32-35] An infusion of norepinephrine lasting at least one hour, even at a small dose, assigns 3 points to the

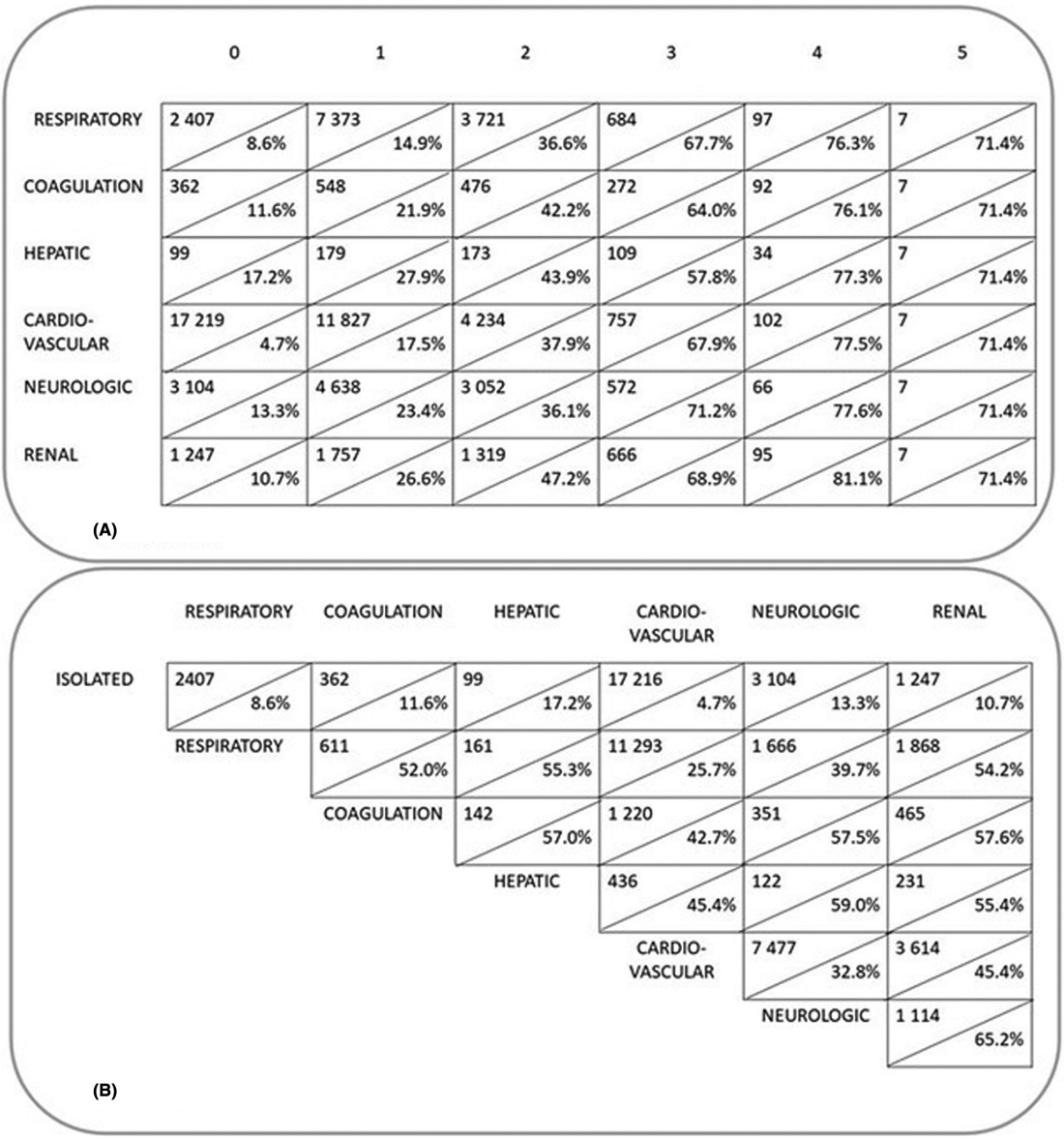

FIGURE 7 Mortality according to the number of failing organ systems and mortality in groups with at least two simultaneous organ failures. In panel a, each organ failure is represented by a line. In each box, the number above the diagonal line presents the number of patients with the column title-presented number of additional failing organ systems in addition to the organ failure of that line. The percentage below the diagonal line presents the in-hospital mortality of these patients. Panel b presents the number of patients with a combination of at least two organ failures. The percentage below the diagonal line shows the in-hospital mortality in patients with that particular combination

cardiovascular component of the SOFA score. Moreover, an infusion rate exceeding 0.1 μg kg^{-1}·min^{-1}, which is not a particularly high dose in contemporary intensive care, gives four points. Because of this change in clinical practice, the cardiovascular SOFA score seems to have suffered from inflation. This could also explain the divergence of our findings from those made by Moreno et al.[3] more than two decades ago.

Risk of death increases with an increasing amount of failing organ systems.[36,37] Our findings imply that some OFs are more likely to occur concurrently than other failures. Moreover, mortality was dependent

on which organ systems were failing. The Working group on sepsis-related problems demonstrated a pattern for concurrently occurring OFs by means of principal components analysis.[3] The group identified two common OF combinations. The first combination comprised respiratory, cardiovascular and neurologic OFs, whereas the second comprised coagulation, hepatic and renal OFs. In our study, this first combination of respiratory, cardiovascular and neurologic OFs was also the most common of the combinations with three OFs, affecting 37% of patients with at least three concurrently failing organ systems. The in-hospital and 12-month mortalities associated with this particular combination were 41% and 55%, respectively, whereas in-hospital and 12-month mortalities of patients with other triple OF combinations ranged between 56%–82% and 63%–88%, respectively.

We found that the second combination, which comprised coagulation, hepatic and renal OFs, occurred 44 times more often than one would have expected by observing merely the frequency of these OFs in the whole study population.

There is growing interest in using the SOFA score as a surrogate endpoint for mortality in clinical trials.[17] Our findings suggest that this may not be without problems. Regarding risk of death, weights of different SOFA component scores are different, and the prognosis of patients with multi-OF is dependent on which organ systems fail. In particular, cardiovascular SOFA scores do not signal cardiovascular dysfunction of equivalent severity to dysfunctions reflected by similar scores of other organ systems. The scoring criteria of cardiovascular dysfunction/failure may need an update. However, we must be aware that changing even one of the SOFA components would practically create a second version of the SOFA score. This might improve the measurement of organ dysfunctions but also mean that we would lose comparability to previous studies that have used the original SOFA score.

4.1 | Strengths and limitations of the study

Our study population consisted of a large unselected group of patients treated in Finnish ICUs. The data were retrieved from a high-quality national database with all Finnish general ICUs participating. Therefore, our study population is well representative of adult ICU patients in Finland. We do not know whether the results are generalizable to other countries. However, early use of norepinephrine has become more common in other countries as well,[31] and it is likely that the relation between cardiovascular SOFA scores and mortality may have weakened also in other countries.

A major limitation of our study is that the SOFA scores were based only on measurements during the first 24 h after admission to the ICU, whereas previous studies have shown that a change in SOFA score over time is the most reliable predictor of mortality.[38,39]

5 | CONCLUSION

All SOFA components are associated with mortality. However, high cardiovascular SOFA scores did not mean as high a risk of death as high scores of other SOFA components. Moreover, OF combinations including cardiovascular failure were associated with lower mortality than other OF combinations. OFs are likely to occur concurrently. The scoring of cardiovascular dysfunction needs to be updated.

ACKNOWLEDGEMENTS

We are grateful to the doctors and nurses in the ICUs participating in the Finnish Intensive Care Consortium for careful data documentation and to TietoEVRY for managing the database.

CONFLICTS OF INTEREST

None.

AUTHOR CONTRIBUTIONS

MR presented the first idea of the study. AP analysed and interpreted the data, supported by MR and PP. AP drafted the first version of the manuscript and created the figures. TS contributed in statistical analysis. JT helped in interpreting the results. All authors revised the manuscript and read and approved the final manuscript.

ORCID

Anssi Pölkki https://orcid.org/0000-0001-8777-5287
Pirkka T. Pekkarinen https://orcid.org/0000-0002-8203-3409
Matti Reinikainen https://orcid.org/0000-0001-6878-3740

REFERENCES

1. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med*. 1996;22:707-710.
2. Vincent JL, de Mendonça A, Cantraine F, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units. *Crit Care Med*. 1998;26:1793-1800.
3. Moreno R, Vincent J, Matos R, et al. The use of maximum SOFA score to quantify organ dysfunction/failure in intensive care. Results of a prospective, multicentre study. *Intensive Care Med*. 1999;25:686-696.
4. Ferreira FL, Bota DP, Bross A, Melot C, Vincent JL. How changes in SOFA score can predict outcome. *Crit Care Med*. 1999;27:A50.
5. Jentzer JC, Bennett C, Wiley BM, et al. Predictive value of individual Sequential Organ Failure Assessment sub-scores for mortality in the cardiac intensive care unit. *PLoS One*. 2019;14:e0216177.
6. Karakike E, Kyriazopoulou E, Tsangaris I, Routsi C, Vincent JL, Giamarellos-Bourboulis EJ. The early change of SOFA score as a prognostic marker of 28-day sepsis mortality: analysis through a derivation and a validation cohort. *Crit Care*. 2019;23:387.
7. Innocenti F, Tozzi C, Donnini C, et al. SOFA score in septic patients: incremental prognostic value over age, comorbidities, and parameters of sepsis severity. *Intern Emerg Med*. 2019;13:405-412.
8. Kajdacsy-Balla Amaral AC, Andrade FM, Moreno R, Artigas A, Cantraine F, Vincent JL. Use of the Sequential Organ Failure Assessment score as a severity score. *Intensive Care Med*. 2005;31:243-249.
9. Peres Bota D, Melot C, Lopes Ferreira F, Nguyen BV, Vincent JL. The Multiple Organ Dysfunction Score (MODS) versus the Sequential Organ Failure Assessment (SOFA) score in outcome prediction. *Intensive Care Med*. 2002;28:1619-1624.
10. Raith EP, Udy AA, Bailey M, et al. Prognostic accuracy of the SOFA score, SIRS criteria, and qSOFA score for in-hospital mortality among adults with suspected infection admitted to the intensive care unit. *JAMA*. 2017;317:290-300.

11. Jones AE, Trzeciak S, Kline JA. The Sequential Organ Failure Assessment score for predicting outcome in patients with severe sepsis and evidence of hypoperfusion at the time of emergency department presentation. *Crit Care Med*. 2009;37:1649-1654.
12. Basile-Filho A, Lago AF, Menegueti MG, et al. The use of SAPS 3, SOFA, and Glasgow Coma Scale to predict mortality in patients with subarachnoid hemorrhage. *Medicine*. 2018;97:e12769.
13. Pekkarinen PT, Bäcklund M, Efendijev I, et al. Association of extracerebral organ failure with 1-year survival and healthcare-associated costs after cardiac arrest: an observational database study. *Crit Care*. 2019;23:67.
14. Santacruz CA, Pereira AJ, Celis E, Vincent JL. Which multicenter randomized controlled trials in critical care medicine have shown reduced mortality? *A Systematic Review. Crit Care Med*. 2019;47:1680-1691.
15. Vincent JL. We should abandon randomized controlled trials in the intensive care unit. *Crit Care Med*. 2010;38:S534-S538.
16. Vincent JL. Endpoints in sepsis trials: more than just 28-day mortality? *Crit Care Med*. 2004;32:S209-S213.
17. Lambden S, Laterre PF, Levy MM, Francois B. The SOFA score—development, utility and challenges of accurate assessment in clinical trials. *Crit Care*. 2019;23:374.
18. Guideline on clinical investigation of medicinal products for the treatment of sepsis. Edited by European Medicine Agency. CHMP/EWP/4713/03 2006.
19. Reinikainen M, Mussalo P, Hovilehto S, et al. Finnish Intensive Care Consortium, Association of automated data collection and data completeness with outcomes of intensive care. A new customised model for outcome prediction. *Acta Anaesthesiol Scand*. 2012;56:1114-1122.
20. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med*. 1985;13:818-829.
21. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA*. 1993;270:2957-2963.
22. Gupta T, Puskarich MA, DeVos E, et al. Sequential organ failure assessment component score prediction of in-hospital mortality from sepsis. *J Intensive Care Med*. 2020;35:810-817.
23. Reintam Blaser A, Regli A, De Keulenaer B, et al. Incidence, risk factors, and outcomes of intra-abdominal hypertension in critically ill patients-a prospective multicenter study (IROI Study). *Crit Care Med*. 2019;47:535-542.
24. Reintam Blaser A, Padar M, Mändul M, et al. Development of the gastrointestinal dysfunction score (GIDS) for critically ill patients—a prospective multicenter observational study (iSOFA study). *Clin Nutr*. 2021;40:4932-4940.
25. Bachmann KF, Arabi YM, Regli A, Starkopf J, Reintamm BA. Cardiovascular SOFA score may not reflect current practice. *Intensive Care Med*. 2021;. doi:10.1007/s00134-021-06536-6
26. Ponikowski P, Voors AA, Anker SD, et al. ESC Scientific Document Group, 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J*. 2016;37:2129-2200.
27. Møller MH, Granholm A, Junttila E, et al. Scandinavian SSAI clinical practice guideline on choice of inotropic agent for patients with acute circulatory failure. *Acta Anaesthesiol Scand*. 2018;62:420-450.
28. Mundt HM, Yard BA, Krämer BK, Benck U, Schnülle P. Optimized donor management and organ preservation before kidney transplantation. *Transpl Int*. 2016;29:974-984.
29. Jentzer JC, Wiley B, Bennett C, et al. Temporal trends and clinical outcomes associated with vasopressor and inotrope use in the cardiac intensive care unit. *Shock*. 2020;53:452-459.
30. Thongprayoon C, Cheungpasitporn W, Harrison AM, et al. Temporal trends in the utilization of vasopressors in intensive care units: an epidemiologic study. *BMC Pharmacol. Toxicol*. 2016;17:19.
31. Scheeren TWL, Bakker J, De Backer D, et al. Current use of vasopressors in septic shock. *Ann Intensive Care*. 2019;9:20.
32. Marik PE, Linde-Zwirble WT, Bittner EA, Sahatjian J, Hansell D. Fluid administration in severe sepsis and septic shock, patterns and outcomes: an analysis of a large national database. *Intensive Care Med*. 2017;43:625-632.
33. Pittard MG, Huang SJ, McLean AS, Orde SR. Association of positive fluid balance and mortality in sepsis and septic shock in an Australian cohort. *Anaesth Intensive Care*. 2017;45:737-743.
34. Boyd JH, Forbes J, Nakada T-A, Walley KR, Russell JA. Fluid resuscitation in septic shock: a positive fluid balance and elevated central venous pressure are associated with increased mortality. *Crit Care Med*. 2011;39:259-265.
35. Maheshwari K, Nathanson BH, Munson SH, et al. The relationship between ICU hypotension and in-hospital mortality and morbidity in septic patients. *Intensive Care Med*. 2018;44:857-867.
36. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. Prognosis in acute organ-system failure. *Ann Surg*. 1985;202:685-693.
37. Hebert PC, Drummond AJ, Singer J, Bernard GR, Russell JA. A simple multiple system organ failure scoring system predicts mortality of patients who have sepsis syndrome. *Chest*. 1993;104:230-235.
38. Levy MM, Macias WL, Vincent JL, et al. Early changes in organ function predict eventual survival in severe sepsis. *Crit Care Med*. 2005;33:2194-2201.
39. de Grooth H, Geenen IL, Girbes AR, Vincent JL, Parienti J-J, Oudemans-Van SH. SOFA and mortality endpoints in randomized controlled trials: a systematic review and meta-regression analysis. *Crit Care*. 2017;21:38.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Pölkki A, Pekkarinen PT, Takala J, Selander T, Reinikainen M. Association of Sequential Organ Failure Assessment (SOFA) components with mortality. *Acta Anaesthesiol Scand*. 2022;66:731–741. doi:10.1111/aas.14067

II

Vasoactive Inotropic Score compared to the Sequential Organ Failure Assessment cardiovascular score in intensive care

Pölkki A, Pekkarinen PT, Lahtinen P, Koponen T, Reinikainen M

Acta Anaesthesiol Scand. 2023;67:1219-1228

Received: 26 January 2023 | Revised: 8 May 2023 | Accepted: 22 May 2023

DOI: 10.1111/aas.14287

RESEARCH ARTICLE

acta Anaesthesiologica Scandinavica

Vasoactive Inotropic Score compared to the sequential organ failure assessment cardiovascular score in intensive care

Anssi Pölkki[1] | Pirkka T. Pekkarinen[2] | Pasi Lahtinen[3] | Timo Koponen[4] | Matti Reinikainen[1]

[1]Department of Anaesthesiology and Intensive Care, Kuopio University Hospital and University of Eastern Finland, Kuopio, Finland

[2]Division of Intensive Care Medicine, Department of Anaesthesiology, Intensive Care and Pain Medicine, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[3]Department of Anaesthesiology and Intensive Care, Kuopio University Hospital, Kuopio, Finland

[4]Department of Anaesthesiology and Intensive Care, North Karelia Central Hospital, Joensuu, Finland

Correspondence
Anssi Pölkki, Department of Anaesthesiology and Intensive Care, Kuopio University Hospital, PO Box 100, FI 70029 Kuopio, Finland.
Email: anssi.polkki@pshyvinvointialue.fi

Funding information
Einar och Karin Stroems stiftelse; Finska Läkaresällskapet; Kuopio University Hospital; University of Eastern Finland

Abstract

Background: The cardiovascular component of the sequential organ failure assessment (cvSOFA) score may be outdated because of changes in intensive care. Vasoactive Inotropic Score (VIS) represents the weighted sum of vasoactive and inotropic drugs. We investigated the association of VIS with mortality in the general intensive care unit (ICU) population and studied whether replacing cvSOFA with a VIS-based score improves the accuracy of the SOFA score as a predictor of mortality.

Methods: We studied the association of VIS during the first 24 h after ICU admission with 30-day mortality in a retrospective study on adult medical and non-cardiac emergency surgical patients admitted to Kuopio University Hospital ICU, Finland, in 2013–2019. We determined the area under the receiver operating characteristic curve (AUROC) for the original SOFA and for $SOFA_{VISmax}$, where cvSOFA was replaced with maximum VIS (VIS_{max}) categories.

Results: Of 8079 patients, 1107 (13%) died within 30 days. Mortality increased with increasing VIS_{max}. AUROC was 0.813 (95% confidence interval [CI], 0.800–0.825) for original SOFA and 0.822 (95% CI: 0.810–0.834) for $SOFA_{VISmax}$, $p < .001$.

Conclusion: Mortality increased consistently with increasing VIS_{max}. Replacing cvSOFA with VIS_{max} improved the predictive accuracy of the SOFA score.

KEYWORDS
cardiovascular organ failure, cardiovascular SOFA score, mortality prediction, prognostication, sequential organ failure assessment, Vasoactive Inotropic Score

Editorial Comment

In this Finnish registry study, the authors studied the impact of a Vasoactive Inotropic Score (VIS) on mortality. They found mortality to increase with increasing VIS which could hence be used in a revised sequential organ failure assessment score.

1 | INTRODUCTION

The sequential organ failure assessment (SOFA) score[1] is commonly used to quantify the severity of organ dysfunctions in intensive care (Table S1). Recent studies have shown that cardiovascular SOFA (cvSOFA) scores may no longer be appropriate measures of cardiovascular failure in modern intensive care.[2–5] Thus, scoring of cardiovascular dysfunction should be updated.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. *Acta Anaesthesiologica Scandinavica* published by John Wiley & Sons Ltd on behalf of Acta Anaesthesiologica Scandinavica Foundation.

Acute cardiovascular failure is a life-threatening condition affecting approximately one third of patients admitted to intensive care.[6–8] Vasopressors, such as norepinephrine, epinephrine, and vasopressin, along with inotropes, including dobutamine, dopamine, milrinone, and levosimendan, provide the foundation of vasopressor/inotropic support aiming to maintain sufficient perfusion of vital organs. Norepinephrine is currently the most used vasopressor, although there is no definitive evidence of superiority of any one agent over others.[9,10]

The amount of vasopressor/inotropic support needed may be used as an indirect measure of cardiovascular failure. Vasoactive Inotropic Score (VIS) is a scoring system which quantifies the amount of pharmacologic cardiovascular support.[11–13] Using conversion factors to make doses of different drugs comparable, VIS reflects the sum of doses of most common intravenous drugs used for acute cardiovascular failure. Previous studies have demonstrated the association of higher VIS with increased mortality and unfavourable events in neonatal and paediatric cardiac surgery patients and in paediatric patients with septic shock.[14–18] Moreover, the value of VIS as an outcome predictor has been confirmed in adult cardiac surgery patients.[19,20] However, the usefulness of VIS in general adult intensive care unit (ICU) patients is unknown.

In this study, we evaluated the association of VIS with mortality in a general adult ICU population and examined whether replacing the cvSOFA score with a VIS-based score improves the predictive accuracy of the total SOFA score.

2 | METHODS

2.1 | Study population and eligibility criteria

We performed a single-centre retrospective study on patients admitted to the ICU at Kuopio University Hospital in Kuopio, Finland. We extracted data on all patients admitted between January 1st 2013 and December 31st, 2019. The research ethics committee of the Northern Savo hospital district reviewed and approved the study protocol, and permission to use the data for research was granted by the hospital administration (reference number: 478/2021). The need for consent of the patients was waived due to the retrospective nature of the study.

TABLE 1 VIS agents, units and coefficients as suggested by Gaies et al. and Favia et al.[12,13] converted to norepinephrine equivalent dosages.

Agent	Coefficient
Norepinephrine	1
Epinephrine	1
Dobutamine	0.01
Dopamine	0.01
Milrinone	0.1
Levosimendan	0.5
Vasopressin	100

Abbreviation: VIS, Vasoactive Inotropic Score.

We extracted data on the doses of inotropes and vasopressors administered during the first 24 h after ICU admission from the electronic health record system. Data on clinical characteristics, the SOFA score recorded during the first 24 h after ICU admission, and length of stay (LOS) in the ICU and hospital had been prospectively validated and documented in the Finnish ICU quality register, Finnish Intensive Care Consortium.[21]

We excluded all elective (non-emergency) postoperative admissions and all cardiac surgery patients. Regarding the 24-h SOFA score, we excluded all patients with missing data on any SOFA score component (respiratory, cardiovascular, renal, coagulation, and neurological) from the first 24 h after ICU admission except for hepatic function, which we assumed to be normal in case of missing data. The reasoning for this assumption of normality was that determination of hepatic score is based on the serum concentration of bilirubin, which is not routinely measured at KUH for all patients with no clinical suspicion of hepatic problems. For patients with recurrent ICU admissions during the same hospitalisation, only the first ICU admission was included. Patients admitted for the sole purpose of possible organ donation after brain death were excluded.

TABLE 2 The cutoff values for VIS_{max} and VIS_{mean}.

Category	VIS_{max}	VIS_{mean}
0	0	0
1	>0	>0
2	>0.15	>0.05
3	>0.3	>0.1
4	>0.45	>0.15

Abbreviations: VIS, Vasoactive Inotropic Score; VIS_{max}, maximum VIS; VIS_{mean}, mean VIS.

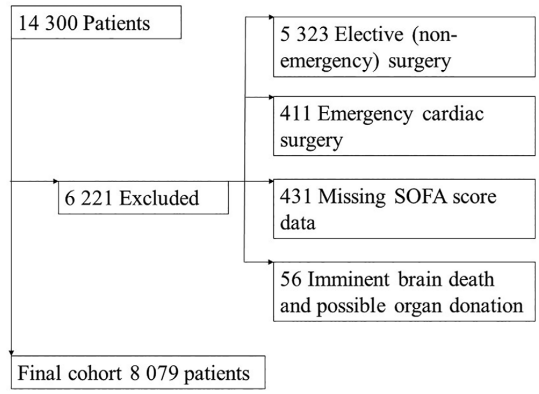

FIGURE 1 Flowchart of patient selection.

TABLE 3 Characteristics of the study population, stratified according to 30-day outcome.

	Overall	Alive at 30 days	Dead at 30 days	*P* value
Number of patients (%)	8079	6972 (86.3%)	1107 (13.7%)	
Age (years)	61 (49–70)	60 (47–69)	68 (60–75)	<.001
Sex (male)	5041 (62.4%)	4326 (62.0%)	715 (64.6%)	.10
VIS_{max} (μg/kg/min)	0 (0–0.10)	0 (0–0.07)	0.097 (0–0.26)	<.001
VIS_{mean} (μg/kg/min)	0 (0–01.6)	0 (0–0.0088)	0.014 (0–0.099)	<.001
SOFA score[a]	5 (3–8)	5 (2–7)	9 (6–11)	<.001
Cardiovascular SOFA	1 (0–3)	1 (0–3)	3 (1–4)	
VIS_{max}-based cardiovascular SOFA	0 (0–1)	0 (0–1)	1 (0–2)	
VIS_{mean}-based cardiovascular SOFA	0 (0–1)	0 (0–1)	1 (0–2)	
Respiratory SOFA	2 (1–2)	2 (1–2)	2 (1–3)	
Coagulation SOFA	0 (0–1)	0 (0–1)	0 (0–1)	
Hepatic SOFA	0 (0–0)	0 (0–0)	0 (0–1)	
Renal SOFA	0 (0–0)	0 (0–0)	0 (0–1)	
Neurologic SOFA	1 (0–2)	0 (0–2)	2 (0–4)	
APACHE II score	18 (12–24)	17 (12–22)	27 (22–33)	<.001
SAPS II score	32 (23–44)	22 (30–40)	53 (40–65)	<.001
Postoperative admissions	1730 (21.4%)	1499 (21.5%)	213 (20.9%)	.66
LOS ICU (median, SD)	1.6 (0.9–3.1)	1.6 (0.9–3.0)	1.8 (0.8–3.9)	.47
LOS ICU (mean, SD)	3.1 (4.4)	3.1 (4.2)	3.4 (4.0)	.005
LOS hospital (median, IQR)	6.2 (3.1–12.0)	6.5 (3.4–12.4)	3.6 (1.0–8.8)	<.001
LOS hospital (mean, SD)	8.5 (33.7)	8.8 (34.9)	5.4 (21.1)	<.001
Mechanical ventilation				
During the first 24 h	3754 (46.5%)	2922 (42.0%)	832 (75.2%)	<.001
During the ICU stay	4269 (52.9%)	3348 (48.0%)	921 (83.2%)	<.001
Admission diagnosis by APACHE IV				<.001
Neurologic diseases (nonop)	1568 (19.4%)	1402 (20.1%)	166 (15.0%)	
Cardiovascular diseases (nonop)	1300 (16.1%)	969 (13.9%)	331 (29.9%)	
Respiratory diseases (nonop)	910 (11.3%)	751 (10.8%)	159 (14.4%)	
Metabolic diseases	803 (9.9%)	763 (10.9%)	40 (3.6%)	
Trauma (nonop)	739 (9.1%)	682 (9.8%)	57 (5.1%)	
Neurologic diseases (postop)	616 (7.6%)	499 (7.2%)	117 (10.6%)	
Gastroenterological diseases (nonop)	503 (6.2%)	425 (6.1%)	78 (7.0%)	
Other nonoperative diseases	383 (4.7%)	362 (5.2%)	21 (1.9%)	
Cardiovascular diseases (postop)	347 (4.3%)	308 (4.4%)	39 (3.5%)	
Trauma (postop)	277 (3.4%)	242 (3.5%)	35 (3.1%)	
Gastroenterological diseases (postop)	237 (2.9%)	202 (2.9%)	35 (3.2%)	
Gynaecologic diseases (postop)	95 (1.2%)	94 (1.3%)	1 (0.1%)	
Renal diseases (nonop)	95 (1.2%)	84 (1.2%)	11 (1.0%)	
Respiratory diseases (postop)	93 (1.2%)	91 (1.3%)	2 (0.2%)	
Not defined	44 (0.5%)	42 (0.6%)	2 (0.2%)	
Haematologic diseases	42 (0.5%)	29 (0.4%)	13 (1.2%)	
Orthopaedic diseases (postop)	25 (0.3%)	25 (0.4%)	0 (0%)	
Renal diseases (postop)	2 (0%)	2 (0%)	0 (0%)	

Note: Data presented as absolute numbers and percentages, as means and standard errors, or as medians with interquartile ranges. Mann–Whitney *U* test for continuous variables and Fisher's exact test or Pearson Chi-Square test for categorical variables was used in the comparison of survivors and non-survivors.
Abbreviations: APACHE, Acute Physiology And Chronic Health Evaluation; ICU, intensive care unit; LOS, length of stay; nonop, non-surgical admission; postop, postoperative surgical admission; SAPS Simplified Acute Physiology Score; SD, standard error; SOFA, sequential organ failure assessment; VIS, Vasoactive Inotropic Score; VIS_{max}, maximum VIS; VIS_{mean}, mean VIS.

[a]Hepatic SOFA score for the first 24 h was missing with 1725 (21.4%) patients.

2.2 | VIS-based variables

We defined two separate VIS-based variables: the maximum (VIS_{max}) and mean (VIS_{mean}) values. We used data from the first 24 h after the ICU admission to calculate these values. To make doses of different vasopressor/inotropic medications comparable, we multiplied the doses of medications with inotropic/vasopressor-specific coefficients as previously suggested[12,13] also including levosimendan as suggested by Favia et al.[13] to capture most vasopressors and inotropes commonly used in clinical practice. We calculated VIS as the coefficient-adjusted sum of intravenously administered vasopressor/inotropic medications. Unlike Gaies et al., who reported the vasopressor/inotropic medications in dopamine equivalents,[12] we used norepinephrine equivalents in reporting (Table 1). We defined VIS_{max} as the sum of the highest coefficient-adjusted infusion rates per kilogram per minute of each vasopressor/inotrope. To calculate VIS_{mean}, we first calculated the total amount of each medication administered during the first 24 h after ICU admission. We then divided this coefficient-adjusted sum of infusions by 1440 to get the mean dose in μg/kg/min. For vasopressin, we used U/kg/min as the unit of infusion rate. The VIS_{mean} value for patients with LOSs less than 24 h was calculated for the true length of ICU stay.

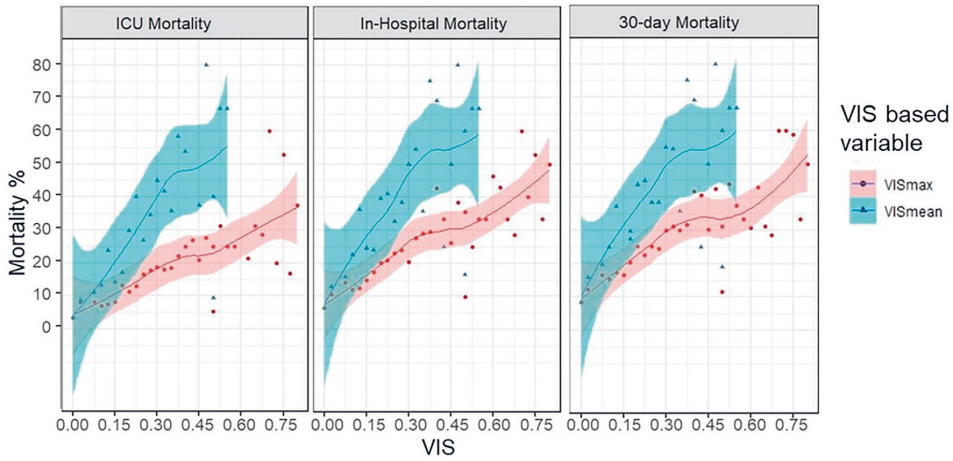

FIGURE 2 A scatterplot presenting the relationship between VIS and mortality. The turquoise line represents the mean (VIS_{mean}) and red line represents the maximum (VIS_{max}) VIS value recorded during the first 24 h at the ICU. The turquoise and red areas represent the standard error. ICU, intensive care unit; VIS_{max}, maximum Vasoactive Inotropic Score; VIS_{mean}, mean Vasoactive Inotropic Score.

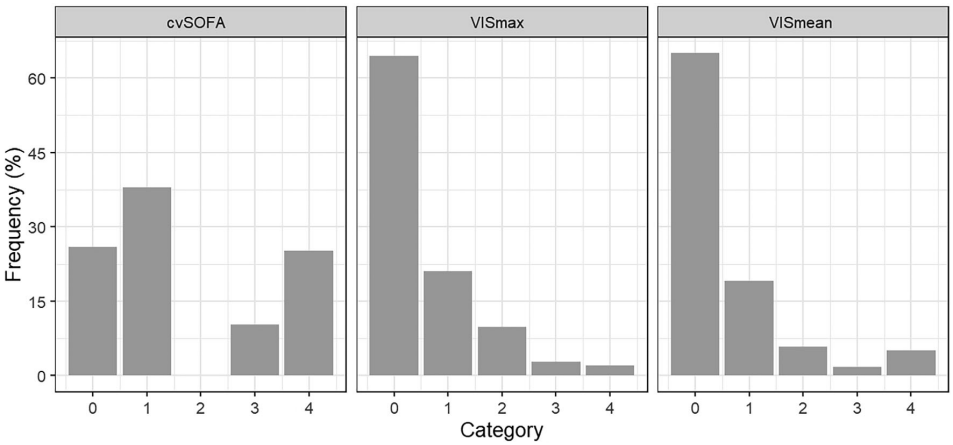

FIGURE 3 The distribution in VIS categories and cvSOFA scores. cvSOFA, cardiovascular sequential organ failure assessment; VIS, Vasoactive Inotropic Score.

2.3 | Categorising VIS

We determined VIS_{max} and VIS_{mean} as presented in Table 2. The cutoff values for VIS_{max} and VIS_{mean} were determined by visually inspecting a scatterplot of the relationship between mortality and VIS together with the resulting frequencies in the VIS categories. We aimed at cutoff values that (1) are easy to remember and thus practical in clinical practice, (2) result in a rather linear VIS-mortality relationship and (3) result in categories with a comparable distribution of patients compared to other SOFA component scores (i.e., decreasing proportions of patients in more severe organ failure categories).[5] Moreover, we searched for 'natural' cutoffs with cubic spline analysis (Figures S3 and S4).

2.4 | Statistical analysis

For statistical analyses, we used R, Version 4.1.1, Vienna, Austria and IBM SPSS Statistics for Windows, Version 26.0, Armonk, NY: IBM Corp.

We compared differences of each baseline variable in groups of survivors and non-survivors at 30 days with the non-parametric Mann–Whitney test for continuous variables and Fisher's exact test or Pearson chi-square test, as appropriate, for categorical variables.

The increased risk of death with increasing steps in VIS categories or increasing cardiovascular SOFA scores was assessed by age and sex-adjusted binary logistic regression analysis.

With logistic regression, we calculated six different predicted probabilities of death for each patient: We calculated probabilities based on age, sex, and severity of circulatory failure, measured with cvSOFA score, VIS_{max} and VIS_{mean} categories. In addition, we calculated predicted probabilities of death based on age, sex, and severity of organ failures, measured with the conventional total SOFA score and with VIS_{max}- and VIS_{mean}-based alternative SOFA scores. Then, we assessed the AUROC for these predicted probabilities to evaluate the discrimination ability of cvSOFA score, VIS_{max} and VIS_{mean} categories, SOFA score, and alternative VIS_{max}- and VIS_{mean}-based SOFA scores. In the alternative SOFA scores, the cvSOFA was replaced with

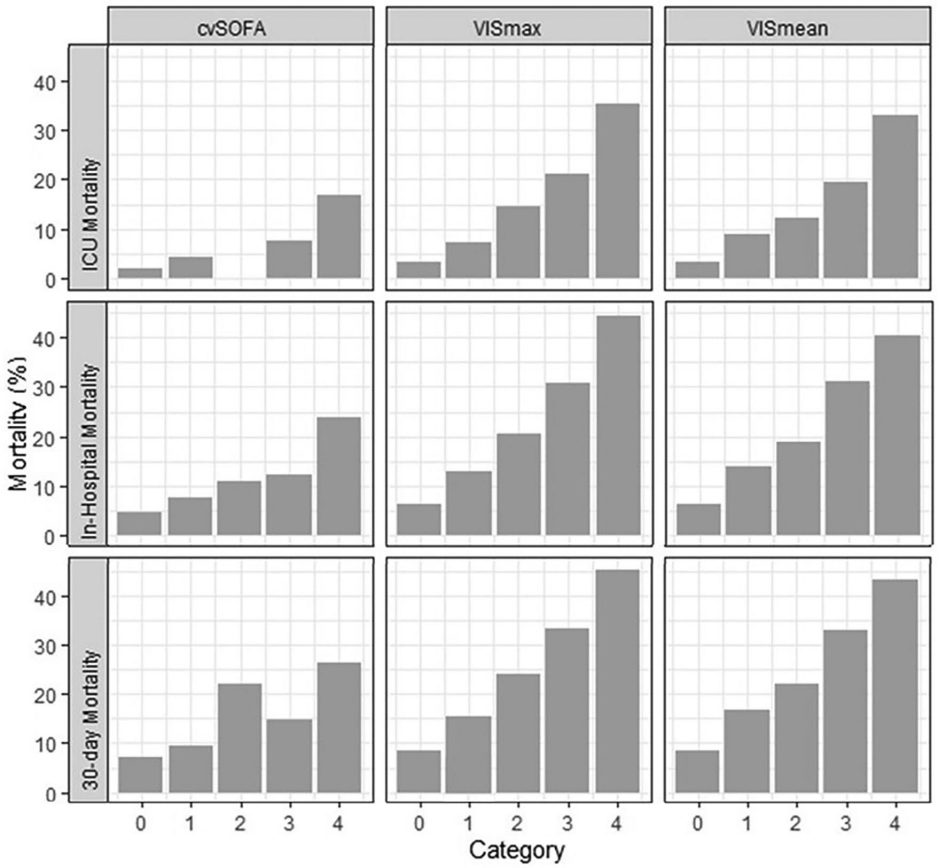

FIGURE 4 The 30-day, ICU and in-hospital mortalities in VIS_{max}, VIS_{mean} and cvSOFA score categories. cvSOFA, cardiovascular sequential organ failure assessment; ICU, intensive care unit; VIS_{max}, maximum Vasoactive Inotropic Score; VIS_{mean}, mean Vasoactive Inotropic Score.

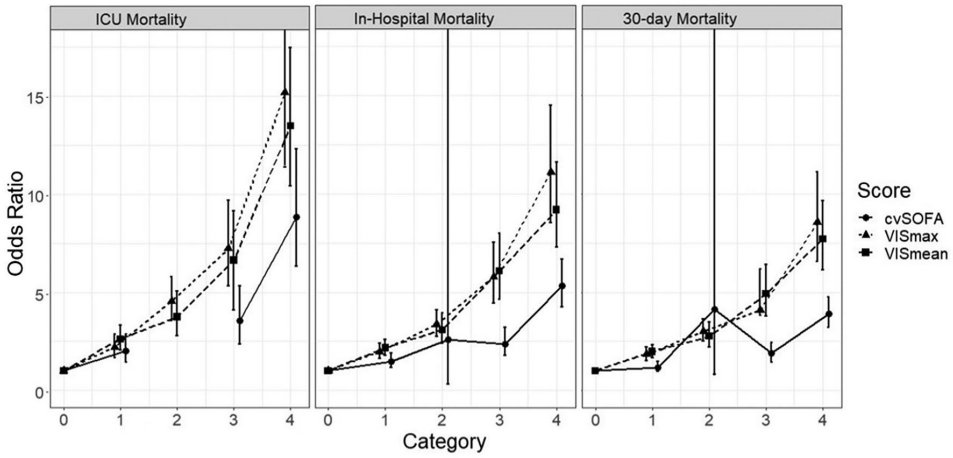

FIGURE 5 The age and sex-adjusted odds ratios for 30-day mortality, ICU mortality and in-hospital mortality for increasing cardiovascular (cv) SOFA and VIS categories. Category 0 of each score was used as reference. The solid line (balls) represents the odds ratios for increasing cvSOFA categories, dotted line (triangles) for increasing VIS_{max} categories and dashed line (squares) for increasing VIS_{mean} categories. The error bars represent 95% confidence intervals. The odds of death were consistently higher for higher VIS categories but not for higher cvSOFA categories. The odds ratio for ICU mortality in cvSOFA category 2 was not calculable because this category contained only eight patients who all survived to ICU discharge. ICU, intensive care unit; SOFA, sequential organ failure assessment; VIS_{max}, maximum Vasoactive Inotropic Score; VIS_{mean}, mean Vasoactive Inotropic Score.

VIS_{max}- or VIS_{mean}-based scores, as follows: patients in the VIS category 0 score 0 points and patients in other categories score points corresponding to each category, the maximum score being 4 points for category 4. The difference between these scoring methods in discrimination was assessed with DeLong test.

P value under 0.05 was considered as statistically significant in all statistical tests.

2.5 | Outcomes

Our primary outcome was a change in AUROC depending on whether we used cvSOFA- or VIS-based scores as the cardiovascular component in SOFA score predicting 30-day mortality. The secondary outcomes were differences in AUROC of these scores predicting ICU and in-hospital mortality.

3 | RESULTS

3.1 | Study population

The number of ICU patients meeting the inclusion criteria was 8079 (Figure 1). The median age was 61 (interquartile range [IQR] 49–70) years, and the majority of the patients (62%) were male. ICU mortality was 7.0%, in-hospital mortality 11.4% and 30-day mortality 13.7%, respectively. The most frequent causes of admission were neurologic non-operative causes (19.4%), including intracranial haemorrhage, subarachnoid haemorrhage, central nervous system infection and seizures, followed by non-operative cardiovascular causes and non-operative respiratory causes (Table 3). The data on ICU, in-hospital and 30-day mortality were complete.

TABLE 4 Area under the receiver operating characteristic curve (AUROC) of original SOFA score and SOFA scores where the cardiovascular component score has been replaced with VIS-based variables. The ability to predict ICU mortality, in-hospital mortality and 30-day mortality was assessed.

Score	AUROC	95% CI
ICU mortality		
VIS_{max}-based SOFA score	0.853	0.834–0.868
VIS_{mean}-based SOFA score	0.846	0.830–0.861
Original SOFA score	0.842	0.827–0.858
In-hospital mortality		
VIS_{max}-based SOFA score	0.826	0.812–0.840
VIS_{mean}-based SOFA score	0.820	0.807–0.834
Original SOFA score	0.818	0.804–0.832
30-day mortality		
VIS_{max}-based SOFA score	0.822	0.809–0.834
VIS_{mean}-based SOFA score	0.816	0.803–0.828
Original SOFA score	0.813	0.800–0.825

Abbreviations: CI, confidence interval; ICU, intensive care unit; SOFA, sequential organ failure assessment; VIS, Vasoactive Inotropic Score; VIS_{max}, maximum VIS; VIS_{mean}, mean VIS.

3.2 | VIS and mortality

During the first 24 h in the ICU, 2809 (34.7%) patients were administered vasopressor or inotropic infusions. Almost all (99.1%) of these patients received norepinephrine. Epinephrine was administered to 4.5%, dobutamine to 10.4%, dopamine to 0.4%, levosimendan to 3.2%, milrinone to 1.9% and vasopressin to 1.8% of patients receiving at least one of these medications.

The median (IQR) for VIS_{max} was 0 (0–0.10) and mean value 0.08, whereas the median for VIS_{mean} was 0 (0–0.008) and mean value 0.0025, respectively, in the overall population. For patients receiving any vasopressor/inotropic infusions, the median for VIS_{max} was 0.16 (0.093–0.28) and the mean value was 0.25, whereas the median for VIS_{mean} was 0.04 (0.013–0.096) and the mean value was 0.072. Both VIS_{max} and VIS_{mean} median values were statistically significantly higher in 30-day non-survivors compared to survivors (VIS_{max}: 0.097 vs. 0, $p < .001$; VIS_{mean} 0.014 vs. 0, $p < .001$) (Table 3). The 30-day mortality increased with increasing VIS values (Figure 2).

TABLE 5 The area under the receiver operating characteristic curve (AUROC) of categorised VIS_{max}, VIS_{mean} and cvSOFA score predicting ICU mortality, in-hospital mortality and 30-day mortality.

Score	AUROC	95% CI
ICU mortality		
VIS_{max}	0.751	0.735–0.768
VIS_{mean}	0.746	0.729–0.762
cvSOFA	0.738	0.722–0.754
In-hospital mortality		
VIS_{max}	0.775	0.754–0.795
VIS_{mean}	0.767	0.747–0.787
cvSOFA	0.759	0.739–0.778
30-day mortality		
VIS_{max}	0.750	0.735–0.765
VIS_{mean}	0.746	0.731–0.746
cvSOFA	0.737	0.722–0.752

Abbreviations: CI, confidence interval; cvSOFA, cardiovascular sequential organ failure assessment; ICU, intensive care unit; VIS_{max}, maximum Vasoactive Inotropic Score; VIS_{mean}, mean Vasoactive Inotropic Score.

3.3 | Mortality in different VIS categories

The proportions of patients in different VIS categories and cvSOFA score categories are presented in Figure 3. Notably, only 0.1% of patients received 2 cvSOFA points. The 30-day, ICU and in-hospital mortalities according to cvSOFA and VIS categories are presented in Figure 4.

The odds for 30-day mortality increased with increasing VIS_{max} and VIS_{mean} categories in age- and sex-adjusted multivariable regression analysis. The increase was consistent for both VIS-based variables. The increase in the risk of death was not consistent with increasing conventional cvSOFA categories. Compared to patients with 0 cvSOFA points, the increase in odds of death was statistically significant for patients with 3 or 4 cvSOFA points (Figure 5). In addition, increasing age and male sex were associated with increased mortality.

3.4 | The predictive value of conventional and VIS-based SOFA scores

The original SOFA score had decent discrimination ability for mortality risk (AUROC 0.813; 95% CI: 0.800–0.825). The accuracy of this score, however, improved as we replaced cvSOFA with VIS_{max}-based scores (AUROC for $SOFA_{VISmax}$ 0.822; 95% CI: 0.809–0.834; difference compared to the original SOFA score 0.009; $p < .001$). Replacing cvSOFA with VIS_{mean}-based criteria also improved the SOFA score's AUROC (AUROC for $SOFA_{VISmean}$ 0.816; 95% CI: 0.803–0.828; difference compared to the original SOFA score 0.003; $p = .004$). The discrimination ability of VIS_{max}-based SOFA was superior to VIS_{mean}-based SOFA ($p < .001$; Table 4).

Both categorised VIS_{max} ($p < .001$) and categorised VIS_{mean} ($p < .001$) outperformed cvSOFA in discrimination ability of mortality risk, as measured by AUROC. The discrimination ability of VIS_{max} was better than that of VIS_{mean} ($p = .02$; Table 5). The ROC curves for the original SOFA score, VIS-based SOFA score, VIS_{max}, VIS_{max} and cvSOFA are described in Supporting Information (Figures S1 and S2).

4 | DISCUSSION

In our retrospective study on 8079 general adult ICU patients, higher VIS scores were associated with higher odds of death, and the ability of VIS to discriminate between 30-day survivors and non-survivors was good. VIS was superior to cvSOFA score in predicting mortality. Replacing the cardiovascular component of the SOFA score with a VIS-based score improved the accuracy of the total SOFA score in predicting the 30-day mortality.

VIS_{max} does not require calculating mean doses. Being more straightforward compared to VIS_{mean}, VIS_{max} is likely easier to use in clinical practice. VIS_{max}-based SOFA score also outperformed the VIS_{mean}-based one in discrimination ability. Therefore, VIS_{max} may fulfil better the original requirements for a SOFA score variable: 'limited number of simple but objective variables that are easily and routinely measured in every institution'.[1]

The criteria of the SOFA score have remained unchanged since its introduction in 1996.[1] Since 1990s, the clinical practice of treating cardiovascular failure has changed remarkably.[10] The administration of norepinephrine, which gives the patients at least 3 cardiovascular points to the SOFA score, has become more routine instead of aggressive resuscitation with fluids.[22,23] The cardiovascular component of SOFA may no longer adequately reflect the severity of cardiovascular failure and its usefulness has been questioned in the ongoing debate over the need to update the SOFA score.[4,5,24]

Our study confirms that a VIS-based cardiovascular score is a better alternative to the conventional cardiovascular component of the SOFA score. According to our results, either VIS_{max} or VIS_{mean} could be used for this purpose. Since VIS_{max} requires less calculation, it would be a simpler choice. An additional advantage of the VIS-based SOFA score would be the adjustability in the case of introduction of new vasopressor/inotropic agents to clinical practice in the future: in that case the whole scoring does not need to be changed, but only a conversion factor for the new agent needs to be determined.[25]

Mean arterial pressure (MAP) below or above 70 mmHg is the current distinctive criterion between 0 and 1 cvSOFA points. However, brief episodes of MAP below 60–70 mmHg do not significantly increase the risk of death in intensive care patients, whereas mortality increases in patients with MAP below 55 mmHg, which often triggers the initiation of vasopressor/inotropic agents in clinical practice.[26–28] Hence, it would make sense to score 0 cardiovascular SOFA points to patients who do not receive any vasopressor/inotropic agents and to use the scale of 1–4 points to categorise the magnitude of vasopressor/inotropic medication needed. In our study, one third of patients received any vasopressor/inotropic support which is in line with previous studies done in ICU patients without preceding cardiac surgery.[29,30]

Previously, several proposals have been made to replace cvSOFA with an alternative score. Vacheron et al. proposed a cvSOFA component score based on cumulative dosages of vasopressor/inotropic agents (limited to norepinephrine, epinephrine and dopamine) as part of a suggested full renewal of the SOFA score.[31] The discrimination ability of their fully renewed SOFA score proposal was significantly better compared to the conventional SOFA score in prediction of 28-day mortality. They used only three categories for each of the six organ systems included in SOFA, but their findings are in line with ours, supporting the feasibility of an approach based on vasopressor/inotropic dosages in generating an updated version of the cvSOFA.

Yadav et al.[32] proposed a modified cvSOFA based on Shock Index, lactate measures and infusion rates of dopamine, epinephrine and norepinephrine. In their observational cohort study, they found a statistically significant difference in discrimination between conventional SOFA score and modified SOFA score with substituted cvSOFA as predictors of 28-day mortality (AUROC 0.822 vs. 0.836; difference = 0.014). However, the scoring should be ideally based on a simple measurement instead of a mix of laboratory results, physiologic findings, and drug infusion dosages. In a retrospective study by Bosch et al.,[33] the ratio of norepinephrine equivalent doses (NEQ) and MAP was associated with in-hospital mortality. The patients were given 0–2 points according to the NEQ/MAP ratio. The NEQ/MAP was more accurate than the conventional cvSOFA score and the modified cvSOFA score by Yadav et al.[32] in predicting in-hospital mortality. In this study, the scoring had, however, three categories and it was not implementable in the full SOFA score.

Previous studies on VIS have focused on paediatric or neonatal patients or on cardiac surgical patients.[11,12,14–20] Our findings confirm the good predictive value of VIS in mixed ICU population excluding the post-cardiac surgery and elective surgery patients. In our study population, no more than one third of the patients needed cardiovascular support with infusion of intravenous vasopressor/inotropic medications during the first 24 h. This is far less than in patients with preceding cardiac surgery. However, our results are comparable to those of Koponen et al.,[19] who studied cardiac surgical patients: VIS is associated with outcome both in general ICU patients and in patients admitted after cardiac surgery.

4.1 | Strengths and limitations

Our study population was a heterogenic group of critically ill adult patients treated in a mixed medical-surgical ICU. This increases the generalisability of the results. The data on administered infusions and 30-day mortality were complete.

The main limitation is that the study was a single-centre study. Differences in local practices in administering vasopressor/inotropic agents in different countries and healthcare systems may affect the VIS values. Some of the medications included in VIS in this study have not been included in all previous VIS studies. Levosimendan has been part of VIS only in studies on adult patients and the coefficient used for its dose is disputable.[13,19] However, the controversies of conversion factors concern not only levosimendan but also other vasopressor/inotropic agents to some degree.[34]

For VIS_{max}, we were unable to track the highest concomitant infusion rates of the vasopressors and inotropes. Instead, we used the highest infusion rates of each drug during the first 24 h at the ICU. The use of concomitant rates might be a more accurate measure for cardiovascular dysfunction but less useful in clinical practice: it would require calculations from multiple time points to determine where the highest VIS score is reached. Moreover, our study did not take into account possible mechanical cardiovascular support (e.g. veno-arterial extracorporeal membrane oxygenation or intra-aortic balloon pump). However, during the study period (2013–2019), these interventions were seldom used in our hospital in other patient groups than post-cardiac surgical patients. The increasing use of mechanical cardiovascular support devices must be taken into consideration if an update of the cardiovascular SOFA score is planned in the future.[24]

One major limitation is that the observation is limited to the first 24 h after ICU admission. Future studies should focus on the dynamic nature of SOFA score and observe how daily increase/decrease in VIS-based cvSOFA is associated with outcomes.

To the best of our knowledge, this study is the first one to evaluate the association of increasing VIS with mortality in a general ICU population. It is obvious that more studies validating the usefulness of VIS are needed before it can be considered for use in clinical practice.

5 | CONCLUSION

In contemporary intensive care, the cardiovascular component of the SOFA score is no longer an optimal metric of cardiovascular failure.

VIS-based variables reflect the severity of cardiovascular failure, and the predictive accuracy of the total SOFA score improves when the original cardiovascular component is replaced by VIS_{max} or VIS_{mean}.

AUTHOR CONTRIBUTIONS

Anssi Pölkki: Methodology; software; validation; formal analysis; investigation; resources; data curation; writing—original draft; visualisation; project administration. **Pirkka T. Pekkarinen**: Methodology; writing—review and editing. **Pasi Lahtinen**: Conceptualization; methodology; writing—review and editing. **Timo Koponen**: Methodology; writing—review and editing. **Matti Reinikainen**: Conceptualization; methodology; writing – review & editing; supervision.

ACKNOWLEDGMENTS

The authors thank Petteri Kisonen for help with data extraction from the electronic clinical information management system. Dr Pölkki was supported by institutional funding from Kuopio University Hospital and University of Eastern Finland. Dr Pekkarinen was supported by grants from Finska Läkaresällskapet and Einar och Karin Stroems stiftelse. The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data set analysed during the current study includes sensitive patient information. Legal restrictions prohibit us from making the data publicly available.

ORCID

Anssi Pölkki https://orcid.org/0000-0001-8777-5287
Pirkka T. Pekkarinen https://orcid.org/0000-0002-8203-3409
Pasi Lahtinen https://orcid.org/0000-0001-5062-184X
Timo Koponen https://orcid.org/0000-0001-8279-2643
Matti Reinikainen https://orcid.org/0000-0001-6878-3740

REFERENCES

1. Vincent JL, Moreno R, Takala J, et al. 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;22:707-710. doi:10.1007/BF01709751
2. Reintam Blaser A, Regli A, De Keulenaer B, et al. Incidence, risk factors, and outcomes of intra-abdominal hypertension in critically ill patients-a prospective multicenter study (IROI Study). *Crit Care Med*. 2019;47:535-542. doi:10.1097/CCM.0000000000003623
3. Reintam Blaser A, Padar M, Mändul M, et al. Development of the gastrointestinal dysfunction score (GIDS) for critically ill patients—a prospective multicenter observational study (iSOFA study). *Clin Nutr*. 2021;40:4932-4940. doi:10.1016/j.clnu.2021.07.015
4. Bachmann KF, Arabi YM, Regli A, Starkopf J, Reintamm BA. Cardiovascular SOFA score may not reflect current practice. *Intensive Care Med*. 2022;48:119-120. doi:10.1007/s00134-021-06536-6
5. Pölkki A, Pekkarinen PT, Takala J, Selander T, Reinikainen M. Association of Sequential Organ Failure Assessment (SOFA) components with mortality. *Acta Anaesthesiol Scand*. 2022;66:731-741. doi:10.1111/aas.14067
6. Varpula M, Tallgren M, Saukkonen K, Voipio-Pulkki LM, Pettila V. Hemodynamic variables related to outcome in septic shock. *Intensive Care Med*. 2005;31:1066-1071. doi:10.1007/s00134-005-2688-z
7. Vincent JL, De Backer D. Circulatory shock. *N Engl J Med*. 2013;369:1726-1734. doi:10.1056/NEJMra1208943
8. Marchick MR, Kline JA, Jones AE. The significance of non-sustained hypotension in emergency department patients with sepsis. *Intensive Care Med*. 2009;35:1261-1264. doi:10.1007/s00134-009-1448-x
9. Uhlig K, Efremov L, Tongers J, et al. Inotropic agents and vasodilator strategies for the treatment of cardiogenic shock or low cardiac output syndrome. *Cochrane Database Syst Rev*. 2020;11:CD009669. doi:10.1002/14651858.CD009669.pub4
10. Jentzer JC, Coons JC, Link CB, Schmidhofer M. Pharmacotherapy update on the use of vasopressors and inotropes in the intensive care unit. *J Cardiovasc Pharmacol Ther*. 2015;20:249-260. doi:10.1177/1074248414559838
11. Wernovsky G, Wypij D, Jonas RA, et al. Postoperative course and hemodynamic profile after the arterial switch operation in neonates and infants. A comparison of low-flow cardiopulmonary bypass and circulatory arrest. *Circulation*. 1995;92:2226-2235. doi:10.1161/01.cir.92.8.2226
12. Gaies MG, Gurney JG, Yen AH, et al. Vasoactive-inotropic score as a predictor of morbidity and mortality in infants after cardiopulmonary bypass. *Pediatr Crit Care Med*. 2010;11:234-238. doi:10.1097/PCC.0b013e3181b806fc
13. Favia I, Vitale V, Ricci Z. The vasoactive-inotropic score and levosimendan: time for LVIS? *J Cardiothorac Vasc Anesth*. 2013;27:e15-e16. doi:10.1053/j.jvca.2012.11.009
14. Davidson J, Tong S, Hancock H, Hauck A, da Cruz E, Kaufman J. Prospective validation of the vasoactive-inotropic score and correlation to short-term outcomes in neonates and infants after cardiothoracic surgery. *Intensive Care Med*. 2012;38:1184-1190. doi:10.1007/s00134-012-2544-x
15. Sanil Y, Aggarwal S. Vasoactive-inotropic score after pediatric heart transplant: a marker of adverse outcome. *Pediatr Transplant*. 2013;17:567-572. doi:10.1111/petr.12112
16. Gaies MG, Jeffries HE, Niebler RA, et al. Vasoactive-inotropic score is associated with outcome after infant cardiac surgery: an analysis from the Pediatric Cardiac Critical Care Consortium and Virtual PICU System Registries. *Pediatr Crit Care Med*. 2014;15:529-537. doi:10.1097/PCC.0000000000000153
17. Haque A, Siddiqui NR, Munir O, Saleem S, Mian A. Association between vasoactive-inotropic score and mortality in pediatric septic shock. *Indian Pediatr*. 2015;52:311-313. doi:10.1007/s13312-015-0630-1
18. McIntosh AM, Tong S, Deakyne SJ, Davidson JA, Scott HF. Validation of the vasoactive-inotropic score in pediatric sepsis. *Pediatr Crit Care Med*. 2017;18:750-757. doi:10.1097/PCC.0000000000001191
19. Koponen T, Karttunen J, Musialowicz T, Pietiläinen L, Uusaro A, Lahtinen P. Vasoactive-inotropic score and the prediction of morbidity and mortality after cardiac surgery. *Br J Anaesth*. 2019;122:428-436. doi:10.1016/j.bja.2018.12.019
20. Yamazaki Y, Oba K, Matsui Y, Morimoto Y. Vasoactive-inotropic score as a predictor of morbidity and mortality in adults after cardiac surgery with cardiopulmonary bypass. *J Anesth*. 2018;32:167-173. doi:10.1007/s00540-018-2447-2
21. Reinikainen M, Mussalo P, Hovilehto S, et al. Finnish Intensive Care Consortium, Association of automated data collection and data completeness with outcomes of intensive care. A new customised model for outcome prediction. *Acta Anaesthesiol Scand*. 2012;56:1114-1122. doi:10.1111/j.1399-6576.2012.02669.x
22. Marik PE, Linde-Zwirble WT, Bittner EA, Sahatjian J, Hansell D. Fluid administration in severe sepsis and septic shock, patterns and outcomes: an analysis of a large national database. *Intensive Care Med*. 2017;43:625-632. doi:10.1007/s00134-016-4675-y
23. Pittard MG, Huang SJ, McLean AS, Orde SR. Association of positive fluid balance and mortality in sepsis and septic shock in an Australian

cohort. *Anaesth Intensive Care*. 2017;45:737-743. doi:10.1177/0310057X1704500614
24. Moreno R, Rhodes A, Piquilloud L, et al. The Sequential Organ Failure Assessment (SOFA) Score: has the time come for an update? *Crit Care*. 2023;27:15. doi:10.1186/s13054-022-04290-9
25. Belletti A, Lerose CC, Zangrillo A, Landoni G. Vasoactive-inotropic score: evolution, clinical utility, and pitfalls. *J Cardiothorac Vasc Anesth*. 2021;35:3067-3077. doi:10.1053/j.jvca.2020.09.117
26. Lamontagne F, Richards-Belle A, Thomas K, et al. Effect of reduced exposure to vasopressors on 90-day mortality in older critically ill patients with vasodilatory hypotension: a randomized clinical trial. *JAMA*. 2020;323:938-949. doi:10.1001/jama.2020.0930
27. Vincent JL, Nielsen ND, Shapiro NI, et al. Mean arterial pressure and mortality in patients with distributive shock: a retrospective analysis of the MIMIC-III database. *Ann Intensive Care*. 2018;8:107. doi:10.1186/s13613-018-0448-9
28. Lamontagne F, Meade MO, Hébert PC, et al. Higher versus lower blood pressure targets for vasopressor therapy in shock: a multicentre pilot randomized controlled trial. *Intensive Care Med*. 2016;42:542-550. doi:10.1007/s00134-016-4237-3
29. Hunter S, Manias E, Hirth S, Considine J. Intensive care patients receiving vasoactive medications: a retrospective cohort study. *Aust Crit Care*. 2022;35:499-505. doi:10.1016/j.aucc.2021.07.003
30. Sakr Y, Reinhart K, Vincent J-L, et al. Does dopamine administration in shock influence outcome? Results of the Sepsis Occurrence in Acutely Ill Patients (SOAP) Study. *Crit Care Med*. 2006;34:589-597. doi:10.1097/01.CCM.0000201896.45809.E3
31. Vacheron CH, Friggeri A, Iwaz J, et al. A new simplified and accurate sa-SOFA score. *J Crit Care*. 2020;57:240-245. doi:10.1016/j.jcrc.2019.11.009
32. Yadav H, Harrison AM, Hanson AC, Gajic O, Kor DJ, Cartin-Ceba R. Improving the accuracy of cardiovascular component of the sequential organ failure assessment score. *Crit Care Med*. 2015;43:1449-1457. doi:10.1097/CCM.0000000000000929
33. Bosch NA, Teja B, Wunsch H, Walkey AJ. Characterization and validation of a novel measure of septic shock severity. *Intensive Care Med*. 2020;46:135-137. doi:10.1007/s00134-019-05837-1
34. Kotani Y, Di Gioia A, Landoni G, et al. An updated "norepinephrine equivalent" score in intensive care as a marker of shock severity. *Crit Care*. 2023;27:29. doi:10.1186/s13054-023-04322-y

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Pölkki A, Pekkarinen PT, Lahtinen P, Koponen T, Reinikainen M. Vasoactive Inotropic Score compared to the sequential organ failure assessment cardiovascular score in intensive care. *Acta Anaesthesiol Scand*. 2023;67(9):1219-1228. doi:10.1111/aas.14287

III

Noradrenaline dose cutoffs to characterise the severity of cardiovascular failure: data-based development and external validation

Pölkki A, Pekkarinen PT, Hess B, Reintam Blaser A, Bachmann KF, Lakbar I, Hollenberg SM, Lobo SM, Rezende E, Selander T, Reinikainen M

Acta Anaesthesiol Scand.
In Press

IV

The Influence of Potential Organ Donors on Standardized Mortality Ratios and ICU Benchmarking

Pölkki A & Moser A, Raj R, Takala J, Bendel S, Reinikainen M & Jakob S

Crit Care Med. 2024;52:387-395

The Influence of Potential Organ Donors on Standardized Mortality Ratios and ICU Benchmarking

Pölkki, Anssi MD[1,2]; Moser, André PhD[3]; Raj, Rahul MD, PhD[4]; Takala, Jukka MD, PhD[5]; Bendel, Stepani MD, PhD[1]; Jakob, Stephan M. MD, PhD[5]; Reinikainen, Matti MD, PhD[1,2]

[1]Department of Anaesthesiology and Intensive Care, Kuopio University Hospital, Kuopio, Finland, [2]University of Eastern Finland, Kuopio, Finland [3]CTU Bern, University of Bern, Bern, Switzerland, [4]Department of Neurosurgery, Helsinki University Hospital and University of Helsinki, Helsinki, Finland, [5]University of Bern, Bern, Switzerland.

Critical Care Medicine 52(3):p 387-395, March 2024. | *DOI:* 10.1097/CCM.0000000000006098

Abstract

OBJECTIVES:

The standardized mortality ratio (SMR) is a common metric to benchmark ICUs. However, SMR may be artificially distorted by the admission of potential organ donors (POD), who have nearly 100% mortality, although risk prediction models may not identify them as high-risk patients. We aimed to evaluate the impact of PODs on SMR.

DESIGN:

Retrospective registry-based multicenter study.

SETTING:

Twenty ICUs in Finland, Estonia, and Switzerland in 2015–2017.

PATIENTS:

Sixty thousand forty-seven ICU patients.

INTERVENTIONS:

None.

MEASUREMENTS AND MAIN RESULTS:

We used a previously validated mortality risk model to calculate the SMRs. We investigated the impact of PODs on the overall SMR, individual ICU SMR and ICU benchmarking. Of the 60,047 patients admitted to the ICUs, 514 (0.9%) were PODs, and 477 (93%) of them died. POD deaths accounted for 7% of the total 6738 in-hospital deaths. POD admission rates varied from 0.5 to 18.3 per 1000 admissions across ICUs. The risk prediction model predicted a 39% in-hospital mortality for PODs, but the observed mortality was 93%. The ratio of the SMR of the cohort without PODs to the SMR of the cohort with PODs was 0.96 (95% CI, 0.93–0.99). Benchmarking results changed in 70% of ICUs after excluding PODs.

CONCLUSIONS:

Despite their relatively small overall number, PODs make up a large proportion of ICU patients who die. PODs cause bias in SMRs and in ICU benchmarking. We suggest excluding PODs when benchmarking ICUs with SMR.

KEY POINTS

Question: Does admission of potential organ donors (PODs) to ICUs affect standardized mortality ratios (SMRs)?

Findings: In this retrospective study on 60,047 ICU patients from three countries, PODs made up 0.9% of all ICU admissions (range across ICUs, 0.05–1.8%), but accounted for 7% of all in-hospital deaths. PODs had a much higher observed than risk model-predicted mortality (93% vs. 39%), and therefore they increased the SMRs. SMR-based benchmarking results changed for 70% of ICUs after PODs were excluded.

Meaning: PODs cause bias in SMR calculations and ICU benchmarking.

INTRODUCTION

Severity- and case-mix-adjusted mortality prediction models allow for the calculation of the standardized mortality ratio (SMR). The SMR, which represents the ratio of observed to expected mortality, is an important component of quality benchmarking of ICUs and is routinely applied by many ICU registries[1]. The SMR enables comparisons of the performance of ICUs with different case mixes. Many risk prediction models have also been modified or recalibrated to improve their performance in national or regional registries[2-12].

Limiting treatment upon ICU admission is associated with an increased risk of death ([13]). Patients admitted for evaluation as potential organ donors (PODs) represent an extreme treatment limitation: death is anticipated and accepted, and the goal of ICU admission of the POD is not to save the patient but to protect organs for possible donation[14,15]. Accordingly, the expected mortality should approach 100%. However, because PODs often have no other major organ dysfunctions besides severe brain injury, risk prediction models may give them erroneously low probabilities of death. In fact, most mortality prediction models do not address this concern.

The impact of POD admissions on the SMRs of ICUs and their implications on benchmarking results are unknown. We recently published a risk prediction model that excludes PODs in predicting the risk of death in ICU patients[6].

STUDY AIMS

The aim of this study was to assess how the inclusion of PODs impacts the overall SMR and the SMRs of individual ICUs. Additionally, we investigated whether the inclusion or exclusion of PODs affected the benchmark rankings of the ICUs.

MATERIALS AND METHODS

Data Extraction, Patient Selection, and Exclusion Criteria

In this secondary analysis, we used the SMR study population described by Takala et al[16]. In brief, Takala et al[16] used data from the Finnish Intensive Care Consortium (FICC) database, encompassing 168,108 admissions between 2008 and 2017. Data regarding possible treatment limitations upon ICU admission are recorded in the database. Since 2015, these recordings have captured

patients who are admitted for the sole purpose of possible organ donation. Thus, data on PODs were available for 2015–2017. Therefore, we restricted the SMR study population from Takala et al[16] to 2015–2017, which yielded a total of 60,047 patients from 20 ICUs in three nations—Finland (18 ICUs), Estonia (one ICU), and Switzerland (one ICU) (**eFig. 1**, for flowchart).

Ethical Considerations

The data management plan, database contents, and study process were approved by the Finnish Institute for Health and Welfare (THL/1524/5.05.00/2017; THL/1173/05/00/2018; THL/3795/14.06.00/2021). According to regulations in Finland, Estonia, and Switzerland, no ethics committee approval was needed.

Identification of Potential Organ Donors

We identified PODs if a recording of "admission because of possible organ donation" was registered at the time of ICU admission.

Calculation of the Standardized Mortality Ratio

The SMR was defined as the number of observed deaths divided by the number of predicted deaths. We used the model described by Moser et al ([6]) to calculate the predicted mortality risk. This model was based on age, a modified Simplified Acute Physiology Score (SAPS) II score ([17]) (excluding age and admission type), admission type (elective vs. emergency and surgical vs. nonsurgical admission), and premorbid functional status determined using a modified Eastern Cooperative Oncology Group (ECOG) classification ([18]). Importantly, in the model creation and validation, PODs were excluded. In this study, we estimated the effect of PODs on SMR by calculating the SMR in the study population with and without PODs.

Statistical Methods

We report frequencies (*n*), percentages (%), median values, and interquartile ranges. For group differences between cohorts of PODs and admissions for other causes, we report *p* values using a chi-square or Wilcoxon rank-sum test. We calculated SMRs for the overall cohort and each ICU. We calculated the ratio of the SMR of the cohort without PODs to the SMR of the cohort with PODs (with 95% CIs) using multivariable Poisson regression models with a cohort-specific indicator adjusted for calendar year for the overall cohort and separately for each ICU. To assess the impact of case-mix, we modeled the

overall SMR in a model with indicators for the two cohorts and hospital typology. First, we tested for an interaction effect between the two predictors. In case of a nonsignificant interaction effect, we model the two predictors additively. All *p* values were two-sided, and *p* values smaller than 0.05 were considered statistically significant.

For the statistical analyses, we used R Version 4.1.2 (R Core Team, Vienna, Austria).

RESULTS

Study Population

We included 60,047 patients from 20 ICUs: eight university ICUs, six large nonuniversity ICUs, and six small nonuniversity ICUs. Totally 514 patients (0.85%) were admitted as PODs (0.5–18.3 per 1000 admissions across the ICUs). The frequencies of PODs admitted to each ICU for each study year are illustrated in **eFigure 2**. The overall in-hospital mortality for all years was 6,738 of 60,047 (11.2%). The etiology of the brain damage of the PODs was intracerebral hemorrhage (ICH) in 44%, trauma in 22%, subarachnoid hemorrhage in 15%, hypoxemic brain injury in 5%, ischemic stroke in 4%, and miscellaneous etiology in 10% of the cases.

After excluding PODs, the predicted number of deaths was 6324 and the actual number of deaths was 6261, indicating 63 fewer deaths were observed than predicted. However, when the PODs were included, the predicted number of deaths was 6479 but the observed number of deaths was 6738, resulting in 259 more deaths observed than predicted. The PODs had a predicted in-hospital mortality risk of 39% but the observed mortality was 93%. The deaths of PODs accounted for 7% of all deaths in the study population during the hospital stay.

We found no association between the frequency of POD admissions and calendar year ($p = 0.60$) or hospital size (small nonuniversity hospital, large nonuniversity hospital, or university hospital) ($p = 0.44$). Furthermore, there was no statistically significant difference in the median age between POD patients and non-POD patients ($p = 0.05$; Table 1).

TABLE 1. Baseline Characteristics and Hospital Mortality

Characteristic	Admission Cause Other Than Being	Potential Organ Donors, $n = 514$[a]	p [b]

	Potential Organ Donor, n = 59,533[a]		
Year			0.6
2015	19,321 (32%)	157 (31%)	
2016	20,034 (34%)	182 (35%)	
2017	20,178 (34%)	175 (34%)	
ICU class			0.4
Nonuniversity (large)	9,547 (16%)	77 (15%)	
Nonuniversity (small)	5,981 (10%)	60 (12%)	
University	44,005 (74%)	377 (73%)	
Age	63 (49–73)	65 (54–72)	0.053
Operative	23,486 (39%)	24 (4.7%)	< 0.001
Emergency	47,436 (80%)	513 (100%)	< 0.001
Simplified Acute Physiology Score II score	15 (7–27)	43 (39–51)	< 0.001
Hospital mortality	6,261 (11%)	477 (93%)	< 0.001
Acute Physiology and Chronic Health Evaluation-III diagnosis group			
Nonoperative: Cardiovascular	8,599 (14%)	20 (3.9%)	
Nonoperative: Respiratory	6,012 (10%)	2 (0.4%)	
Nonoperative: Gastrointestinal	3,150 (5.3%)	0 (0%)	
Nonoperative: Neurologic	9,664 (16%)	346 (67%)	
Nonoperative: Trauma	3,708 (6.2%)	104 (20%)	
Nonoperative: Metabolic	2,392 (4.0%)	0 (0%)	
Nonoperative: Hematologic diseases	214 (0.4%)	0 (0%)	
Nonoperative: Renal	700 (1.2%)	0 (0%)	
Nonoperative: Other	1,749 (2.9%)	16 (3.1%)	
Operative: Cardiovascular	4,606 (7.7%)	2 (0.4%)	
Operative: Respiratory	1,685 (2.8%)	0 (0%)	
Operative: Gastrointestinal	6,730 (11%)	0 (0%)	
Operative: Neurologic	7,261 (12%)	15 (2.9%)	
Operative: Trauma	1,421 (2.4%)	9 (1.8%)	
Operative: Urology/gynecology	1,101 (1.8%)	0 (0%)	
Operative: Other	541 (0.9%)	0 (0%)	

[a] n (%), median (interquartile range).

[b] Pearson's χ^2 test.

"Operative" means being admitted to the ICU from operation theater after surgery; "Emergency" means an unscheduled ICU admission for an acute reason.

SMRs in Cohorts With and Without PODs

The SMR without PODs was 0.99 (95% CI, 0.97–1.02) but it increased to 1.04 (95% CI, 1.01–1.06) when the PODs were included. The ratio of the SMR in the cohort without PODs to the SMR in the cohort with PODs was 0.96 (95% CI, 0.93–0.99). We found no evidence for an interaction effect between the two cohorts and hospital typology ($p = 0.89$). In an additive model without an interaction effect, hospital typology was strongly associated with a change in SMR ($p < 0.001$). Small nonuniversity hospitals showed a ratio of SMR of 1.14 (95% CI, 1.08–1.20), compared with university hospitals (**eTables 1-3**). The adjusted ratio of the SMR in the population without PODs to the SMR in the population with PODs was 0.96 (95% CI, 0.93–0.99). Calendar year-adjusted ratios comparing the SMRs of the cohort without PODs and the cohort with PODs in individual ICUs ranged from 0.92 to 1.00 (**Fig. 1**). The annual SMRs for each ICU in the cohort with and without PODs are presented in the **eFigure 3**. SMRs ranged from 0.89 to 1.51 in the cohort with PODs and 0.86 to 1.47 in the cohort without PODs. The impact of POD exclusion on the individual ICU level is illustrated in **Figure 2**.

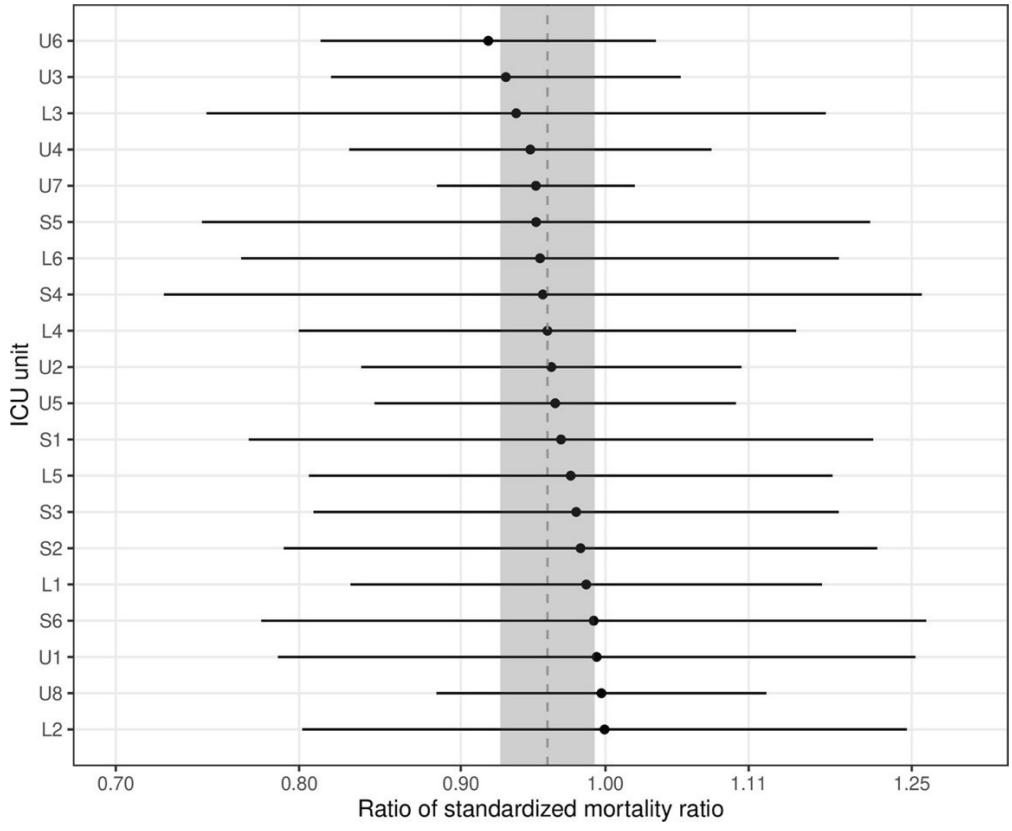

Figure 1. Ratios of standardized mortality ratios (SMRs) comparing the cohort without potential organ donors (PODs) to the cohort with PODs (SMRPOD excluded/SMRPOD included) in each ICU during the entire study period. The error bars represent the 95% CIs. The vertical dashed line represents the average ratio in the whole study population (0.96). The dark gray area represents the 95% CI (0.93–0.99).

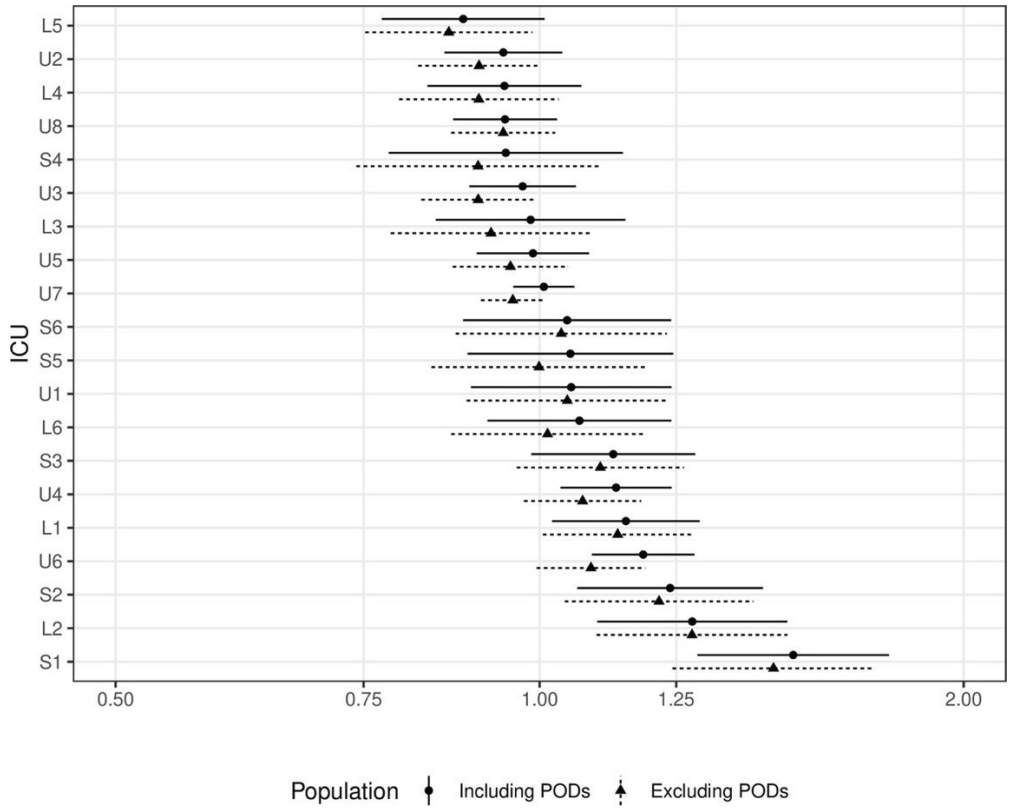

Figure 2. Impact of exclusion of potential organ donors (PODs). The filled circles represent the standardized mortality ratios (SMRs) of each ICU during the whole study period with PODs included. The triangles represent the SMRs of the ICUs with PODs excluded. The error bars represent the 95% CIs. The ICUs listed on the y-axis are arranged by increasing SMRs with PODs included. L1–L6 represent the ICUs of large nonuniversity hospitals, S1–S6 those of small nonuniversity hospitals, and U1–U8 those of university hospitals.

Alterations to Benchmark Rankings

Including PODs affected ICUs' benchmarking rankings. Rankings were altered in 70% (14/20) of the ICUs by exclusion vs. inclusion of PODs (**Fig. 3**).

There was no difference in mean ranking change between ICUs of large nonuniversity hospitals, small nonuniversity hospitals, and university hospitals (p = 0.45). There was a weak trend toward the improved ranking of the ICUs admitting more PODs after excluding the PODs from the whole study population (**eFig. 4**).

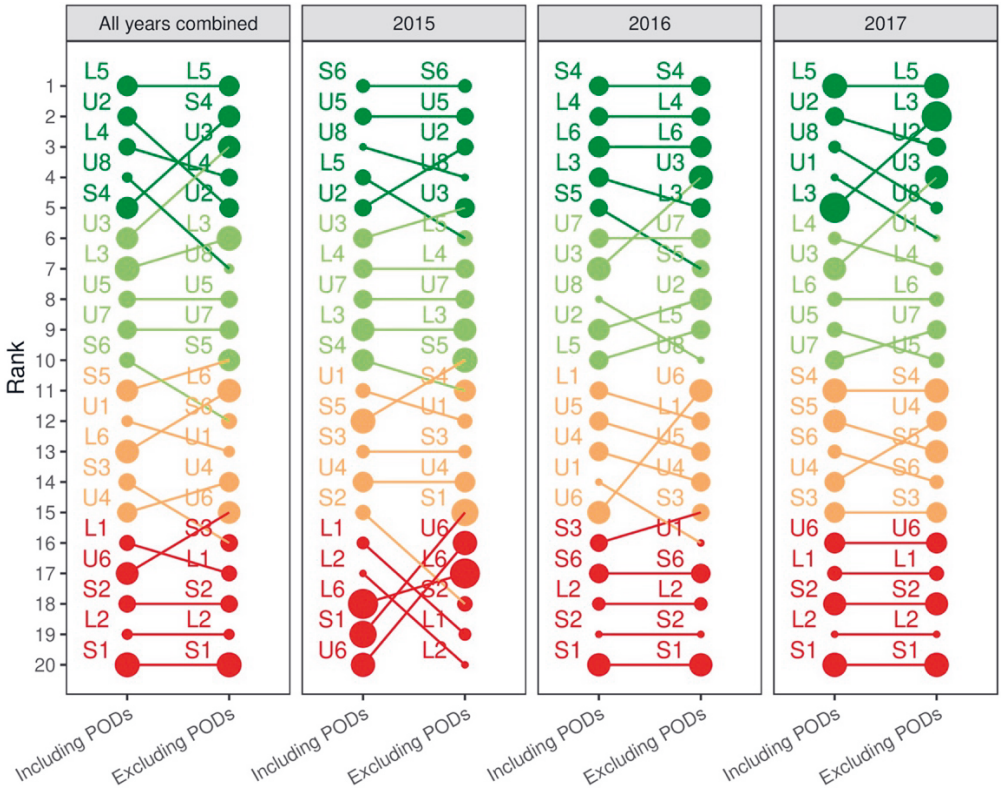

Figure 3. Alterations to standardized mortality ratio-based ranking of ICUs caused by excluding the potential organ donors (PODs) during the whole study period (first panel, left), and alterations during each study year separately (second to fourth panel). The size of the symbol indicates the proportion of PODs of all admissions in the ICU.

DISCUSSION

In this registry-based study on 60,047 ICU patients, PODs accounted for 0.9% of all patients but 7% of all in-hospital deaths of ICU patients. PODs had a statistically significant impact on SMRs: excluding PODs decreased the SMR in the whole population. The effect was consistent over the 3 study years and ICU categories. SMR-based ranking positions changed for 70% of the ICUs after POD exclusion.

If PODs are carefully selected, their in-hospital mortality will be close to 100%. This was the case in our study, whereas the predicted risk of death was substantially lower. As a result, this patient group has an erroneously high SMR. This discrepancy explains the higher SMR in the whole cohort if PODs are not excluded. Because the common risk prediction models do not detect the true expected high risk of death in PODs, we propose that PODs should be excluded (and analyzed separately) when performing ICU benchmarking.

In benchmarking, both absolute performance and performance with respect to the other members of the consortium are important. Our risk prediction model excludes PODs due to their potential SMR confounding effects. This effect was clearly demonstrated in our study. Including PODs increased the SMR in the overall patient population, but for individual ICUs, ranking positions could change in either direction, depending on case-mix and POD admission frequency. This was caused by different magnitudes of the effect of PODs in different ICUs. Although all ICUs in our benchmark consortium treated PODs, the rate varied between one and 18 of 1000 admissions, with no differences between the three groups of hospitals. If POD treatment had been centralized in specific centers, the impact might have been much larger in these ICUs.

The current mortality prediction model is based on data from 2015 to 2017. Mortality outcomes tend to improve over time, and it is inevitable that the model needs to be recalibrated in the future. In addition to the FICC benchmarking program, the prediction model used by the Intensive Care National Audit & Research Centre takes PODs into account, by excluding PODs from the model[8].

In the prediction model used in the current study, the only measure of neurologic condition is the Glasgow Coma Scale (GCS) score, which is included in the SAPS II score. Although the GCS score is relatively highly weighted in SAPS II, it does not alone capture the dismal prognosis associated with POD. GCS is known to be prone to interobserver variability[19]. To improve the accuracy of neurologic evaluation in predictive models, it may be valuable to incorporate more objective variables, such as pupil reactivity and CT scan findings. There are several disease-specific prediction models for critically ill neurologic ICU patients, such as the International Mission for Prognosis and Analysis of Clinical Trials for traumatic brain injury patients, the Subarachnoid Hemorrhage International Trialists model for aneurysmal subarachnoid hemorrhage, and the ICH score for ICH patients[20–22]. However, it is unlikely that incorporating these scores would eliminate the need for a more accurate identification of PODs for benchmarking purposes.

ICU ranking lists based on SMR should be interpreted with caution. League table rankings contain uncertainty, and random variation is high[23–25]. For example, in 16 cardiothoracic centers in the Netherlands, ranking lists demonstrated considerable reordering during 3 consecutive years, but with very wide 95% CIs of adjusted mortalities[25]. We also found wide and overlapping CIs in the SMRs.

The prognostic scores used in benchmarking are best suited to comparing and interpreting the risk-adjusted outcomes of patient groups (external benchmarking)[26].

According to the European Society of Intensive Care Medicine, monitoring, reporting, and analyzing SMRs is a useful method for improving the quality and safety of intensive care[27]. The effect of PODs is neglected in the prediction models of most benchmarking programs[9–12]. The influence of PODs on performance quality benchmarking has not been previously investigated. The much higher observed than predicted mortality in PODs is plausible in prediction models with high weight on physiologic abnormalities. The impact of exclusion vs. inclusion of PODs on the overall mortality and the SMRs was confirmed in our study. An alternative to excluding the PODs from SMR calculations would be to create prediction models giving high expected mortality to PODs.

The need for transplantable organs is increasing worldwide[28]. Compared with the patients admitted for other causes, the goal of the treatment of PODs is utterly different and their probability of in-hospital survival is extremely low. Therefore, including PODs in SMR calculations can result in wrong interpretations of an ICU's performance.

A strength of our study is that data in the FICC database were prospectively collected and validated. Second, the multinational patient cohort increases the generalizability of the results. Third, the used mortality prediction model has been validated, with good discrimination and calibration[6].

Our study has some limitations. PODs represented a very small proportion of all admissions and the annual number of PODs in individual ICUs and between the ICUs was highly variable. Due to this variability, there were only 514 PODs out of more than 60,000 admissions during 3 years. The low number of PODs might have resulted in an underestimation of the impact of PODs on SMR. Some patients may become candidates for organ donation later during their ICU stay but this is not recorded in the FICC database. In 2015–2017, nonheart-beating organ donations were not established in the participating ICUs. Their impact on SMRs should be considered in the future. In general, benchmarking SMR is associated with several confounding variables, such as differences in admission and discharge policies[26], setting treatment limitations[13], and data completeness and sampling frequency[7]. Despite the standardization of data collection, we cannot estimate the possible impact of these common confounders.

The SMR may be susceptible to differences in case-mixes. In our study, the impact of PODs on the SMRs was consistent across different ICU typologies, regardless of varying SMRs. However, the study was performed in ICUs located in high-income countries. Therefore, the findings may not be generalizable to low- and middle-income countries.

CONCLUSIONS

PODs make up a small number of all ICU admissions, but their mortality is high, which is not captured by mortality risk prediction models. This causes bias in SMR calculations and consequently benchmarking results. Therefore, we propose

identifying, documenting, and excluding POD admissions from SMR calculations to improve the accuracy of ICU benchmarking.

REFERENCES

1. Salluh JIF, Soares M, Keegan MT: Understanding intensive care unit benchmarking. Intensive Care Med. 2017; 43:1703–1707

2. Zimmerman JE, Kramer AA, McNair DS, et al.: Acute Physiology and Chronic Health Evaluation (APACHE) IV: Hospital mortality assessment for today's critically ill patients. Crit Care Med. 2006; 34:1297–1310

3. Metnitz PG, Moreno RP, Almeida E, et al.; SAPS 3 Investigators: SAPS 3--from evaluation of the patient to evaluation of the intensive care unit. Part 1: Objectives, methods and cohort description. Intensive Care Med. 2005; 31:1336–1344

4. Moreno RP, Metnitz PG, Almeida E, et al.; SAPS 3 Investigators: SAPS 3--from evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. Intensive Care Med. 2005; 31:1345–1355

5. Higgins T, Teres D, Nathanson B, et al.: Updated mortality probability model - MPM0-III. Chest. 2005; 128(4 Suppl):348S–348S

6. Moser A, Reinikainen M, Jakob SM, et al.: Mortality prediction in intensive care units including premorbid functional status improved performance and internal validity. J Clin Epidemiol. 2022; 142:230–241

7. Reinikainen M, Mussalo P, Hovilehto S, et al.; Finnish Intensive Care Consortium: Association of automated data collection and data completeness with outcomes of intensive care. A new customised model for outcome prediction. Acta Anaesthesiol Scand. 2012; 56:1114–1122

8. Harrison DA, Parry GJ, Carpenter JR, et al.: A new risk prediction model for critical care: The Intensive Care National Audit & Research Centre (ICNARC) model. Crit Care Med. 2007; 35:1091–1098

9. McClean K, Mullany D, Huckson S, et al.: Identification and assessment of potentially high-mortality intensive care units using the ANZICS Centre for Outcome and Resource Evaluation clinical registry. Crit Care Resusc. 2017; 19:230–238

10. Paul E, Bailey M, Kasza J, et al.: The ANZROD model: Better benchmarking of ICU outcomes and detection of outliers. Crit Care Resusc. 2016; 18:25–36

11. van de Klundert N, Holman R, Dongelmans DA, et al.: Data resource profile: The Dutch National Intensive Care Evaluation (NICE) registry of admissions to adult intensive care units. Int J Epidemiol. 2015; 44:1850–1850h

12. Engerström L, Kramer AA, Nolin T, et al.: Comparing time-fixed mortality prediction models and their effect on ICU performance metrics using the Simplified Acute Physiology Score 3. Crit Care Med. 2016; 44:e1038–e1044

13. Rubio O, Arnau A, Cano S, et al.: Limitation of life support techniques at admission to the intensive care unit: A multicenter prospective cohort study. J Intensive Care. 2018; 6:24

14. Wood KE, Coursin DB: Intensivists and organ donor management. Curr Opin Anaesthesiol. 2007; 20:97–99

15. Lytle FT, Afessa B, Keegan MT: Progression of organ failure in patients approaching brain stem death. Am J Transplant. 2009; 9:1446–1450

16. Takala J, Moser A, Raj R, et al.: Variation in severity-adjusted resource use and outcome in intensive care units. Intensive Care Med. 2022; 48:67–77

17. Le Gall J, Lemeshow S, Saulnier F: A new Simplified Acute Physiology Score (SAPS II) based on a European/North American Multicenter Study. JAMA. 1993; 270:2957–2963

18. Oken MM, Creech RH, Tormey DC, et al.: Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol. 1982; 5:649–656

19. Reith FC, Synnot A, van den Brande R, et al.: Factors influencing the reliability of the Glasgow Coma Scale: A systematic review. Neurosurgery. 2017; 80:829–839

20. Steyerberg EW, Mushkudiani N, Perel P, et al.: Predicting outcome after traumatic brain injury: Development and international validation of prognostic scores based on admission characteristics. PLoS Med. 2008; 5:e165; discussion e165

21. Jaja BNR, Saposnik G, Lingsma HF, et al.; SAHIT collaboration: Development and validation of outcome prediction models for aneurysmal subarachnoid haemorrhage: The SAHIT multinational cohort study. BMJ. 2018; 360:j5745

22. Hemphill JC, Bonovich DC, Besmertis L, et al.: The ICH score: A simple, reliable grading scale for intracerebral hemorrhage. Stroke. 2001; 32:891–897

23. van Dishoeck AM, Looman CWN, van der Wilden-van Lier ECM, et al.: Displaying random variation in comparing hospital performance. BMJ Qual Safety. 2011; 20:651–657

24. Jacobs R, Goddard M, Smith PC: How robust are hospital ranks based on composite performance measures? Med Care. 2005; 43:1177–1184

25. Siregar S, Groenwold RH, Jansen EK, et al.: Limitations of ranking lists based on cardiac surgery mortality rates. Circ Cardiovasc Qual Outcomes. 2012; 5:403–409

26. Kramer A, Zimmerman JE, Knaus WA: Severity of illness and predictive models in society of critical care medicine's first 50 years: A tale of concord and conflict. Crit Care Med. 2021; 49:728–740

27. Rhodes A, Moreno RP, Azoulay E, et al.; Task Force on Safety and Quality of European Society of Intensive Care Medicine (ESICM): Prospectively defined indicators to improve the safety and quality of care for critically ill patients: A report from the Task Force on Safety and Quality of the European Society of Intensive Care Medicine (ESICM). Intensive Care Med. 2012; 38:598–605

28. Lewis A, Koukoura A, Tsianos GI, et al.: Organ donation in the US and Europe: The supply vs demand imbalance. Transplant Rev (Orlando). 2021; 35:100585

ANSSI PÖLKKI

The Sequential Organ Failure Assessment (SOFA) score is commonly used to assess the severity of organ failure in intensive care. This study showed that the cardiovascular SOFA score component is outdated. Replacing the cardiovascular score by Vasoactive Inotropic Score (VIS)-based criteria enhanced the SOFA score accuracy. Noradrenaline dose alone with cutoffs 0.2 and 0.4 µg/kg/min was useful to categorise the need for vasoactive support that reflects the severity of circulatory failure. Potential organ donors caused bias in standardised mortality ratio (SMR) calculations and ICU benchmarking.

UNIVERSITY OF
EASTERN FINLAND

uef.fi

**PUBLICATIONS OF
THE UNIVERSITY OF EASTERN FINLAND**
Dissertations in Health Sciences

ISBN 978-952-61-5288-2
ISSN 1798-5706