CHALMERS
UNIVERSITY OF TECHNOLOGY

Performance of a new Risk Score in a Patient Monitoring System in Low-Resource Settings in a Hospital in Zomba, Malawi

Data analysis on an observational study of introducing the Risk Score in the pediatric High-Dependency Unit in Zomba Central Hospital

Master's thesis in Biomedical Engineering

KELLY MORRENHOF

DEPARTMENT OF ELECTRICAL ENGINEERING

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2025
www.chalmers.se

Master's thesis 2025

Performance of a new Risk Score in a Patient Monitoring System in Low-Resource Settings in a Hospital in Zomba, Malawi

Data analysis on an observational study of introducing the Risk Score in the pediatric High-Dependency Unit in Zomba Central Hospital

KELLY MORRENHOF

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Chalmers University of Technology
Gothenburg, Sweden 2025

Performance of a new Risk Score in a Patient Monitoring System in Low-Resource Settings in a Hospital in Zomba, Malawi
Data analysis on an observational study of introducing the Risk Score in the pediatric High-Dependency Unit in Zomba Central Hospital
KELLY MORRENHOF

Supervisors: Bart Bierling, GOAL 3
Stefan Candefjord, Department of Electrical Engineering

Examiner: Stefan Candefjord, Department of Electrical Engineering

Master's Thesis 2025
Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Screenshot of the GOAL 3 app from the Patient Monitoring system with the Risk Score

Typeset in LaTeX
Gothenburg, Sweden 2025

Performance of a new Risk Score in a Patient Monitoring System in Low-Resource Settings in a Hospital in Zomba, Malawi
Data Analysis on an observational study on the risk score in the pediatric High Dependency Unit in Zomba Central Hospital
Kelly Morrenhof
Department of Electrical Engineering
Chalmers University of Technology

Abstract

In Low- and Middle-Income countries (LMIC), it is hard to provide adequate healthcare to critically ill children due to limited resources in staff and equipment. Within these Low-Resource settings (LRS), pediatric mortality and the occurrence of Critical Illness Events (CIE) are high. Continuous patient monitoring of vital signs can be particularly helpful in LRS, especially for critically ill children. GOAL 3 designed a Continuous Patient Monitoring System (CPMS), service, and training fit for LRS, IMPALA, and implemented it in the Pediatric High-Dependency Unit of Zomba Central Hospital in Malawi, Africa. To this IMPALA CPMS, a newly developed Risk Score was added. This Risk Score, adjusted from an earlier developed physiological score, can potentially assign a health risk level with a score between 0-100, show the progress of the patient over time, and function as an early warning system. In this study, the performance of the Risk Score is assessed, based on its distinctive ability to assign a high-risk level to patients who eventually passed away and to patients who had one or more CIE. The study consisted of 192 participants under 5 years of age, whose clinical and score data were analyzed. With a mortality rate of 12% and 140 CIE in total, there were many critically ill children. That group showed the highest average Risk Score of 64 in the death subset, and 37 in all patients with one or more CIE, significantly different from an average score of 23 in participants who had no CIE and survived. With an AUROC in the range of comparable literature, it shows that the Risk Score can distinguish and potentially predict patients who pass away moderately to well, especially in the last 6-8 hours before discharge. Observations and data analysis revealed challenges with the Risk Score availability, which results in an average score coverage of 82%, mainly due to problems with the oxygen saturation probe and nurses not implementing the score system in their routines at the beginning of the study. Optimization and more research have to be done to improve these challenges and validate the Risk Score.

Keywords: Low-Resource settings, Patient Monitoring, Malawi, Risk Score, Early Warning Score.

Acknowledgements

In this section, I would like to express my gratitude and share more about everything I learned and valued about this whole Master's Thesis experience. The findings of this study are not the only things I learned; I learned a lot more from my experience of three months in Malawi and being involved with all the people and work of GOAL 3, the IMPALA consortium, and TRUE.
It all started with me reaching out to GOAL 3, after dreaming of being able to apply my medical engineering knowledge and skills in countries such as Malawi. When I learned that Chalmers offered a Global Mentorship and Scholarship program for students who do their thesis abroad in a low- or middle-income country, I reached out to GOAL 3 because I found their work fascinating after coming across their LinkedIn page when I was still studying in the Netherlands. Their response was helpful and enthusiastic right from the start, and together we worked towards a project for me. It became a data analysis project on the new IMPALA 3.0 study, and I would spend three months in Zomba, Malawi. With the scholarship in the pocket, I flew from Gothenburg to Blantyre and was received amazingly by the lovely people of Blend Lodge in Zomba, where I would spend my first few weeks. The other weeks were spent in Paphiri house, with Bart, Martine, Pam, and Rianne. I am so grateful for the amazing time we have spent there, from complaining but also admiring the monkeys in the garden (and sometimes in the kitchen too, unfortunately), to all the good dinner nights with conversations and discussions from which I learned so much. It was very interesting to hear Bart and Martine share their experiences of living in Malawi, working with start-ups in the medical field, and all the (bureaucratic) challenges that come with that, always with this thriving passion of really wanting to improve the health sector. But also Pam and Rianne, who could tell me more about the clinical experiences, both the beautiful and sad stories. Living with the Zomba Plateau as your backyard also has its terms with hiking, cycling, and swimming in that beautiful nature.
I got to work in the TRUE office at the Zomba Central Hospital every day with amazing colleagues there, working hard on their projects but also taking the time to learn each other more about Malawi, the Netherlands, and Sweden. From market trips to find screws that fitted the monitors, to lunch breaks at Bhongororo bar with stony rice or nsima that burned your hands, it was never boring working there. Taking the minibus or sketchy taxis to and from the hospital was an eventful experience, and every time you would walk, you couldn't count the "sistah motorbike" on one hand.
I have a lot more to tell, a lot more insights to share, and met and learned from a lot more people than I have mentioned now, but this report has to get finished at some point, at least.
I would like to thank Chalmers University and the Kullman family with the Global Mentorship Program for enabling this experience as a part of my masters. With that also came my mentor Karin Dahllöf, who really supported me during this period and after, encouraging me to even apply for jobs in Swedish, and look where I ended up now.

I would also like to thank all the people at GOAL 3 for really involving me in the company and teaching me so much, and for having (a lot of) patience with me. Both the people I worked a bit closer with, Bart and Eline, but also the other team members teaching me about the more technical aspects of the monitor system or strategies, among many other things. GOAL 3 works so hard to improve so many human lives all over the world, and it is fascinating to see how they find their way through all the challenges that come with that. Also, all the people working on the IMPALA project are so driven; their work ethic, experience, and enthusiasm are infectious.
Lastly, I would have to thank all the people who really made my time in Malawi the experience it became. From all the random people I met on the street, to all the dutchies we gathered from all over Malawi, the people at Blend Lodge, TRUE, ZCH, and so many more.
I will cherish this experience forever and could not be more grateful for this opportunity.

Kelly Morrenhof, Gothenburg, July 2025

Acronyms

AUROC Area Under ROC Curve (AUC). 14, 25, 31
AVPU Alert Verbal Pain Unresponsive. 6, 16, 25, 29

CIE Critical Illness Events. xi, 1, 2, 9, 22, 25, 30
CPMS Continuous Patient Monitoring System. v, 1, 2
CPO Central Patient Overview. 7

ECM Estimated Crash Point. 14, 22
EWS Early Warning Score. 2, 7

HCWs Healthcare Workers. 2, 7
HDU High-Dependency Unit. 1, 2, 11, 29
HIC High-Income Countries. 1, 2
HR Heart Rate. 6, 12

LMIC Low- and Middle-Income Countries. 1, 2, 11, 31
LRS Low-Resource Settings. 1, 2

NIBP Non-Invasive Blood Pressure. 6, 12, 25

PDS Patient Deterioration Score. 2, 31
PEWS Pediatric Early Warning Score. 2, 31
PI Perfusion Index. 12
PICU Pediatric Intensive Care Unit. 1, 31

qSOFA quick Sequential (sepsis-related) Organ Failure Assessment. 2

ROC Receiver Operating Characteristic. 14, 25
RR Respiratory Rate. 6, 12

SBP Systolic Blood Pressure. 7, 13, 16, 29
SD Standard Deviation. 14
SpO2 Oxygen Saturation. 6, 12, 17, 29

ZCH Zomba Central Hospital. 1, 11, 29

Contents

List of Figures

List of Tables

1
Introduction

In Low- and Middle-Income Countries (LMIC), it is hard to provide sufficient healthcare for critically ill children due to limited resources [1]. Hospitals in LMIC often struggle with finances, adequate healthcare service delivery, infrastructure, knowledge, human resources, influence of beliefs and practices, and social, geographical, and environmental factors, classifying the setting as a Low-Resource Settings (LRS) by the definition in Van Zyl et al. [2]. Although decreasing over the past years, pediatric mortality and the occurrence of Critical Illness Events (CIE) are still high within these LRS [3]. For example, the Under-five mortality rate in low-income countries is 13 times higher than the Under-five mortality rate in high-income countries [4].
The lack of resources, compared to High-Income Countries (HIC), mainly consists of a lack of high-quality equipment, well-trained staff, and proper documentation. One of the most essential principles for providing adequate critical care is monitoring a patient's condition, but monitoring systems used in HIC are often not affordable or applicable in LMIC [3]. In Zomba Central Hospital (ZCH) in Malawi, a recent study showed that continuous monitoring with those systems is impossible due to power outages, ward design, and lack of functioning equipment and human resources [5].

Because of the lack of resources, staff, and appropriate monitoring systems, patients in LRS experience a delay in receiving critical healthcare in the wards where they need it the most, the High-Dependency Unit (HDU) and Pediatric Intensive Care Unit (PICU) [1], [6]. This delay can be decreased by providing an essential monitoring system adjusted to LRS that can help with providing an earlier diagnosis and intervention [1]. GOAL 3 designed a Continuous Patient Monitoring System (CPMS), service, and training fit for LRS, called the IMPALA system. Monitoring systems decrease the detection time and number of missed events, even more when visual cues are added to the monitoring display [7]. The use of the IMPALA patient monitoring system has previously been studied in ZCH in the IMPALA 2.0 study and showed a significant decrease in CIE occurrence [8].

Diagnosing based on just the current observation of a single physiological variable (one of the vital signs) is relatively hard. That is why it helps clinicians to bring the vital signs into context, by showing the trend of the recordings and an alarm when that trend in observations has changed [9]. Within patient monitoring environments, a user-centered design-based system with visual cues such as alternating shapes, colors, and a visual patient, contributes to faster decision-making and decreased perceived workload [10].
Combining the benefits of providing visual cues and context/patterns of vital signs observations, and using that for improving care in critically ill patients, is done in the form of Patient Deterioration Score (PDS) and/or Early Warning Score (EWS). Usage of those scores, such as the Pediatric Early Warning Score (PEWS), is proven to reduce (pediatric) mortality and CIE in hospitals in HIC [11]–[13], and it is suggested for use in critical care in LRS too [3], [14], [15]. However, existing early warning or patient deterioration scores like PEWS, quick Sequential (sepsis-related) Organ Failure Assessment (qSOFA), and Acute Physiological Score (APS, APACHE) are not fitted to work in LRS due to requirements that are hard to meet in LRS, like certain vital signs or well-trained Healthcare Workers (HCWs) [12], [16]. Several types of PEWS and qSOFA were tested in LMIC and showed promising results, but most studies were retrospective and did not observe the effect of the score in real time [17]–[19]. One study developed a PEWS for LRS, the PEWS-RL, and studied it in real time, but that study only used the score during admission and consisted of 6 equally weighted variables [17]. This shows the need for a PDS adjusted to LRS with fewer variables, that can be used at all times during hospital admission, and is observed in real time with a large study population. The development of a new Risk Score answered that need.

1.1 Aim of Study

The overall aim of this study was to observe and evaluate the performance of a newly developed pediatric Patient Deterioration Score (Risk Score) integrated in a Continuous Patient Monitoring System in Low-Resource Settings and compare its performance within different patient subgroups using data analysis.
More specifically, this study aims to show how the Risk Score is performing:

1. around Critical Illness Events (CIE);
2. in the first hours after HDU admission;
3. in the last hours before death or discharge;
4. when one or more vital input variables are missing, and which input variable is missing the most.

Another specific aim is to highlight and explain the specific cases where the Risk Score does not perform as expected, and to provide their frequency of occurrence.

According to the overall aim, the following main research question was established.

To what extent can a patient monitoring system with an additional Risk Score distinguish the overall state of a patient and their likelihood of Critical Illness Events in low-resource settings?

The main research question is supported by the following sub-questions, according to the specific aims.

1. How does the score act 0-4 hours around CIE?
2. How does the score behave in the 8 hours after study admission compared to the last 8 hours before discharge or death?
3. How does the score act in the 8 hours before death or discharge for the different subgroups (CIE/no CIE and death/survive)?
4. In which cases is the score missing?
5. In which cases does the score give results that are not as expected?

The Risk Score is expected to be good at distinguishing patients who will pass away compared to patients who are normally discharged. It is also expected that the score will be higher at and around CIE and that patients who do not get any CIE will have the lowest average score.

2

Theory

The Risk Score is integrated into the GOAL 3 IMPALA patient monitoring system. The entire IMPALA system has already been proven to reduce pediatric mortality, detection time, and missed events in the IMPALA 2.0 pilot [8], [20]. In this section, the IMPALA system and its relevant parts and functions will be explained, as well as the Risk Score itself.

2.1 IMPALA monitoring system

The IMPALA system consists of:

- Bedside monitor (see 1 in Figure 2.1)
- Bedside tablet (see 2 in Figure 2.1)
- Vital sign sensors, leads, and wires
- Software (GOAL 3 app) (see Figure 2.2)
- Tablet at nurses station (see Figure 2.2)
- Server
- Service, support, and training

Figure 2.1: The main part of the Impala system. On top (1) is the bedside monitor displaying the vital signs with attached wires, and on the bottom (2) is the bedside tablet with the GOAL 3 app.

First, the bedside monitor, which shows the real-time vital sign values and waveforms, patient information, alarms, and vital signs history. The monitor is operated with a touchscreen, to pause or acknowledge an alarm or change the settings, or patient information. The vital signs that can be measured and displayed on the monitor are Heart Rate (HR), Respiratory Rate (RR), Oxygen Saturation (SpO2), and Non-Invasive Blood Pressure (NIBP). From these vital signs, the HR, RR, and SpO2 are measured continuously with ECG leads and SpO2 sensor, which are connected with wires to the monitor. The NIBP is measured with an NIBP cuff connected to the monitor and has to be initiated manually. The monitoring system described until now is the system used in the IMPALA 2.0 pilot. The IMPALA system referred to in this current IMPALA 3.0 pilot study includes the tablet with the app and Risk Score as described below, alongside the monitor.

Next to the bedside monitor is the bedside tablet, which is connected through a WiFi module and the server. On the bedside tablet, the GOAL 3 app is installed and displayed. The screen that is displayed is the individual patient overview, which can be seen on the left side in Figure 2.2. This view shows the same vital signs as the monitor (HR, RR, SpO2, NIBP), as well as the temperature and level of responsiveness (Alert Verbal Pain Unresponsive (AVPU)) on the left side of the screen. In the middle of the screen, the Risk Score is displayed, both in textual cues and visual cues in the shape of a plot. Under the plot, it shows the SpotCheck button, through which the AVPU can be logged, among other things that will be explained later on.

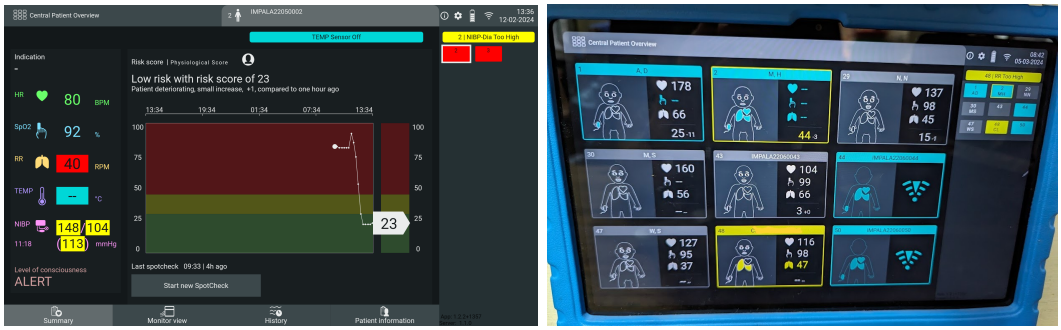

Figure 2.2: GOAL 3 app screens on bedside tablet (left) and nurses station tablet (right).

In addition to the tablets by every bed and monitor, there is also one tablet in the nurses' station. It is the same tablet with the same GOAL 3 app, but it shows a different screen. The nurse's tablet shows the Central Patient Overview (CPO), see the right picture in Figure 2.2. This overview enables nurses to get an idea of how patients are doing even though they are not in clear sight, through visual and textual cues of the three continuous vital signs, Risk Score, and alarms. Here, the blue color indicates a technical alarm (sensor/leads off or network problem), yellow indicates a medium priority alarm, and red indicates a high priority alarm.

Besides the hardware and software, the IMPALA system also includes a local server with WiFi modules for communication, as well as training of the HCWs on how to use the system, and other services.

2.2 Risk Score

As mentioned in the introduction, the Risk Score (also called IC-Len score or physiological score) is inspired by other types of EWS. However, this Risk Score was specifically developed to fit this setting and system. This Risk Score is based on the physiological score developed using various African and UK cohorts of critically ill children to see the influence of a saline or albumin fluid bolus [21]. The only difference from that physiological score is that the capillary refill time (CRT) is removed due to inaccuracy. Based on the three continuous vital signs (HR, RR, SpO2), Systolic Blood Pressure (SBP), and AVPU, combined with the age-adjusted mean HR, RR, and SBP, the three components of the score are calculated. These three components are Cardiovascular (eq. 2.1), Respiratory (eq. 2.2), and Neurological (eq. 2.3), and are calculated with the equations below.

$$\text{Cardiovascular score} = (HR - \text{mean HR for age})^* + (\text{mean SBP for age} - SBP)^* \tag{2.1}$$

$$\text{Respiratory score} = (RR - \text{mean RR for age})^* + \left(5 \times \left(100 - \underset{(\geq 60\%)}{\text{SpO}_2}\right)\right) \tag{2.2}$$

$$\text{Neurological score} = (\text{SBP} - \text{mean SBP for age})^* + (\text{mean HR for age} - HR)^* + \begin{cases} 50 + 25 \times \text{AVPU} & \text{if AVPU} > 0 \\ 0 & \text{otherwise (AVPU} = 0) \end{cases} \tag{2.3}$$

$$\text{Combined Score} = \text{Cardiovascular score} + \text{Respiratory score} + \text{Neurological score} \tag{2.4}$$

All parts marked with * can not go below 0, so if the difference is below zero, that part becomes 0. The oxygen saturation (SpO_2) can not go below 60%, all values below that become 60. The values of the HR, RR, and SpO2 are the mean values of the latest available fifteen minutes before the time point of when the score is calculated. For the value of the SBP, the most recent measurement is used. The AVPU scale (level of responsiveness) is numerized as: Alert = 0; responds to Verbal stimulus = 1; responds to Painful stimulus = 2; Unresponsive = 3 [21], [22]. For the mean HR, RR, and SBP for age, the values in table 2.1 are used, derived and adjusted from earlier published tables [21].

Vital Sign and Age Group	0–1 mo	1–12 mo	12–48 mo	4–12 yr	12–18 yr	18+ yr
Heart Rate (BPM)	140	135	125	100	80	80
Respiratory Rate (BPM)	40	35	30	18	15	15
Systolic BP (mmHg)	60	70	75	85	100	120

Table 2.1: Mean age-adjusted HR, RR, and SBP values by age group

All three components are summed to obtain the Combined Score (eq. 2.4). After that, the Combined Score is scaled to a 0-100 scale with an adjusted sigmoid function to obtain the final Risk Score (eq. 2.5).

$$\text{Risk Score} = \frac{103}{1 + e^{-a \cdot (x-b)}} - 3 \tag{2.5}$$

In this sigmoid function (eq. 2.5), x is the Combined Score (eq. 2.4), and a and b are parameters to adjust the steepness and midpoint of the sigmoid curve, respectively. In all the different types of scores in this report $a = 0.02$. Later in this report, the Risk Score gets compared to the score without one or both of the manual inputs: score without SBP and score without SBP and AVPU (also referred to as score without AVPU or score without manual input). These scores get calculated the same way, but without either the SBP and/or AVPU part in formulas 2.1 and 2.3. The formula for the score without any manual inputs, for example, becomes then eq. 2.6. To adjust these two other scores to match the range and distribution of the Risk Score, different values for b are used in the equation 2.5.

$$\begin{aligned} \text{Score without manual input} = \max &\left|(HR - \text{mean HR for age})^*\right| \\ &+ (RR - \text{mean RR for age})^* \\ &+ \left(5 \times \left(100 - \underset{(\geq 60\%)}{\text{SpO}_2}\right)\right) \end{aligned} \tag{2.6}$$

Within the 0-100 range of the Risk Score, there are three Risk Levels:

- 0-29: Low Risk (green)
- 30-44: Medium Risk (yellow)
- 45-100: High Risk (red)

So the higher the score, the more critically ill the patient is, in theory. The Risk Score value and the color of its Risk Level are shown in the GOAL 3 app on the tablet. Every 15 minutes, the score is calculated, as long as the needed vital signs and AVPU are available. For continuous vital signs (HR, RR, SpO2), the buffer is one hour, and for SBP and AVPU, it is 16 hours. This means that if one of the continuous vital signs is not measured for more than one hour (because the sensor fell off, e.g.), the score can not be calculated. The buffer for the SBP and AVPU is longer because they require manual input, which can be missing during the night, for example.

2.3 Critical Illness Events (CIE)

In research, everyone has their own definition of what critical illness events are. Within this study, they are defined as one of the following events in five categories, based on the WHO ETAT (Pediatric Emergency and Triage Assessment and Treatment) updated manual [23].

- Cardiovascular CIE
  - Resus
  - CPR
- Respiratory CIE
  - Oxygen Support
  - CPR
- Neurological CIE
  - Convulsions
  - CPR
- Infections CIE
  - Resus
  - CPR
- Other CIE
  - Resus
  - CPR

The type of CIE, its time point, and more information are logged in REDCap.

3
Methods

This research used the data of the IMPALA 3.0 pilot study. It is part of the bigger IMPALA-UIR study ("Innovative Monitoring in Paediatrics in Low resource setting: an Aid to save lives", https://www.projectimpala.org). IMPALA-UIR is a European Commission's EDCTP programme project and has a goal to develop a pediatric monitoring system for LMIC that is easy and sustainable to use. However, only methods related to this particular research are mentioned here, not from the complete IMPALA project.

3.1 Study setting

All data used in this research comes from Zomba Central Hospital. Located in Zomba, Malawi, ZCH is a large tertiary district hospital with many admissions that are often referred from smaller local clinics in the rural area. Around 1200 of those admissions every year are critically ill children who get admitted to the pediatric High-Dependency Unit (HDU) [24]. During the rainy season, around 10 children are admitted each day to the HDUs on average, which decreases to 6-8 children during the dry season. This study took place from the end of January until the beginning of June 2024, so mainly during the rainy season, which lasts until April. Most of the children get admitted to the HDU for malaria, sepsis, and pneumonia [5]. In the HDU, nine IMPALA bedside monitors and tablets were installed. In the case of more than nine patients, nurses prioritized the most critical patients to connect to the monitors, using a flowchart. It often happened that several patients were in one bed, and one or more guardians per patient were in the room as well. Two government nurses and research nurses were on site and worked with the IMPALA monitoring system. At admission, guardians were told by the nurses how to respond and act when there is no or a false vital sign reading. Ward rounds were done by the government nurses 3-4 times a day, during which they completed a SpotCheck in the GOAL 3 app using the bedside tablet.

3.2 Participants

One group of pediatric patients admitted to the HDU was recruited by the research nurses. Inclusion criteria were that the patient was connected to an IMPALA monitor and had an age between 28 days and 60 months old. The caregiver of the child gave informed consent.

3.3 Datasets

All data management, analysis, and visualization were done using Python 3.11.5 and NumPy, Pandas, Matplotlib, and scikit-learn packages in the Jupyter Notebook environment in the Visual Studio Code (version 1.94) program. ChatGPT, a large language model developed by OpenAI, was utilized to generate and inspire code for data analysis, visualization, and statistical computations.
Data was collected from the monitor and tablet, and combined with the collected clinical study data from REDCap electronic data capture tools hosted at Training and Research Unit of Excellence (TRUE), Malawi [25], [26], using the common factors between the datasets (patient-ID, monitor-ID, date, and time). REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies, providing 1) an intuitive interface for validated data capture; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for data integration and interoperability with external sources. The collected data from the monitor consisted of all vital signs (HR, RR, SpO2, Perfusion Index (PI), NIBP(lower, upper, mean)) every second. From the tablet, the Risk Score data was collected, including the mean values of the vital signs used to calculate the score every 15 minutes. The REDCap dataset included all the clinical data, such as admission and discharge dates, diagnoses, but also all logged information about the CIE. All relevant information from the REDCap dataset was extracted and put into a smaller Patient Summary dataset to make it easier to create functions such as 'time until discharge or CIE'. Patients who were included in the study but did not get any score data in the end were removed from all three datasets. Study data of patients who passed away after HDU/study discharge were not removed but were not included in the data analysis.

3.4 Data analysis and visualization

First, the data was summarized to give an overview of the quantity and quality of the data. This gave insights into the study population and missing values of certain components, such as the sensors and the Risk Score itself. Among those insights were the coverages (in percentage %) and averages of the individual components, and more.

To see the progression score in each individual patient, the vital signs data from the monitor and score data were combined to visualize the individual patient graphs. The graphs were marked with the patient ID and had information about the duration of admission, age, and cause of death of the patient, if relevant. In these individual patient graphs, there were five subplots that were aligned at the study admission date and HDU discharge or time of death. The first three subplots consisted of vital signs (HR, RR, oxygen saturation) measured by the ECG or SpO2 sensor from the monitor dataset, together with its same vital sign from the score dataset.

These three subplots also contained the duration of the alarms from the monitor regarding each vital sign, plotted as bars. The alarm duration bars were marked by who responded (guardian, nurse, doctor, or other) to that alarm if reported in REDCap. The fourth subplot contained the two manual inputs: the Systolic Blood Pressure and AVPU scale, including indications at which time the spot check (manual input) took place. To make it visually more clear, the number of the enumerated AVPU scale was multiplied by 50, such that the values 50 and 150 represented 'verbally responsive' and 'unresponsive' respectively. The last subplot showed the Risk Score. In all subplots, all CIEs that occurred during the patient's study admission were included by a vertical line at the time point of the CIE. The color of the line showed the CIE category, and a text box clarified the type of CIE. All subplots also have the threshold ranges (Table 2.1 for the vital signs; three Risk Levels for the last subplot) as colors in the background (green, yellow, red). These get automatically adjusted to the age group of each individual patient.
Since these individual patient graphs gave a full overview and combined all data (vital signs and alarms from the monitor, score from the tablet, and clinical data from REDCap) from the patient, they were used to review individual cases. Besides that, the graphs were also used to keep track of certain exceptions, cases where the score did not perform as expected, and highlight certain challenges during and after the study period. The graphs also gave a good overview of what happened to the data when the nurses reported an unexpected event. All cases in which the score did not perform as expected were collected and reported, both for the survive group and the group of patients who passed away.

Eight subsets were formed for the rest of the data analysis, with overlap between the subsets. The first set was just all patients except patients who passed away after being discharged, and the subsets consisted of patients who:

1. survived;
2. passed away;
3. had one or more CIE;
4. had no CIE;
5. survived and had one or more CIE;
6. survived and had no CIE;
7. passed away and had one or more CIE;
8. passed away and had no CIE

For the main set, the coverage of all components and score averages was calculated. For the subsets, it was just the coverages of the Risk Score and its risk zones that were calculated. Coverages were calculated as the percentage of the duration of the data that was covered from the total duration. This gave insight into which vital sign, sensor, or score from a certain subset had more missing data than others. Score averages showed which subsets had the highest or lowest mean Risk Score. The average was calculated over the entire study duration, the first and last 8 hours. These averages were compared to each other according to the statistical analysis below.

The distribution of the score for different subsets and time was visualized using histograms and plots over time, to give the score averages more context. These graphs were supported by the statistical analysis on the distinguishing capacities of the score for different subsets and times.
Lastly, an analysis of the Risk Score and the different categories of CIE was performed. The score was averaged for the CIE subset 4 hours before and after the CIE and compared to the same time around the Estimated Crash Point (ECM) of the non-CIE subset. This ECM-point was based on when, on average, the first CIE occurred after hospital admission in the CIE subset. The comparison was done for all CIE and the individual CIE categories and visualized in a graph, which showed the single CIE types as well in the background. The analysis was mainly concluded on a visual basis, not statistically.

3.5 Statistical Analysis

Statistical Analyses were executed using the SciPy and scikit-learn Python packages. Descriptive statistics were used to summarize the study population with means and Standard Deviation (SD) for continuous variables and percentages for coverages. Risk Score means were compared between subsets using Welch's t-tests, and within subsets, the paired samples t-test was used. To correct for multiple testing in all subsets, the Bonferroni correction was applied. With a confidence interval of 95%, full dataset, and 8 subsets, the new alpha value for which the p-value has to be lower to show significance is $0.05/36 = 0.00139$.
The performance of the distinguishing capacities of the score for different times and subsets was assessed using Receiver Operating Characteristic (ROC) curves and Area Under ROC Curve (AUC) (AUROC).

3.6 Ethical statement

Ethical approval was obtained from the University of Malawi's College of Medicine Research and Ethics Committee (COMREC) for protocol (P.01/22/3552) of the IMPALA 3.0 study by amendment of the original grant on the 15th of November 2023 [24]. Written informed consent was given by the caregiver of the pediatric patient. The study complied with the General Data Protection Regulation (GDPR).

4

Results

4.1 Study Population

Over the period of four months, a total of 192 pediatric patients were recruited into the study. Two patients were excluded because they did not have any Risk Score data by the time of discharge, which makes a total study population of 190 pediatric patients, of whom 51% were male. The total mortality rate was 12.6%, but two patients died after discharge from the study. These patients are included in the study population but are excluded from the data and statistical analysis. During the study, 22 study participants passed away, resulting in a study mortality rate of 11.6%. The study population had a mean age of 26.3 (± 17.1) months. The main reason for death was severe Malaria.

There was a wide variation in HDU- and study admission duration, with values ranging from 6.8 - 554.8 (mean: 82.7) hours for the HDU admission and 1.4 - 524.9 (mean: 64.7) hours for the study admission duration. On average, in 81.6% of the time of all patients' study admission, the Risk Score was available (total Score coverage).

4.1.1 Critical Illness Events

In total, 140 CIEs were registered in 78 patients. The occurrence distribution of the CIE categories and types is shown in the pie charts in Figure 4.1.

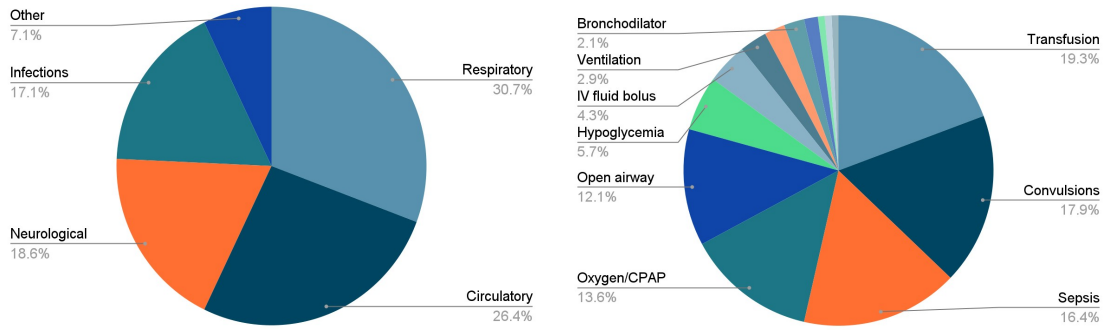

Figure 4.1: Distribution of occurrence of all registered CIEs by categories (left) and type (right).

4.2 Individual patient graphs and situations

The most frequent and expected situation of a patient who is discharged normally (survive) is shown in the first individual patient graph in Figure 4.2. It shows a patient who enters with a Risk Score in the red high-risk zone and is discharged with a low-risk score. In 12.7% of the cases in the survive group, there was a high Risk Score at discharge. However, this graph also contains an occurrence that is not very frequent and expected. What can be observed in this first graph is that there is no Risk Score at the end after the last manual input of SBP and AVPU (blue dotted vertical lines in the fourth subplot) exceeded the 16-hour buffer. In the second individual patient graph in Figure 4.2, a missing Risk Score can also be observed. However, in this case, it is due to the absence of one or more continuous vital signs, with SpO2 being the most problematic. The total percentage of all patients who had missing score data at the end ($>$ 2 hours) due to any reason was 21.6%. The second graph also shows a missing Risk Score in the first hours, because AVPU and/or SBP were not registered in the system at study admission. This occurred in 11.6% of all cases, mainly at the beginning of the study, where there was no Risk Score for at least the first three hours.

For the group of patients who did not survive, there are two situations represented in the individual patient graphs in Figure 4.3. The patient in the first graph shows a characteristic progression for patients who pass away. In the beginning, the Risk Score is already in the high-risk zone, with some ups and downs, one or more CIE. In the last few hours, the score increased (closely) to the maximum value of 100, like the frequencies of CIE and alarms are increasing too. In the second graph on the right in Figure 4.3, a more unusual situation shows a patient who passed away with a low-risk score. This happened in 11.8% of all death cases. There are also almost no alarms and no CIE at the end.

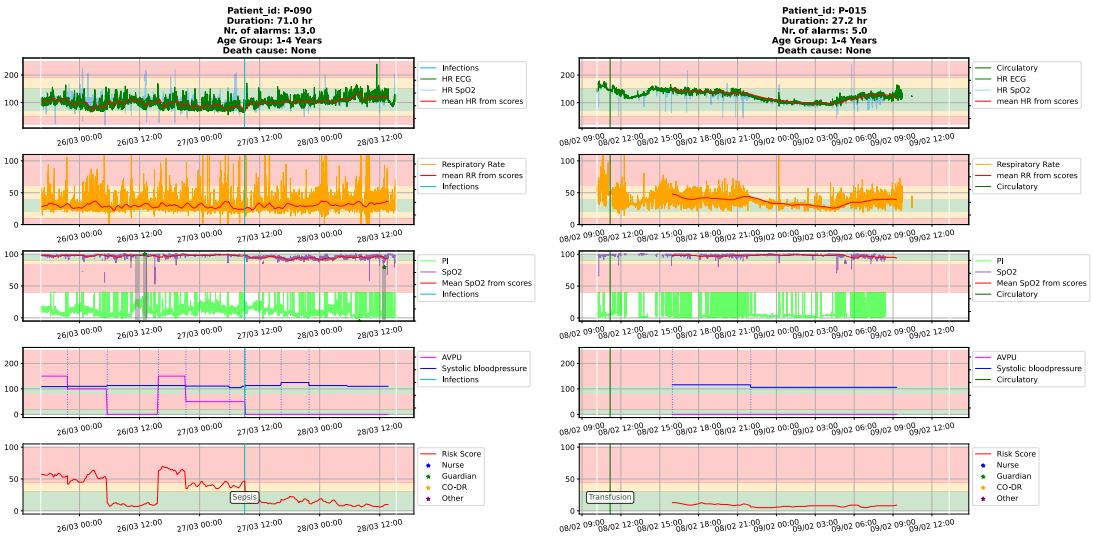

Figure 4.2: Individual patient plots from patients who survived. The left plot is of an expected situation, and on the right is a case with a missing Risk Score.

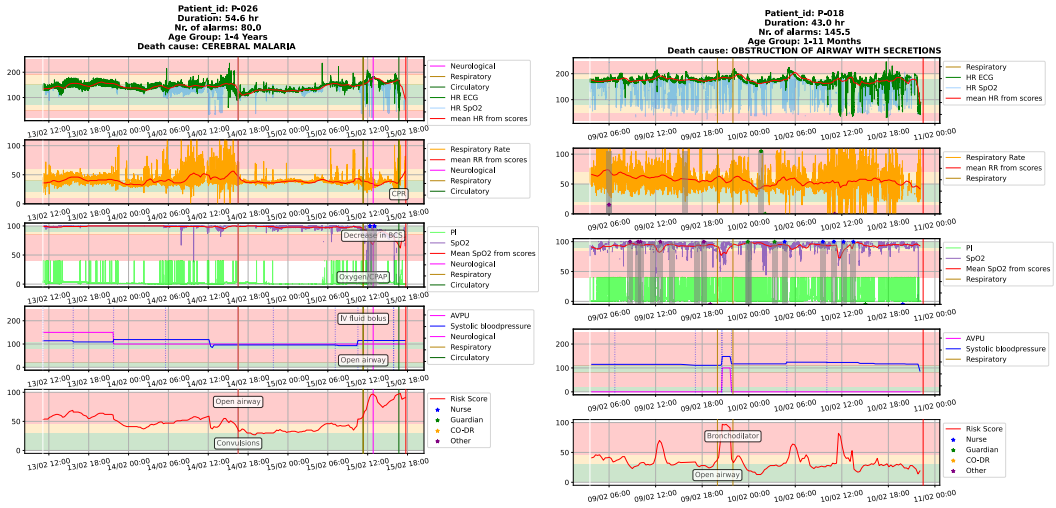

Figure 4.3: Individual patient plots from patients who passed away. The left plot is of an expected situation, and on the right is a case of a low Risk Score at the end.

4.3 Risk Score coverages

In Table 4.1, the Risk Score coverages for each risk level and vital sign sensors are shown. Most of the Risk Score measurements were in the low-risk area (61.8%). The results of these Risk Score coverages per subset can be found in Table 4.2.
The vital sign that recorded the fewest measurements is the SpO2 of the oxygen saturation (63.9%) and heart rate (63.3%) sensors. Blood pressure was present for 95.2% of the total duration of the study, making it the most prevalent vital sign.

Type of Coverage	**Coverage (%)**
All Risk Scores	81.61
Low-Risk Score	61.80
Medium-Risk Score	13.05
High-Risk Score	25.16
ECG Heart Rate	89.75
ECG Respiratory Rate	91.31
SpO_2 Oxygen Saturation	63.85
SpO_2 Heart Rate	63.28
Blood Pressure	95.15

Table 4.1: Coverage of risk scores and vital signs in percentages

In Table 4.2, it shows that the death subset had the highest total score coverage (83.51%), and the No CIE - death subset the lowest (76.74%). The CIE - death subset had the highest prevalence of high-risk score (78.32%), and the No CIE - survive subset had the most low-risk score (73.07%).

Subset	Size	Total Score Coverage (%)	High-Risk (%)	Medium-Risk (%)	Low-Risk (%)
Total study group	190	81.61 ± 17.39	25.16 ± 32.27	13.05 ± 15.69	61.80 ± 35.64
CIE all	78	82.06 ± 17.41	37.13 ± 35.31	13.74 ± 14.80	49.12 ± 36.31
CIE - survive	59	82.48 ± 14.35	24.45 ± 26.17	13.78 ± 13.35	61.77 ± 31.14
CIE - death	17	85.51 ± 17.38	78.32 ± 29.32	14.66 ± 19.86	7.02 ± 13.62
No CIE	112	81.29 ± 17.44	16.82 ± 27.16	12.56 ± 16.33	70.62 ± 32.50
No CIE - survive	107	81.50 ± 16.67	14.53 ± 24.67	12.39 ± 16.26	73.07 ± 30.61
No CIE - death	5	76.74 ± 32.45	65.70 ± 34.98	16.15 ± 19.36	18.15 ± 29.68
Death	22	83.51 ± 21.09	75.45 ± 30.29	15.00 ± 19.29	9.55 ± 18.21
Survive	166	81.85 ± 15.85	18.06 ± 25.58	12.88 ± 15.26	69.06 ± 31.18

Table 4.2: Coverage statistics for different subsets of the dataset. Risk Score coverages for the whole study group and subsets over the complete study duration in Mean ± SD.

4.4 Risk Score averages

In Table 4.3, the Risk Score averages throughout the study period and its first and last eight hours are displayed for the entire study group and its subsets. The subset with the highest score average is the group that had one or several CIE and passed away, with a Risk Score average of 68.01 ± 17.84. The lowest score average is for the opposite subset; the patients who had no CIE and survived, 22.52 ± 13.23. When comparing the total score average between groups, the following subsets showed a significant difference: Total-CIE death, Total-No CIE survive, Total-Death, Death-Survive, Death-CIE survive, Survive-CIE all, Survive-CIE death, CIE all-CIE death, CIE all-No CIE, CIE all-No CIE survive, No CIE-CIE death, CIE survive-CIE death, CIE death-No CIE survive. The table with all p-values can be found in Appendix 1, Table A.1.

Subset	Size	Total Score Average	First 8h	Last 8h
Total study group	190	29.26 ± 19.82	35.68 ± 23.27	25.07 ± 21.98
CIE all	78	37.12 ± 23.14	44.47 ± 24.19	31.98 ± 27.90
CIE - survive	59	27.98 ± 15.71	39.48 ± 21.98	19.71 ± 15.22
CIE - death	17	68.01 ± 17.84	62.15 ± 24.35	72.17 ± 23.97
No CIE	112	23.79 ± 14.96	29.55 ± 20.58	20.26 ± 15.02
No CIE - survive	107	22.52 ± 13.23	28.12 ± 19.49	19.00 ± 13.14
No CIE - death	5	51.13 ± 24.40	60.20 ± 21.56	47.18 ± 27.12
Death	22	64.17 ± 20.20	61.70 ± 23.26	66.49 ± 26.32
Survive	166	24.46 ± 14.36	32.16 ± 21.06	19.25 ± 13.87

Table 4.3: Risk Score averages for the whole study group and subsets over the complete study duration (total), first eight hours, and last eight hours of the study in Mean ± SD.

Comparing the significance between the subsets in the first 8 hours of the study with the last 8 hours of the study gives different results. These p-values are shown in Appendix 1, Tables A.2 and A.3, respectively. There, it shows a significant difference in Total - No CIE survive in the first 8 hours, but not in the last 8 hours. The same goes for CIE all - CIE death, CIE all - Death, CIE survive - CIE death, and No CIE survive - CIE survive.
The opposite (significant difference in the last 8 hours, but not the first) holds for CIE all - Death only. Total - no CIE survive is not significant in the first and last 8 hours, while it is for the whole study duration. On the other hand, CIE all - CIE death and CIE survive - CIE death are significantly different in the last 8 hours of the study and the total study period, but not in the first 8 hours.
The difference between the first to the last 8 hours within each subset was also statistically tested, and the results are shown in Table 4.4. Here, it can be observed that there was a significant difference in the Risk Score values between the first and last 8 hours for the Total study group, CIE all, CIE survive, No CIE, No CIE survive, and Survive subsets. In all these data sets, the Risk Score is lower in the last 8 hours compared to the first 8 hours. The Risk Score increased in the Death and CIE-death subsets, but this was not significant.

Subset	**P-value (First 8h vs Last 8h)**
Total study group	0.00000
CIE all	0.00004
CIE - survive	0.00000
CIE - death	0.07874
No CIE	0.00000
No CIE - survive	0.00001
No CIE - death	0.23793
Death	0.34982
Survive	0.00000

Table 4.4: P-values of significance tests comparing Risk Score values between the first 8 hours and last 8 hours within the total study group and each subset. Significant threshold established at $\alpha = 0.00556$.

4.4.1 Risk Score Distribution

The results of the Risk Score averages presented above can also be reflected in the distribution of the Risk Score, which is shown in the histograms below. Now that the Risk Score averages for the last 8 hours are known, the focus can be shifted towards what is happening within those last 8 hours.

Survive vs Death

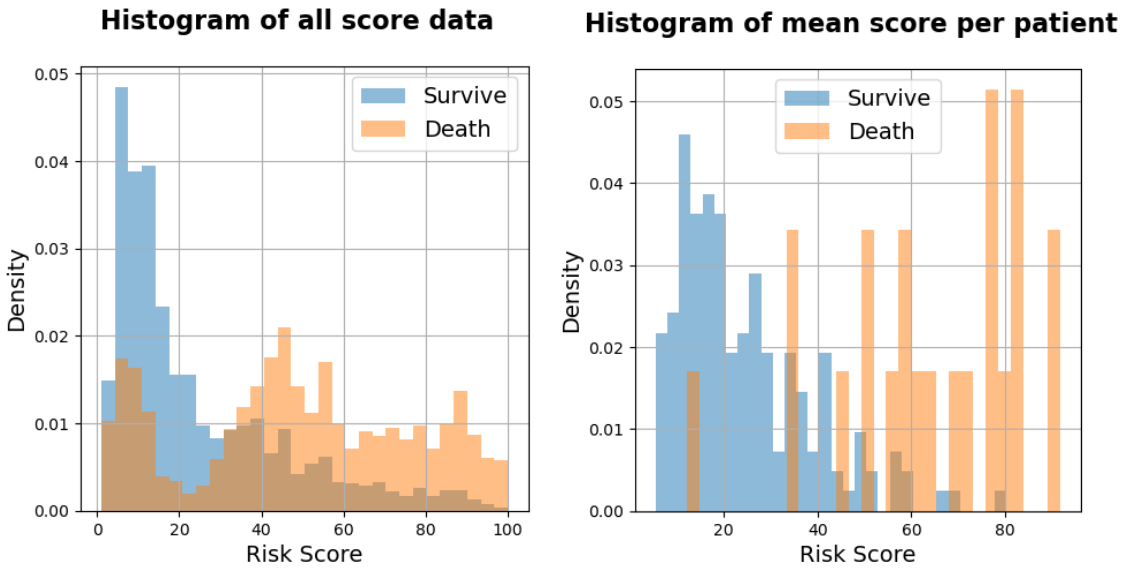

Figure 4.4: Histograms to visualize the total distribution of the Risk Score for the Survive and Death subset.

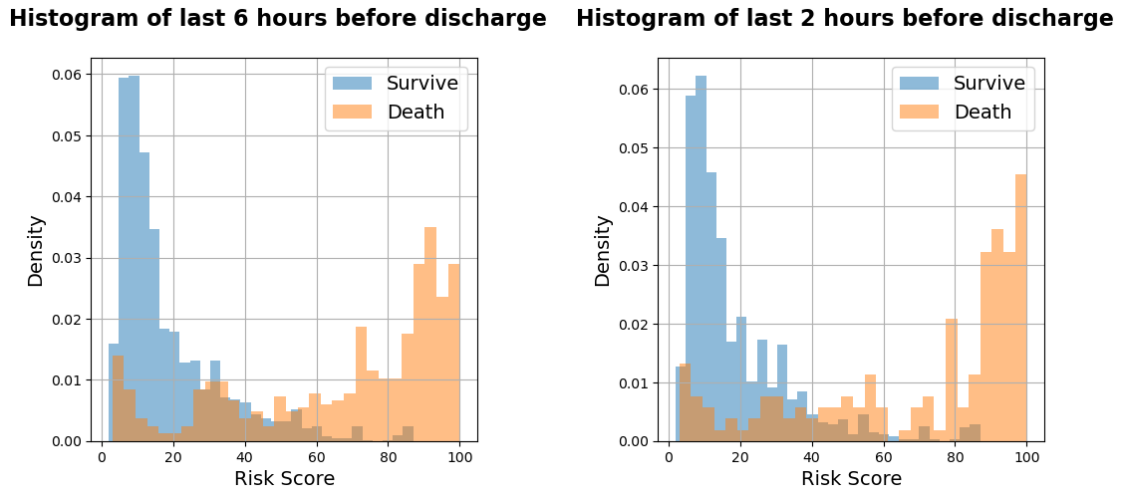

Figure 4.5: Histograms to visualize the distribution of the Risk Score for the Survive and Death subsets in the last six and two hours before discharge/time of death.

Comparing the Survive subset to the Death subset in Figure 4.4, the density of lower Risk Scores is higher in the Survive group. There is overlap, but that overlap decreases closer to the time of discharge/death as seen in Figure 4.5. There is more overlap between the CIE and no CIE subsets as observed in Figures 4.6 and 4.7.

No CIE vs CIE

No CIE compared to CIE gives less distinguishable results, with a great overlap that only decreases slightly closer to the time of discharge.

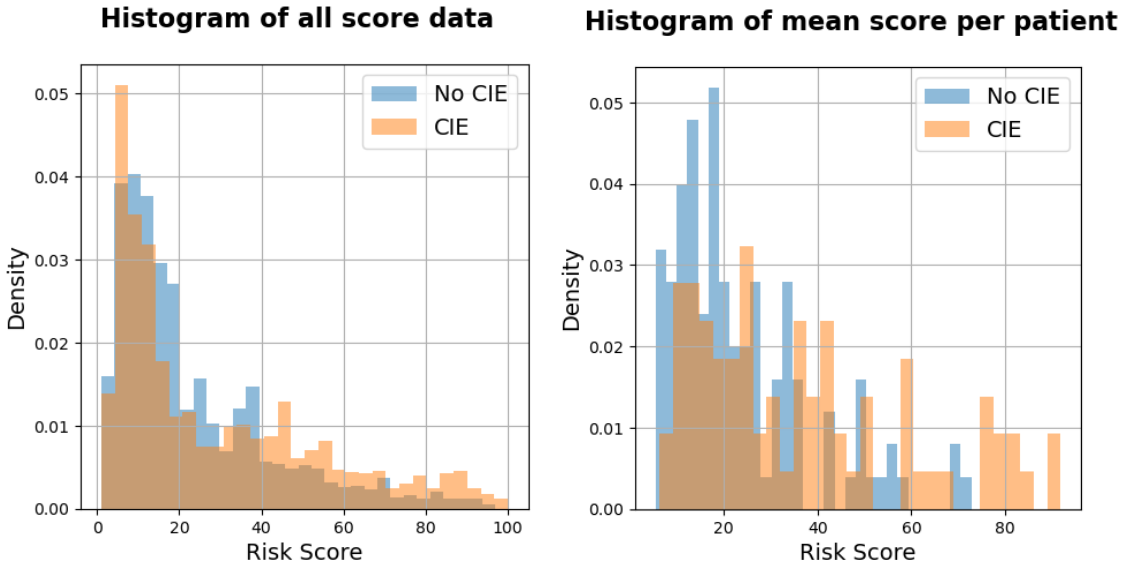

Figure 4.6: Histograms to visualize the total distribution of the Risk Score for the No CIE and CIE subsets.

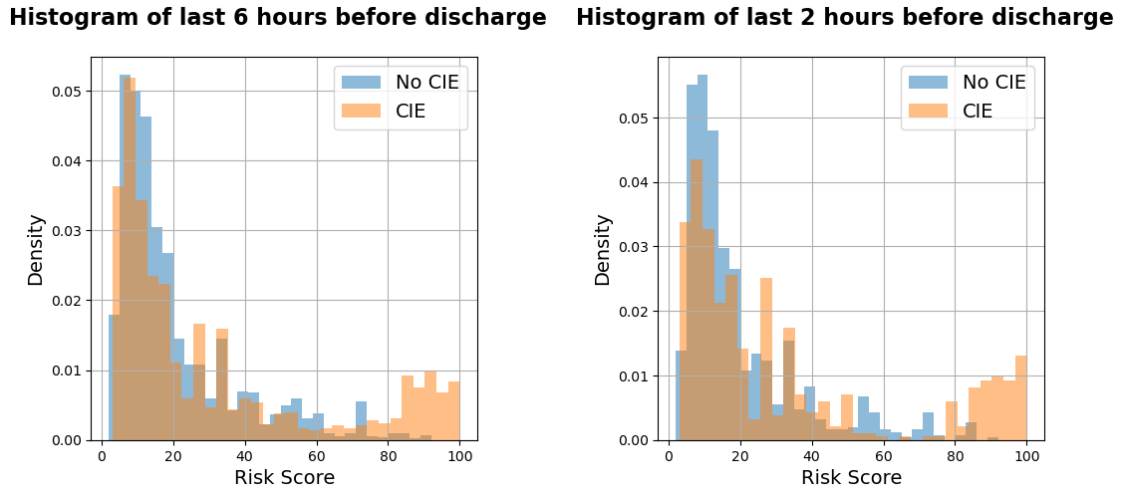

Figure 4.7: Histograms to visualize the distribution of the Risk Score for the No CIE and CIE subsets in the last six and two hours before discharge/time of death.

4.5 Risk Score around CIE

In Figures 4.8 and 4.9 it shows the Risk Score around the time point of all CIE or ECM (t=0). Figure 4.8 compares the CIE group with the group without CIE, which has a significant difference ($p < 0.001$) from one hour before the CIE or ECM to one hour after. The other time points have not been tested against significance, but the figure shows a clear difference between the means for the entire five hours before and after the CIE or ECM. In the background of the figure, the individual patient Risk Score progressions are visualized in different colors per CIE category. Each of these categories is highlighted in Figures 4.10-4.14 below.

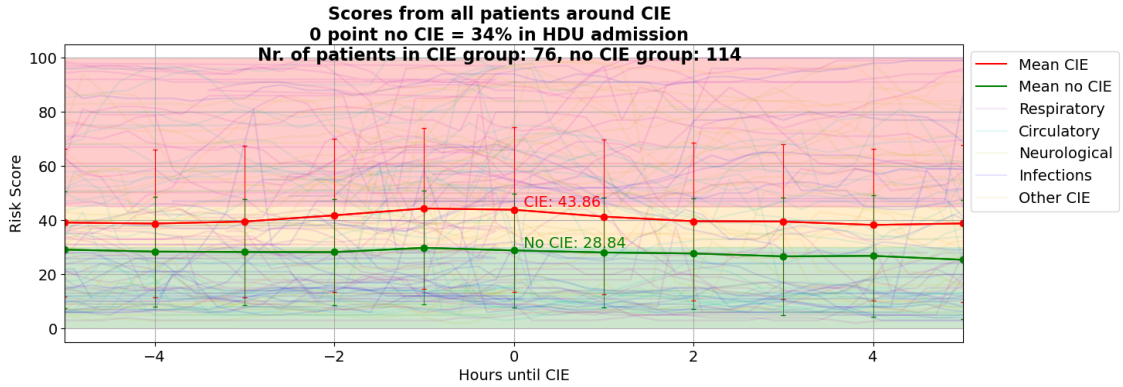

Figure 4.8: Visualizing the Risk Score around the time point of CIE or ECM for the CIE group (red) and non-CIE group (green) respectively. In the background are all individual Risk Score lines around the same time point.

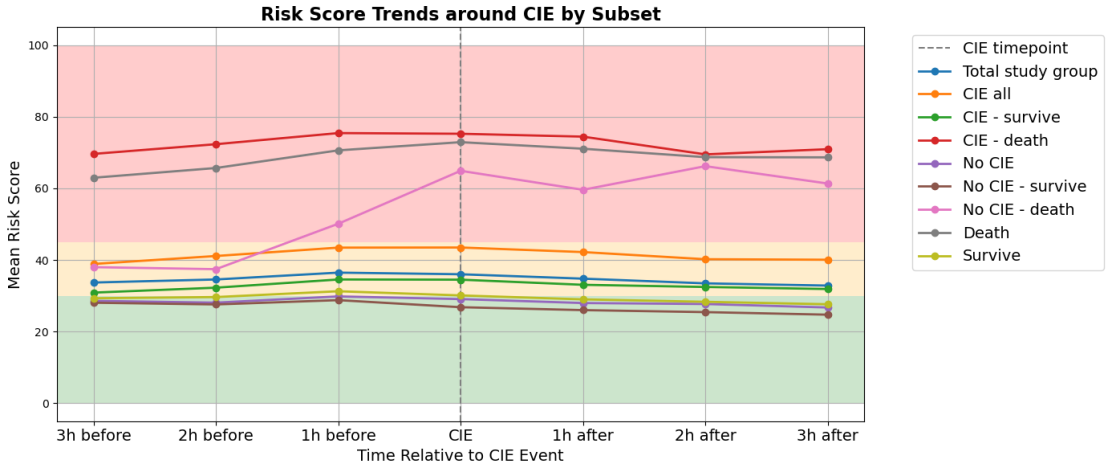

Figure 4.9: Visualizing the Risk Score around the time point of CIE or ECM for each subset.

Figure 4.9 visualizes the mean Risk Score around the CIE or ECM per subset. Like the score averages, the Risk Score is highest in the CIE - death subset and lowest in the No CIE - survive.

Figures 4.10-4.14 show the Risk Score around each CIE category, from the largest group (respiratory CIE) to the smallest (other CIE), with their subcategories as different colors in the background. All categories are compared with the same No CIE mean, where the neurological type shows the greatest difference at the point of CIE, although it was not significantly tested. Infections CIE show the lowest difference compared to the No CIE.

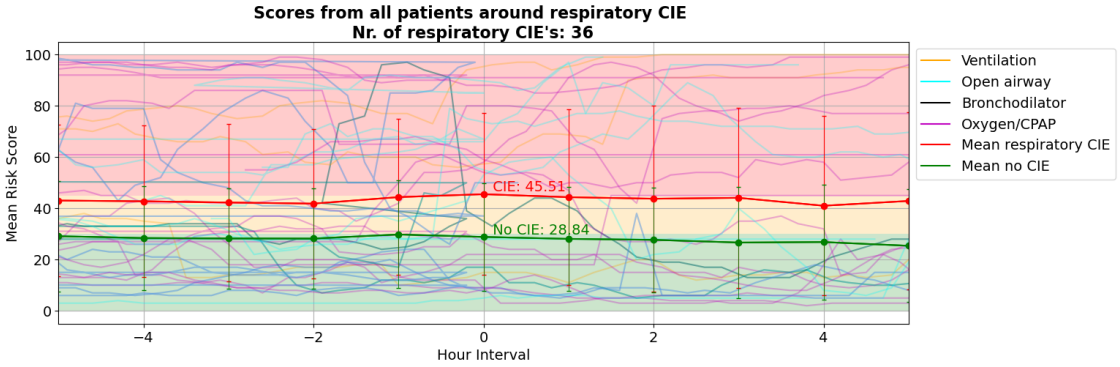

Figure 4.10: Visualizing the Risk Score around the time point of a respiratory CIE or ECM for the respiratory CIE group (red) and non-CIE group (green) respectively. In the background are all individual Risk Score lines around the same time point.

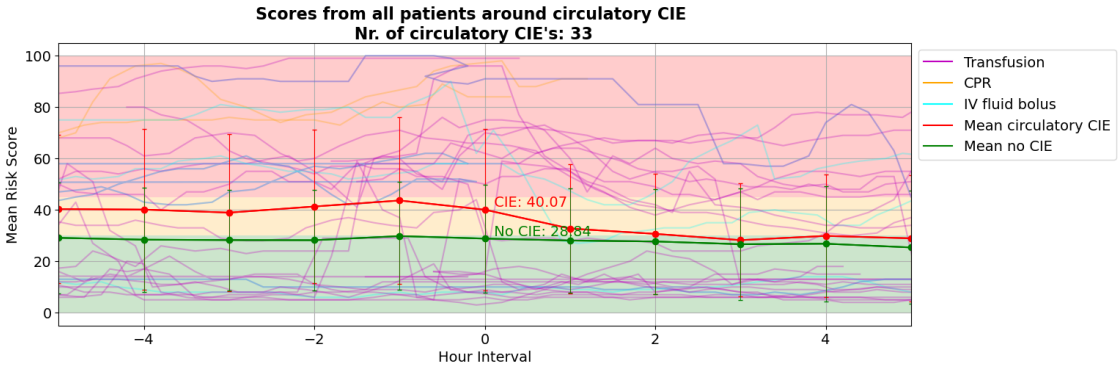

Figure 4.11: Visualizing the Risk Score around the time point of a circulatory CIE or ECM for the circulatory CIE group (red) and non-CIE group (green) respectively. In the background are all individual Risk Score lines around the same time point.

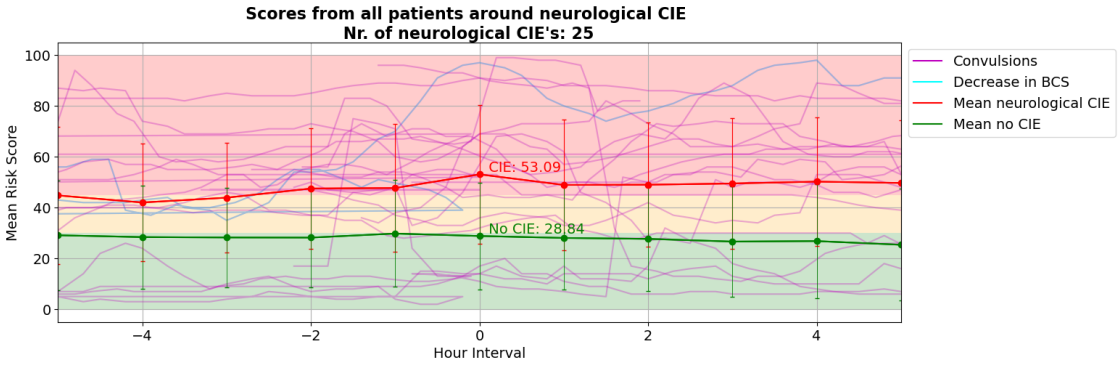

Figure 4.12: Visualizing the Risk Score around the time point of a neurological CIE or ECM for the neurological CIE group (red) and non-CIE group (green) respectively. In the background are all individual Risk Score lines around the same time point.

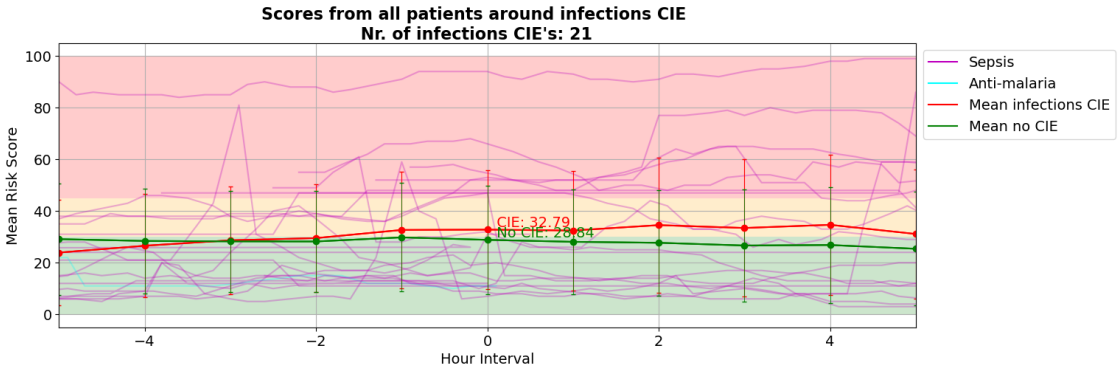

Figure 4.13: Visualizing the Risk Score around the time point of an infections CIE or ECM for the infections CIE group (red) and non-CIE group (green), respectively. In the background are all individual Risk Score lines around the same time point.

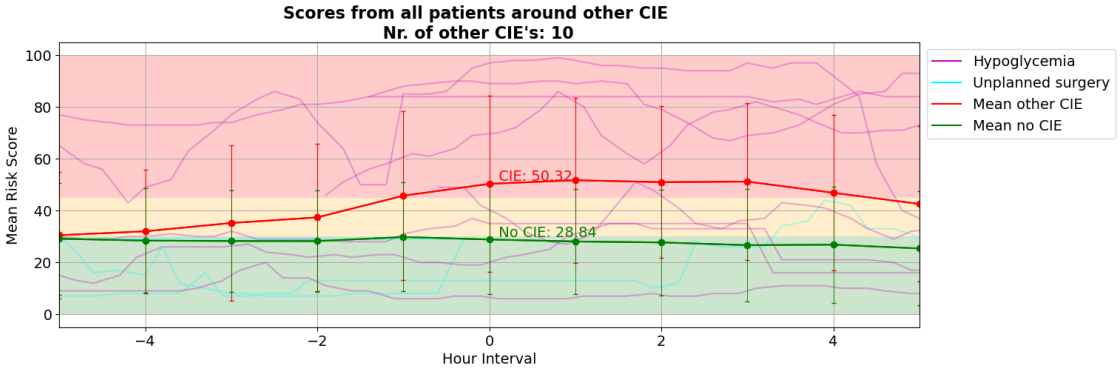

Figure 4.14: Visualizing the Risk Score around the time point of another CIE or ECM for the other CIE group (red) and non-CIE group (green) respectively. In the background are all individual Risk Score lines around the same time point.

4.6 Distinguishing capacities of three types of Risk Score

Three types of Risk Score were tested for their distinguishability using the ROC curve, AUROC, confusion matrix, and accuracy. The ROC curves and the AUROC are shown in Figures 4.15,4.16 and 4.17, the accuracies in Table 4.5 and the confusion matrices are in Appendix 1 in Figures A.4-A.9. RiskScore (blue line) refers to the standard Risk Score as used throughout this entire report, risk_score _adjusted_noNIPB_norm (orange line) refers to the Risk Score without the blood pressure (NIBP) that is normalized to match the standard Risk Score, and risk_score _adjusted_noAVPU_norm (green line) is the Risk Score calculated without any manual input, so no NIBP and no AVPU, and also normalized.

Almost all of these graphs include all the Risk Score points for all patients who passed away as expected high-risk score, except for the second graph in Figure 4.15, which only includes the Risk Score points from the last eight hours before death. All graphs in Figures 4.17 and 4.16 include, on top of all death points, also the Risk Score points from different hours around the CIE. So, all Risk Score points from all patients who passed away, plus the points either 4, 2, or 1 hour(s) before and after the CIE, or all scores from all patients who had one or more CIE.

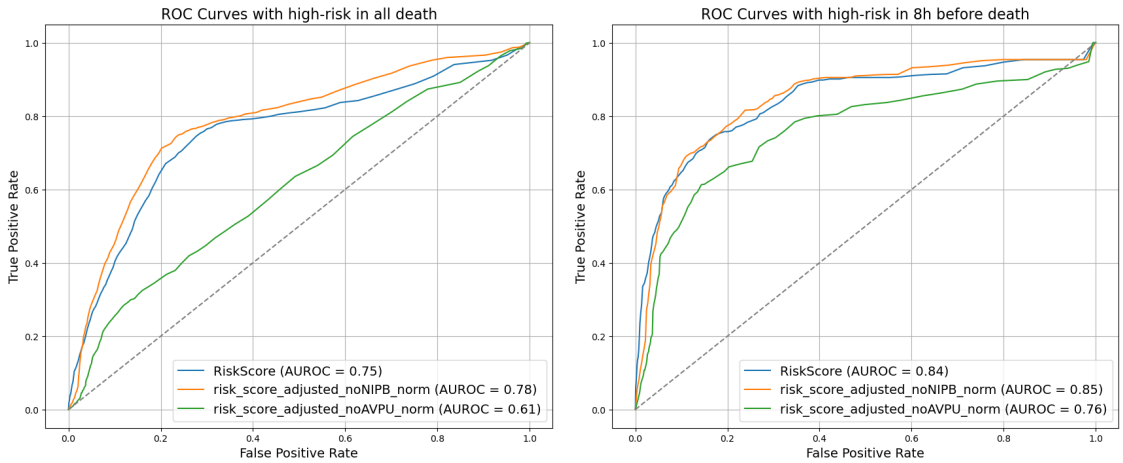

Figure 4.15: ROC curves and AUROC of all three different types of Risk Score for three different sets of expected high-risk score. The first graph shows all score points of all patients who passed away, the second graph only contains their score points in the last 8 hours before death.

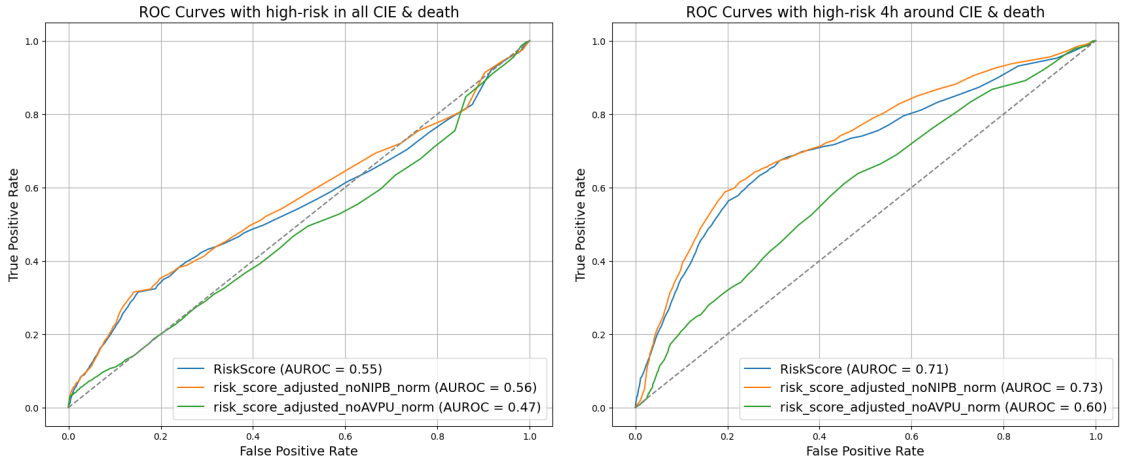

Figure 4.16: ROC curves and AUROC of all three different types of Risk Score for three different sets of expected high-risk score. First, all death scores (the first graph in 4.15) plus all scores from all patients who had one or more CIE. The second graph shows only the 4 hours around the CIE, plus all the death points.

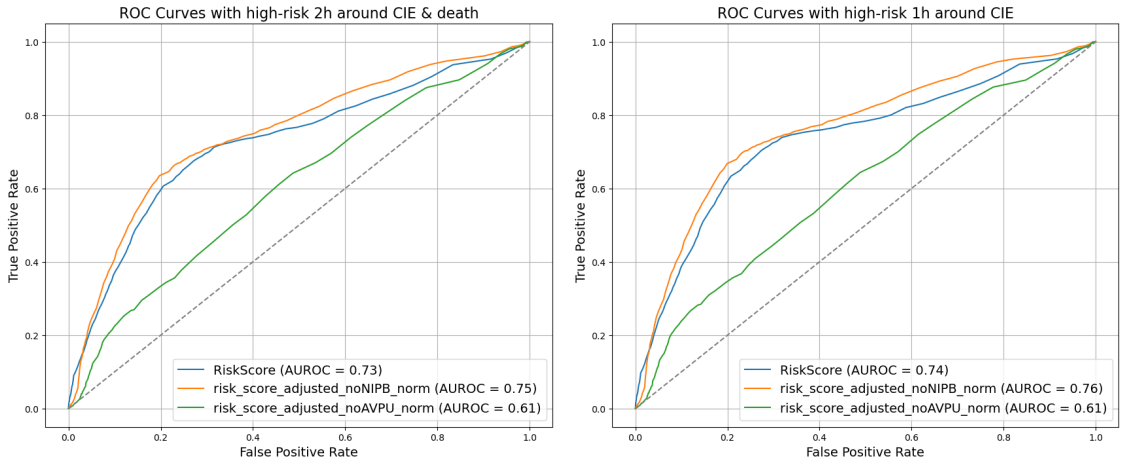

Figure 4.17: ROC curves and AUROC of all three different types of Risk Score for three different sets of expected high-risk score. All these graphs include all score points from all patients who passed away, plus different hours around CIE, from ± 2 hours on the left, and ± 1 hour in the graph on the right.

Risk Score	**All CIE + Death (%)**	**CIE ± 4h (%)**	**CIE ± 2h (%)**	**CIE ± 1h (%)**	**All Death (%)**	**8h before Death (%)**
Normal RS	50.16	79.45	80.28	80.69	81.02	80.56
Without BP	52.88	72.82	73.24	73.43	73.63	70.04
No BP + AVPU	43.52	74.76	75.93	76.62	77.30	80.44

Table 4.5: Accuracy of the three different types of Risk Score in percentage for different sets of when the score is expected to be high-risk. All points of the patients who passed away are included in all sets except for the last set, 8 hours before death.

What can be observed in these graphs is that the standard Risk Score and the score without NIBP show similar results in the shape of their ROC curve, with the Score without BP having a higher AUROC than the standard score everywhere. However, the accuracies are more in favor of the standard score. The highest AUROC (0.85) was reached for the score without NIBP, when only considering the last eight hours before death as expected high-risk scores. This set (last eight hours) gives the highest AUROC of all three different Risk Scores, but not the highest accuracies. Only the score without NIBP and AVPU has its highest accuracy (80.44%) in the last 8h set, while the other two scores have their highest accuracy for the 'all death scores' set (81.02 and 73.63%). The lowest AUROC (0.47) comes from the Risk Score without NIBP and AVPU in the 'all CIE and all death scores' set, which also has the lowest accuracies for the standard Risk Score and the score without NIBP and AVPU. The lowest accuracy for the score without NIBP is in the last 8h set.

5
Discussion

The Risk Score was developed to assign a risk level to patients and track their deterioration or progression. This study aimed to evaluate the performance of the Risk Score and observe the challenges in the use of the Risk Score in the HDU in the pediatric ward in Zomba, Malawi. The study population was 192 children between 1 month and 5 years old with an almost equal distribution of males/females.
In the IMPALA 2.0 study in 2023, in the same hospital and period as IMPALA 3.0, the average HDU admission time was 64.8 hours, which is 17.9 hours shorter than the HDU admission time of this study. Mortality also increased from 11 to 12% from IMPALA 2.0 to 3.0. One of the possible reasons for this is the increase in Malaria cases from 2022 onward. Malaria cases have increased in Malawi from 4.6 million in 2022, to 9.4 million in 2024. The number of deaths due to Malaria also increased from 1829 in 2022 to 2252 in 2024, as reported by the Malawi Ministry of Health [27]. This increase was also reported by nurses and clinicians in ZCH, who saw a longer-lasting peak than usual in the rainy season. Due to this, it is hard to say something about the effect of the addition of the Risk Score on mortality, CIE occurrence, and HDU duration. The CIE occurrence is also hard to compare to IMPALA 2.0 since 2.0 also takes the CIE mentioned at study admission into account, and it is not clear if they counted death as CIE.
In the individual patient plots, a lot is happening: noisy vital sign signals, missing score, CIE types, alarms, and their responses. These plots were observed and discussed every week during the study with the project group. What was noticed is that noisy signals are common, especially in pediatric patients who move a lot, but did not appear to influence the Risk Score much. However, missing Risk Score was a more common problem, with 18.6% of the score missing. This occurred both at admission (11.6%) and towards discharge/death (21.6%). The missing score at admission occurred less and less after mentioning it to the research nurses, who paid more attention to measuring and filling in the manual inputs (AVPU, SBP) properly. This resulted in a coverage of 95.2% for the SBP, which is the best of all vital signs. The missing score towards the end had different causes, in addition to the missing AVPU and/or SBP, like at admission. Although the buffer for the manual inputs is 16 hours, it is much shorter for the continuous vital signs, where it is only one hour. Especially SpO2 turned out to be problematic, which was missing 36.2% of the time. This was a difficult issue to solve, as the SpO2 sensor is difficult to keep on a child, especially in dark-skinned patients [28]. Guardians were trained by the nurses to adjust the SpO2 sensor when they saw that the SpO2 had a technical alarm and/or the sensor fell off. Most of the alarms also came from the oxygen saturation. The response to all alarms was done mostly by guardians.

This shows that the role of guardians is important in the use of the monitoring system; however, the role the guardians play in the Risk Score is yet unclear and will be researched.
Another thing noticed in the individual patient plots was that the deterioration right before death was so fast that the score could not pick up on that deterioration with the 15-minute interval. It is unclear if it would help if the score could report that deterioration faster, or if it is too late to save the patient by then. Other than that, the 15-minute time interval at which the score is calculated did not show any problems.
The last thing observed in the individual plots from patients who passed away is that sometimes vital signs were showing up after the reported time of death. This can stress the nurses or guardians, so it is important that the patient is immediately disconnected from the monitor after death is reported. However, feedback from the nurses said that sometimes it took some time for a clinician to come to the patient and confirm death, leaving the patient on the monitor after suspected passing. This is just reality and cannot really be prevented, just a reality of why there were vital signs and score data after the reported time of death.
Then the score coverage in general, which was highest (85.5%) in the group who had both one or more CIE and passed away. This makes sense since that is the group of patients who are most sick and require the most care, in theory. It is thus good to see that the patients who needed the most care were monitored most accurately. This group also had the highest average score, with the highest amount of high-risk score. The lowest score coverage is found in the five patients who had no CIE but passed away, 76.7%. There are only two patients in this group who did not have a Risk Score for the entire time, P-041 and P-144. For P-041, the SpO2 and thus the Risk Score were missing in the last 12 hours before passing away, where the last recorded score was low-risk. For the other patient, 144, it is suspected that there is some score data missing, since the data gives very inconsistent results. This makes the No CIE - survive the group with the lowest score coverage, which matches the theory named before about the opposite group, which had the highest coverage. This also holds for the score average and the lowest amount of high-risk score.
Looking at the score distribution, it was observed that most patients who passed away had a strong increase in Risk Score from the last six hours before the time of death, and even more in the last two hours. Comparison of the first to the last 8 hours showed an increase in the group of patients who died. This was not significant, possibly due to the few patients who passed away with a low or missing score. All surviving subsets and the total study group did show significance there, with a Risk Score lower in the last 8 hours than in the first. This is a good sign to confirm that the patients were discharged after their Risk Score decreased. The CIE group did not show much different results when it comes to the score distribution compared to the no CIE group. However, when looking at the score around the point of CIE, it shows more prominent results. There, it showed a significant difference between the average Risk Score of the CIE group at one hour before and after the CIE, compared to the Estimated Crash Point in the no CIE group. However, it is not likely that this significant difference is only caused by the CIE.

The group that passed away is included in the CIE group, and in Figure 4.9 it can be observed that death has a much higher impact and that the subset of patients who had one or more CIE but survived showed even a lower average Risk Score than the Total study group. These results do not show the isolated results of the influence of just the CIE accurately. Within the CIE results, some types or categories show more differences than others. The neurological CIEs have the biggest difference before, during, and after the CIE. These were mainly convulsions (seizures), matching the symptoms of severe/cerebral Malaria. It is also the only CIE category that clearly peaks at the point of CIE. The second largest difference came from 'Other' CIE, mainly hypoglycemia. These few cases (10) showed an increase in Risk Score from two hours before the CIE and kept rising even after the CIE, on the same level as the neurological CIE. Respiratory CIE also shows a relatively high difference, but does not show any peak or strong incline or decrease. Circulatory CIE has its peak one hour before CIE, and the score decreases after that. Since circulatory CIEs are mainly blood transfusions or IV fluids, it is good to see that those events influence the score, since that was one of the events the original physiological score research was based on. Infections CIE showed little to no difference from the no CIE group. All these different CIE results make it difficult to draw one conclusion or see the effect of the Risk Score during CIEs in general. Some of the CIEs are not just short events; for example, oxygen support, ventilation, infections, and hypoglycemia are events that have an impact for a longer time. Maybe without those long-lasting events, the results will get clearer and more accurate to view around the time point of the CIE.

Lastly, we will look at the distinguishing capacities of the Risk Score and its two other types with one or more fewer input. As a reference, other PDS in LMIC such as PEWS and pSOFA reached overall AUROC values of 0.81-0.86 when applied in the PICU [29]. However, it is hard to compare the results of this study with those of other studies, as circumstances, validation, scale, variable input, and outcome parameters are different. The highest AUROC achieved in this study was 0.85, when testing the distinguishability of the score without blood pressure for surviving patients and the last eight hours before death. The Risk Score had an AUROC of 0.84 for the same situation as the highest AUROC. This means that both those scores are as good at distinguishing as other studies from the literature. However, a more realistic and desired scenario would be a score that would perform well in distinguishing patients who will pass away from the beginning. For this situation, the AUROC for the Risk Score is 0.75, and for the score without BP, 0.78. This is below the desired performance, so the scores are not yet ready to potentially predict death from admission. However, the Risk Score achieves the highest accuracy overall in the 'all death' situation, with 81.0%. This means that in 81% of all score cases, there is a high-risk score for the patients who passed away and a low- or medium-risk score for patients who survived: the 'correct' expectations. This accuracy is slightly higher compared to only the last 8h before death scores accuracy (80.6%). It is not clear why the accuracy and AUROC do not show the same result. The score without BP and AVPU performs the worst out of all three scores in this situation, both in AUROC and accuracy.

When also including the scores around CIE as expected high-risk scores, the score without BP and AVPU starts performing better than the score without BP in the accuracy, but not in the AUROC. The closer the interval of scores around the CIE gets, the better all three scores are performing. The AUROC for the Risk Score, when expecting the scores of 1 hour before and after CIE and all death scores to be high-risk, is 0.74 and 0.76 for the score without BP. This is very close to the performance of only all death scores, also with the accuracies. Maybe a combination of the last 8 hours before death and one hour before and after CIE could give the best overall results, but that situation has not been tested. In general, it can be said that the Risk Score performs similarly to the existing literature when expecting the last 8 hours before death as high-risk. It also shows potential for the scores without one or more manual inputs, especially the score without BP. The score without BP and AVPU has potential if CIEs are included, but it needs fine-tuning. For example, the parameters in the sigmoid function (eq. 2.5) can be adjusted to improve scaling and distinguishing capabilities.

5.1 Limitations

There are many directions in which this study could have gone, but due to knowledge, tools, and time restrictions, it had to be limited in some way. The study population is limited by the choices made for the inclusion and exclusion criteria mentioned in the Methods section, resulting in a study population that focuses on an age group below five years in one hospital during a certain period of the year. These conditions were very similar to earlier research.
On top of that, the statistical analysis was insufficient, if even correct. Multilevel analysis should have been applied in the statistical analyses involving the subsets. One patient is in several subsets, requiring an adjusted statistical analysis compared to what is done now. This multilevel statistics was not applied due to its time extensiveness. The difference in Risk Scores in the first and last 8 hours of the study of was now tested with a group-level averages t-test, instead we could have looked at the difference in scores per individual patient, and then computing the average of those differences. This might show a significance in the difference in the Death subset, which shows a clear increase in Risk Score, but not significant in the way it was computed now. In addition to that, the statistics applied to the Risk Score around CIE were not extensive.

5.2 Future research

Since this is the first study on the Risk Score, there is a lot of future research that can be done. First of all, the study data can be elaborated by including other hospitals, extending the study duration, or moving it to a different period in the year, or broadening the age inclusion criteria. Since the study was conducted during the rainy season, when the occurrence of Malaria is higher than in the rest of the year, the mortality rate is probably relatively higher than in other periods of the year.

It would be interesting to see how the Risk Score performs when the mortality and occurrence of malaria change. When the size of the data set increases, the potential for machine learning arises. With a large data set and machine learning, the Risk Score might be able to be used as a predictive tool. However, more research and changes are needed before a prediction model can be the aim. The SpO2 sensor could be improved for future research. It is a challenging issue to tackle, but as Al-Halawani mentions in his study, it is possible to optimize the sensor calibration algorithm, for example [28].
To see more impact of the Risk Score during CIE, future research can change which CIE to include, and which subsets to compare. Now, the comparison of the CIE subset vs the no CIE group subset was studied, but to really see the influence of just the CIE, it could be worth it to compare the CIE survive vs the no CIE survive subset. In that case, the score is not higher in the CIE group due to the already higher score of the death subset within that group. It is also an option to look into only researching the short-term CIE, such as convulsions.
Lastly, future research can be done on the scores without one or more manual inputs.

6

Conclusion

The Risk Score performed well in the last 6-8 hours before discharge in the distinction of patients between survive and death. In other situations, the Risk Score performed moderately, with the lowest performance found around CIE. There was a clear difference in the behavior of the score in the first hours after admission and the last hours before discharge in the surviving patient group. With further optimization on Risk Score availability and more research to validate it, the Risk Score can be a powerful tool for clinicians to guide them in the treatment of surviving patients, high-risk patients, and certain CIE.

Bibliography

[1] T. M. Slusher, A. W. Kiragu, L. T. Day, A. R. Bjorklund, A. Shirk, C. Johannsen, *et al.*, "Pediatric critical care in resource-limited settings—overview and lessons learned," *Frontiers in Pediatrics*, vol. 6, 2018. DOI: 10.3389/fped.2018.00049.

[2] C. Van Zyl, M. Badenhorst, S. Hanekom, and M. Heine, "Unravelling 'low-resource settings': a systematic scoping review with qualitative content analysis," *BMJ Global Health*, vol. 6, no. 6, p. 5190, 2021. DOI: 10.1136/BMJGH-2021-005190.

[3] L. I. Losonczy, A. Papali, S. Kivlehan, E. J. Calvello Hynes, G. Calderon, A. Laytin, *et al.*, "White Paper on Early Critical Care Services in Low Resource Settings," *Annals of Global Health*, vol. 87, no. 1, pp. 105–106, 2021. DOI: 10.5334/AOGH.3377.

[4] United Nations Children's Fund (UNICEF), "Levels Trends in Child Mortality Report 2023," UNICEF, Tech. Rep., 2023, p. 96.

[5] D. Mwale, L. Manda-Taylor, J. Langton, A. Likumbo, M. B. van Hensbroek, J. Calis, *et al.*, "The role of healthcare providers and caregivers in monitoring critically ill children: A qualitative study in a tertiary hospital, southern malawi," *BMC Health Services Research*, vol. 24, pp. 1–11, 1 Dec. 2024. DOI: 10.1186/S12913-024-11050-8/FIGURES/2.

[6] G. Laucaityte, F. W. Fahnehjelm, D. Akongo, E. Tenywa, K. Hildebrand, M. Kyangwa, *et al.*, "Capacity for delivery of paediatric emergency care and the current use of emergency triage, assessment and treatment in health facilities in the busoga region, uganda—a mixed methods study," *PLOS Global Public Health*, vol. 4, 9 Sep. 2024. DOI: 10.1371/JOURNAL.PGPH.0003666.

[7] J. M. Tappan, J. Daniels, B. Slavin, J. Lim, R. Brant, and J. M. Ansermino, "Visual cueing with context relevant information for reducing change blindness," *Journal of Clinical Monitoring and Computing*, vol. 23, no. 4, pp. 223–232, 2009. DOI: 10.1007/s10877-009-9186-8.

[8] G. Riebeek, "The effect of an improved monitoring system on the quality of care at a pediatric high-dependency unit in malawi," Vrije Universiteit Amsterdam, AIGHD, Tech. Rep., Jun. 2023.

[9] J. M. Ansermino, J. P. Daniels, R. T. Hewgill, J. Lim, P. Yang, C. J. Brouse, *et al.*, "An evaluation of a novel software tool for detecting changes in physiological monitoring," *Anesthesia and Analgesia*, vol. 108, no. 3, pp. 873–880, 2009. DOI: 10.1213/ane.0b013e318193ff87.

[10] G. Gasciauskaite, J. Lunkiewicz, T. R. Roche, D. R. Spahn, C. B. Nöthiger, and D. W. Tscholl, *Human-centered visualization technologies for patient mon-*

itoring are the future: A narrative review, Dec. 2023. DOI: 10.1186/s13054-023-04544-0.

[11] S. L. Chong, M. S. L. Goh, G. Y. K. Ong, J. Acworth, R. Sultana, S. H. W. Yao, *et al.*, "Do paediatric early warning systems reduce mortality and critical deterioration events among children? a systematic review and meta-analysis," *Resuscitation plus*, vol. 11, Sep. 2022. DOI: 10.1016/J.RESPLU.2022.100262.

[12] A. H. S. Fang, W. T. Lim, and T. Balakrishnan, "Early warning score validation methodologies and performance metrics: a systematic review," *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, pp. 1–7, 2020. DOI: 10.1186/S12911-020-01144-8/FIGURES/3.

[13] V. Lambert, A. Matthews, R. MacDonell, and J. Fitzsimons, "Paediatric early warning systems for detecting and responding to clinical deterioration in children: A systematic review," *BMJ Open*, vol. 7, no. 3, 2017. DOI: 10.1136/bmjopen-2016-014497. eprint: https://bmjopen.bmj.com/content/7/3/e014497.full.pdf.

[14] T. Baker, "Critical care in low-income countries," *Tropical medicine international health : TM IH*, vol. 14, no. 2, pp. 143–148, 2009. DOI: 10.1111/J.1365-3156.2008.02202.X.

[15] S. R. Brown, D. Martinez Garcia, and A. Agulnik, "Scoping review of pediatric early warning systems (pews) in resource-limited and humanitarian settings," *Frontiers in Pediatrics*, vol. 6, 2019. DOI: 10.3389/fped.2018.00410.

[16] S. Sridhar, A. Schmid, F. Biziyaremye, S. Hodge, N. Patient, and K. Wilson, "Implementation of a pediatric early warning score to improve communication and nursing empowerment in a rural district hospital in rwanda," *Global health, science and practice*, vol. 8, pp. 838–845, 4 Dec. 2020. DOI: 10.9745/GHSP-D-20-00075.

[17] S. L. Rosman, V. Karangwa, M. Law, M. C. Monuteaux, C. D. Briscoe, and N. McCall, "Provisional validation of a pediatric early warning score for resource-limited settings," *Pediatrics*, vol. 143, 5 May 2019. DOI: 10.1542/PEDS.2018-3657.

[18] N. Kemps, N. Holband, N. P. Boeddha, A. Faal, A. E. Juliana, G. A. Kavishe, *et al.*, "Validation of the emergency department-paediatric early warning score (ed-pews) for use in low- and middle-income countries: A multicentre observational study," *PLOS global public health*, vol. 4, e0002716, 3 Mar. 2024. DOI: 10.1371/JOURNAL.PGPH.0002716.

[19] M. Kumwenda, R. Assies, I. Snik, G. Chatima, J. Langton, Y. Chimalizeni, *et al.*, "Identifying critically ill children in malawi: A modified qsofa score for low-resource settings," *PLOS global public health*, vol. 4, e0002388, 1 Jan. 2024. DOI: 10.1371/JOURNAL.PGPH.0002388.

[20] GOAL 3. "The first results of the implementation in malawi." (2023), [Online]. Available: https://www.goal3.org/article/the-first-results-of-the-implementation-in-malawi (visited on 06/12/2024).

[21] M. Levin, A. J. Cunnington, C. Wilson, S. Nadel, H. J. Lang, N. Ninis, *et al.*, "Effects of saline or albumin fluid bolus in resuscitation: Evidence from re-analysis of the feast trial," *The Lancet Respiratory Medicine*, vol. 7, pp. 581–593, 7 Jul. 2019. DOI: 10.1016/S2213-2600(19)30114-6.

[22] F. Hoffmann, M. Schmalhofer, M. Lehner, S. Zimatschek, V. Grote, and K. Reiter, "Comparison of the avpu scale and the pediatric gcs in prehospital setting," *Prehospital Emergency Care*, vol. 20, no. 4, pp. 493–498, 2016, PMID: 26954262. DOI: 10.3109/10903127.2016.1139216. eprint: https://doi.org/10.3109/10903127.2016.1139216.

[23] World Health Organization, *Medical devices: Managing the mismatch: An outcome of the priority medical devices project*, Jul. 2010.

[24] AIGHD, "ISRCTN - ISRCTN71392921: Innovative monitoring in paediatrics in low-resource settings: an aid to save lives?" *COMREC*, no. November, 2023.

[25] P. A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, and J. G. Conde, "Research electronic data capture (redcap)—a metadata-driven methodology and workflow process for providing translational research informatics support," *Journal of Biomedical Informatics*, vol. 42, no. 2, pp. 377–381, 2009. DOI: https://doi.org/10.1016/j.jbi.2008.08.010.

[26] P. A. Harris, R. Taylor, B. L. Minor, V. Elliott, M. Fernandez, L. O'Neal, *et al.*, "The redcap consortium: Building an international community of software platform partners," *Journal of Biomedical Informatics*, vol. 95, p. 103 208, 2019. DOI: https://doi.org/10.1016/j.jbi.2019.103208.

[27] C. Tinto, *Malawi records 9.4 million malaria cases, 2,252 deaths in 2024 – capital radio malawi*, Apr. 2025.

[28] R. Al-Halawani, P. H. Charlton, M. Qassem, and P. A. Kyriacou, "A review of the effect of skin pigmentation on pulse oximeter accuracy," *Physiological Measurement*, vol. 44, 5 May 2023. DOI: 10.1088/1361-6579/ACD51A.

[29] D. A. van den Brink, I. S. de Vries, M. Datema, L. Perot, R. Sommers, J. Daams, *et al.*, "Predicting clinical deterioration and mortality at differing stages during hospitalization: A systematic review of risk prediction models in children in low- and middle-income countries," *The Journal of Pediatrics*, vol. 260, p. 113 448, 2023. DOI: https://doi.org/10.1016/j.jpeds.2023.113448.

A

Appendix 1

A.1 P-values Risk Score averages

	Total Study Group	CIE all	CIE survive	CIE death	No CIE	No CIE survive	No CIE death	Death	Survive
Full Dataset	-	0.00965	0.60902	0.00000	0.00708	0.00053	0.11548	0.00000	0.00863
CIE all	0.00965	-	0.00680	0.00000	0.00002	0.00000	0.27314	0.00000	0.00002
CIE-survive	0.60902	0.00680	-	0.00000	0.09482	0.02554	0.10076	0.00000	0.13367
CIE-death	0.00000	0.00000	0.00000	-	0.00000	0.00000	0.20648	0.53363	0.00000
No CIE	0.00708	0.00002	0.09482	0.00000	-	0.50359	0.06584	0.00000	0.71209
No CIE-survive	0.00053	0.00000	0.02554	0.00000	0.50359	-	0.05815	0.00000	0.25331
No CIE-death	0.11548	0.27314	0.10076	0.20648	0.06584	0.05815	-	0.31388	0.07054
Death	0.00000	0.00000	0.00000	0.53363	0.00000	0.00000	0.31388	-	0.00000
Survive	0.00863	0.00002	0.13367	0.00000	0.71209	0.25331	0.07054	0.00000	-

Table A.1: P-values for comparisons between study subgroups in total Risk Score averages of the total study period. Significant threshold established at $\alpha = 0.00139$.

	Total Study Group	CIE all	CIE survive	CIE death	No CIE	No CIE survive	No CIE death	Death	Survive
Full Dataset	-	0.00711	0.25437	0.00039	0.01816	0.00310	0.06277	0.00004	0.13541
CIE all	0.00711	-	0.21059	0.01223	0.00002	0.00000	0.18141	0.00444	0.00018
CIE-survive	0.25437	0.21059	-	0.00207	0.00491	0.00124	0.09765	0.00043	0.02853
CIE-death	0.00039	0.01223	0.00207	-	0.00004	0.00003	0.86796	0.95459	0.00011
No CIE	0.01816	0.00002	0.00491	0.00004	-	0.59740	0.03202	0.00000	0.30591
No CIE-survive	0.00310	0.00000	0.00124	0.00003	0.59740	-	0.02772	0.00000	0.10670
No CIE-death	0.06277	0.18141	0.09765	0.86796	0.03202	0.02772	-	0.89397	0.04265
Death	0.00004	0.00444	0.00043	0.95459	0.00000	0.00000	0.89397	-	0.00001
Survive	0.13541	0.00018	0.02853	0.00011	0.30591	0.10670	0.04265	0.00001	-

Table A.2: P-values for comparisons between study subgroups in the Risk Score average of the first 8 hours in the study. Significant threshold established at $\alpha = 0.00139$.

	Total Study Group	CIE all	CIE survive	CIE death	No CIE	No CIE survive	No CIE death	Death	Survive
Full Dataset	-	0.05332	0.03676	0.00000	0.02497	0.00316	0.14262	0.00000	0.00270
CIE all	0.05332	-	0.00130	0.00000	0.00100	0.00024	0.28429	0.00001	0.00024
CIE-survive	0.03676	0.00130	-	0.00000	0.82123	0.76479	0.08566	0.00000	0.84049
CIE-death	0.00000	0.00000	0.00000	-	0.00000	0.00000	0.11279	0.48659	0.00000
No CIE	0.02497	0.00100	0.82123	0.00000	-	0.50976	0.09041	0.00000	0.57249
No CIE-survive	0.00316	0.00024	0.76479	0.00000	0.50976	-	0.08059	0.00000	0.88038
No CIE-death	0.14262	0.28429	0.08566	0.11279	0.09041	0.08059	-	0.19988	0.08253
Death	0.00000	0.00001	0.00000	0.48659	0.00000	0.00000	0.19988	-	0.00000
Survive	0.00270	0.00024	0.84049	0.00000	0.57249	0.88038	0.08253	0.00000	-

Table A.3: P-values for comparisons between study subgroups in the Risk Score average of the last 8 hours in the study. Significant threshold established at $\alpha = 0.00139$.

	Total Study Group	CIE all	CIE survive	CIE death	No CIE	No CIE survive	No CIE death	Death	Survive
Total study group	-	0.06309	0.70508	0.00000	0.02560	0.00242	0.00475	0.00000	0.03805
CIE all	0.06309	-	0.06377	0.00001	0.00058	0.00006	0.01380	0.00000	0.00082
CIE-survive	0.70508	0.06377	-	0.00000	0.18762	0.05875	0.00232	0.00000	0.26263
CIE-death	0.00000	0.00001	0.00000	-	0.00000	0.00000	0.19961	0.71804	0.00000
No CIE	0.02560	0.00058	0.18762	0.00000	-	0.46956	0.00166	0.00000	0.73220
No CIE-survive	0.00242	0.00006	0.05875	0.00000	0.46956	-	0.00133	0.00000	0.25610
No CIE-death	0.00475	0.01380	0.00232	0.19961	0.00166	0.00133	-	0.28578	0.00225
Death	0.00000	0.00000	0.00000	0.71804	0.00000	0.00000	0.28578	-	0.00000
Survive	0.03805	0.00082	0.26263	0.00000	0.73220	0.25610	0.00225	0.00000	-

Table A.4: P-values for the average score at point of CIE/ECM comparisons between subsets. Significant threshold established at $\alpha = 0.00139$.

	Total study group	CIE all	CIE survive	CIE death	No CIE	No CIE survive	No CIE death	Death	Survive
Total study group	-	0.06879	0.61207	0.00000	0.02903	0.01253	0.12080	0.00000	0.06148
CIE all	0.06879	-	0.05171	0.00000	0.00071	0.00030	0.40005	0.00000	0.00145
CIE survive	0.61207	0.05171	-	0.00000	0.22762	0.14371	0.08703	0.00000	0.37754
CIE death	0.00000	0.00000	0.00000	-	0.00000	0.00000	0.01746	0.40059	0.00000
No CIE	0.02903	0.00071	0.22762	0.00000	-	0.74574	0.04419	0.00000	0.62438
No CIE survive	0.01253	0.00030	0.14371	0.00000	0.74574	-	0.03814	0.00000	0.40404
No CIE death	0.12080	0.40005	0.08703	0.01746	0.04419	0.03814	-	0.03786	0.05581
Death	0.00000	0.00000	0.00000	0.40059	0.00000	0.00000	0.03786	-	0.00000
Survive	0.06148	0.00145	0.37754	0.00000	0.62438	0.40404	0.05581	0.00000	-

Table A.5: P-values for the average Risk Score one hour before point of CIE/ECM comparisons between subsets. Significant threshold established at $\alpha = 0.00139$.

	Total study group	CIE all	CIE survive	CIE death	No CIE	No CIE survive	No CIE death	Death	Survive
Total study group	-	0.07810	0.53445	0.00000	0.02814	0.01975	0.80046	0.00000	0.06997
CIE all	0.07810	-	0.04698	0.00000	0.00082	0.00057	0.75160	0.00005	0.00200
CIE survive	0.53445	0.04698	-	0.00000	0.27322	0.22387	0.65875	0.00000	0.47164
CIE death	0.00000	0.00000	0.00000	-	0.00000	0.00000	0.03631	0.24625	0.00000
No CIE	0.02814	0.00082	0.27322	0.00000	-	0.87261	0.43482	0.00000	0.58511
No CIE survive	0.01975	0.00057	0.22387	0.00000	0.87261	-	0.41340	0.00000	0.47748
No CIE death	0.80046	0.75160	0.65875	0.03631	0.43482	0.41340	-	0.06242	0.50763
Death	0.00000	0.00005	0.00000	0.24625	0.00000	0.00000	0.06242	-	0.00000
Survive	0.06997	0.00200	0.47164	0.00000	0.58511	0.47748	0.50763	0.00000	-

Table A.6: P-values for the average Risk Score two hours before point of CIE/ECM comparisons between subsets. Significant threshold established at $\alpha = 0.00139$.

	Total study group	CIE all	CIE survive	CIE death	No CIE	No CIE survive	No CIE death	Death	Survive
Total study group	-	0.16227	0.43488	0.00000	0.08982	0.06475	0.72839	0.00002	0.10551
CIE all	0.16227	-	0.06745	0.00001	0.00853	0.00611	0.94265	0.00038	0.00946
CIE survive	0.43488	0.06745	-	0.00000	0.54404	0.45888	0.57613	0.00001	0.65889
CIE death	0.00000	0.00001	0.00000	-	0.00000	0.00000	0.05715	0.34944	0.00000
No CIE	0.08982	0.00853	0.54404	0.00000	-	0.87028	0.46445	0.00000	0.80393
No CIE survive	0.06475	0.00611	0.45888	0.00000	0.87028	-	0.44212	0.00000	0.67205
No CIE death	0.72839	0.94265	0.57613	0.05715	0.46445	0.44212	-	0.10551	0.49556
Death	0.00002	0.00038	0.00001	0.34944	0.00000	0.00000	0.10551	-	0.00000
Survive	0.10551	0.00946	0.65889	0.00000	0.80393	0.67205	0.49556	0.00000	-

Table A.7: P-values for the average Risk Score three hours before CIE comparisons between subsets. Significant threshold established at $\alpha = 0.00139$.

	Total study group	CIE all	CIE survive	CIE death	No CIE	No CIE survive	No CIE death	Death	Survive
Total study group	-	0.05060	0.62936	0.00000	0.02106	0.00256	0.00325	0.00000	0.02933
CIE all	0.05060	-	0.03750	0.00002	0.00033	0.00004	0.01372	0.00001	0.00043
CIE survive	0.62936	0.03750	-	0.00000	0.16561	0.05263	0.00141	0.00000	0.23752
CIE death	0.00000	0.00002	0.00000	-	0.00000	0.00000	0.05641	0.63950	0.00000
No CIE	0.02106	0.00033	0.16561	0.00000	-	0.51204	0.00087	0.00000	0.71312
No CIE survive	0.00256	0.00004	0.05263	0.00000	0.51204	-	0.00068	0.00000	0.27283
No CIE death	0.00325	0.01372	0.00141	0.05641	0.00087	0.00068	-	0.10061	0.00139
Death	0.00000	0.00001	0.00000	0.63950	0.00000	0.00000	0.10061	-	0.00000
Survive	0.02933	0.00043	0.23752	0.00000	0.71312	0.27283	0.00139	0.00000	-

Table A.8: P-values for the average Risk Score one hour after CIE comparisons between subsets. Significant threshold established at $\alpha = 0.00139$.

	Total study group	CIE all	CIE survive	CIE death	No CIE	No CIE survive	No CIE death	Death	Survive
Total study group	-	0.08111	0.78768	0.00006	0.04682	0.00401	0.00796	0.00000	0.05077
CIE all	0.08111	-	0.09296	0.00059	0.00177	0.00018	0.01619	0.00004	0.00185
CIE survive	0.78768	0.09296	-	0.00004	0.21528	0.06234	0.00532	0.00000	0.25600
CIE death	0.00006	0.00059	0.00004	-	0.00001	0.00001	0.73941	0.92624	0.00001
No CIE	0.04682	0.00177	0.21528	0.00001	-	0.43272	0.00399	0.00000	0.82382
No CIE survive	0.00401	0.00018	0.06234	0.00001	0.43272	-	0.00331	0.00000	0.27381
No CIE death	0.00796	0.01619	0.00532	0.73941	0.00399	0.00331	-	0.78186	0.00469
Death	0.00000	0.00004	0.00000	0.92624	0.00000	0.00000	0.78186	-	0.00000
Survive	0.05077	0.00185	0.25600	0.00001	0.82382	0.27381	0.00469	0.00000	-

Table A.9: P-values for the average Risk Score two hours after CIE comparisons between subsets. Significant threshold established at $\alpha = 0.00139$.

	Total study group	CIE all	CIE survive	CIE death	No CIE	No CIE survive	No CIE death	Death	Survive
Total study group	-	0.06197	0.79465	0.00002	0.04093	0.00492	0.07112	0.00000	0.05116
CIE all	0.06197	-	0.07198	0.00024	0.00098	0.00012	0.14539	0.00008	0.00114
CIE survive	0.79465	0.07198	-	0.00001	0.17811	0.05647	0.06410	0.00000	0.23607
CIE death	0.00002	0.00024	0.00001	-	0.00000	0.00000	0.49706	0.78794	0.00000
No CIE	0.04093	0.00098	0.17811	0.00000	-	0.50623	0.04054	0.00000	0.75030
No CIE survive	0.00492	0.00012	0.05647	0.00000	0.50623	-	0.03420	0.00000	0.28433
No CIE death	0.07112	0.14539	0.06410	0.49706	0.04054	0.03420	-	0.59369	0.04436
Death	0.00000	0.00008	0.00000	0.78794	0.00000	0.00000	0.59369	-	0.00000
Survive	0.05116	0.00114	0.23607	0.00000	0.75030	0.28433	0.04436	0.00000	-

Table A.10: P-values for the average Risk Score three hours after CIE comparisons between subsets. Significant threshold established at $\alpha = 0.00139$.

A.2 Risk Score around CIE/ECM per age group

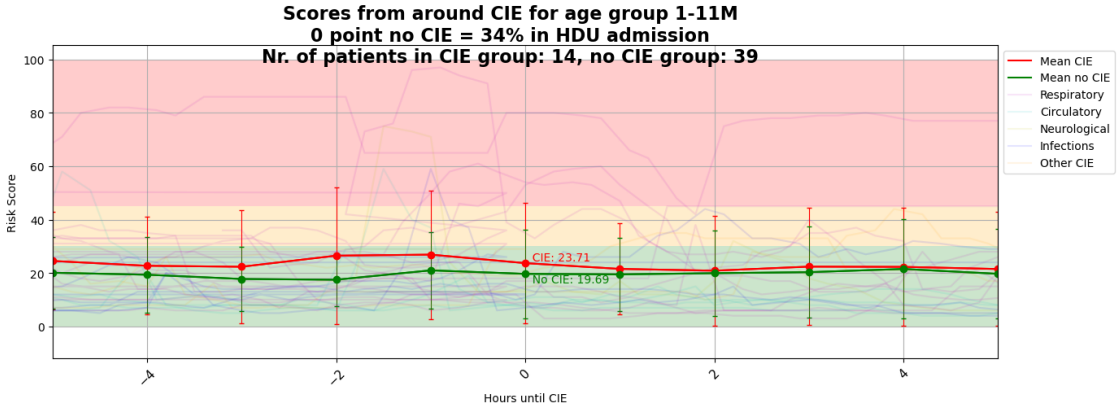

Figure A.1: Visualizing the Risk Score around the time point of a CIE or ECM for the CIE group (red) and non-CIE group (green) respectively of all the patients in the 1 month - 11 months age group. In the background are all individual Risk Score lines around the same time point.

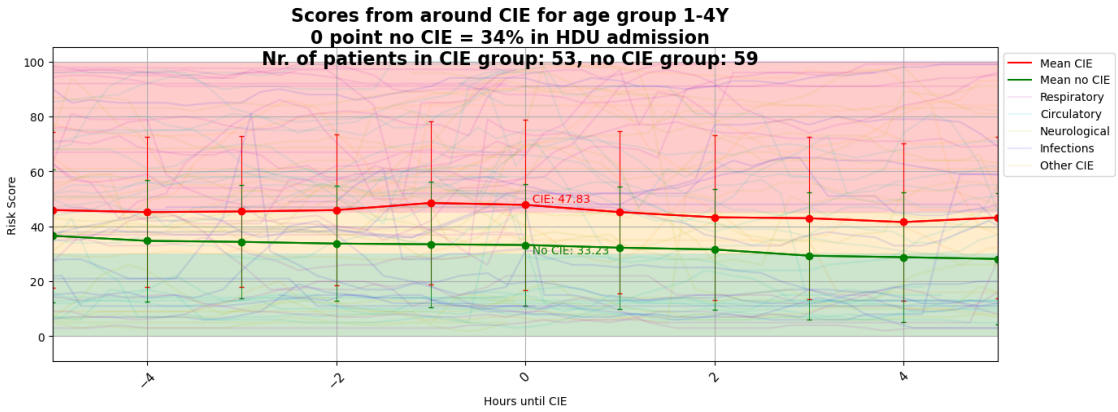

Figure A.2: Visualizing the Risk Score around the time point of a CIE or ECM for the CIE group (red) and non-CIE group (green) respectively of all the patients in the 12 months - 4 years age group. In the background are all individual Risk Score lines around the same time point.

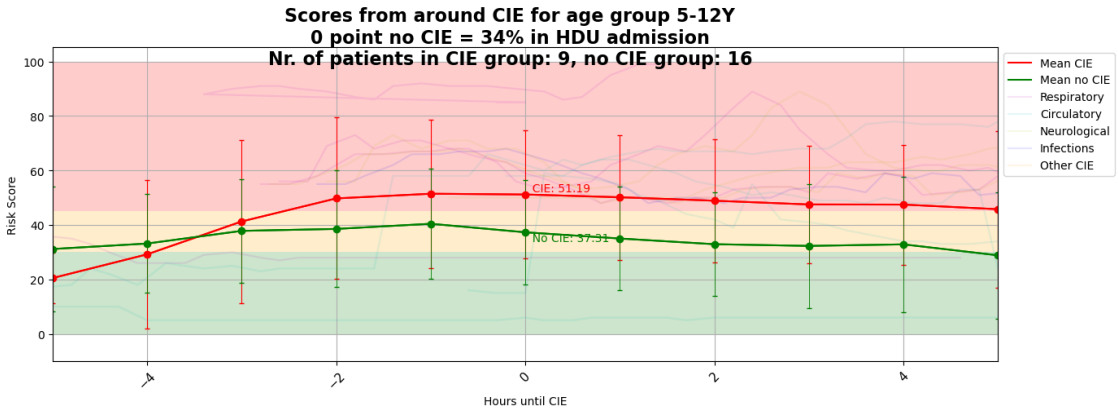

Figure A.3: Visualizing the Risk Score around the time point of a CIE or ECM for the CIE group (red) and non-CIE group (green) respectively of all the patients in the 5 - 12 years age group. In the background are all individual Risk Score lines around the same time point.

A.3 Accuracy, AUROC and Confusion Matrices from three different types of Risk Score

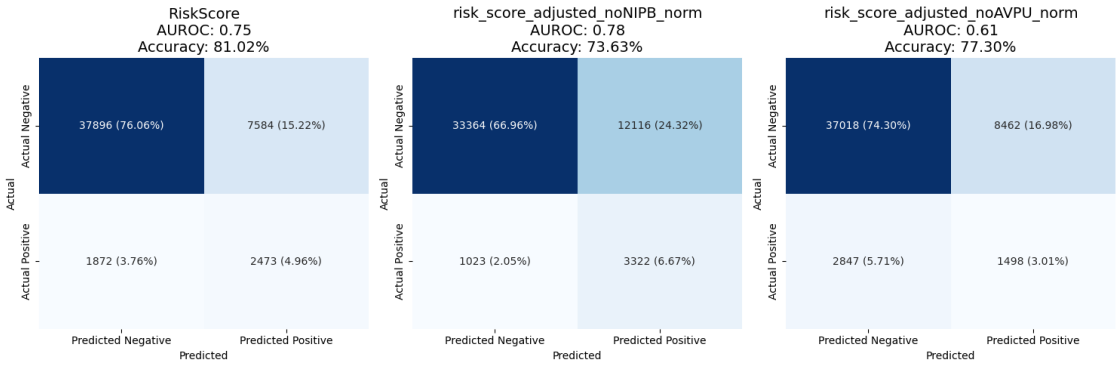

Figure A.4: AUROC, Accuracies, and Confusion Matrices for the three different types of Risk Score for when all score points of all patients who passed away are counted as expected high-risk.

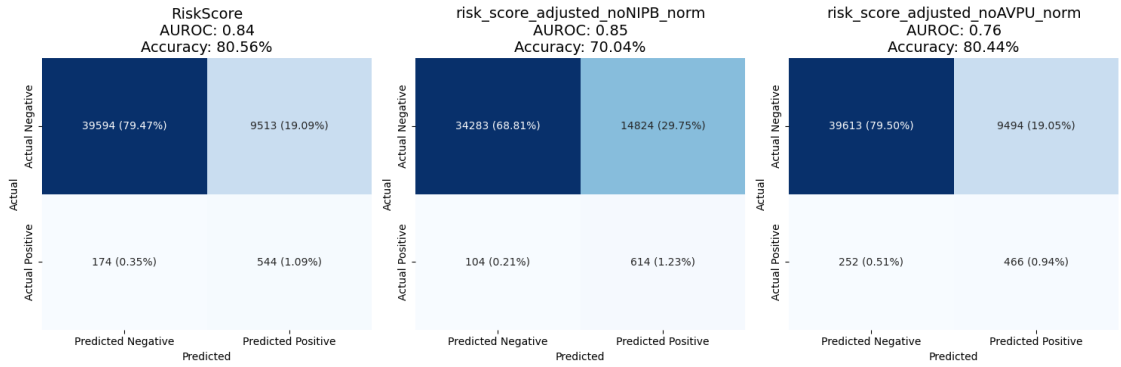

Figure A.5: AUROC, Accuracies, and Confusion Matrices for the three different types of Risk Score for when all score points of the last 8 hours before passing away are counted as expected high-risk.

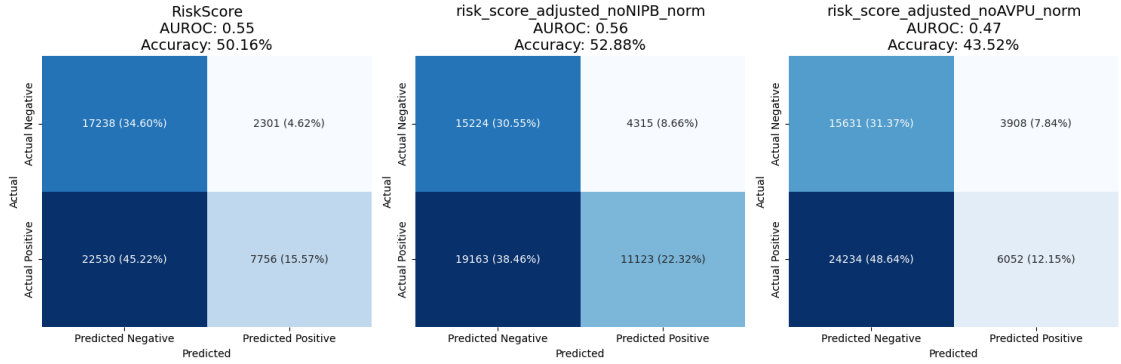

Figure A.6: AUROC, Accuracies, and Confusion Matrices for the three different types of Risk Score for when all score points of all patients who passed away plus all patients who had one or more CIE, are counted as expected high-risk.

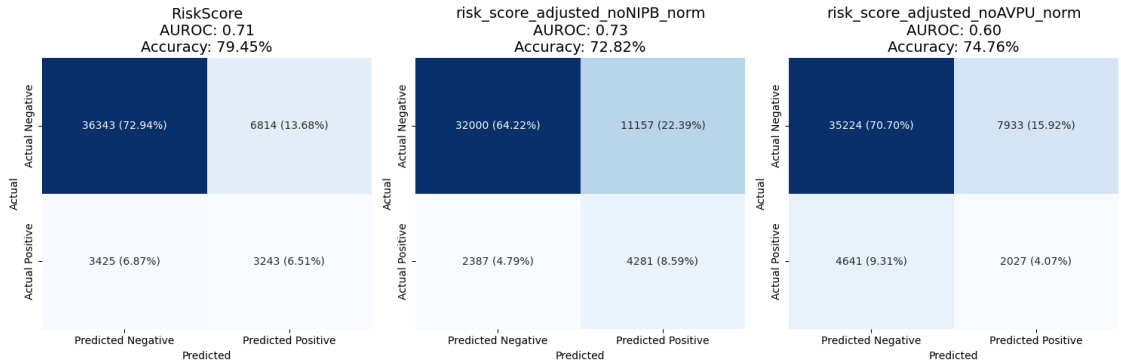

Figure A.7: AUROC, Accuracies, and Confusion Matrices for the three different types of Risk Score for when all score points of all patients who passed away plus 4 hours before and after the CIE, are counted as expected high-risk.

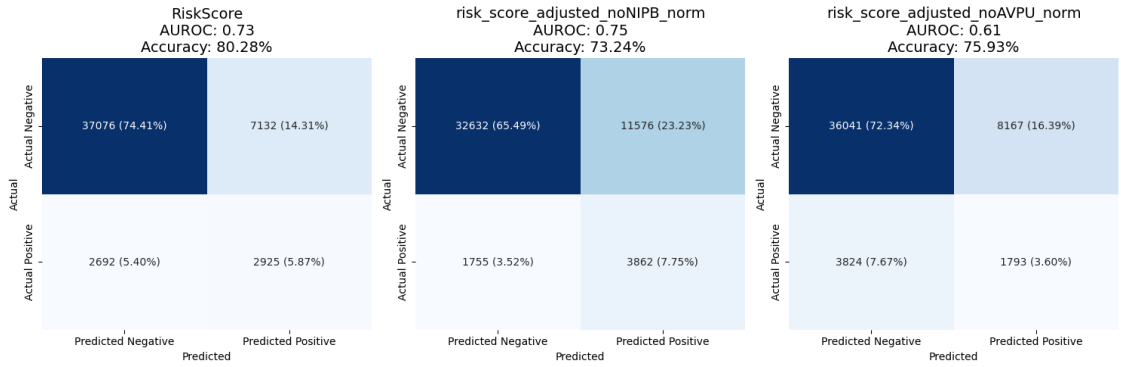

Figure A.8: AUROC, Accuracies, and Confusion Matrices for the three different types of Risk Score for when all score points of all patients who passed away plus 2 hours before and after the CIE, are counted as expected high-risk.

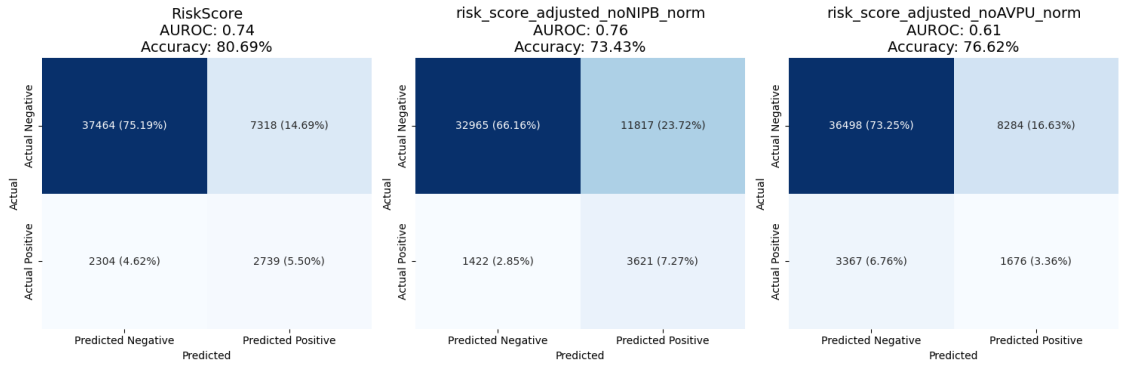

Figure A.9: AUROC, Accuracies, and Confusion Matrices for the three different types of Risk Score for when all score points of all patients who passed away plus 1 hour before and after the CIE, are counted as expected high-risk.

DEPARTMENT OF ELECTRICAL ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY