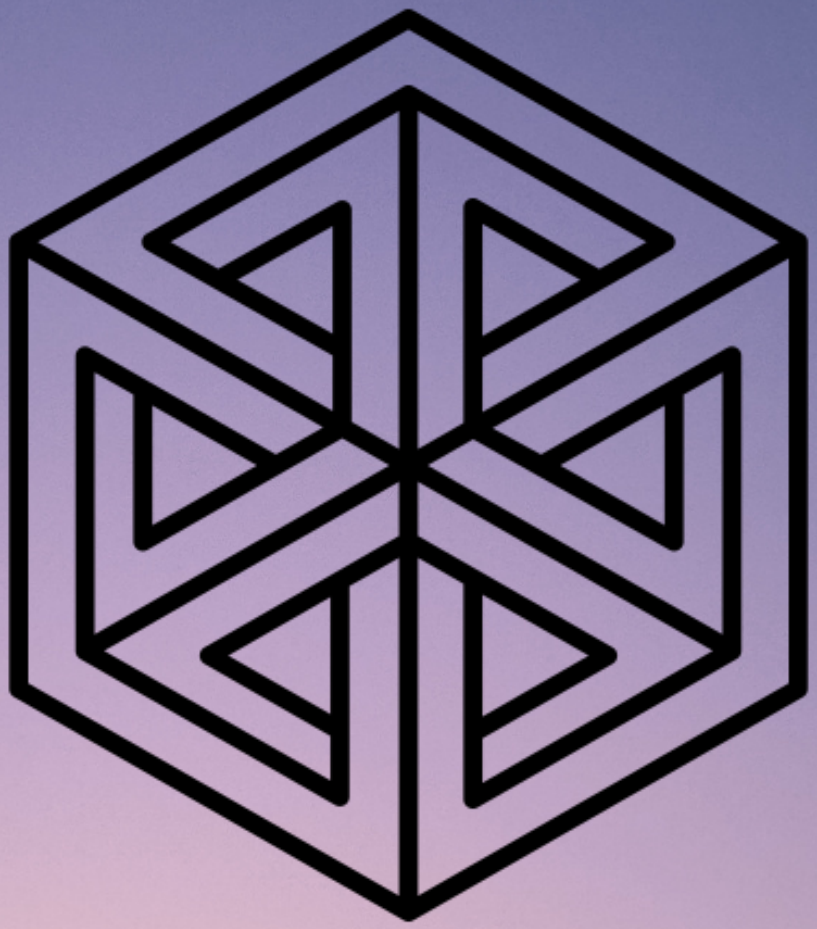
Janno Schouten

The ICU-Recover Box



Master Thesis

27-03-2023

The ICU-Recover Box

- Using Smart Technology For Monitoring Health Status After ICU Admission -

Janno Schouten
Student number: 4476565
Monday 27th March, 2023

Thesis in partial fulfilment of the requirements for the joint degree of Master of Science in

Technical Medicine

Leiden University ; Delft University of Technology ; Erasmus University Rotterdam

Master thesis project (TM30004 ; 35 ECTS)
Dept. of Biomechanical Engineering, TU Delft
November 2021 – March 2023
Supervisor(s):

Dr. Sesmu Arbous
Prof. dr. ir. Marcel Reinders
Drs. Tina van Hemel-Rintjap

Thesis committee members:

Prof. dr. ir. M.J.T. Reinders (Chair)
Dept. of Pattern Recognition and Bioinformatics
TU Delft

Dr. M.S. Arbous
Anaesthesiologist-Intensivist
LUMC

Drs. J.D. van Hemel-Rintjap
Anaesthesiologist-Intensivist
LUMC

Dr. ir. A. Albayrak (Independent committee member)
Dept. of Human-Centered Design
TU Delft

Universiteit Leiden

TU Delft

Erasmus University Rotterdam

Preface

It is with great pleasure that I present this Master thesis in Technical Medicine, called The ICU-Recover Box. This thesis marks the end of a journey that started in November 2021 and during which I have explored the entire process of conducting clinical research.

This thesis consists of two independent parts. Throughout part one of the report, I aim to demonstrate what adjustments need to be made before implementing home monitoring for increased ICU involvement in follow-up care for patients discharged from the ICU. In part two of this thesis, I will discuss an attempt to predict ICU mortality based on data acquired on the ICU.

I would like to express my gratitude to Sesmu Arbous and Tina van Hemel, my medical and daily supervisors. They offered me the opportunity to finalize my Master's on the ICU, a department where I really enjoyed my second year internship. Together, we came up with an approach for the pilot study, set up an extensive protocol for the medical ethics committee and eventually conducted the study successfully. We had some setbacks and unforeseen delays, but we always remained positive. I wish Sesmu and Tina the best of luck with finalizing the pilot study and I hope that the ICU-Recover Box will one day improve follow-up care for patients who were admitted to the ICU.

I would also like to express my gratitude to my technical supervisor, Marcel Reinders. Not only for being my thesis supervisor, but especially for recommending changing our approach for the data analysis. This changed approached meant that I could gain a better understanding of big data analyses using a machine learning approach. Something I really enjoyed doing.

For questions regarding the data analysis, one of Marcel's PhD students, Jim Smit, was available for questions. I am very grateful to Jim for being willing to sacrifice some of his time. His advice definitely helped raise the level of the analysis.

As mentioned before, the past year also had its setbacks that asked a lot of my endurance and resilience. Sometimes it was difficult to remain positive, but several people helped me through these periods. I would like to thank my fellow students on "het hok", Floor, Imane, Melissa, Rowan, and Suus for all the laughter during the sometimes too lengthy lunches and afternoon walks.

Finally, there is three people in particular that I would like to express my gratitude to. Without these three people, this thesis would probably never have been finished. Firstly, I would like to thank Jeanette Wigbers for her efforts getting the pilot study started and organizing all required forms. Secondly, I would like to thank Mike del Prado for his efforts, especially with his help during the period Jeanette was on holiday. However, I am most grateful to both for always having their door wide open and having time for all questions, requests, or needs for a conversation I had during the past year. Finally, I would like to thank Roger Snijder. His ability to quickly find solutions where others would take several days or weeks, made this thesis possible. I don't know where this thesis would currently be without his solutions for problems with the ICU-Recover Box. Without Roger, I doubt I would have gotten all the required data for the second part of this thesis within the short amount of time that it took. I would also like to thank Roger for always being available for questions and for a laugh.

I hope you enjoy reading this Master thesis.

Janno Schouten
Delft, March 2023

Summary

Patients who are admitted to the Intensive Care Unit (ICU) are extremely ill and at high risk of organ failure and death. Being admitted to the ICU is known to cause long lasting physical, cognitive, and physiological symptoms, which is called Post-Intensive Care syndrome (PICS). To provide better management of PICS, early recognition and intervention are required. Two LUMC intensivists are exploring methods to increase ICU involvement in follow-up care for ICU patients, and thereby improve outcome related to PICS. Home monitoring may provide a method feasible for the ICU. To better assess feasibility, a pilot study, called "The ICU-Recover Box", was conducted. Additionally, since patients admitted to the ICU vary widely, it is essential to determine which patients are most likely to benefit from increased ICU involvement in follow-up care. An attempt was made to determine whether measurements recorded during ICU admission contain valuable information to predict ICU mortality. When this data contains valuable information, it may also be used to predict which patients should receive the ICU-Recover Box.

Fifteen patients who were admitted to the ICU received a weight scale, blood pressure monitor and smartwatch at hospital discharge. During the twelve week follow-up period, these patients would monitor their health status, which could be monitored remotely. Based on the preliminary results of six patients, a list of recommendations for large-scale clinical research and eventual implementation of the ICU-Recover Box into standard ICU practices, was formulated. The most important recommendations were to reduce the dependency on third parties and to substitute the home monitoring devices for devices that are medically CE-marked and have better ergonomics.

One recommendation was to assess which patients are most likely to benefit from home monitoring before implementing the ICU-Recover Box into standard ICU procedures. An attempt was made to predict ICU outcome using a machine learning approach. Data of 1364 patients, admitted to the LUMC ICU in 2017 and 2018, was collected and labeled "Survivor" or "ICU-Death", based on the ICU mortality. Features consisted of patient characteristics, hemodynamic status, and medication administrations. An Extreme Gradient Boosting (XGB) classifier was selected and optimized. Training and testing of this classifier was iterated 100 times. Mean AUC over all iterations was 0.809 with a standard deviation of 0.036. Mean sensitivity was 0.298 and mean specificity was 0.964, with standard deviations of 0.067 and 0.011 respectively. The mean corresponding F1 score was 0.851 with a standard deviation of 0.020. These results indicated that data recorded on the ICU seems to contain valuable information to predict ICU mortality. To identify whether this data could be used to determine which patients should receive the ICU-Recover Box, further research is required.

Nomenclature

Abbreviations

Abbreviation	Definition
AI	Artificial Intelligence
APACHE	Acute Physiology and Chronic Health Evaluation
ATC-code	Anatomical Therapeutic Chemical Code
AUC	Area Under the ROC Curve
BPM	Blood Pressure Monitor
CABG	Coronary Artery Bypass Graft
ECG	Electrocardiogram
ED	Emergency Department
HRQoL	Health Related Quality of Life
ICU	Intensive Care Unit
KNN	K-Nearest Neighbors
LR	Logistic Regression
LUMC	Leiden University Medical Center
MI	Myocardial Infarction
ML	Machine Learning
NICE	Nationale Intensive Care Evaluatie
PDMS	Patient Data Management System
PICS	Post Intensive Care Syndrome
PPG	Photoplethysmography
RBF	Radial Basis Function
RF	Random Forest
ROC	Receiver Operating Characteristics
S_pO_2	Peripheral Oxygen Saturation
SVM	Support Vector Machine
WHO	World Health Organization
XGB	Extreme Gradient Boosting

List of Figures

List of Tables

Contents

1

Introduction

1.1. Intensive Care

Intensive care, or critical care, is a multidisciplinary medical specialty dedicated to the management of patients suffering from acute, life threatening organ dysfunction. Patients requiring critical care are admitted to the Intensive Care Unit (ICU), where a wide array of technologies is available to support or completely replace failing organ systems, in particular the lungs, cardiovascular system, and kidneys. [1–3] Patients admitted to the ICU are severely ill and at extremely high risk of organ failure and death, causing the ICU to be one of the most resource demanding and stressful areas of the hospital. [1] A key difference between standard ward-based care and the ICU is the ability to monitor patients continuously, providing the main guide for hemodynamic support. [3, 4] Although critical illness has high mortality rates, many patients survive. These patients face new challenges that strongly influence their daily lives. [5] Many of these challenges are caused by symptoms that are known to be related to being admitted to the ICU.

1.2. Post Intensive Care Syndrome

After ICU admission patients may suffer from different symptoms. These include physical, cognitive, and physiological symptoms such as muscle weakness, memory loss, and depression. Approximately 50% of patients who have been admitted to an ICU suffer from a collection of these symptoms, which is referred to as Post-Intensive Care Syndrome (PICS). PICS can have a negative impact on the health-related quality of life (HRQoL) of both patients and their relatives, leading to increased healthcare utilization. The precise causes of PICS remain largely unknown, but it is thought to result from a combination of the effects of critical illness, prolonged ICU stay, and the adverse effects of mechanical ventilation, medication, and other treatments. Management of PICS requires early recognition and intervention to reduce impact on HRQoL and improve outcome. [6–10]

1.3. Healthcare Utilization After Intensive Care Admission

Since approximately 50% of patients suffer from PICS after ICU admission, it was deemed important to assess healthcare utilization of patients discharged from an ICU. For a complete overview, a systematic review, was conducted. This systematic review is included in Section A.1 on page 27. High heterogeneity in study designs and populations caused widely varying outcomes which proved difficult to interpret. However, in general, patients admitted to the ICU required increased healthcare utilization in the first year after ICU admission. Hospital readmission rates, Emergency Department (ED) visits and costs were increased and nearly all patients visited an outpatient clinic at least once in the year after ICU admission. All these results were within expectations since patients admitted to the ICU suffer from severe illness and can suffer from problems related to PICS. Increased comorbidity or severity of illness were associated with increased healthcare utilization when comparing different ICU patient populations. [11] A limited exploratory sample test, conducted on patients admitted to the Leiden University Medical Center (LUMC) ICU, indicated that readmission rates and number of ED visits within the year after ICU admission were very similar to those reported in the systematic review. Hospital readmission rates for patients admitted to the LUMC ICU seemed as high as 40% while nearly 25% of included patients had to visit the ED at least once in the year following ICU admission.

1.4. Home Monitoring

Recent technological developments have provided the opportunity to better monitor patients' health status in the extramural environment. The consumer market for personal health-monitoring devices and systems is developing rapidly and patients frequently provide physicians with data obtained using these devices. [12] Many vital physiological parameters can be measured and monitored using wearable health monitoring systems. Most modern systems combine multiple sensors to measure multiple different parameters continuously without interrupting the user's daily activities. These parameters include cardiovascular parameters such as electrocardiogram (ECG), heart rate, heart rate variability, and peripheral oxygen saturation (S_pO_2). Many wearable health monitoring systems include activity monitoring by recognizing walking patterns and use pedometry to count the number steps as a measure of activity. [13]

Within the LUMC, the departments of Cardiology and Cardiothoracic Surgery have successfully implemented home monitoring devices into standard healthcare procedures. Their "Box" program, - after the box in which the devices are distributed to patients - has been part of the patient journey for several years, which started for patients with myocardial infarction (MI). [14] Currently, thirteen different boxes are composed for various patient populations. Different Boxes contain varying devices and come with different instructions based on a specific patient population's patient journey. [15]

Home monitoring trough health monitoring devices provides continuous care for patients while it also reduces burden on healthcare professionals. [16] Additionally, remote monitoring reduces the frequency at which patients have to visit an outpatient clinic, which is beneficial for both patient and physician. [17] While it is known that many patients suffer from PICS after ICU admission, ICU staff is not involved in patient follow-up care after ICU discharge. It is believed that increased involvement of ICU specialists in the follow-up care of patients admitted to the ICU may improve care and therefore patient outcome, especially considering symptoms related to PICS. The aim of this research is to investigate the opportunities home monitoring devices created to increase ICU involvement in follow-up care while reducing workload for ICU staff to a minimum.

1.5. Thesis Sub-Studies

This thesis report is divided into two sub-studies. Firstly, Chapter two describes the execution and evaluation of a pilot study in which fifteen ICU patients received home monitoring devices in the so called ICU-Recover Box. Lessons learned during this pilot study resulted in a list of recommendations for future large scale clinical research and eventual implementation of the ICU-Recover Box. Lastly, Chapter three describes a research aiming to determine whether physiological parameters, measured on the ICU, provide an indication of patient outcome. A model to predict ICU mortality was designed to evaluate whether acquired data contains useful information that could potentially be used to determine which patients sustain most ICU related symptoms and could therefore benefit most of the ICU-Recover Box.

2

Implementation of the ICU-Recover Box

Abstract

Background: Following ICU admission, patients may experience a variety of symptoms, including physical, cognitive, and psychological complaints. Post-Intensive Care Syndrome refers to the combination of these symptoms. While these problems are specific to ICU patients, the LUMC ICU provides very limited follow-up care. To improve ICU patients' outcomes and health-related quality of life, a pilot study, called the ICU-Recover Box, was established to introduce home monitoring for ICU patients.

Objective: The aim of this study was to deliver a list of recommendations for future research concerning the ICU-Recover Box, and for implementation in standard LUMC healthcare procedures.

Methods: The pilot study protocol was conducted and evaluated to determine feasibility. Preliminary findings and necessary adjustments were provided. Additionally, practices from various departments within the LUMC that successfully implemented home monitoring, were evaluated. Based on the findings, a list of recommendations for future research and implementation into standard healthcare procedures was developed.

Results: Out of 520 screened patients, eleven patients were eligible to participate in the study, of which six patients provided informed consent. Since inclusion was not completed and no patient had completed the three month follow-up period, no study results were available yet. The protocol was adjusted twice. The first adjustments were required to start the study, while later, an amendment was set up in an attempt to increase inclusion rates. Current practices of the departments of Cardiology and Cardiothoracic Surgery were evaluated since these departments were the firsts to successfully implement home monitoring. All findings formed the base for two lists of recommendations. The most important recommendations were to reduce the dependency on third parties and to substitute the home monitoring devices for devices that are medically CE-marked and have better ergonomics.

Conclusion: The ICU-Recover Box could potentially improve follow-up care for ICU patients. Future research will be feasible with significant efforts from the research team. To fully realize the benefits of home monitoring, improvements and additions to current methods must be made before the ICU-Recover Box can be integrated into standard ICU practices.

2.1. Introduction

2.1.1. Post Intensive Care Syndrome

After Intensive Care Unit (ICU) discharge, patients can suffer from multiple symptoms. Post-Intensive Care Syndrome (PICS) is a collection of physical, cognitive, and psychological symptoms, which can include physical weakness, difficulty with memory and concentration, depression and anxiety, and changes in personality, The combination of physical trauma and stress of being in the ICU, as well as the medications and interventions that are used to treat critical illness are thought to be causes for PICS. PICS can have a significant impact on a patient's health related quality of life (HRQoL) and their ability to return to normal activities. [6–10]

2.1.2. Current Follow-Up Care

Currently, follow-up care for patients who were admitted to the ICU is similar to that of patients who were not admitted to the ICU. Patients who have been in the ICU for at least five days are approached by one of the ICU nurses who are a part of the "IC-nazorg" group, which focuses on ICU follow-up care. The nurse explains the events that occurred during the patient's ICU stay and answers any questions the patient may have. The nurse also provides the patient with a brochure that contains information about the potential physical and mental consequences of an ICU admission, including PICS. After three months, patients are contacted by phone to participate in a short questionnaire. Additionally, patients are offered the opportunity to visit the ICU, which is reported to have a positive impact on their ability to process the experience of being admitted to the ICU. [18]

Two LUMC intensivists, M.S. Arbous and J.D. van Hemel-Rintjap, believe that better ICU-specific follow-up care could significantly improve outcome for ICU patients, particularly with regards to PICS and HRQoL. However, they recognize that this will not be easy to achieve given the current staffing challenges and the lack of a dedicated follow-up care team consisting of medical staff. Additionally, the ICU is one of few departments within the LUMC without an outpatient clinic. To overcome these challenges, the intensivists are exploring alternative methods within the LUMC to provide ICU-specific follow-up care.

2.1.3. Home Monitoring

Within the LUMC, several departments have been experimenting with incorporating home monitoring into standard healthcare procedures. The departments of Cardiology and Cardiothoracic Surgery have been particularly successful in this area, having implemented home monitoring into multiple different patient journeys. These different methods are combined in an initiative called "The Box," as patients receive all devices in a special box. For example, patients receive a Box as part of standard care after experiencing a myocardial infarction (MI) or Coronary Artery Bypass Graft surgery (CABG). [14, 15]

Given that home monitoring offers the opportunity to provide follow-up care without the necessity for an outpatient clinic, it was deemed a feasible method to overcome the challenges present on the ICU. To assess feasibility of large scale clinical research, a pilot study was set up, called the ICU-Recover Box. The adjustments that need to be made were determined by conducting and evaluating the pilot study.

2.1.4. Aims and Objectives

The aim of this study was to deliver a list of recommendations for both future research concerning the ICU-Recover Box, and for future implementation into standard ICU healthcare procedures.

2.2. Methods

In order to produce these recommendations, multiple steps had to be completed. Firstly, a pilot study was conducted and evaluated. The complete pilot study protocol is included in Section B.1 on page 29. The methodology of the protocol is summarized below. Additionally, operating procedures from multiple LUMC departments that successfully implemented home monitoring into standard healthcare procedures were assessed. Two categories of recommendations were obtained; a list of recommendations for future large scale clinical research using the ICU-Recover Box, and recommendations for implementation of the Box in standard ICU healthcare procedures.

2.2.1. Pilot Study

Study Design

Patients who were admitted to the LUMC ICU received the ICU-Recover Box at hospital discharge. The Box contained multiple home monitoring devices which are described below. During a three month follow-up period, patients tracked their health status using these devices. Every four weeks, patients had to conduct two questionnaires via a mobile app. One questionnaire to evaluate the health related quality of life (HRQoL) and one questionnaire to gain insight in the healthcare utilization of patients after ICU admission. A schematic overview of the study is shown in Figure 2.1.

Study Population

Funding sufficient for fifteen ICU-Recover Boxes was available for the pilot study. Therefore, the first fifteen patients admitted to the LUMC ICU that met the inclusion criteria and provided informed consent, were included in the study.

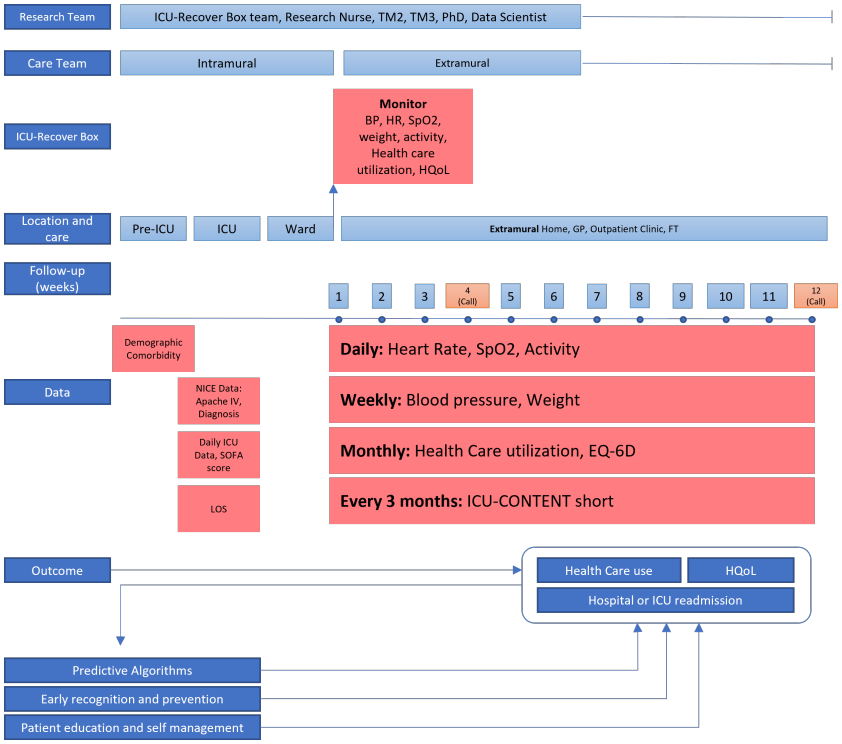


Figure 2.1: Schematic overview of the pilot study design

Inclusion criteria:

- Patient has been admitted to the ICU of the LUMC for $>$ 48 hours.
- Patient has received mechanical ventilation for $>$ 24 hours.
- Patient masters the English or Dutch language.
- Patient is able and capable to use smart technology at home. (i.e. Wi-Fi available, sufficient comprehension of smart technology).
- Patient can be contacted and informed about the ICU-Recover box on one of the clinical wards of the LUMC.
- Patient is discharged from a ward within the LUMC to home.

Exclusion criteria

- Patient is $<$ 18 years old.
- Patient is pregnant.
- Patient breastfeeds during the course of the study
- Patient underwent cardiothoracic surgery (as they will receive the Cardiothoracic Box in another study).
- Patient is discharged for palliative care.
- Patient is considered an incapacitated adult.
- Patient is unwilling to sign the informed consent form.
- Patient is discharged to another hospital.

Investigational Devices

Withings BPM Connect

The Withings BPM Connectis a bluetooth enabled blood pressure cuff which can be placed around the patient's upper arm. It is CE-marked for medical use. After activation, inflation and deflation of the cuff will lead to a systolic and diastolic blood pressure, as well as a heart rate, which are send to and shown

in the Withings Health Mate App. The Withings BPM Connect is battery powered and electrically safe.

Withings Body
The Withings Body is a bluetooth enabled weight scale which measures body weight. It is not CE-marked for medical use. To measure weight, patients must stand on the scale. The Withings Body shows weight, eight day trend and calculated BMI. Data are send to and shown in the Withings Health Mate App. The Withings Body is battery powered and electrically safe.

Withings ScanWatch
The Withings ScanWatch is a bluetooth enabled watch which measures activity, heart rate, peripheral oxygen saturation (S_pO_2) and can record an electrocardiogram (ECG). The software used to determine the above mentioned measurements is embedded within the ScanWatch and is called the Scan Monitor. Thus, the Withings ScanWatch works in concordance with the Withings Scan Monitor. The software is CE-marked and clinically validated with risk classification IIa. The Withings ScanWatch is not CE-marked for medical use. The watch contains a pedometer, GPS for path tracking, sleep pattern recognition technology and electrodes to record an ECG. The watch contains a multi wavelength photoplethysmography (PPG) heart rate/S_pO_2 sensor in the back of the case, which measures S_pO_2. The watch also contains two electrodes in the back of the case and a ring electrode around the watch. To record an ECG, a finger of the opposite hand must be placed on the ring electrode to close the electrical loop required to make an ECG. Data are send to and shown in the Withings Health Mate App. The Withings ScanWatch is battery powered and electrically safe.

Patient Instructions
All participants were instructed to measure their weight and blood pressure once a week during the follow-up period. Heart rate and activity were measured automatically when wearing the ScanWatch, which patients were instructed to do daily. S_pO_2 measurements, which needed to be started manually and take approximately 30 seconds to complete, also had to be performed daily.

Every four weeks, patients had to conduct two questionnaires via a mobile app. One questionnaire on HRQoL and one on their healthcare utilization during the previous four weeks. Each questionnaire could be completed in several minutes.

For technical support, a member of the research team was available by phone three days a week, or every day via email. If patients had concerns about performed measurements or any health related questions, an intensivist was available 24 hours a day, seven days a week, by phone.

After twelve weeks, patients were contacted by an intensivist to finalize the study and to conduct a short evaluation questionnaire. Questions concerned patients' experiences during the study, including experienced workload and personal thoughts on advantages and disadvantages of the ICU-Recover Box.

2.2.2. Evaluation of the Pilot Study
Besides the outcomes resulting from the pilot study, different parameters are valuable for future research and eventually implementation of the ICU-Recover Box. The required workload was assessed and all necessary protocol adjustments and researchers' personal experiences over the course of the study were collected to evaluate the protocol and find areas of improvement for future research and large-scale implementation of the ICU-Recover Box.

2.2.3. Home Monitoring in Different LUMC Departments
As stated before, different departments within the LUMC successfully implemented hoe monitoring into standard healthcare procedures. Front runners in this area are the departments of Cardiology and Cardiothoracic Surgery. Operating procedures of these departments were a great source of inspiration for the pilot study protocol. These operating procedures were summarized and evaluated to formulate recommendations for implementation of the ICU-Recover Box by the ICU.

2.3. Results

2.3.1. Pilot Study

Out of the first 520 patients screened, eleven patients were eligible for inclusion. Nine of the eleven patients were asked to provide informed consent, of which six agreed to participate in the study. One test subject decided to withdraw before completing a measurement. Results of the pilot study were unavailable since patient inclusion had not yet been completed and no patient had completed the entire follow-up period. The definition of feasibility, provided below and in the study protocol in Section B.1, could however, be tested on this first group of patients.

- We were able to acquire data from the devices in the ICU-Recover Box.
- We were able to store the acquired data in a safe manner.
- We were able to analyse the acquired data.
- > 80% of the persons that were discharged with an ICU-Recover Box contributed for three months to post-ICU data.

In order to meet the first three requirements of feasibility, several protocol adjustments were required. These changes are described in the following section. With these changes, secure data collection, storage, and analysis were accomplished. Due to the fact that no patient completed the entire follow-up period, the last item could not be evaluated. It is possible to conclude that all test subjects made contributions for at least the first few weeks following hospital discharge. All patients were able to deliver heart rate and activity data, while S_pO_2 measurements were performed least consistently. After the final patient has finished the follow-up period, complete and detailed results will be provided.

2.3.2. Evaluation of the Pilot Study

Protocol adjustments

The protocol was adjusted twice after approval was granted by the medical ethics committee. The first changes were required to initiate the study. According to the original protocol, data would be visible in the electronic patient filing system called EPD-Vision, which is used by the departments of Cardiology and Cardiothoracic surgery. Before data from ICU patients would be visible in the application, certain additions had to be made by the IT staff from these departments. However, due to underemployment and high workload, these additions would take too long to complete. Therefore, it was decided not to use EPD-Vision, but import data directly from the Withings Health Mate app. This adjustment was easily accomplished since patients would already receive anonymous Withings accounts. The second protocol adjustment required to start the study, was to conduct all questionnaires by phone instead of mobile app. The app developer charged hundreds of euros to implement the questionnaires in the app. Since only fifteen patients would participate in the study and they would already be contacted after four and twelve weeks, only fifteen extra phone calls had to be performed.

During the course of the study, a protocol amendment was produced in an attempt to increase inclusion rates. Instead of an ICU admission lasting more than 48 hours in which a patient had to be mechanically ventilated for at least 24 hours, patients were eligible to participate if they were admitted to the ICU for at least 24 hours in which they were mechanically ventilated for an indefinite period. Also, it became apparent that Cardiothoracic surgery patients who were referred from another hospital did not receive a Box from the Cardiothoracic surgery department. These patients were therefore also eligible for participating in this study.

Workload

One researcher was responsible for patient screening and inclusion, and for distribution, installation, and explanation of the devices. This researcher was also responsible for inspecting incoming data and was available for providing technical support three mornings per week by phone and the entire workweek by email. Two intensivists were responsible for conducting all questionnaires and for 24/7 availability of the provided phone number. Patient screening cost the researcher a maximum of 2 hours on Mondays, since all admissions and discharges from the weekend had to be processed, to at least half an hour on other days. Obtaining informed consent cost varying amounts of time, but at least one hour per patient that decided to participate in the study. Distribution, installation and explanation of

the devices cost one hour per patient. Finally, checking incoming data cost at least one hour weekly and being available for technical support had only a slight contribution in workload since other duties could be performed while being available. The workload for the two intensivists consisted of in total 45 phone conversations, which would take on average approximately half an hour per conversation, and providing 24/7 availability over the course of the study. The intensivists alternated the responsibility for the availability weekly.

2.3.3. Home Monitoring in Different LUMC Departments

Home monitoring using similar Box initiatives has been implemented in the LUMC for several years already. The first departments to successfully implement home monitoring were the departments of Cardiology and Cardiothoracic Surgery. Together, thirteen different boxes, for different patient populations were established. Multiple different populations receive different devices and measurement instructions. One of the responsible physicians determines which patients require a Box and places a request in EPD-Vision. Several employees, exclusively assigned the logistic tasks concerning the Boxes, receive these requests and distribute multiple Boxes daily. These employees are also available for technical support and work in a centrally located office in the LUMC called "The Box Office". Incoming measurement data is inspected weekly by a specialist nurse, or by a physician before a scheduled visit to the outpatient clinic. [15]

Since these departments are so well advanced in the application of home monitoring, most departments considering similar initiatives adopt their established methods. The departments of Neurology, Nephrology, and Hematology started clinical research concerning home monitoring using Withings devices, using the data transfer to EPD-Vision. This remains remarkable since none of these departments were familiar with EPD-Vision.

2.3.4. Recommendations on the ICU-Recover Box

Based on all findings and experiences, a list of recommendations was formulated. The recommendations were categorized into two categories. Recommendations required for future large scale clinical research concerning the ICU-Recover Box, and recommendations on the implementation of the Box in current standard healthcare procedures on the ICU.

Recommendations for future research

- Reduce dependency on third parties
- Include devices that are CE-marked for medical applications
- Include devices with improved ergonomics
- Conduct questionnaires digitally
- Abandon the 24/7 availability

Recommendations on implementing the ICU-Recover Box into standard healthcare procedures

- Make at least one employee responsible for all tasks regarding the ICU-Recover Box
- Determine which patient populations are most likely to benefit from home monitoring
- Reduce dependencies on third parties to an absolute minimum
- Determine how to handle incoming data and how to act when measurements deviate

2.4. Discussion

2.4.1. Recommendations on the ICU-Recover Box

Recommendations for two different objectives were formulated. The first objective was to prepare a protocol for future research regarding the ICU-Recover Box. The second objective was implementation of the Box into standard healthcare procedures. Although the ICU-Recover Box is unlikely to be implemented before further research is conducted, lessons learned during this study could already prove usefu. All recommendations are elaborated below.

Recommendations for future research

Reduce dependency on third parties

In order to start the pilot study, data acquisition as described in the protocol had to be adjusted. It would take too long before data would be available via EPD-vision. Therefore it was decided to import data directly from the Withings Health Mate app via the anonymous accounts provided to each patient. To view and import measurement data, researchers had to use the provided accounts to login for each patient individually. Researchers did not have any administrator rights over all accounts combined. For large scale implementation, data inspection and import should be simplified. The ICU could adopt EPD-Vision. However, ICU staff is unfamiliar with this program and, additionally, EPD-Vision is managed by the IT staff of the Cardiology and Cardiothoracic Surgery departments, which may cause difficulties when other departments request technical support. Although time consuming, it might be preferable to build a data transfer into Metavision or HiX, two programs ICU staff is familiar with.

Include devices that are CE-marked for medical applications

Since the Withings Body and ScanWatch are not medically CE-marked, the medical ethics committee instructed the researchers to include several safety precautions in the protocol. The 24/7 available phone number, as well as extensive patient instructions were required before approval of the protocol was granted. To simplify the implementation of the Box and to lose the extra safety precautions, devices that are CE-marked for medical applications should be used.

Include devices with improved ergonomics

Over the course of the study, it became apparent that a large proportion of the targeted study population did not meet the criterion of being capable of using the home monitoring devices. Nearly all patients with neurological problems, causing cognitive or neuromuscular impairment, were not eligible to participate in the study. The S_pO_2 measurements proved especially difficult to perform when suffering from these issues, due to the very small control button on the side of the ScanWatch. To be able to also include this patient population, ergonomics of the included devices should be assessed. It should be considered to replace the ScanWatch with a similar device that performs all measurements automatically, thereby eliminating difficulties performing measurements manually.

Conduct questionnaires digitally

Implementing all questionnaires in the mobile application used to digitally conduct questionnaires proved to be too expensive. However, in future large scale clinical research, conducting questionnaires by phone is too time consuming. Since the costs of implementing the questionnaires in the designated mobile app are one-time fixed costs, use of this mobile app will become cost-effective with an increased amount of participants.

Abandon the 24/7 availability

The burden of being available 24/7 is too high. Additionally, no patient contacted the available phone number. Why patients did not contact the number is unknown, but they may have experienced a certain barrier. It is also possible that patients did not have any health related questions or concerns, or that they reached out to a different healthcare provider. When all included devices are CE-marked for medical use, 24/7 availability is not required by the medical ethics review committee. It remains important for patients to have easy access to a professional when facing difficulties or concerns. However, if these concerns are so severe that someone must be contacted immediately, the research team is not the correct healthcare provider to contact.

Recommendations on implementing the ICU-Recover Box into standard healthcare procedures

Make at least one employee responsible for all tasks regarding the ICU-Recover Box

Workload for the researcher was not excessively high, but too high to combine with many other responsibilities. Especially Box distribution, installation, and explanation was time consuming. In order to implement the Box on a larger scale, these tasks should be delegated to someone hired exclusively for this purpose, similar to the Box Office from the Cardiology and Cardiothoracic surgery departments.

Determine which patient populations are most likely to benefit from home monitoring
The ICU population is very diverse, with patients from many different medical specialties. It is expected that certain patient populations are more likely to benefit from home monitoring than others. To optimize the use of resources and minimize unnecessary costs and workload, it is crucial to carefully select the patient population that will receive the Box. This could be based on factors such as severity of illness, comorbidities, age, referring specialist, or other relevant parameters. Future research should aim to identify specific patient populations that are most likely to benefit from the Box in post-ICU care.

Reduce dependencies on third parties to an absolute minimum
While this recommendation was previously discussed for future research, it becomes increasingly important to reduce dependencies on third parties when home monitoring is implemented in standard ICU procedures. Being dependent on third parties, generally means that changes are implemented slowly. For research purposes, a protocol must be followed, meaning that there is very limited space for adjustments. However, when home monitoring is implemented within standard care, being able to make quick adjustments becomes essential. To minimize reliance on external parties, it is crucial to establish internal processes and systems for data acquisition and management. This includes reducing dependency on third-party apps and programs, and instead utilizing internal systems that are familiar to ICU staff and can be easily adjusted and managed by the ICU IT staff. This will not only improve the efficiency of the home monitoring process, but also increase the flexibility to make quick adjustments and adapt to changing needs.

Determine how to handle incoming data and how to act when measurements deviate
During the pilot study, one member of the research team inspected measurement data for all patients weekly. For the included number of patients, this took only limited time. However, if the ICU-Recover Box is implemented on large scale, the amount of data to inspect will increase drastically. It is essential to determine how this data will be inspected and who is responsible. In addition, since measurement data is known, the ICU takes a certain level of responsibility. There must be a way to intervene when patients' health status appears to deteriorate.

2.4.2. Additional Recommendations

The objective of the pilot study was to assess the feasibility of using the ICU-Recover Box for future research. The primary objective of the research team, however, is to enhance follow-up care for patients discharged from the ICU. While home monitoring using the Box presents a promising approach, it is not the sole method for improving follow-up care for ICU patients. Possible adjustments to regular care, without the addition of the Box, may also aid in achieving the overall objective. A list of potential modifications to standard care was created to this end.

Additional recommendations to improve follow-up care for patients discharged from the ICU

- *Differentiate the care of patients admitted to the ICU from those who were not by raising awareness of PICS.*

  One potential adjustment to regular care is to enhance hospital-wide understanding of the symptoms that ICU patients may experience. By being aware of PICS, physicians can better recognize and understand certain symptoms, and adjust care accordingly. They should also prioritize addressing these specific issues.

- *Evaluate and demonstrate the importance of ICU involvement in post-discharge care.*

  Within current ICU procedures, there is very limited follow-up care. ICU staff is not equipped with the task of follow-up care, and the ICU is one of few departments within the LUMC without an outpatient clinic. Several significant additions to staff responsibilities and tasks are required before the ICU could provide follow-up care. These changes should be supported department-wide as they will have to be conducted and funded. This may only be achievable when the importance is further assessed and demonstrated.

- *Expand the "IC-Nazorg" group.*

  The "IC-Nazorg" group is currently a small team of ICU staff focused on post-ICU care. To increase effectiveness and responsibilities, it would be beneficial to expand the group. This might even include the addition of one or two intensivists, to provide additional medical expertise. This could be an important first step in establishing and highlighting the significance of ICU follow-up care.

The ICU-Recover Box could potentially improve follow-up care for patients discharged from the ICU, but it may be difficult to implement without additional adjustments. To fully utilize the potential of home monitoring, the Box needs to be integrated into existing infrastructure, which the ICU currently lacks. Conducting large scale clinical research using the Box would require a significant amount of work from the research team, but is feasible. Future integration of the ICU-Recover Box into standard healthcare procedures would necessitate substantial additions to current ICU practices. Given the current challenges of staffing, this will be an ambitious undertaking.The above-mentioned suggestions may need to be considered before the ICU-Recover Box can be established.

2.4.3. Limitations

The most significant limitation of this research is that the pilot study was not completed when recommendations were formulated. This had multiple consequences. Firstly, patient measurement data was not completed and analyzed, making it uncertain whether measurements contained any valuable information. However, it could be argued that the small sample size of fifteen patients would always leave limited data to draw any conclusions. The aim of the pilot study was to assess feasibility of the Box for large scale clinical research. Data analysis from this large scale research will indicate whether the measurement data provides any useful information. Secondly, the types and frequencies of measurements were not evaluated and, since no patient had reached the end of the three month follow-up period, patients' experiences using the Box had not been questioned yet. Patients' experiences may provide additional recommendations and should therefore be evaluated thoroughly.

2.5. Conclusion

The ICU-Recover Box could potentially improve follow-up care for ICU patients. Large-scale clinical research is required to determine whether measurements provide useful information and to determine which patient populations are most likely to benefit from home monitoring. This research will be feasible but difficult since it will necessitate significant efforts from the research team. To fully realize the benefits of home monitoring, improvements and additions to current methods must be made before implementation into standard ICU practices.

3

Predicting ICU Mortality Using Machine Learning

Abstract

Background: Patients admitted to the ICU differ widely. As concluded in Chapter 2 of this Master thesis report, it is essential to identify which patients are most likely to benefit from home monitoring using the ICU-Recover Box. An attempt was made to predict ICU mortality using data acquired during ICU admission. If this data contains predictive information, it may be a future guide to determine which patients are most likely to benefit from receiving the ICU-Recover Box.

Objective: The objective of this study is to determine whether data acquired during an ICU admission contains valuable information to predict ICU mortality.

Methods: All patients admitted to the LUMC ICU in 2017 and 2018, who's admission exceeded one day, were included. Data from three main databases was used for feature engineering. A grid search was conducted using a Random Forest, Extreme Gradient Boosting, and K-Nearest Neighbors classifier, Support Vector Machine and Logistic Regression model. Univariate feature selection was used to select different numbers of features by applying an ANOVA F-test. The best performing classifier was further optimized using a new grid search. The resulting classifier was trained and tested 100 times to produce a general view of model performance.

Results: Data of 1364 patients was used to derive 917 features. The Extreme Gradient Boosting classifier performed best in the primary grid search with number of estimators set to 100 and 457 selected features. After further optimization and 100 iterations of training and testing, mean AUC was 0.809 with a standard deviation of 0.036. Mean sensitivity was 0.298 and mean specificity was 0.964, with standard deviations of 0.067 and 0.011 respectively. The mean corresponding F1 score was 0.851 with a standard deviation of 0.020.

Conclusion: Data recorded on the ICU seemed to contain valuable information to predict ICU outcome. To identify whether this data could be used to determine which patients should receive the ICU-Recover Box, further research is required.

3.1. Introduction

3.1.1. Background

Intensive care is a multidisciplinary medical specialty [2], meaning that patients from all different departments could be admitted to the Intensive Care Unit (ICU). Therefore, demographics, admission diagnoses, severity of illness and many other parameters vary widely between patients. [19] As concluded in Chapter 2 of this thesis report, it is essential to determine which patients are most likely to benefit from the ICU-Recover Box to successfully implement home monitoring into ICU procedures. The final chapter of this thesis report describes the attempt to predict ICU mortality using data acquired during ICU admission. An attempt was made to evaluate whether data collected during the ICU admission contains useful information to predict patient outcome, and could therefore be a guide to determine which patients may benefit from ICU follow-up care using home monitoring.

3.1.2. Measurements on the ICU

One of the key differences between the ICU and standard ward-based care is the possibility to continuously monitor all patients. [3] The objective of hemodynamic monitoring is to provide a guide for optimization of end organ tissue oxygenation and combat tissue hypoxia, shock, and multi-organ failure. The vital parameters that are monitored continuously include heart rate, blood pressure, peripheral oxygen saturation (S_pO_2) and electrocardiogram (ECG). [4] These measurements are not only displayed on a bedside monitor, which is present in all ICU rooms, but are also stored securely. These high amounts of data provide the opportunity to apply extensive analyses. Given the high amounts of available data, Artificial Intelligence (AI), and in particular Machine Learning (ML), seems well-suited for these analyses. [20]

3.1.3. Machine learning

Machine Learning is a field that focuses on the learning aspect of AI by developing algorithms that best represent a set of data. ML uses subsets of data to generate an algorithm that may use novel or

different combinations of features and weights. [21] Recently developed machine learning applications in healthcare have primarily served as a support tool for physicians or analysts to identify healthcare trends, and develop disease prediction models. [22] Application of ML in healthcare presents several challenges, such as availability of sufficient high quality training and test data, legal approval, and difficulties with interpretation and applicability of the results given the lack of transparency in decision making. However, it also provides the opportunities to provide more effective treatments, with increased quality, speed, and precision. [22] Using ML to predict outcome of patients who are admitted on the ICU seems particularly appropriate since the ICU continuously collects high granular data of each patient. [3]

3.1.4. Objectives

The objective of this study is to determine whether a patients physiological status on the ICU, expressed through many different monitoring methods, could be used to predict ICU outcome. If the acquired data contains predictive information, it may also be used to determine which patients are most likely to benefit from home monitoring through the ICU-Recover Box.

3.2. Methods

3.2.1. Study Design

Data from all patients admitted to the Leiden University Medical Center (LUMC) ICU is stored. This data is of high granularity and provides the opportunity to build a prediction model. In this study, we aimed to determine whether cardiovascular status during the first 48 hours of ICU admission, described through hemodynamic measurements and different medication administrations, combined with patient characteristics, contains information that could predict ICU outcome. ML is especially suited for handling these amounts of data. [23] Figure 3.1 shows a schematic overview of the typical ML workflow. Using this approach, multiple machine learning classifiers, namely a Random Forest (RF), Extreme Gradient Boosting (XGB) and K-Nearest Neighbors (KNN) classifier and a Support Vector Machine (SVM), and a more classical approach, namely logistic regression (LR) were evaluated. Firstly, raw data was acquired and features, deemed relevant by two medical specialists, were derived. A model that suited the data best was then selected and further optimized. After optimization, training and testing of the model was iterated 100 times to provide the model performance through the Receiver Operating Curves (ROC), Area Under the ROC Curve (AUC), sensitivity, specificity and F1 score. All required steps are performed using Python, with extensive use of the Scikit-Learn modules. [24] Python scripts are available using the following link:

`https://git.lumc.nl/jsschouten/icu-recover-box.git`

3.2.2. Patient Selection

Each year, a list of patients admitted to the ICU is maintained. This document, which is stored securely on a LUMC server, contains patient names, gender, identification number, admission- and discharge date, and referring specialist. A special note is added to all patients who died during the ICU admission. Using these documents, all identification numbers of patients admitted to the ICU in the years 2017 and 2018 were retrieved. All patients that were admitted to the ICU for more than one day were included in this study. Patients who had died on the ICU were given the label "ICU death", while all other patients were labeled "Survivor".

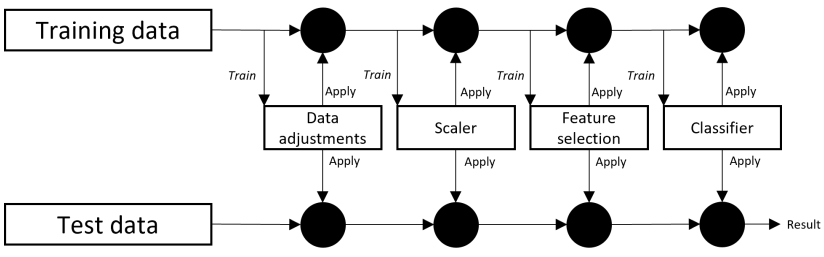


Figure 3.1: Schematic overview of typical machine learning workflow

Table 3.1: Features from the NICE-Database

Feature	Measure
Gender	Male/Female^a
Age	Years
Length	Centimeters
Weight	Kilograms
APACHE II Score	integer 0 - 71
Admission type	integer 1 - 4^a
Planned Admission	0/1
Chronic Renal Insufficiency	0/1
Chronic Dialysis	0/1
COPD	0/1
Respiratory Insufficiency	0/1
Cardiovascular Insufficiency	0/1
Cirrhosis	0/1
Neoplasm	0/1
Hematologic Malignancy	0/1
Aids	0/1
Immunological Insufficiency	0/1
CPR	0/1
Burns	0/1
Gastrointestinal Bleeding	0/1
Diabetes	0/1
Mechanical Ventilation at admission	0/1
Mechanical Ventilation after 24 hours	0/1

a : *Adjusted in preprocessing*

3.2.3. Data Acquisition and Feature Engineering

Data acquired during ICU admission is securely stored on three main databases. Patient demographics and characteristics at admission are stored on the NICE (Nationale Intensive Care Evaluatie) database. All data from the hemodynamic monitor is stored within the LUMC ICU specific Patient Data Management System (PDMS), called Metavision, and medication administrations are registered and saved on specific LUMC servers . These different databases are accessible via different departments. These departments were provided the compiled list of patient identification numbers and delivered the corresponding raw data.

Information stored on the NICE-database provides basic patient information such as gender, weight and date of birth. It also contains all information required to calculate the Acute Physiology and Chronic Health Evaluation (APACHE) II score. [25] Since the score itself is not included in the database, it had to be calculated. The calculation can be found in Section C.1 on page 31. All variables required for calculating the APACHE II score were also included as features for the prediction model. All included features engineered from data retrieved from the NICE-database can be found in Table 3.1.

Measurements displayed on the hemodynamic monitor are sent to Metavision, providing physicians insight in a patient's status remotely. Data from the monitor is also saved on the Metavision servers. The ICU IT staff delivered data of different parameters, from all patients, with a sample frequency of one value per minute, for the entire duration of the admission. The parameters that were included were systolic, diastolic, and mean blood pressure, and heart rate. Features were derived per hour for the first 48 hours of the admission. From the mentioned parameters, the average, minimum, maximum, and variance per hour were calculated.

In agreement with 2 intensivists, a list of 22 different medications that affect the hemodynamic status and could be administered on the ICU, was compiled. Using the corresponding Anatomical Therapeutic Chemical (ATC) codes, which are proposed to be used in drug utilization studies by the World Health Organization (WHO) [26], medications were identified. Besides an ATC-code, each medication carries

an identifier value. One type of medication can carry multiple identifier values, identifying different concentrations or administration methods. All identifier values corresponding to the list provided by the two intensivists, totaling 52 unique values, were included. A complete overview of all included medication, including corresponding ATC-codes and identifier values, is included in Section C.2 on page 32. For each identifier values. the following features were derived: Did the patient receive that medication (yes/no) (1), the total summed dose (2), and average dose per day (3) of that medication over the first three days of admission.

3.2.4. Preprocessing

Data Adjustments

Not all engineered features were directly suited for model training and application. Firstly, categorical data had to be transformed. Gender, which was expressed as M or F, was transformed into the feature "Male" with a 1 or 0 identifying men and women. Admission type, which was categorized using the numbers 1 to 4, was binary encoded using one-hot encoding. [27] Next, data was split into a training and test set using a 80-20 ratio. Missing values, infinite values or strings in both the training and test sets where replaced with the median of the corresponding column in the training set.

Scaling

All features that were not binary were scaled using the Scikit-Learn standard scaler. This scaler standardizes features by removing the mean and scale to unit variance. Centering and scaling is performed individually for each feature by computing the relevant statistics on the samples in the training set. The mean and standard deviation are stored to be used on the test set. [24]

Feature Selection

All features were derived from the raw data in agreement with 2 intensivist, who expected all features to provide a certain predictive value. However, not all features necessarily contained discriminating information. Therefore, features were selected using univariate feature selection. An ANOVA F-test was performed on all features, resulting in a list of features ordered on level of importance. [28] The number of features to be included in the prediction model was determined using a grid search in which all model performances were evaluated using 10, 50, 100, 200, or half of all features. The grid search was also performed with all features included.

3.2.5. Model Selection

The RF, XGB, and KNN classifiers, as well as the SVMs and LR model were evaluated for all number of selected features. The ML based models were evaluated with a set of different hyperparameter settings. For the RF and XGB classifiers, the number of estimators was set to 10, 50, 100, 200 and 500. The number of neighbors in the KNN classifier was set to 3, 11, 15 and 21. Finally, the SVM was evaluated using the kernel settings linear, Radial Basis Function (RBF), and polynomial, and slack set to 0.5, 0.1, and 0.05. Default settings were used for the LR model. [24] A schematic view of this process is provided in Section C.3 on page 33.

An exhaustive grid search was performed in which all models were trained and validated, using all different selected number of features, through five-fold cross validation. The combination of classifier with set hyperparameter(s) and selected number of features that produced the highest AUC was selected for further optimization.

3.2.6. Model Optimization

Exploratory analyses indicated that RF and XGB classifiers performed better than the SVMs and KNN classifiers. Since it was strongly expected that one of these tree based models would be selected, further optimization of these models was set up. Table 3.2 contains the different settings for hyperparameters that were evaluated for the selected model. A new grid search was performed with all possible combinations of these hyperparameters. The different models were evaluated using five-fold cross validation, which delivered the best performing combination of settings. [24, 29] The performance of the selected model with these optimized settings was then tested. An overview of this process is provided in Section C.4 on page 33.

Table 3.2: Tree based model optimization hyperparameter settings

Hyperparameter	Evaluated Inputs RF	Evaluated Inputs XGB
criterion	gini, entropy	friedman_mse, squared_error
max_features	auto, log2	auto, log2
n_estimators	from model selection	from model selection
max_depth	3, 5, 10, 15, 20	3, 5, 10, 15, 20
min_samples_leaf	3, 5, 10	3, 5, 10
min_samples_split	3, 5, 10	3, 5, 10

3.2.7. Model Evaluation

Model optimization produced a model with a set of hyperparameters. Since data is split randomly, each time the model is trained and tested, performance will differ slightly. To provide a general picture of model performance, each step from splitting, scaling, and adjusting data, feature selection, and model training and testing, was repeated 100 times. The mean, standard deviation, and 95% confidence intervals of the outcome measures, which were AUC, sensitivity, specificity, and F1 score were calculated. A schematic overview of this process is provided in Figure 3.2.

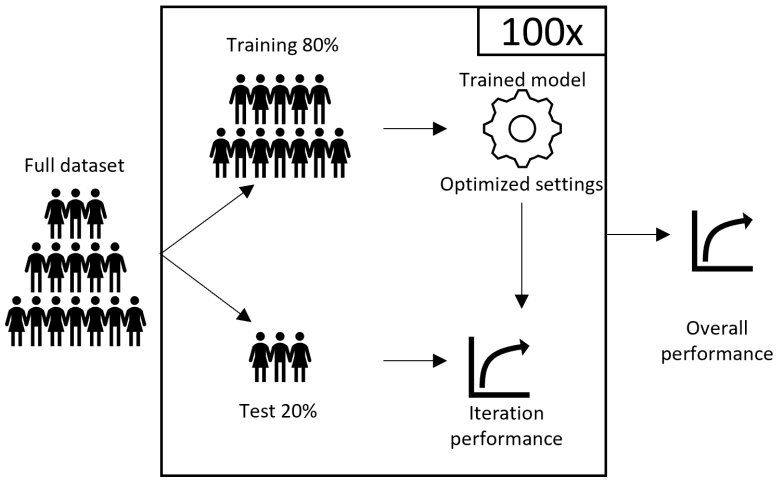


Figure 3.2: Schematic overview of model evaluation methods

3.3. Results

3.3.1. Data Acquisition and Feature Engineering

In total, data of 1364 patients admitted to the LUMC ICU in the years 2017 and 2018 was collected. Of these patients, 1131 survived the included ICU admission, while 233 died on the ICU. Feature engineering produced 917 features for each patient.

3.3.2. Model Selection

As shown in Table 3.3, the best performing classifier in the initial grid search was the XGB classifier. With the number of estimators set to 100 and 458 selected features, the mean AUC over the five-fold cross validation was 0.831. This classifier was further optimized during model optimization.

3.3.3. Model Optimization

The new grid search, using all possible combinations of hyperparameter settings shown in Table 3.2, resulted in the hyperparameter settings provided in Table 3.4.

Table 3.3: AUC scores of broad grid search

Classifier/No. of Features	10	50	100	200	458	917
Random Forest 10[a]	0,743	0,746	0,727	0,714	0,732	0,721
Random Forest 50[a]	0,772	0,767	0,780	0,752	0,757	0,773
Random Forest 100[a]	0,785	0,775	0,779	0,768	0,779	0,768
Random Forest 200[a]	0,784	0,776	0,783	0,773	0,781	0,790
Random Forest 500[a]	0,786	0,776	0,785	0,777	0,785	0,788
XG Boosting 10[a]	0,798	0,798	0,802	0,795	0,797	0,792
XG Boosting 50[a]	0,799	0,792	0,810	0,807	0,829	0,819
XG Boosting 100[a]	0,787	0,789	0,808	0,806	0,831	0,816
XG Boosting 200[a]	0,770	0,772	0,799	0,796	0,830	0,819
XG Boosting 500[a]	0,753	0,749	0,777	0,786	0,817	0,806
KNN 3[b]	0,684	0,667	0,635	0,654	0,632	0,621
KNN 11[b]	0,767	0,730	0,715	0,707	0,676	0,705
KNN 15[b]	0,785	0,747	0,730	0,717	0,689	0,705
KNN 21[b]	0,807	0,752	0,732	0,721	0,707	0,708
SVM linear[c] 0.5[d]	0,789	0,765	0,764	0,727	0,670	0,695
SVM linear[c] 0.1[d]	0,792	0,778	0,768	0,752	0,676	0,705
SVM linear[c] 0.05[d]	0,795	0,782	0,773	0,757	0,694	0,697
SVM rbf[c] 0.5[d]	0,761	0,756	0,769	0,750	0,540	0,499
SVM rbf[c] 0.1[d]	0,760	0,756	0,769	0,749	0,540	0,502
SVM rbf[c] 0.05[d]	0,761	0,756	0,769	0,749	0,540	0,502
SVM poly[c] 0.5[d]	0,730	0,715	0,732	0,709	0,544	0,500
SVM poly[c] 0.1[d]	0,721	0,737	0,731	0,710	0,550	0,500
SVM poly[c] 0.05[d]	0,715	0,742	0,721	0,719	0,551	0,500
Logistic Regression	0,822	0,798	0,775	0,756	0,543	0,500

a : *number of estimators*
b : *number of neighbors*
c : *kernel*
d : *slack*

Table 3.4: Optimized model and hyperparameters

Hyperparameter	Optimized Inputs
classifier	Extreme Gradient Boosting
criterion	squared_error
max_features	auto
n_estimators	100
max_depth	5
min_samples_leaf	10
min_samples_split	3

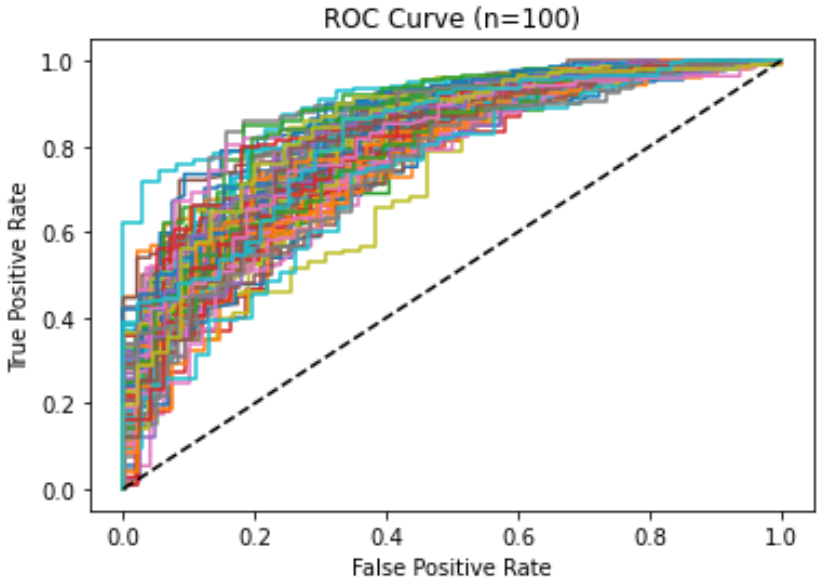


Figure 3.3: ROC curves for 100 optimized model iterations

Table 3.5: Optimized model performance

Outcome	Mean	Standard Deviation	95% Confidence Interval
AUC	0.809	0.036	0.802 - 0.816
Sensitivity	0.298	0.067	0.284 - 0.311
Specificity	0.964	0.011	0.962 - 0.966
F1 Score	0.851	0.020	0.848 - 0.855

3.3.4. Model Performance

Figure 3.3 contains the individual ROC-curves of all iterations. Model performance is provided in Table 3.5. Mean AUC was 0.809 with a standard deviation of 0.036. Mean sensitivity was 0.298 and mean specificity was 0.964, with standard deviations of 0.067 and 0.011 respectively. The mean corresponding F1 score was 0.851 with a standard deviation of 0.020. Model performance per individual iteration is provided in Section C.5 on page 34.

3.4. Discussion

3.4.1. Model Selection and Optimization

Tree based classifiers and the LR model appeared to perform best in exploratory research. The tree based classifiers performed better with more selected features while the LR model performed best with only ten included features and performed increasingly worse with added features. This is somewhat counter intuitive since LR could reduce the weight of non valuable features while increasing the weight of valuable features. The high amounts of features most likely caused overfitting of the model. Eventually, the XGB classifier with 100 estimators was selected and further optimized using an exhaustive grid search. Only limited hyperparameter settings were evaluated in this grid search, given the amount of time calculations would take. Using a different approach, such as gradient descent optimization, may have been a better method to achieve optimal model performance. [30]

3.4.2. Model Performance

Mean AUC over 100 iterations was decent, specificity was excellent, but sensitivity was poor. The model performed very well in predicting ICU survival, but was much less successful in predicting ICU mortality. Correctly predicting ICU mortality may be more valuable than correctly predicting ICU survival. Recommendations on how to improve the current model are listed below.

3.4.3. Limitations

Several patients were admitted to the ICU more than once during the inclusion period. All one-day admissions were automatically excluded. However, for patients who had multiple admissions exceeding one day, only the final admission was included. Given the time it would take to separate data from multiple admissions, and the limited time available for this research, it was decided to only include the final admission. Only few admissions were lost and the amount of included admissions was deemed sufficient for model training and evaluation. More importantly, by only including the final ICU admission, all admissions in which patients died on the ICU remained included.

Data from the NICE-database contained information required to calculate the APACHE II score. However, the LUMC ICU uses not the APACHE II, but the more widely used APACHE IV score. This score is very similar to the APACHE II, but requires more information on ICU admission diagnosis. The admission diagnosis could be in one of 116 categories, each with a different addition to the overall score. [31] Although the scoring system is not complicated and all required information was available, building the scoring system would be too time consuming. It was decided to calculate the APACHE II, since this would require much less time.

Only heart rate and blood pressure were included to quantify the hemodynamic status of a patient. Although these are important, they are not the only important parameters. It may have been useful to include more advanced hemodynamic parameters such as cardiac index or cardiac output. Unfortunately, the ICU IT staff was only able to deliver these basic hemodynamic parameters within the time available for this research.

Since many types of medication are administered continuously on the ICU, it was not possible to determine precisely what dose patients received of a certain medication during the first 48 hours of the ICU admission. To assure all medication administered during this period was included in the feature engineering process, all administrations within the first three days of admission were included. For patients that required increased or decreased hemodynamic support during the third day of admission, numbers may therefore be under- or overestimated.

In retrospect, used methods contained a minor omission. Before model selection and optimization, scaling, data adjustments and feature selection had to be executed, requiring the data to be split into a training and test set. However, since the resulting classifiers were never applied to these test sets, it may have been better to perform model selection and optimization using the complete data set. Smaller data sets are more likely to cause overfitting. [32] However, given the high number of included admissions, and the similar performance during model evaluation compared to the performance during cross validation, it was considered unlikely that the models were overfitted.

3.4.4. Recommendations

Although model performance expressed in AUC and F1 score was decent, the model's sensitivity was poor. It seems that the model performed very well in predicting which patients would survive till ICU discharge but performed poorly in predicting who would not survive. This resulted from the used decision threshold, which was set to the default value 0.5. For clinical application, it may be more useful for the model to better predict poor ICU outcome. By optimizing this decision threshold, the sensitivity of the model could be improved. Future research should include this threshold optimization.

Additionally, although predicting ICU mortality is valuable, being able to predict long-term outcome may be more beneficial. Especially with implementation of the ICU-Recover Box in mind. Determining long-term outcome was not feasible within this study, but may be in future studies. Future researchers should attempt to recover long-term outcome for each patient and label data accordingly. Patients who are predicted to have poor outcome, based on their status during ICU admission, are probably most likely to benefit from increased involvement of the ICU in the post-ICU care. The ICU-Recover Box may be a useful tool to improve outcome for these patients.

3.5. Conclusion

Data recorded on the ICU seemed useful to predict ICU outcome. The presented Extreme Gradient Boosting classifier performed exceptionally well in predicting which patients would survive the ICU admission, but performed less well in predicting ICU mortality. Decision threshold optimization could provide a better balance between sensitivity and specificity. Additionally, future researchers should attempt to retrieve long-term outcome for the included patients to better identify whether the collected data is useful to determine which patients are most likely to benefit from the ICU-Recover Box.

References

[1] Ervin JN, Kahn JM, Cohen TR, Weingart LR. Teamwork in the intensive care unit. Am Psychol. 2018;73(4):468-77.

[2] Marshall JC, Bosco L, Adhikari NK, Connolly B, Diaz JV, Dorman T, et al. What is an intensive care unit? A report of the task force of the World Federation of Societies of Intensive and Critical Care Medicine. J Crit Care. 2017 Feb;37:270-6.

[3] Pinsky MR. Hemodynamic evaluation and monitoring in the ICU. Chest. 2007 Dec;132(6):2020-9.

[4] Kipnis E, Ramsingh D, Bhargava M, Dincer E, Cannesson M, Broccard A, et al. Monitoring in the Intensive Care. Crit Care Res Pract. 2012;2012:473507.

[5] Iwashyna TJ. Survivorship will be the defining challenge of critical care in the 21st century. Ann Intern Med. 2010 Aug;153(3):204-5.

[6] Ramnarain D, Aupers E, den Oudsten B, Oldenbeuving A, de Vries J, Pouwels S. Post Intensive Care Syndrome (PICS): an overview of the definition, etiology, risk factors, and possible counseling and treatment strategies. Expert Rev Neurother. 2021 Oct;21(10):1159-77.

[7] Langerud AK, en T, stuen MC, Kongsgaard U, Stubhaug A. Health-related quality of life in intensive care survivors: Associations with social support, comorbidity, and pain interference. PLoS One. 2018;13(6):e0199656.

[8] LaBuzetta JN, Rosand J, Vranceanu AM. Review: Post-Intensive Care Syndrome: Unique Challenges in the Neurointensive Care Unit. Neurocrit Care. 2019 Dec;31(3):534-45.

[9] Needham DM, Davidson J, Cohen H, Hopkins RO, Weinert C, Wunsch H, et al. Improving long-term outcomes after discharge from intensive care unit: report from a stakeholders' conference. Crit Care Med. 2012 Feb;40(2):502-9.

[10] Brown SM, Bose S, Banner-Goodspeed V, Beesley SJ, Dinglas VD, Hopkins RO, et al. Approaches to Addressing Post-Intensive Care Syndrome among Intensive Care Unit Survivors. A Narrative Review. Ann Am Thorac Soc. 2019 Aug;16(8):947-56.

[11] Schouten JS, van Hemel-Rintjap JD, Arbous MS. Healthcare Utilization Within 1-Year After Intensive Care Unit Discharge: a Systematic Review. 2022 Nov.

[12] Treskes RW, Van Der Velde ET, Atsma DE, Schalij MJ. Redesigning healthcare: The 2.4 billion euro question? : Connecting smart technology to improve outcome of patients. Neth Heart J. 2016 Jul;24(7-8):441-6.

[13] Majumder S, Mondal T, Deen MJ. Wearable Sensors for Remote Health Monitoring. Sensors (Basel). 2017 Jan;17(1).

[14] Liem SS, van der Hoeven BL, Oemrawsingh PV, Bax JJ, van der Bom JG, Bosch J, et al. MISSION!: optimization of acute and chronic care for patients with acute myocardial infarction. Am Heart J. 2007 Jan;153(1):1-11.

[15] Hartlongcentrum Leiden. The Box [Internet]. LUMC; 2023. [cited 23-01-2023]. Available from: `https://hartlongcentrum.nl/informatie-voor-patienten/the-box/`.

[16] Breeman LD, Keesman M, Atsma DE, Chavannes NH, Janssen V, van Gemert-Pijnen L, et al. A multi-stakeholder approach to eHealth development: Promoting sustained healthy living among cardiovascular patients. Int J Med Inform. 2021 Mar;147:104364.

[17] Van der Velde ET, Atsma DE, Foeken H, Witteman TA, Hoekstra WH. Remote monitoring of patients with implanted devices: data exchange and integration. Eur J Prev Cardiol. 2013 Jun;20(2 Suppl):8-12.

[18] Maas J, Goedemans-de Graaf SA. Uw herstel na een Intensive Care-opname [Internet]. LUMC; 2019. [cited 23-01-2023]. Available from: https://www.lumc.nl/over-het-lumc/afdelingen/intensive-care/overige-folders/uw-herstel-na-een-intensive-care-opname/.

[19] Seferian EG, Afessa B. Demographic and clinical variation of adult intensive care unit utilization from a geographically defined population. Crit Care Med. 2006 Aug;34(8):2113-9.

[20] Badillo S, Banfai B, Birzele F, Davydov II, Hutchinson L, Kam-Thong T, et al. An Introduction to Machine Learning. Clin Pharmacol Ther. 2020 Apr;107(4):871-85.

[21] Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Campbell JP. Introduction to Machine Learning, Neural Networks, and Deep Learning. Transl Vis Sci Technol. 2020 Feb;9(2):14.

[22] Habehh H, Gohel S. Machine Learning in Healthcare. Curr Genomics. 2021 Dec;22(4):291-300.

[23] Tai AMY, Albuquerque A, Carmona NE, Subramanieapillai M, Cha DS, Sheko M, et al. Machine learning and big data: Implications for disease modeling and therapeutic discovery in psychiatry. Artif Intell Med. 2019 Aug;99:101704.

[24] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12:2825-30.

[25] Akavipat P, Thinkhamrop J, Thinkhamrop B, Sriraj W. Acute Physiology and Chronic Health Evaluation (APACHE) II Score) - The Clinical Predictor In Neurosurgical Intensive Care Unit. Acta Clin Croat. 2019 Mar;58(1):50-6.

[26] Peng Y, Wang M, Xu Y, Wu Z, Wang J, Zhang C, et al. Drug repositioning by prediction of drug's anatomical therapeutic chemical code via network-based inference approaches. Brief Bioinform. 2021 Mar;22(2):2058-72.

[27] Verma Y. A Complete Guide to Categorical Data Encoding [Internet]. Analytics India Magazine; 2021. [cited 20-02-2023]. Available from: https://analyticsindiamag.com/a-complete-guide-to-categorical-data-encoding/.

[28] Brownlee J. How to Choose a Feature Selection Method For Machine Learning [Internet]. Machine Learning Mastery; 2019. [cited 20-02-2023]. Available from: https://machinelearningmastery.com/feature-selection-with-real-and-categorical-data/.

[29] Chauhan A. Random Forest Classifier and its Hyperparameters [Internet]. Medium; 2021. [cited 20-02-2023]. Available from: https://medium.com/analytics-vidhya/random-forest-classifier-and-its-hyperparameters-8467bec755f6.

[30] Sagar R. Top Optimisation Methods In Machine Learning [Internet]. Analytics India Magazine; 2020. [cited 27-02-2023]. Available from: https://analyticsindiamag.com/optimisation-machine-learning-methods-gradient-descent/.

[31] Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. Crit Care Med. 2006 May;34(5):1297-310.

[32] Maheswari J. Breaking the curse of small datasets in Machine Learning: Part 1 [Internet]. Towards Data Science; 2018. [cited 03-03-2023]. Available from: https://towardsdatascience.com/breaking-the-curse-of-small-datasets-in-machine-learning-part-1-36f28b0c044d.

A

Introduction

A.1. Literature Review Healthcare Utilization

This page was intentionally left blank

B

Implementation of the ICU-Recover Box

B.1. Pilot Study Protocol

This page was intentionally left blank

C

Predicting ICU Mortality Using Machine Learning

C.1. APACHE II Score Calculation

Physiologic Variable	Points								
	+4	+3	+2	+1	0	+1	+2	+3	+4
1. Temperature (°C)	≥41	39-40.9		38.5-38.9	36-38.4	34-35.9	32-33.9	30-31.9	≤29.9
2. Mean arterial pressure (mmHg)	≥160	130-159	110-129		70-109		50-69		≤49
3. Heart rate (/min)	≥180	140-179	110-139		70-109		55-69	40-54	≤39
4. Respiratory rate (/min)	≥50	35-49		25-34	12-24	10-11	6-9		≤5
5. Oxygenation (mmHg) a. A-aDO_2 if FiO_2 ≥0.5 b. PaO_2 if FiO_2 <0.5	500	350-499	200-349		<200 >70	61-70		55-60	<55
6. Acid-base balance a. Arterial pH b. Serum HCO_3 (mEq/l) if no arterial blood gas	≥7.7 ≥52	7.6-7.69 41-51.9		7.5-7.59 32-40.9	7.33-7.49 22-31.9		7.25-7.32 18-21.9	7.15-7.24 15-17.9	<7.15 <15
7. Sodium (mEq/l)	≥180	160-179	155-159	150-154	130-149		120-129	111-119	≤110
8. Potassium (mEq/l)	≥7	6-6.9		5.5-5.9	3.5-5.4	3-3.4	2.5-2.9		<2.5
9. Creatinine (mg/dl)	≥3.5	2-3.4	1.5-1.9		0.6-1.4		<0.6		
10. Hematocirt (%)	≥60		50-59.9	46-49.9	30-45.9		20-29.9		<2.5
11. White blood count (×1000/mm^3)	≥40		20-39.9	15.19.9	3-14.9		1-2.9		<1
12. Glasgow Coma Score (GCS)	Score = 15 minus actual GCS								
A. Total Acute Physiology Score (sum of 12 above points)									
B. Age points (years) ≤44=0; 45 to 54=2; 55 to 64=3; 65 to 74=5; ≥75=6									
C. Chronic Health Points*									
Total APACHE II Score (add together the points from A+B+C)									

* Chronic Health Points: If the patient has a history of severe organ system insufficiency or is immune-compromised as defined below, assign points as follows:
5 points for non-operative or emergency post-operative patients
2 points for elective post-operative patinets

Figure C.1: Calculation of the APACHE II Score
[25]

C.2. Included Medication

Table C.1: All included medication with corresponding ATC-codes and identifier values

Name	ATC-code	Identifier value
Isoprenalin	C01CA02	387 6760
Norepinephrine	C01CA03	414 6900
Phenylephrine	C01CA06	373 12140 6669
Dobutamine	C01CA07	363 6278
Midodrine	C01CA17	12581 6843
Epinephrine	C01CA24	334 6343
Milrinone	C01CE02	963 6844
Enoximone	C01CE03	368 6633
Terlipressin	H01BA04	436 6515
Nitroprusside	C02DD01	413 6899
Nitroglycerin	C01DA02	412 6897 14737
Dexmedetomidine	N05CM18	12524 16584 6596
Clonidine	C02AC01	468 6553
Sufentanil	N01AH03	1011 433 6531
Morphine	N02AA01	13592 13921 404 6856
Methadone	N07BC02	398 6827
Propofol	N01AX10	7006 425 12990
Midazolam	N05CD08	403 6842
Sotalol	C07AA07	432 6551
Metoprolol	C07AB02	956 6835 12580
Esmolol	C07AB09	915 6643
Amiodarone	C01BD01	337 6353

C.3. Schematic View of Model Selection Methods

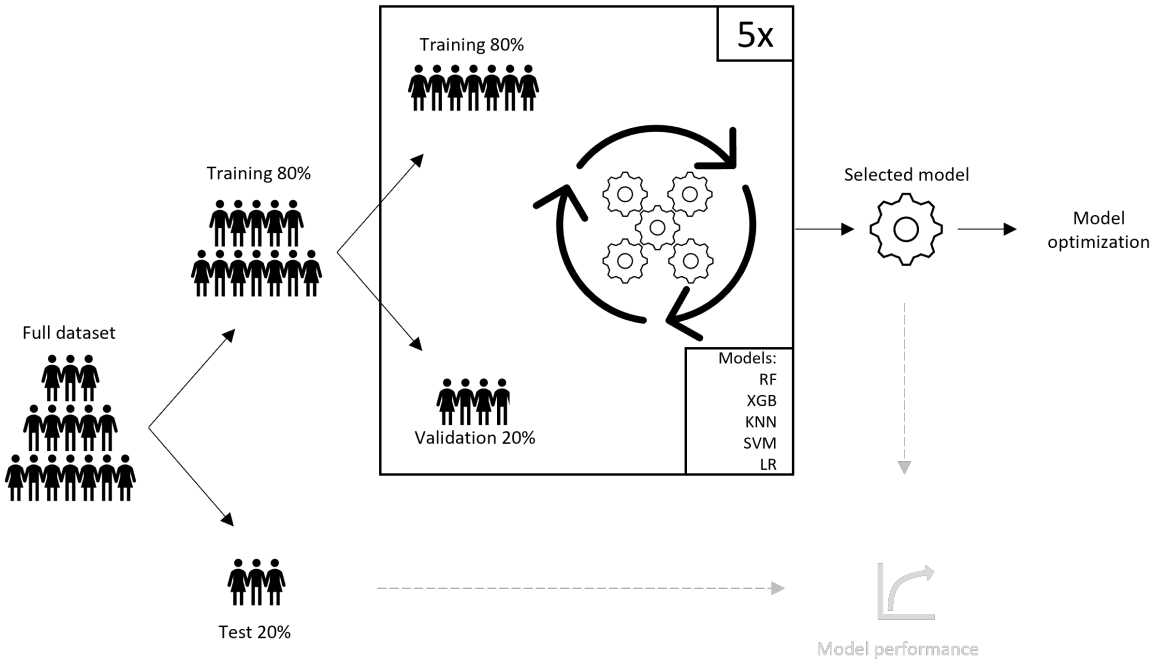


Figure C.2: Schematic view of model selection methods

C.4. Schematic View of Model Optimization Methods

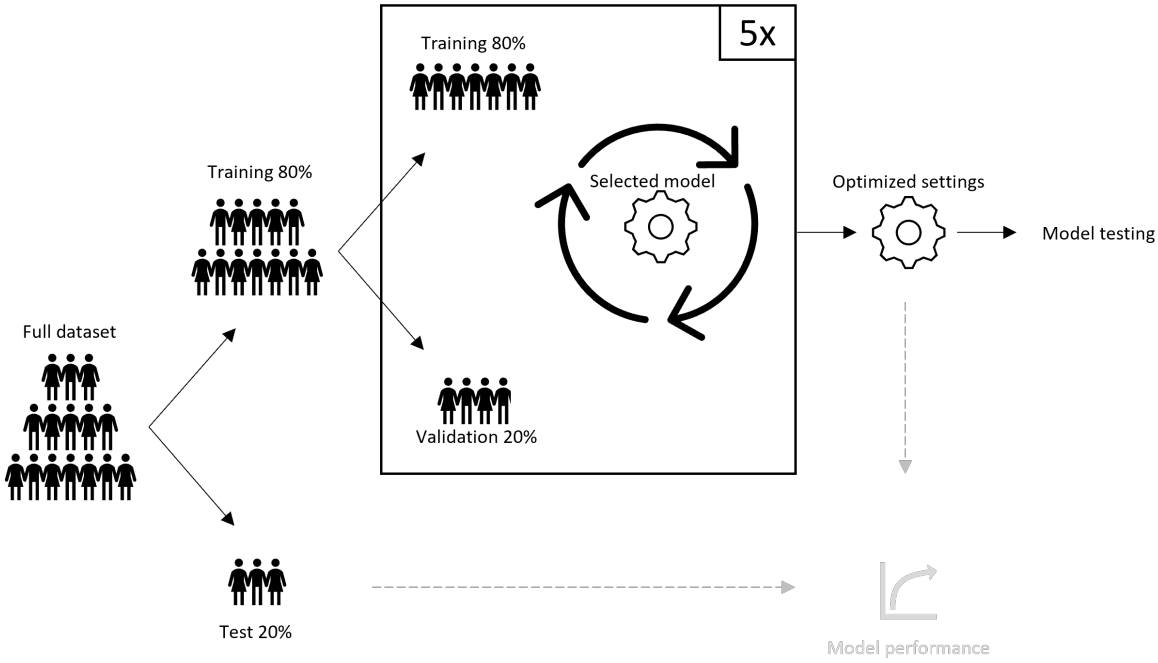


Figure C.3: Schematic view of model optimization methods

C.5. Model Performance Per Iteration

Table C.2: Model performance per iteration

Iteration	AUC	Sensitivity	Specificity	F1 Score
1	0,860	0,286	0,982	0,857
2	0,829	0,313	0,978	0,861
3	0,761	0,255	0,965	0,842
4	0,828	0,302	0,978	0,872
5	0,771	0,245	0,973	0,842
6	0,765	0,304	0,982	0,868
7	0,809	0,289	0,952	0,842
8	0,773	0,283	0,956	0,842
9	0,802	0,333	0,964	0,853
10	0,839	0,333	0,986	0,864
11	0,798	0,302	0,961	0,857
12	0,849	0,229	0,978	0,846
13	0,845	0,286	0,961	0,857
14	0,792	0,200	0,982	0,839
15	0,827	0,375	0,973	0,868
16	0,858	0,324	0,975	0,894
17	0,792	0,367	0,960	0,853
18	0,775	0,200	0,964	0,824
19	0,824	0,333	0,983	0,883
20	0,852	0,404	0,960	0,864
21	0,824	0,326	0,957	0,857
22	0,789	0,229	0,969	0,839
23	0,785	0,333	0,968	0,842
24	0,745	0,279	0,948	0,842
25	0,789	0,286	0,961	0,857
26	0,870	0,447	0,965	0,875
27	0,816	0,273	0,954	0,817
28	0,817	0,362	0,973	0,868
29	0,781	0,286	0,948	0,846
30	0,812	0,300	0,964	0,842
31	0,769	0,227	0,969	0,850
32	0,734	0,293	0,953	0,853
33	0,796	0,396	0,951	0,853
34	0,800	0,310	0,965	0,864
35	0,805	0,298	0,965	0,850
36	0,794	0,244	0,953	0,846
37	0,842	0,313	0,964	0,850
38	0,809	0,400	0,966	0,883
39	0,723	0,231	0,957	0,853
40	0,911	0,405	0,962	0,886
41	0,740	0,213	0,947	0,821
42	0,764	0,306	0,946	0,832
43	0,800	0,231	0,968	0,828
44	0,826	0,302	0,974	0,868
45	0,796	0,286	0,951	0,832
46	0,841	0,367	0,955	0,850
47	0,757	0,357	0,957	0,864
48	0,784	0,189	0,964	0,813
49	0,771	0,263	0,949	0,806
50	0,817	0,389	0,958	0,883

Iteration	AUC	Sensitivity	Specificity	F1 Score
51	0,888	0,315	0,995	0,861
52	0,815	0,347	0,946	0,839
53	0,813	0,204	0,978	0,839
54	0,817	0,327	0,973	0,850
55	0,813	0,269	0,964	0,832
56	0,806	0,220	0,964	0,828
57	0,814	0,259	0,967	0,817
58	0,876	0,184	0,974	0,864
59	0,791	0,189	0,973	0,821
60	0,784	0,310	0,957	0,857
61	0,842	0,367	0,969	0,861
62	0,780	0,205	0,952	0,832
63	0,880	0,455	0,967	0,905
64	0,776	0,347	0,964	0,853
65	0,836	0,288	0,986	0,835
66	0,769	0,275	0,948	0,850
67	0,799	0,276	0,953	0,810
68	0,803	0,222	0,952	0,832
69	0,859	0,375	0,970	0,883
70	0,809	0,163	0,970	0,842
71	0,825	0,367	0,978	0,868
72	0,788	0,364	0,956	0,861
73	0,823	0,339	0,963	0,828
74	0,813	0,229	0,960	0,832
75	0,812	0,200	0,978	0,850
76	0,817	0,304	0,982	0,868
77	0,871	0,316	0,979	0,886
78	0,841	0,268	0,966	0,861
79	0,833	0,388	0,955	0,853
80	0,850	0,297	0,970	0,879
81	0,836	0,500	0,933	0,879
82	0,785	0,295	0,952	0,846
83	0,841	0,304	0,960	0,850
84	0,858	0,333	0,962	0,872
85	0,795	0,255	0,978	0,853
86	0,802	0,306	0,964	0,846
87	0,749	0,260	0,955	0,828
88	0,828	0,472	0,941	0,879
89	0,815	0,350	0,953	0,864
90	0,724	0,239	0,943	0,824
91	0,838	0,333	0,970	0,879
92	0,803	0,239	0,969	0,846
93	0,796	0,209	0,970	0,850
94	0,803	0,271	0,969	0,846
95	0,837	0,333	0,964	0,853
96	0,799	0,205	0,970	0,861
97	0,803	0,229	0,956	0,828
98	0,759	0,224	0,960	0,828
99	0,836	0,356	0,969	0,868
100	0,824	0,292	0,960	0,842