UNIVERSITY
INSTITUTE
OF LISBON

Intelligent Clinical Decision Support System for Managing COPD Patients

José Maria Silva Pereira

Master in **Integrated Business Intelligence Systems**

Supervisor
Doctor João Carlos Amaro Ferreira, Assistant Professor with habilitation
ISCTE-University Institute of Lisbon

Co-Supervisor
Nuno Francisco Castro da Gama Antunes
ISCTE-University Institute of Lisbon

September, 2023

TECNOLOGIAS
E ARQUITETURA

Intelligent Clinical Decision Support System for Managing COPD Patients

José Maria Silva Pereira

Master in **Integrated Business Intelligence Systems**

Supervisor
Doctor João Carlos Amaro Ferreira, Assistant Professor with habilitation
ISCTE-University Institute of Lisbon

Co-Supervisor
Nuno Francisco Castro da Gama Antunes
ISCTE-University Institute of Lisbon

September, 2023

Resumo

Doença Pulmonar Obstrutiva Crónica (DPOC) é a terceira principal causa de morte em todo o mundo. Sistemas de Monitorização Remota de Saúde (SMRS) desempenham um papel crucial na gestão de doentes com DPOC, identificando anomalias em seus sinais biométricos e alertando profissionais de saúde. Ao analisar as relações entre os sinais biométricos e os fatores ambientais, é possível desenvolver modelos de inteligência artificial capazes de inferir os riscos futuros de deterioração da saúde dos doentes. Esta dissertação tem como objetivo desenvolver um Sistema Inteligente de Apoio à Decisão Clínica (SISDC) capaz de fornecer informações precoces sobre a evolução da saúde do paciente e análise de risco para apoiar o tratamento de doentes com DPOC. O SISDC do presente trabalho é composto por dois módulos principais: o Módulo de Previsões de Sinais Vitais e o Módulo de Cálculo do Early Warning Score, que geram informações sobre a saúde do paciente e o risco de deterioração, respectivamente. Além disso, o SISDC gera alertas sempre que uma medição de sinal biométrico estiver fora da intervalo normal de valores para um paciente ou no caso de uma mudança significativa em um valor basal. Finalmente, o sistema foi implementado e avaliado em um caso real e também validado em termos clínicos por meio de um inquérito respondido por profissionais de saúde envolvidos no projeto. Em conclusão, o SISDC demonstra ser uma ferramenta útil e valiosa para profissionais de saúde, permitindo intervenções proativas e facilitando ajustes no tratamento médico dos doentes.

Palavras-chave: Doença Pulmonar Obstrutiva Crónica; Sistema Inteligente de Apoio à Decisão Clínica; Sistema de Monitorização Remota de Saúde; Detecção de Erros em Sinais Biométricos; Escala de Alerta Precoce; Inteligência Artificial; Previsão de Séries Temporais.

Abstract

Chronic obstructive pulmonary disease (COPD) is the third leading cause of death worldwide. Health remote monitoring systems (HRMSs) play a crucial role in managing COPD patients by identifying anomalies in their biometric signs and alerting healthcare professionals. By analyzing the relationships between biometric signs and environmental factors, it is possible to develop artificial intelligence models capable of inferring patients' future health deterioration risks. In this research work, we review recent works in this area and develop an intelligent clinical decision support system (ICDSS) capable of providing early information concerning patient health evolution and risk analysis in order to support the treatment of COPD patients. The present work's ICDSS is composed of two main modules: the vital signs prediction module and the early warning score calculation module, which generate the patient health information and deterioration risks, respectively. Additionally, the ICDSS generates alerts whenever a biometric sign measurement falls outside the allowed range for a patient or in case a basal value changes significantly. Finally, the system was implemented and assessed in a real case and validated in clinical terms through an evaluation survey answered by healthcare professionals involved in the project. In conclusion, the ICDSS proves to be a useful and valuable tool for medical and healthcare professionals, enabling proactive intervention and facilitating adjustments to the medical treatment of patients.

Keywords: Chronic Obstructive Pulmonary Disease; Intelligent Clinical Decision Support System; Health Remote Monitoring Systems; Biometric Signs Errors Detection; Early Warning Score; Artificial Intelligence; Time Series Prediction.

Acknowledgements

I would like to express my gratitude to my supervisors, Professor João Ferreira and Nuno Antunes, for the knowledge they have shared and their collaboration, guidance, and availability not only during the development of this project but also throughout my master's program.

I extend my thanks to INOV INESC Inovação and Pedro Santos for granting me the opportunity to work on this dissertation in collaboration with the institution. A special acknowledgment to my project manager Joana Rosa for her collaboration and guidance.To the team that shared with great moments of companionship in the development of this research, particularly Pedro Águas, Filipe Sousa, Alexandre Vidal, Sérgio Valentim and Daniel Cale, as well as the two new INOV incomers, Catarina Brito and Matilde Ferreira. I am also grateful to Hopecare, particularly Manuel Pereira and José Paulo Carvalho, and to Professor Sandra Mogo from the University of Beira Interior for their collaboration and knowledge during the research work published in the project. To my Master's colleagues, I extend my genuine appreciation for their consistent support and team-based cooperation. I also want to express my appreciation to Professor Luis Elvas, not only for the valuable knowledge he imparted but also for his companionship.

My gratitude to my parents and my brother, who I look up to, for their emotional support during challenging moments and their priceless guidance, empowering me to make better choices throughout this academic journey.

A special thanks to my family, who have always been a great source of inspiration and support. Their continuous support, dedication, and belief in my abilities have inspired me to move forward and accomplish my goals.

To my close friends who were a constant presence during this phase, thank you for your support during challenging moments and the laughter we shared during the good ones.

All these people contributed, directly and indirectly, to the conclusion of this master thesis as well as this academic phase, and without them, none of this would have been possible.

José Maria Silva Pereira

Contents

List of Figures

List of Tables

Abbreviations

ARIMA	**A**utoregressive **I**ntegrated **M**oving **A**verage
BPM	**B**eats **P**er **M**inute
BT	**B**ody **T**emperature
BI-LSTM	**B**idirectional **L**ong **S**hort-**T**erm **M**emory
COPD	**C**hronic **O**bstructive **P**ulmonary **D**isease
DBP	**D**iastolic **B**lood **P**ressure
DSS	**D**ecision **S**upport **S**ystem
EPA	**E**nvironmental **P**rotection **A**gency
EWS	**E**arly **W**arning **S**core
GOLD	**G**lobal **I**nitiative for **O**bstructive **L**ung **D**isease
GRU	**G**ated **R**ecurrent **U**nit
HR_MED	**H**umidity **R**atio **M**edian
HRMS	**H**ealth **R**emote **M**onitoring **S**ystem
ICDSS	**I**ntelligent **C**linical **D**ecision **S**upport **S**ystem
IPMA	**P**ortuguese **I**nstitute for the **O**cean and **A**tmosphere
LSTM	**L**ong **S**hort-**T**erm **M**emory
LightGBM	**L**ight **G**radient **B**oosting **M**achine
NOAA HYSPLIT	**N**ational **O**ceanic and **A**tmospheric **A**dministration **H**ybrid **S**ingle **P**article **L**agrangian **I**ntegrated **T**rajectory
PR_QTD	**P**recipitation **Q**uantity **T**otal **D**aily
RMSE	**R**oot **M**ean **S**quare **E**rror
SBP	**S**ystolic **B**lood **P**ressure
SPO2	**S**aturation of **P**eripheral **O**xygen
T_MED	**M**edian **T**emperature

TVM	**T**riage **V**alidation **M**odule
WHO	**W**orld **H**ealth **O**rganization
XGBoost	**X**treme **G**radient **B**oosting

Chapter 1

Introduction

1.1 Chronic Obstructive Pulmonary Disease

According to the World Health Organization (WHO), Chronic obstructive pulmonary disease (COPD) is one of the most deadly major lung diseases and the third leading cause of death worldwide [1]; the organization further indicates that COPD was responsible for about 3.24 million deaths in 2019. The Portuguese Society of Pulmonology [2] estimates that 5.42% of individuals in Portugal between the ages of 35 and 69 suffer from COPD. According to the Portuguese Lung Foundation [3], COPD was responsible for approximately 2834 fatalities in the country. The same organization estimates that in 2019, this illness cost the economy 1.6 billion euros.

1.1.1 Disease Symptoms

COPD is caused by airway obstruction. The most common symptoms of COPD are coughing, wheezing, and dyspnea (shortness of breath). Patients often seek medical attention only when the disease reaches an advanced stage, as it is a condition that progresses slowly.

Initially, the disease presents as a cough accompanied by increased sputum production. However, as it progresses, it can lead to repeated episodes of acute

bronchitis and respiratory infections. As the disease develops, shortness of breath becomes more frequent, even with seemingly minor tasks, such as talking and performing daily hygiene. Shortness of breath is most noticeable during activities that require physical effort.

1.1.2 COPD Exacerbations and Their Prevention

COPD exacerbations are associated with a worsening of the disease, a deterioration in the patient's health status, and an accelerated decline in the patient's respiratory function. A severe exacerbation of COPD always leads to the need for medical intervention and eventual hospitalization.

A medical professional can determine whether a patient is experiencing an exacerbation through the values of the vital signs using an Early Warning Score system. The Early Warning Score is a protocol that aims to improve the detection and response time to situations of clinical deterioration. Depending on the score given by this protocol, we can detect the level of deterioration of the patient, as shown in Figure 1.1 as an example.

Physiological parameter	Score						
	3	2	1	0	1	2	3
Respiration rate (per minute)	≤8		9–11	12–20		21–24	≥25
SpO_2 Scale 1 (%)	≤91	92–93	94–95	≥96			
SpO_2 Scale 2 (%)	≤83	84–85	86–87	88–92 ≥93 on air	93–94 on oxygen	95–96 on oxygen	≥97 on oxygen
Air or oxygen?		Oxygen		Air			
Systolic blood pressure (mmHg)	≤90	91–100	101–110	111–219			≥220
Pulse (per minute)	≤40		41–50	51–90	91–110	111–130	≥131
Consciousness				Alert			CVPU
Temperature (°C)	≤35.0		35.1–36.0	36.1–38.0	38.1–39.0	≥39.1	

FIGURE 1.1: The current NEWS2 scoring system. This is the version currently recommended by the UK Royal College of Physicians for use in clinical practice [4].

1.2 Motivation

The integration of technology into healthcare has revolutionized patient care, with health remote monitoring systems (HRMSs) emerging as powerful tools [5]. By storing data, such as heart rate (HR) and oxygen saturation (SPO2) levels, HRMSs help medical professionals to treat patients with COPD. These systems offer real-time monitoring and personalized treatment options. However, to maximize the potential of HRMSs, it is crucial to integrate them with well-defined clinical processes, therapeutics, and rules. This integration ensures that the collected measurements are correlated and directly linked to effective patient care, enabling proactive interventions and improving health outcomes.

The Internet of Things plays a crucial and influential role in the successful implementation of HRMSs [6]. Wearable device sensors, videos, and images are essential to gathering valuable patient information. Daily physiological data of the patient is collected and stored by the HRMS through data processing tools, analytics, and artificial intelligence (AI). Recording daily physiological data provides healthcare providers with actionable insights, facilitating proactive and personalized care.

The use of AI by HRMSs to predict patient health deterioration is a significant benefit[7, 8].AI algorithms examine historical patient data to find patterns that might point to higher risks of unfavorable events or health deterioration. These forecasts offer healthcare professionals with insightful information that enables them to intervene early and prevent complications. A more preventive model of care is promoted by this proactive approach, which also enhances patient safety and lowers hospital admissions.

1.3 Objectives

The research question addressed by this study is: “Is it possible to automatically monitor and analyse the risk of potential health deteriorations of COPD patients?”. With this research question in mind, the defined objective is to develop a system

capable of providing early information concerning patient health evolution and exacerbation risk analysis in order to support the treatment of patients with COPD. Additionally, the system allows healthcare professionals to more efficiently manage their time by automatically providing said professionals with alerts, supported by a risk analysis of the patient's COPD health status.

1.4 The HC PSI Project

The Hope Care Intelligent Services Platform (HC PSI) is a P2020 project that involves the participation of Hope Care SA, INOV—INESC Inovação and the University of Beira Interior. Its main objective is to research and develop an intelligent services platform that enables healthcare professionals to make more informed decisions regarding the health conditions of COPD patients, thereby increasing the efficiency of clinical entities.

The components of the HC PSI include a ICDSS, HCAlert platform, and environmental data sources, all geared toward automating the clinical treatment of remotely monitored COPD patients.

This dissertation focuses on the ICDSS developed by INOV—INESC Inovação. The ICDSS assists in making decisions regarding patient treatment. This platform is composed of three modules: an HRMS that provides patients' health information through a mobile application to the ICDSS, a TVM that receives and processes patient risk information from the ICDSS, and a graphical user interface (GUI) that displays relevant clinical information to healthcare professionals.

Figure 1.2 presents the HC PSI architecture, which includes the ICDSS developed by INOV, the HCAlert platform, and other external data sources.

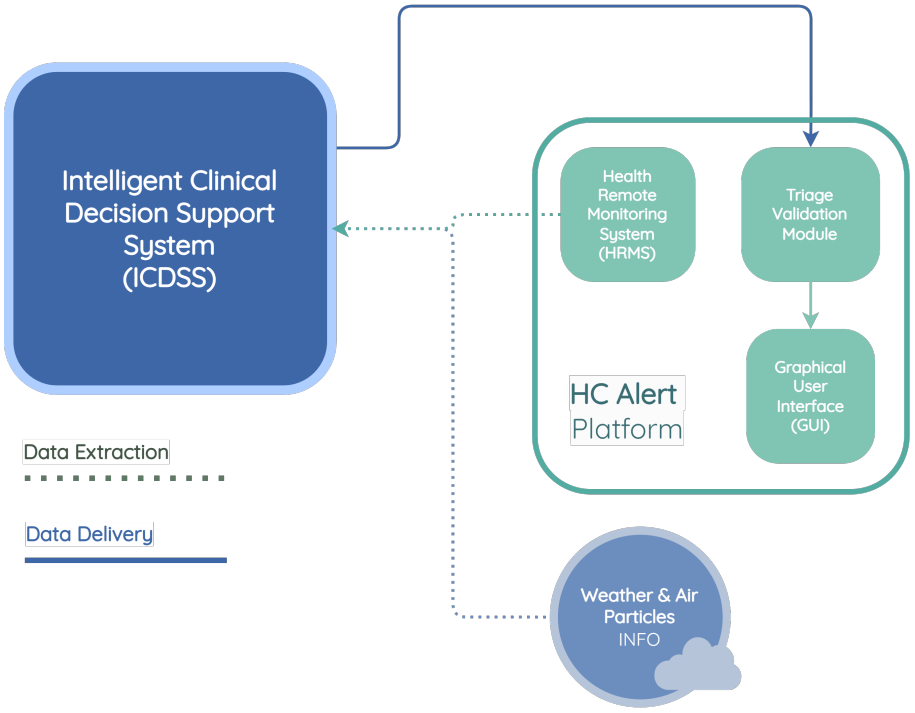

FIGURE 1.2: HC PSI architecture including ICDSS developed by INOV, the HC Alert Platform and other external data sources.

The HCAlert platform was developed by Hope Care SA and includes a mobile application that supports HRMSs and a set of backend services for clinical validation and triage.

In the scope of the HC PSI project, the requirements for the HCAlert mobile application include the collection of patient symptoms and residential data. For the clinical validation and triage backend services, the following requirements are defined:

- Capability to categorize alerts.
- Capability to provide Early Warning Scores and other relevant metrics per patient to healthcare professionals.
- Capability to obtain information about hospital visits internally or from other sources.
- Enabling the clinical team to have an overview of new alerts per patient, including client name, data type and last measurement date.

- Allowing the clinical team to define which relevant health values to display on the dashboard.

1.5 Methodology

The DSR methodology is a research methodology commonly used in the field of information systems; it focuses on the development and evaluation of innovative artifacts, which include cutting-edge framework prototypes, techniques, and algorithms that address present-day challenges. It consists of the following six phases: problem identification, definition of objectives, design and development, demonstration, evaluation, and communication. This methodology focuses on creating and evaluating artifacts based on their effectiveness, quality, and usefulness in addressing real-world problems [9].

In this dissertation, since we developed a Intelligent Clinical Decision Support System (ICDSS) which is an interactive information system that analyzes large volumes of data for informing business decisions, we applied the design science research methodology.

Figure 1.3 presents the iterations within the design science research methodology (DSRM) process.

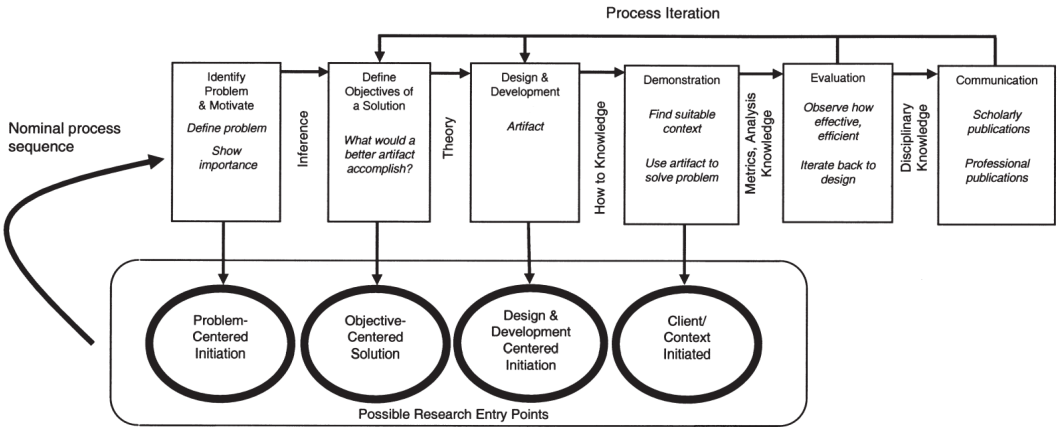

FIGURE 1.3: Iterations represented in the design science research methodology (DSRM) process model; Peffers et al.[10]

1.6 Outline of the Dissertation

Having the objectives and methodology defined, we will have five chapters (Introduction included). The chapters are:

Chapter 2: Outlines a systematic literature review on In-Home Healthcare for COPD, E-Health Care supported by Predictive Analytics, factors related to COPD deteriorations, and Machine Learning for Early Identification of Deterioration, using the PRISMA method.

Chapter 3: Provides a detailed description of our ICDSS. This includes a thorough exploration of each module within the system, covering aspects such as data extraction, system modules, and health information provided to medical professionals.

Chapter 4: Presents a demonstration of the system, showcasing the interaction between modules, including the generation and reception of input and output. Additionally, we highlight the system's evaluation process, ensuring its usefulness and impact in a clinical context.

Chapter 5: Presents the discussion and conclusions of the work developed, where we highlight the contributions and limitations of our efforts.

Chapter 2

Related Work

In this chapter, we present an overview of the systematic review conducted in this article, which followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) Methodology [11]. This chapter also covers the latest advances in managing pulmonary disease patients, particularly on COPD patients. It includes the improvement of effectiveness that remote health monitoring brings to patients' treatment by providing a real-time warning to medical professionals.Additionally, we explore how the integration of predictive analytics in remote health monitoring improves patients' assistance management by offering early warnings of potential patient deterioration risks, thus optimizing effectiveness.

The systematic review also covers factors and biometric signs related to the acute deterioration of COPD and how prediction of biometric signs and subsequent early warning generation provide a risk of patient future deterioration. Table 2.1 presents the topics and the respective queries used to extracted and filtered related works.

TABLE 2.1: Related work topics and the corresponding queries used to filter research papers related to each topic.

Subsection	Query
In-Home Health Care for COPD	("Healthcare Management Systems" AND "Real-time Detection")
E-Health Care supported by Predictive analytics	("Healthcare Management Systems" AND "Early Detection" AND ("Artificial Intelligence" AND "Machine Learning"))
Factors related with COPD deteriorations	("Early Detection" AND "Vital Signs" AND "COPD")
Machine Learning for for Early Identification of a Deterioration	("Early Detection" AND "Vital Signs" AND "Machine Learning")

Table 2.2 presents the eligibility criteria used to filter documents in the related work.

TABLE 2.2: Eligibility criteria to filter research papers

Eligibility criteria	
Inclusion Criteria	Exclusion Criteria
Written in English or Portuguese	Not written in English nor Portuguese
Publication date after/during 2010	Publication date before 2010

We identified 810 documents, with 10 documents removed due to duplication issues. A total of 400 articles not related to healthcare or artificial intelligence (AI) were excluded from further screening based on titles and abstracts. Moreover, 40 articles were excluded as we were unable to access their full versions, leaving 160 articles for full-text screening. A total of 82 articles were removed as they did not fit the eligibility criteria. Finally, 56 articles were excluded as they did not contain relevant information concerning vital signs, time series techniques, and HRMSs.

The selection results, according to the PRISMA flow diagram, are shown in Figure 2.1.

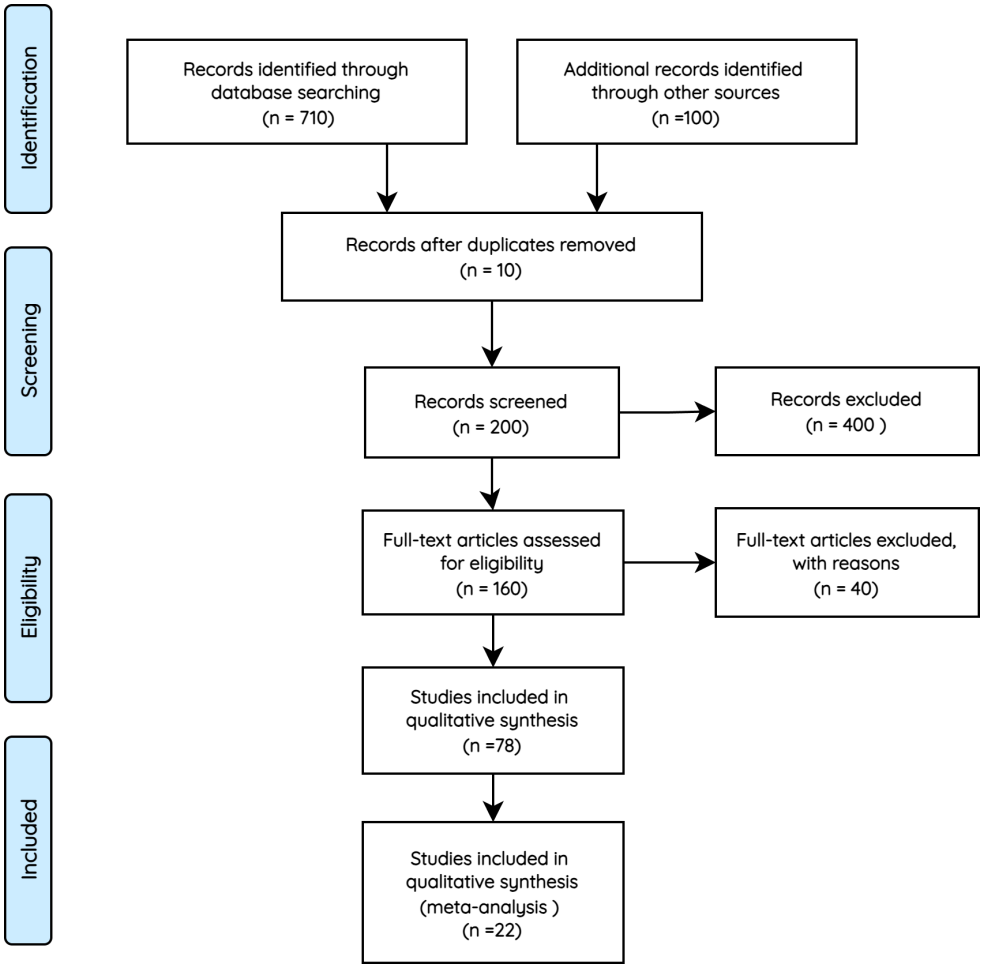

FIGURE 2.1: PRISMA methodology [11]

2.1 In-Home Healthcare for COPD

Home telemonitoring refers to the utilization of audio, video, and various telecommunication technologies to monitor a patient's status from a distance [12]. This method involves remotely monitoring a patient's health parameters, usually within a larger chronic care model. Telemonitoring is a crucial component of telehealth and telemedicine [13], showing potential in aiding patients in managing diseases

and predicting complications [14]. Projects using telemonitoring with patients having pulmonary conditions have demonstrated the capability to detect early changes in a patient's condition, allowing for immediate intervention and the prevention of exacerbation. Patients have shown a positive reception toward telemonitoring as a patient management strategy [12].

A systematic review and meta-analysis have found that telemonitoring interventions can prevent unnecessary visits to the emergency room and potentially reduce severe COPD exacerbations. In a meta-analysis of 20 studies with six-month telemonitoring interventions, it was found that the intervention effectively decreased the number of ER visits (pooled SMD = 0.14, corresponding to a small effect size; 95% CI: 0.28, 0.01) [13].

In a retrospective, population-based cohort study involving 944 individuals using telemonitoring and 9838 control individuals, the total direct medical costs were significantly lower in the telemonitoring group (EUR -895.11, $p = 0.04$). The main factor driving the total cost difference was the reduction in hospitalization costs by EUR -1056.04 ($p = 0.01$). A lower percentage of individuals died in the intervention group than in the control group (3.23 vs. 6.22%, $p < 0.0001$), resulting in a mortality hazard ratio (HR) of 0.51 (95% CI: 0.30–0.86). Over a 12-month period, the proportion of patients hospitalized due to all causes (-15.16%, $p < 0.0001$), due to COPD (-20.27%, $p < 0.0001$), and for COPD-related emergency department (ED) visits (-17.00%, $p < 0.0001$) was consistently lower in telemonitoring patients, leading to fewer all-cause admissions (-0.21, $p < 0.0001$), fewer COPD-related admissions (-0.18, $p < 0.0001$), and fewer COPD-related ED admissions [15].

2.2 E-Healthcare Supported by Predictive Analytics

Telemonitoring has become indispensable in diagnosing and medically intervening for COPD patients. Nowadays, due to better storage of electronic health records

and improved vital sign detection methods, large amounts of patient data are available daily in ICUs [16]. Medical equipment, ranging from hands-free monitors and portable devices to modern wristbands and watch-like monitors, have helped in the collection of biometric data, such as heart rate, blood pressure, physical activity, and sleep information [17].

A remote monitoring system, capable of gathering extensive data and backed by predictive analytics algorithms and techniques for effective data assessment and identifying underlying patterns, provides better efficiency in identifying declining patient health [18]. In the present COPD case study, such systems can reduce emergency room (ER) visits, acute deterioration-related readmissions, days spent in the hospital, and mortality in patients with COPD [19].

Predictive analytics refers to the systematic use of statistical or machine learning methods to make predictions and support decision-making. Predictive analytics applied to healthcare can be divided into two components: the data underlying the model, particularly predictors or features, and machine learning and statistical methods, both based on a set of mathematical techniques applied to data in order to generate an output [20].

Machine learning is a crucial methodology in predictive analytics. Conventional statistical analysis focuses on explaining data and relies on an expert (i.e., human) to formulate and discover cause–effect relationships, driven by a set of predefined assumptions. Machine learning is more data-focused and orientated toward generating hypotheses and building predictive models using algorithms. It has enabled clinical support research and applications to provide actionable insights by utilizing large amounts of intensive care unit patient datasets that are useful in many clinical scenarios [16]. Machine learning can predict in-hospital mortality and the risk of 30-day readmission due to COPD exacerbation [21].

2.3 Factors associated with COPD exacerbations

2.3.1 Biometric signs associated with COPD exacerbations

The prevention of acute exacerbation in COPD requires the identification of factors associated with exacerbation. Most studies have shown that oxygen saturation (SpO2) (p-value < 0.05), respiratory rate (RR), and heart rate (HR) (p-value < 0.05) influence exacerbation events, with SpO2 being the most predictive vital sign. The deterioration in COPD patients has been associated with a slight decrease in oxygen saturation and a slight increase in HR. One article suggested that using multiple vital signs as the inputs of a single classifier could provide better predictions, given that these multiple-input models showed the best AUC results [22].

Although some studies monitored blood pressure in order to determine whether there was a significant correlation with acute exacerbation, there was no sufficient evidence indicating that a change in blood pressure during a COPD exacerbation was a potent predictive factor for exacerbation (p-value > 0.05, i.e., not significant).

Body temperature with a p-value equal to 0.059 could be considered an exacerbation predictor. In the study conducted by Martin-Lesende, changes in body temperature had triggered 27.8% of alerts, of which, 5% were due to temperatures exceeding 37 °C [23].

2.3.2 External factors associated with COPD exacerbations

Most studies have focused on vital signs and internal factors of COPD patients, rather than external ones, despite being equally relevant. Some meteorological data, such as humidity (p-value $= 0.0137$), variation of diurnal temperature (p-value $= 0.0472$), the cumulative lowest temperature 7 days prior to acute deterioration (p-value $= 0.005$), and total rainfall in the 7 days preceding an acute exacerbation (p-value $= 0.0389$) was associated with acute exacerbation in COPD.

Over the recent decades, several epidemiological studies have shown that exposure to particulate matter (PM), including coarse and fine fractions, has a negative influence on health [24, 25, 26, 27]. This particulate matter may originate from either a natural source, like desert dust, or a human-made one, such aerosols produced by burning biomass or burning fossil fuels. The concentration of particles in the atmosphere relies on emission sources, meteorological factors, and transport mechanisms, considering that aerosols can traverse great distances (transported by air masses). Additionally, household activities can be significant sources of fine particles. Particles resulting from cooking and heating can penetrate the respiratory system more deeply, particularly when they are finer. Lee J. [28] conducted a univariate analysis on air pollution and COPD exacerbations, revealing a substantial correlation between PM10 levels one day before a patient's condition deteriorated and acute exacerbation (p-value = 0.0260) [28].

The World Health Organization's data on household air pollution indicates that COPD accounts for 19% of the 3.2 million deaths linked to exposure to household air pollution. In addition, 23% of all COPD-related deaths in adults in low-income and middle-income countries are linked to exposure to household air pollution [29].

The analysis of both internal and external factors with significant correlations to COPD exacerbation revealed that the frequency with which certain variables are measured must also be taken into consideration. The higher the frequency of a vital sign measurement, the better the perception of its association with an exacerbation occurrence. Daily or multi-daily vital sign monitoring improves the analysis of these signs. For example, Pépin J-L [17] mentions that overnight pulse oximetry increases sensitivity, allowing for early detection of deterioration [17].

2.4 Machine learning for Early Identification of a Deterioration

In recent literature, machine learning techniques have attracted attention for predicting the clinical conditions of patients. Time series forecasting models have

been applied successfully in medical applications to predict disease progression, estimate mortality rates, and assess time-dependent risks. These models are able to identify patterns and trends from sequential data collected over time, such as health-related signals [30, 31].

Some traditional machine learning techniques, such as random forest, SVM (support vector machine), Bayesian networks, and logistic regression, have been employed to improve predictive performance in identifying early clinical deterioration [32]. However, these traditional models are not optimized for handling the unique characteristics of time series data, such as autocorrelation, seasonality, and trend patterns [33, 34].

With sufficient data, the development of deep learning models can reduce several preprocessing steps, emphasizing the relationships between the data, without the need to identify the best predictors, leading to better results [35]. For instance, long short-term memory network (LSTM) can learn extended time series dependencies, while a convolutional neural network can generate a compact latent representation.

Gradient boosting models are alternatives to specialized models, such as long short-term memory network (LSTM) and gated recurrent unit (GRU) [36, 37]. Although these models are not ideal for time series forecasting, they are still generally better suited for handling sequential data compared to non-sequential algorithms (such as random forest, SVM, logistic regression, and naive Bayes) [34].

Chapter 3

Intelligent Clinical Decision Support System Design & Development

The ICDSS receives every patient's vital signs, which are remotely monitored by Hope Care SA as inputs. Additionally, it daily incorporates weather forecast conditions and air particle forecasts that are specific to each patient's location. In response, the system provides daily vital sign predictions and early warning scores for each patient for the following five days. It also provides the basal values of each patient and issues an alert whenever a vital sign measurement falls outside the expected parameter range, requiring a reevaluation.

Figure 3.1 illustrates the ICDSS developed by INOV—INESC Inovação, its interactions with weather and air pollution data providers, and the HCAlert platform. The ICDSS comprises five distinct modules, each serving a specific purpose. These modules are as follows:

Communication manager - This module assumes a crucial role within the system, and is responsible for the communication interactions among HC (Hope Care) Alert, weather, air particles API, and the clinical decision support system.

Vital signs prediction module - It is designed to generate forecasts for a five-day period regarding four essential vital signs: oxygen saturation level (SpO2), heart rate, systolic blood pressure (SBP), and body temperature. This module

utilizes various machine learning algorithms to accomplish the predictions. The input data for these models are sourced from the stored vital sign records within the database. Subsequently, the predicted vital signs are stored back in the database for further reference and analysis.

Early warning score calculation module - Within this module, the recorded vital sign predictions from the database play a crucial role in calculating the early warning score for each of the five predicted days. The early warning score is computed using the aforementioned vital sign data and the resulting early warning scores are subsequently stored in the database.

Biometric signal error detection module - The primary objective of this module is to thoroughly analyze and evaluate potential measurement errors and abnormal variations detected within the patient's historical data. The purpose is to promptly alert both the patients themselves and the attending nurse regarding the invalidity or questionable nature of the entered information. By diligently identifying such anomalies, this module serves as a critical mechanism for ensuring data accuracy and reliability within the system.

Basal value monitoring module - The main function is to monitor and continuously and intelligently adjust the patient's baseline values. This adjustment is based on the historical records of vital sign values measured by the patient and documented within the HCAlert platform. The module's purpose is to enhance the precision and effectiveness of the monitoring system by dynamically adapting the baseline values in accordance with the patient's specific health history.

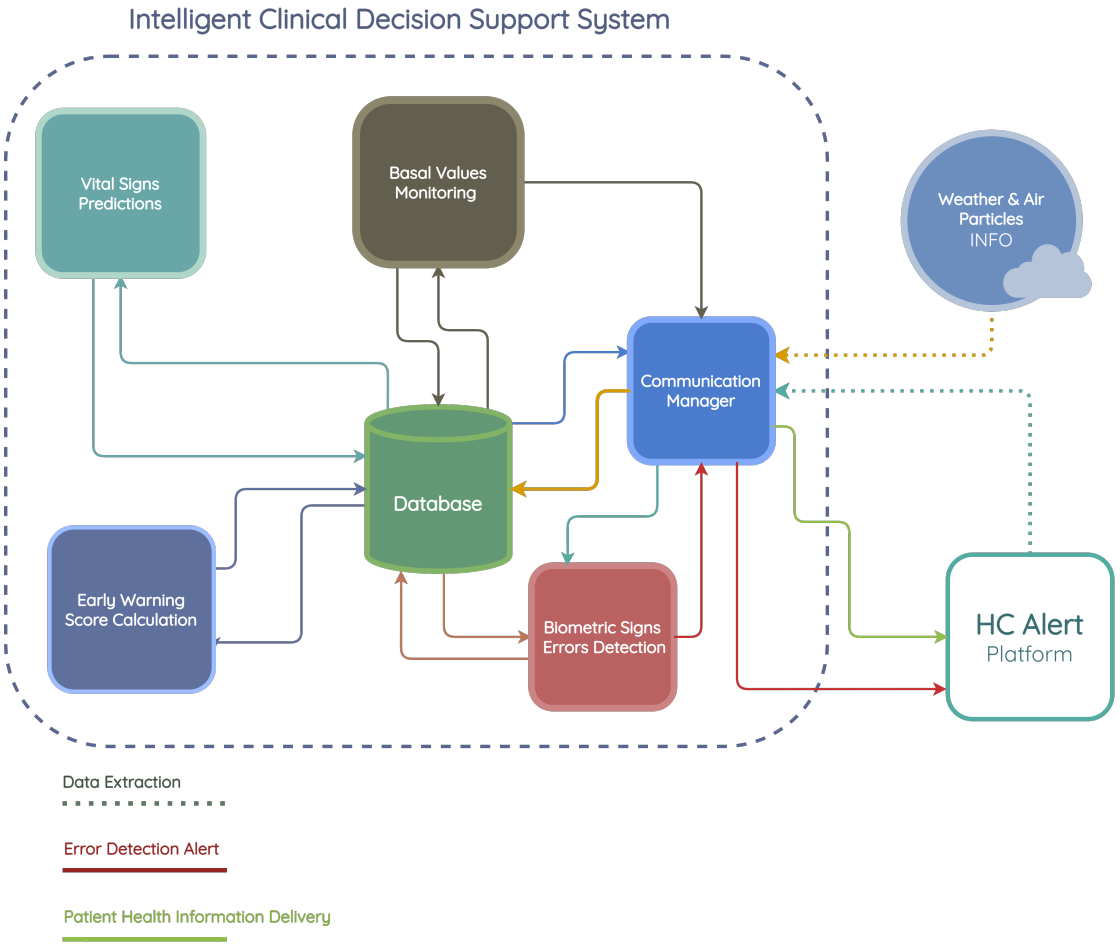

FIGURE 3.1: ICDSS architecture and interactions with external modules

3.1 Requirements

During the initial phase of the HC PSI project, we defined the functional requirements through an interactive and iterative process involving UBI and Hope Care SA. Certain clinical-oriented requirements were specifically delegated based on their domain of expertise. Subsequently, the remaining requirements served as the fundamental basis for the development of the ICDSS discussed in this article. All ICDSS functional requirements have been grouped into system modules, as shown in the following Table 3.1.

TABLE 3.1: Functional requirements associated with each module.

Description	Module
The predictive service should collect environmental data, such as air quality, seasonal infection incidences, and weather conditions	**Vital Signs Prediction**
The predictive service should correlate parameters and detect patterns	
The predictive service should reevaluate the weighting of each parameter depending on the context (e.g., patient, clinical history, etc.)	
The collected data should undergo anonymization (if applicable), normalization, and data fusion	
The predictive service should consider the Early Warning Score to generate alerts	**Communication Manager**
The predictive service should consider the alert classification to detect false positives	
The predictive service should advise the user to take a new measurement and launch inquiries to validate if it's a false positive	**Biometric Signs Error Detection**
The predictive system should apply the Early Warning Score to the clinical protocol and suggest changes to the protocol based on the basal value	**Early Warning Calculation**
The predictive service should calculate the Early Warning Score (defining the correlation weighting of each parameter in the EWS calculation)	
The predictive system should recommend reassessment of the basal value	**Basal Value Monitoring**
The predictive system should take into account changes made to the clinical protocol by the clinical team	
The predictive system should analyze the threshold for advising changes to the applied clinical protocol for the patient	

3.2 Database Architecture

The ICDSS database architecture represented in Figure 3.2 was developed based on MariaDB. Its main purpose was to store all information related to vital signs and external data used as input for the model, as well as the predictions generated by the vital signs predictions module. This includes the error between the predicted values and the real values recorded and provided by the Hopecare API, along with information concerning the early warning score. The database was also intended to record the history of basal values and to store information regarding whether a certain vital sign was valid or not, determined by the biometric signs error detection.

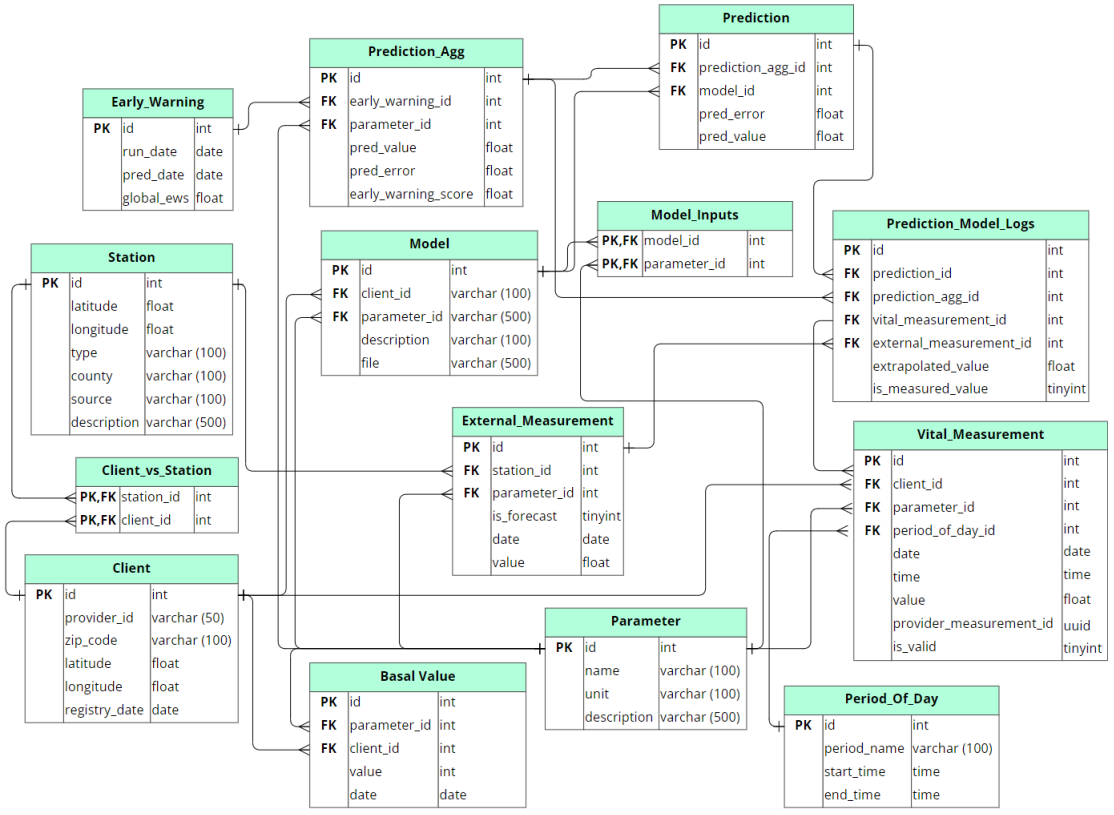

FIGURE 3.2: ICDSS Database architecture

3.3 Communication Manager

This module is composed of four submodules: data extraction, measurement error alert, basal values notification, and the patient's risk information delivery submodule, as is present in Figure 3.3.

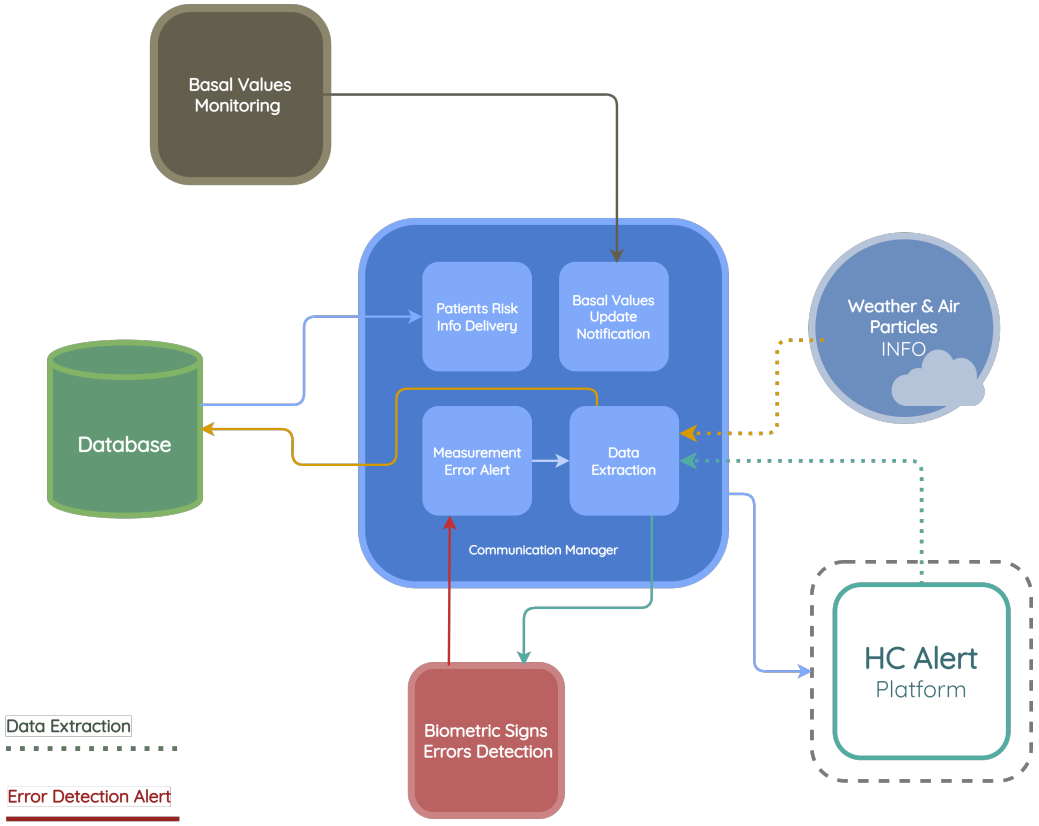

FIGURE 3.3: Communication manager module architecture

3.3.1 Data Extraction

The medical records, which stored the vital signs used as input for the ICDSS, are presented in Table 3.2. Each record is formatted to have one entry per day per parameter. Each record had an ID (idRawMeasurement), the collection date (createdOn), the coordinates where it was collected (latitude and longitude), the measurement type (ProviderMNameStandard), measurement value (value), and the units representing the value (units).

The measurement type could address various factors, including vital signs, such as oxygen saturation level (SpO2), heart rate (HR), body temperature, systolic blood pressure (SBP), and diastolic blood pressure (DBP), as well as other biometric indicators, like the number of steps, body fat, energy burned, weight, and height.

TABLE 3.2: Clinical information extracted from the Hope Care API.

idRawMeasurement	Measurement identifier
createdOn	Measurement creation date
clientID	Identification of the patient to whom the measurement belongs
Latitude	Latitude of the patient
Longitude	Longitude of the patient
ProviderMNameStandard	Standard name of the type of measurement
Value	Measurement value
Unit	Units of measurement (in the dataset are available %, C, bpm, count, mmHg, NA, null and percent)

The weather historical information used as input for the predictive models was provided by the Weatherbit API. Each record had an ID (idWeatherMeasurement), the coordinates of the station (latitude, longitude), date of measurement (columns year, month, day), mean daily temperature (T_MED), and mean relative humidity (HR_MED), as shown in Table 3.3.

Table 3.3: Weather historical information.

idWeatherMeasurement	Measurement identifier
Station ID	Station identifier
Latitude	Latitude of the station
Longitude	Longitude of the station
Year	Year of the collected measurement
Month	Month of the collected measurement
Day	Day of the collected measurement
T_MED	Value of the daily mean temperature in celsius
HR_MED	Value of the daily mean relative humidity in percent

The air pollution historical information used as input for the predictive models was provided by the OpenWeather API. Each record had an ID (idWeatherMeasurement), the coordinates of the station (latitude, longitude), date of the measurement, an average count of 10-micrometer particles (PM10), and an average count of 2.5-micrometer particles (PM2_5), as shown in Table 3.4.

Table 3.4: Air pollution historical information.

idParticlesMeasurement	Measurement identifier
Location	Location of the station
Latitude	Latitude of the station
Longitude	Longitude of the station
Date	Date of the collected measurement
PM10	Value of PM10
PM2_5	Value of PM2.5

3.3.2 Measurement Error Alert

This submodule was designed to receive alerts from the biometric sign error detection module and subsequently send alerts to the HCAlert platform. After a set short duration, it sends a notification to the data extraction submodule to execute the data extraction of biometric signs from HCAlert, concerning the specific patient dataset where the error was found.

3.3.3 Basal Value Monitoring Notification

The basal value update notification submodule was designed to receive notifications from the basal value monitoring module; it subsequently notifies the HCAlert platform with new basal value recommendations for a specific patient.

3.3.4 Patients Risk Information Delivery

The patient risk information delivery submodule extracts information regarding the last five days of vital sign predictions and the calculated early warning scores stored in the database. It then sends this information to the HCAlert platform.

3.4 Biometric Signs Errors Detection

The HCAlert platform's operational efficiency is affected by the patients' inaccurate vital sign measurements, which can result in inaccurate clinical protocol adjustment alerts and future vital sign projections. It is necessary to guarantee that the system receives data that obey certain quality levels.

Prior to the implementation of the current project, measurements are validated by nurses who identified instances of anomalous readings, reporting potential causes, such as deterioration in the patient's condition, measurement errors, cold fingers during measurements, etc.

The biometric sign error detection module consists of three components:

- **Validation of clinical rules**: This component compares the measurements taken by the patient with a set of business rules defined according to Hope Care guidelines. For example, a measurement of oxygen saturation above 100 or below 20 cannot be correct since a percentage value cannot exceed 100, and a value below 20 corresponds to situations of compromised brain function and even comas. The medical team involved in this research work validated all ranges used to filter the vital signs.

- **Patient pattern modelling**: The objective of this component is to approximate a probability density function for each metric in the patient's measurements. These probability density models are then stored in the database, eliminating the need to repeat the function modelling each time a new inference is made. This module runs monthly to create a new probability function that captures the variability of the new measurements entered by the patient.

- **Validation of atypical measurements based on the patient's history**: This module uses the probability density models stored in the database, which are associated with each patient's vital signs, to determine whether a newly recorded measurement falls within the normal patterns for that specific patient. Considering these variations could be due to disease exacerbation, improvements from a new medication, or other factors, these need to be validated by a nurse and, if necessary, by the patients themselves, to determine the true cause of the variation.

The operationalization of this module is presented in Figure 3.5. The system begins with the measurement and input of a vital signal by a patient in the HCAlert application. The measurement is compared and validated based on clinical rules, according to the type of measurement performed. The following clinical rules are defined, where the value is considered erroneous and discarded in the following cases:

- Oxygen saturation above 100 or below 20;
- Body temperature below 30 or above 40;
- Systolic blood pressure below 50 or above 350;
- Diastolic blood pressure below 40 or above 200;
- Pulse rate less than or equal to 30, or greater than 250.

Figure 3.4 presents the architecture of the Biometric sign error detection module.

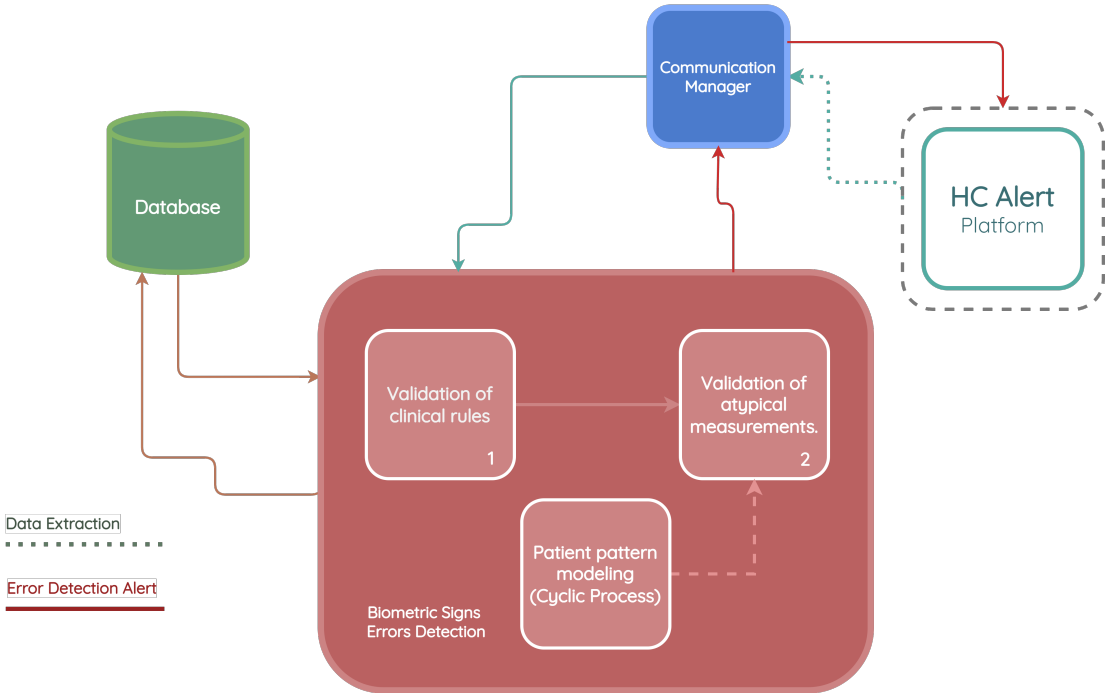

FIGURE 3.4: Biometric sign error detection module architecture.

In the event of an incorrect measurement, a type 1 alert is triggered, recommending a new measurement of the vital signal by the patient.

If there is no inconsistency with the rules, the system then determines if the measurement is atypical for a patient. If it is not considered atypical, the verification process is concluded without any identified errors. If an atypical value is recorded, a type 2 alert is triggered, and human verification of this alert is recommended to a nurse and the patient. This is done to verify whether this value corresponds to a health deterioration, an improvement in the clinical condition, or a measurement error.

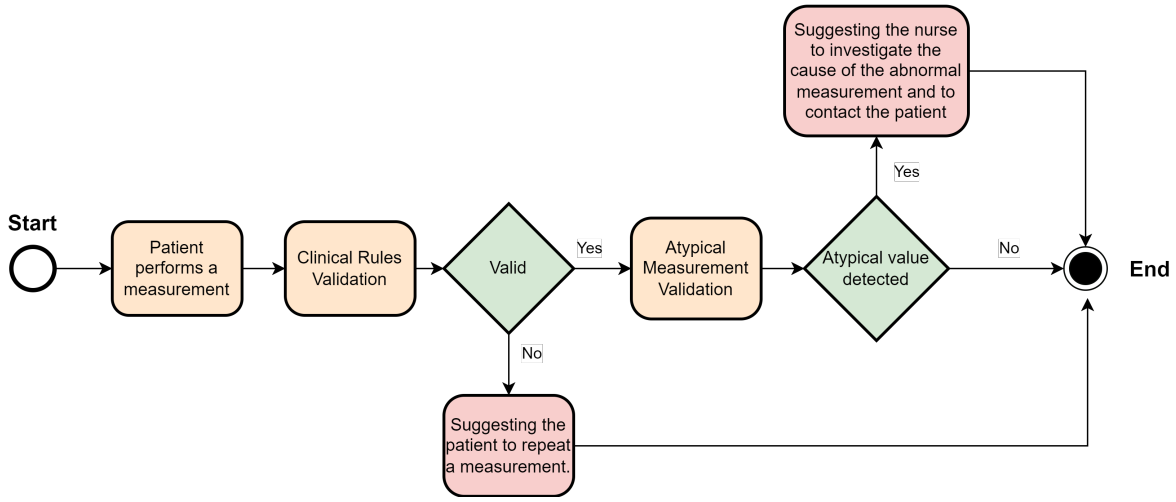

FIGURE 3.5: Biometric sign error detection implementation.

Probability density functions were applied in order to model the pattern of vital signs of each patient and assess the probability that a newly measured value fits the distribution function computed for that specific patient's vital sign. The process of training a model for a given patient begins with the request for all the vital sign measurements made by this patient. From this request, as shown in Figure 3.6, a distribution function is trained and stored in the database for each vital sign recorded, with the following steps:

1. From all the measurements collected for the patient, only the measurements made for specific vital signs in training are used.

2. Existing outliers in the database, prior to modelling, are removed. Outliers are removed based on the standard deviation by calculating the standard score (z-score), which corresponds to the number of standard deviations by which a newly recorded value deviates from the mean of the observed measurements. If the z-score is greater than 3, which corresponds to a value that is three times the standard deviation away from the mean of the data, the value is not used in the modelling.

3. The following distributions are tested: normal, exponential, Pareto, double Weibull, t, generalized extreme value distribution, gamma, lognormal, beta, and uniform. For each distribution, the density and weights of the histogram are computed. Subsequently, an estimation of the function parameters is performed based on the data. The maximum likelihood estimation (MLE) is used to identify the values that best fit the data.

4. The goodness-of-fit is calculated with a test of the sum of squares of the residuals for each distribution found.

5. The model with the best goodness-of-fit, which implies a lower value in the sum of squares of the residuals, is stored for the vital signs of the patient under study.

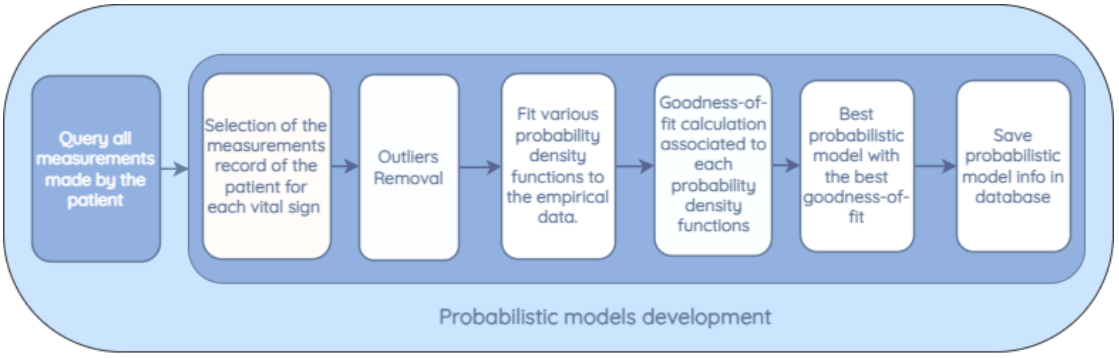

FIGURE 3.6: Biometric sign error detection model development.

The inference starts with the reception of a vital sign measurement taken by a patient and entered into the HCAlert system. The system selects the model

corresponding to the probability density function that models the distribution of the vital signs measured for the patient who entered it into the system, as is present in Figure 3.7.

This model is then used to test the null hypothesis, which corresponds to checking whether the value that has been measured is outside the typical pattern of the patient, based on the selected distribution and the parameters adjusted according to the empirical distribution of the patient. If the p-value is less than 0.05, it implies that the null hypothesis is not rejected, which means that there is a probability that the measurement may correspond to an error, exacerbation, or improvement of the condition. A reminder should be sent to both the nurse and the patient to investigate the situation.

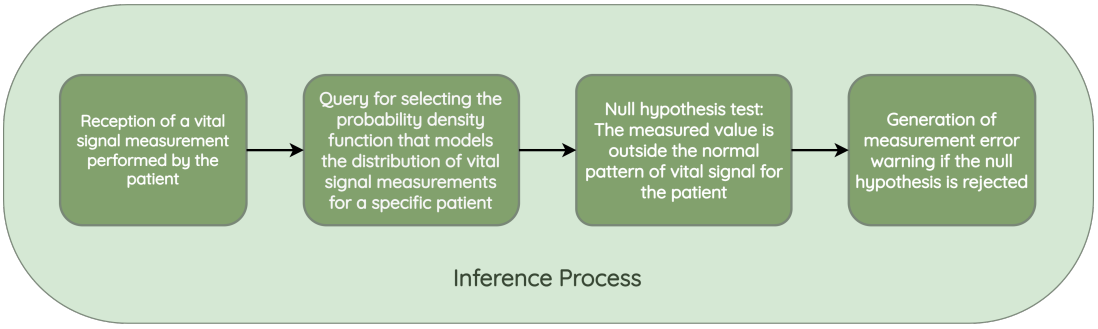

FIGURE 3.7: Biometric sign error detection inference process.

3.5 Basal Value Monitoring

The deterioration or improvement of COPD reflected in the negative or positive evolution of the patient's baseline values may be due to several explanatory factors, such as weather conditions, exposure to particulate matter, a change in medication or lifestyle, among others. The recorded baseline values are indicative of the severity of a condition, as outlined by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) [38] strategy for the diagnosis, management, and prevention of COPD.

Values below or above the standards result in the patient's category changing into one of the GOLD I–GOLD IV [38] categories, depending on the severity of the

patient's condition, with GOLD I being the most severe condition. It is important to identify and monitor any deterioration in a patient's baseline values in order to adjust the clinical protocol and treatment guidelines.

Figure 3.8 presents a clinical protocol defined by the Hope Care SA medical team; it is based on the GOLD strategy and addresses patients whose basal values are within a normal range and, thus, do not belong to categories GOLD I–GOLD IV. Consequently, the range of colors isn't associated with the GOLD categories. The color is associated with the severity of the COPD patient's condition: Category I (red) corresponds to a higher degree of deterioration in their health condition, while Category V (green) corresponds to the lowest or non-deterioration of their health condition. Some fields are filled with the expression "N/D" because there is no defined range of values for that specific category.

	Systolic	Diastolic	Pulse	Oximetry (with oxygen therapy)	Oximetry (without oxygen therapy)	Temperature	Weight	Steps	FEV
Color	mm Hg	mm Hg	bpm	%	%	°C	kg	# steps weekly average	%
I	0 – 70	0 – 30	0 – 50	0 – 85	0 – 85	N/D		N/D	0 – 30
II	70 – 80	30 – 40	N/D	85 – 91	85 – 92	0 – 35		N/D	30 – 50
III	80 – 90	40 – 50	N/D	N/D	N/D	N/D		N/D	50 – 80
IV	N/D	N/D	N/D	N/D	N/D	N/D		N/D	N/D
V	90 – 140	50 – 90	50 – 100	91 – 100	92 – 100	35 – 37.6	Rules to be denied for each patient	N/D	80 – 100
IV	N/D	N/D	N/D	N/D	N/D	N/D		N/D	N/D
III	140 – 160	90 – 100	N/D	N/D	N/D	N/D		0 – 12000	N/D
II	160 – 180	100 – 110	N/D	N/D	N/D	N/D		N/D	N/D
I	180 – 250	110 – 250	100 – 250	N/D	N/D	37.6 – 50		N/D	N/D
Absence of measurements	28h	28h	28h	28h	28h	28h	28h	128h	28h

FIGURE 3.8: Clinical protocol defined by the Hope Care SA Medical Team and based on the GOLD clinical protocols.

3.5.1 Basal Value Monitoring Module Architecture

This module, as shown in Figure 3.9, uses the list of metrics to be monitored and the history of vital signs recorded by each patient as input. Based on these

measurements, the patient's current baseline value and the forecast of the evolution of the same value are determined. In case there is a substantial difference between the most recently recorded value and the historical baseline value, an alert should be triggered, containing the previous baseline value, the newly calculated value, and the difference. The newly calculated baseline value is suggested as a change to the clinical protocol.

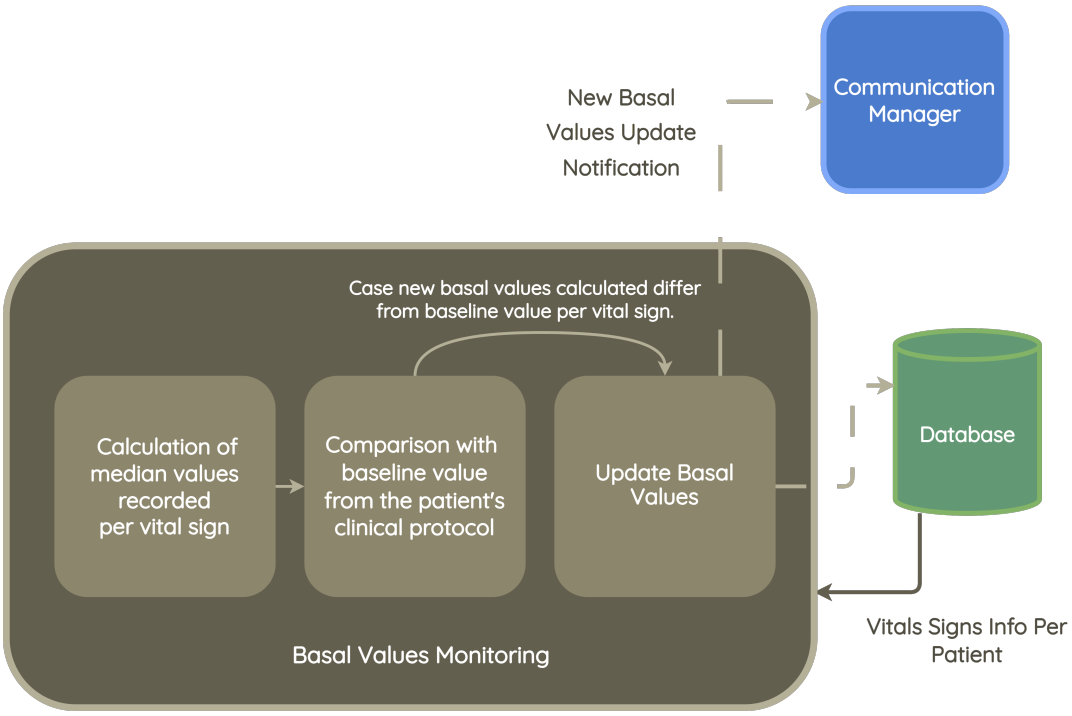

FIGURE 3.9: Basal value monitoring module architecture.

The following variables are also used as input to the module:

- Number of months considered: This indicates the past time window that is analyzed for the baseline calculation. The default value is 3 months, which indicates that when this module runs, the measurements taken from the last 3 months are extracted for the baseline calculation. This value can be configured by rules in the system.

- Minimum number of records: This corresponds to the minimum number of measurements taken by the patient, so that the calculated baseline information is considered reliable. If the patient does not have a satisfactory

number of measurements in the time horizon under study, the module will not provide recommendations. For example, a patient with only five SpO2 measurements over 3 months will not be considered for updating the baseline value. This value is configurable by a rule, and value 50 is used by default in the system.

- Patience: In case the patient does not present enough measurements of a certain parameter in the defined time horizon, the system expands the time horizon of the search to include more months of history until it finds an acceptable amount of records. For example, with a patience of 3 months and a minimum of 50 required measurements, if the patient only has 30 measurements, an additional month will be incorporated into the analysis, and the module will be rerun using the past four months, reducing the patience counter by 1. In case patience reaches zero, and the minimum value of measurements defined is not reached, the system will not provide any recommendation for the given parameter due to the lack of consistency in the measurements. The default value for patience, which can be configurable by a rule, is 3.

The default values in the system are set and adjusted after testing with historical values recorded by patients in the HCAlert platform, provided by Hope Care SA.

3.5.2 Basal Value Monitoring Module Implementation

In this section, we present the implementation details of the basal value monitoring module. Figure 3.10 shows an activity diagram, which represents the operations performed by the module.

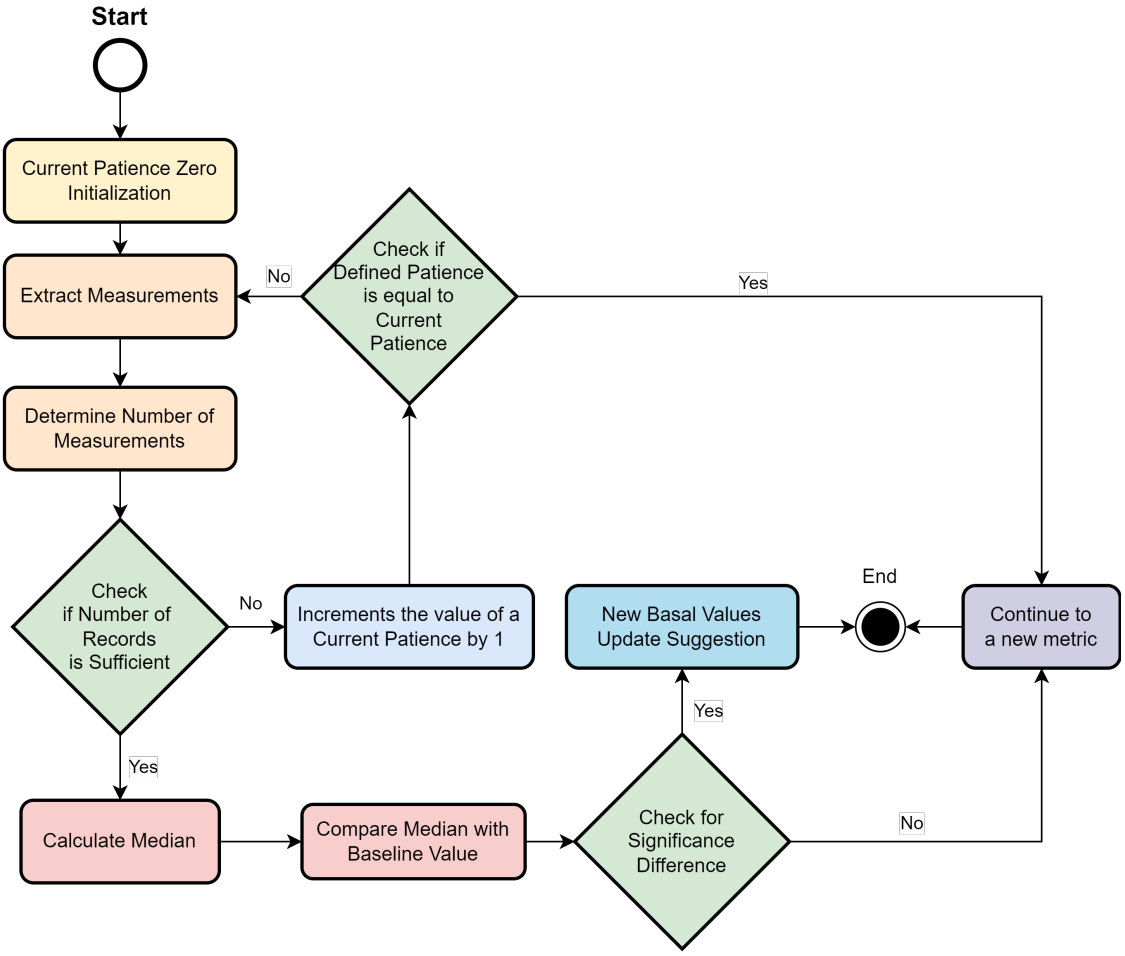

FIGURE 3.10: Basal value monitoring module implementation

As presented and detailed in the previous section, the system inputs are the list of metrics under evaluation, the patient's vital signs history, the number of months to be considered, the minimum number of records, the baseline value of the patient's clinical protocol, and patience.

For each metric under evaluation, the system performs the following process:

1. A flag representing the current patience is initialized to zero.

2. The measurements are related to the period of months corresponding to the last X months from the date of execution of the module, where X is the sum

between the system input "number of months to consider" and the current patience value.

3. The number of measurements performed by the patient is calculated.

   (a) In case the number of measurements is not sufficient, the current patience is incremented by 1.

      (i) If the current patience value is equal to the user-defined patience value, no recommendation is displayed, and the cycle continues to the next measurement in the list.

      (ii) If the current patience value is less than the set patience value, the system summarizes the run from step 2.

   (b) In case the measurements are sufficient, the system summarizes the run in step 4.

4. The median of the patient's measured values of a given vital sign is calculated.

5. The median value is compared with the baseline value recorded in the clinical protocol.

   (a) If the values are very different, a recommendation is made to update the baseline value to reflect the new median value recorded in the time interval under consideration. This recommendation should be evaluated by a medical professional.

   (b) If the values are similar, the baseline value is not adjusted, and the system summarizes in step 1, with a new iteration of a new metric under evaluation.

6. The cycle ends when all metrics in the list have been processed.

This process is run independently for each patient in the system. It is worth noting the use of the median as the metric calculated for the baseline value. This is due to the fact that it better handles extreme values outside of a patient's normal

patterns, such as exacerbation, which should not be considered for the calculation of a baseline value, as it does not correspond to a normal patient pattern.

3.6 Vital Signs Prediction Module

3.6.1 Predictive Model Development

Data Treatment

For the predictive model development and evaluation, 91 patients who were flagged as having COPD were included. Each patient was monitored remotely and provided health status information for tracking their health status. The vital sign information was then gathered by each medical center. These patients were from different districts of the country, such as Aveiro (Anadia), Leiria (Óbidos, Pombal), Santarém (Ourém) Castelo Branco (Fundão), Coimbra (Cantanhede, Cernache, Assafarge, Antanhol, Condeixa-A-Nova, Mira, Almargem Bispo), Lisboa (Amadora, Rinchoa, Queluz, Algueirão, Tapada Das Merces, Rio de Mouro), and Faro (Quarteira, Albufeira, Tavira, Olhão, Loulé, Lagos, Portimão, and Castro Marim).

Meteorological variables (temperature, humidity, wind, and rain) and exterior particle matter concentrations (PM10, PM2.5) were obtained from the nearest IPMA (Portuguese Institute for the Ocean and Atmosphere) and EPA (Environmental Protection Agency) stations. To analyze the source and transport pathways of the air masses and relate the air masses with aerosols, we used the NOAA HYSPLIT model [39, 40].

Information about the weather, air quality, and vital signs was analyzed. The data processing module was divided into four sub-phases: data cleaning, data transformation, patient datasets selection, and environmental data integration, as is present in Figure 3.11.

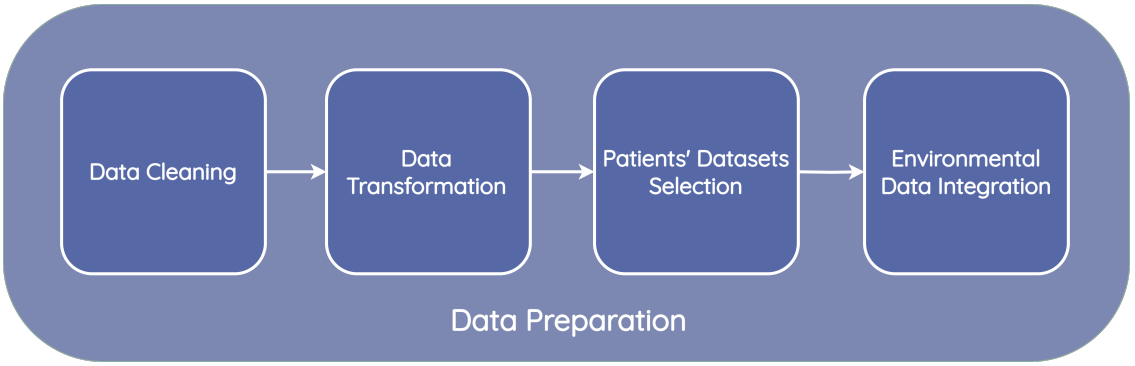

FIGURE 3.11: Data preparation pipeline.

During the data cleaning process, a thorough analysis was conducted on outliers (values that deviated significantly from the rest of the dataset and could potentially introduce anomalies in the results obtained from algorithms and analysis systems) based on the distribution of values in Figures 3.12–3.15, as well as on null values within the vital signs.

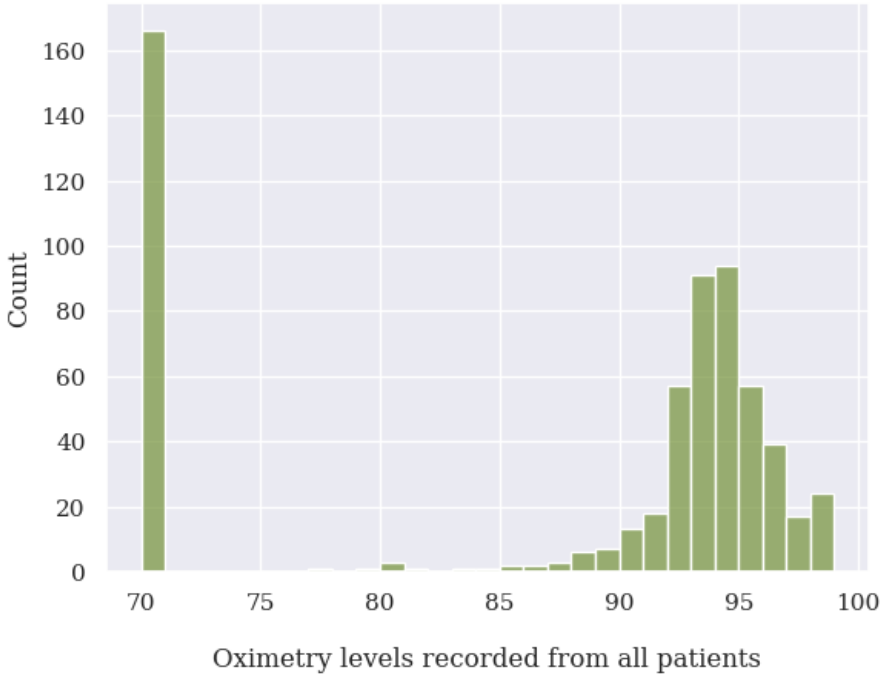

FIGURE 3.12: Oxygen saturation level value distribution of all patients analyzed

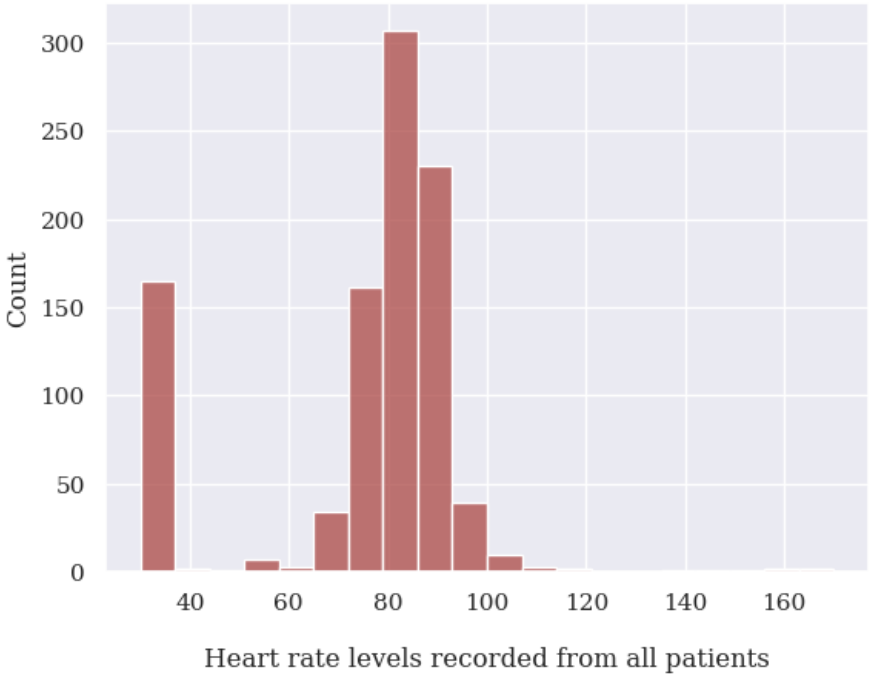

FIGURE 3.13: Heart rate level value distribution of all patients analyzed

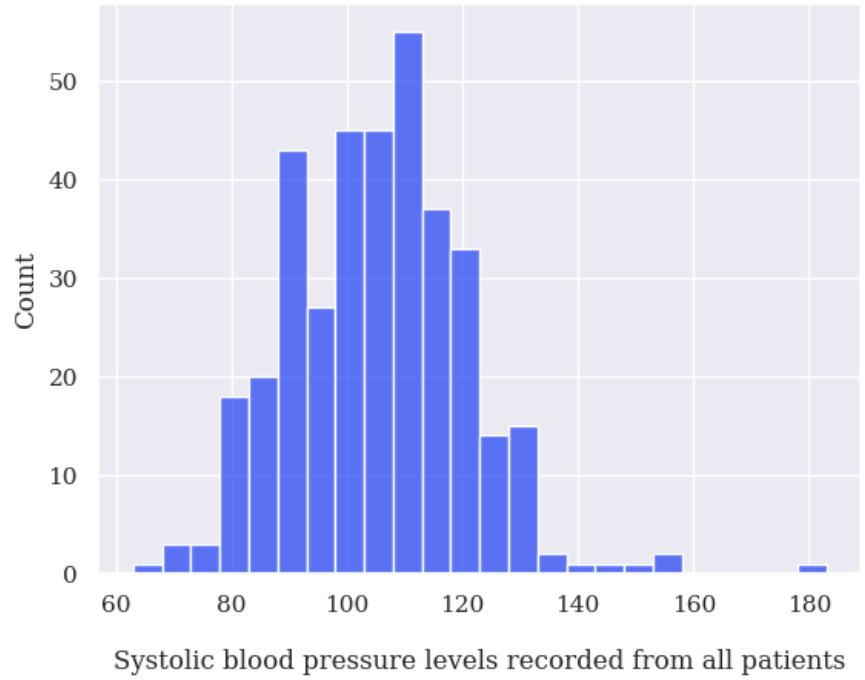

FIGURE 3.14: Systolic blood pressure level value distribution of all patients analyzed

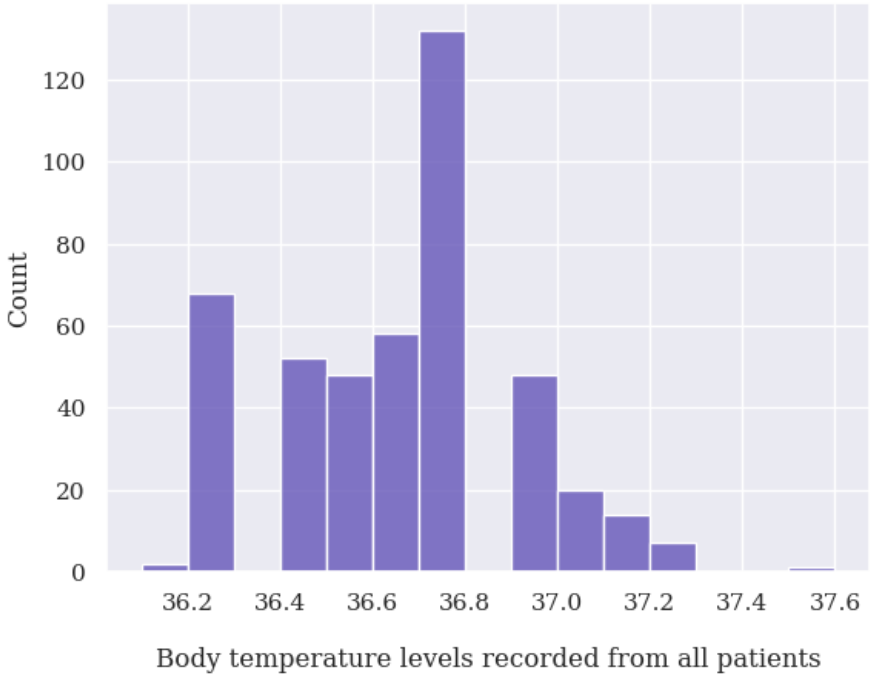

FIGURE 3.15: Body temperature level value distribution of all patients analyzed

Regarding vital signs, any values that met the following criteria were identified as outliers and subsequently removed:

- For oxygen saturation (SpO2), any values below or equal to 70% and above 100%. Since we have detected many measurements at exactly 70%, we suspect these are measurement errors;
- For body temperature, all values below 30 °C and above 40 °C;
- For systolic blood pressure (SBP), any values below 50 mmHg and above 350 mmHg;
- For heart rate (HR), any values below 39 BPM and above 250 BPM.
- For diastolic blood pressure (DBP), any values below 40 mmHg or above 200 mmHg.

In the data transformation process, we adjusted the format of historical records related to the vital sign data of patients. The data, initially in a format of one record per day per parameter, were converted to one record per day with all the

collected vital sign values for that day. Specifically, there was a change in the granularity of each data row from one row per measurement of a specific vital sign at a specific moment in time for a specific patient to one row for each day of measurements taken for a specific patient, with columns representing the measured vital signs (data pivoting). After the format change, every time segment with over 10 consecutive days of missing data was removed and only patients with over 180 records whose vital sign data were fully complete were selected.

In the data integration process, the historical records of each patient's vital signs were supplemented with information regarding weather data (average daily temperature, average relative humidity, and amount of daily precipitation) and air particle data (10 µm particles and 2.5 µm particles, as these two dimensions have a greater impact on the patients' respiratory capacity).

Modelling and Evaluation

Following the data treatment, we modeled the development and evaluation. As a result of the data treatment phase, only 14 datasets were considered for the model training and evaluation phase. Since the ICDSS was designed to assist COPD patients with different health profiles, we developed models using 14 different datasets and incorporated the best models in the system. Figure 3.16 shows the steps of the development and evaluation phase.

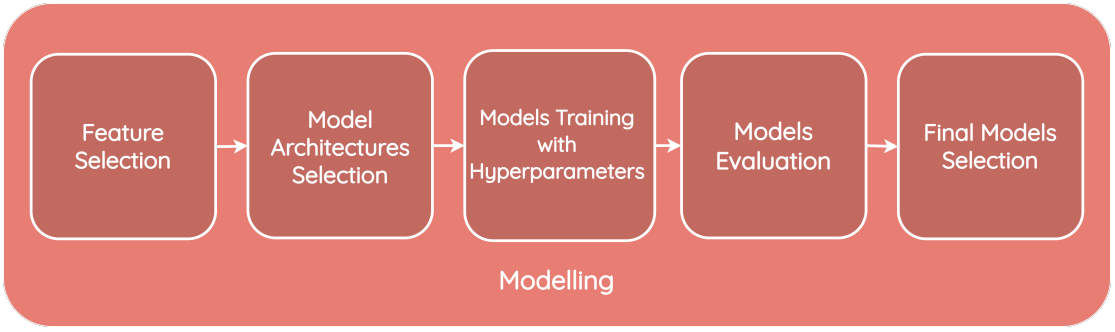

FIGURE 3.16: Modelling and evaluation pipeline

We employed multivariate machine learning models capable of conducting the multi-step-ahead time series prediction of vital signs. Multi-step-ahead forecasting

involves predicting multiple future time steps in a time series [41]. In our case, it would mean predicting the vital sign values for the following 5 days. The vital signs chosen for prediction include SpO2, heart rate, body temperature, and systolic blood pressure, which are utilized in the early warning score calculation module to assess the risk of deterioration.

During the feature selection process, we conducted a comprehensive correlation analysis between vital signs and clinical validation, resulting in the identification of the most relevant vital signs for predicting health variations in COPD patients.

Figure 3.17 shows an example of a correlation between SpO2 values (Spo2_1_day), the pm25 external parameter (PM25), relative humidity (HR_MED), and SpO2 values (SpO2) of the previous day, using the dataset for the patient with ID no. 156.

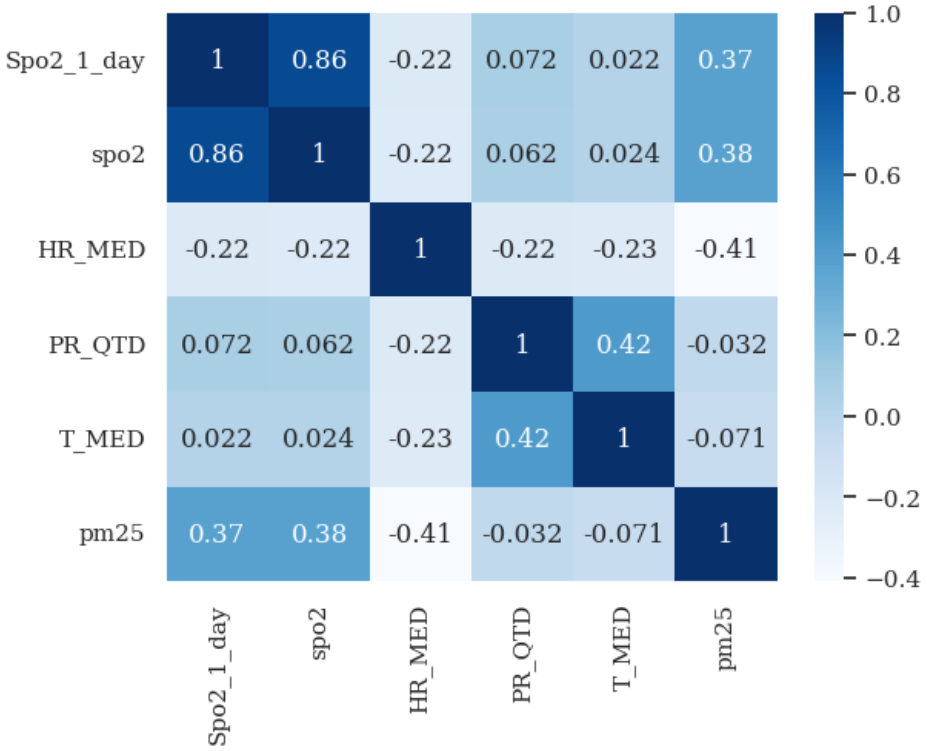

FIGURE 3.17: Correlation matrix between Spo2, Spo2 from the previous day, relative humidity, the levels of precipitation, the pm25 concentration, the external temperature values, and SpO2 level from the previous day, using the dataset for the patient with ID no. 156.

For multi-step-ahead time series prediction, all vital signs receive the previous day's value (n − 1) as input to forecast the value for the current day (n). To predict the value of SpO2, we selected the following inputs: the SpO2 value of the previous day, the relative humidity value of the previous day, the levels of precipitation from the previous day, the pm25 value from the previous day, and the external temperature value from the previous day.

Regarding the other vital signs, based on the analysis of the correlation between the four vital signs analyzed in Figure 3.18, and the clinical insight provided by the Hope Care SA medical team suggesting that SpO2 influences heart rate, body temperature, and systolic blood pressure, we decided to use only the SpO2 value from the previous day and the specific vital sign in question from the previous day as inputs.

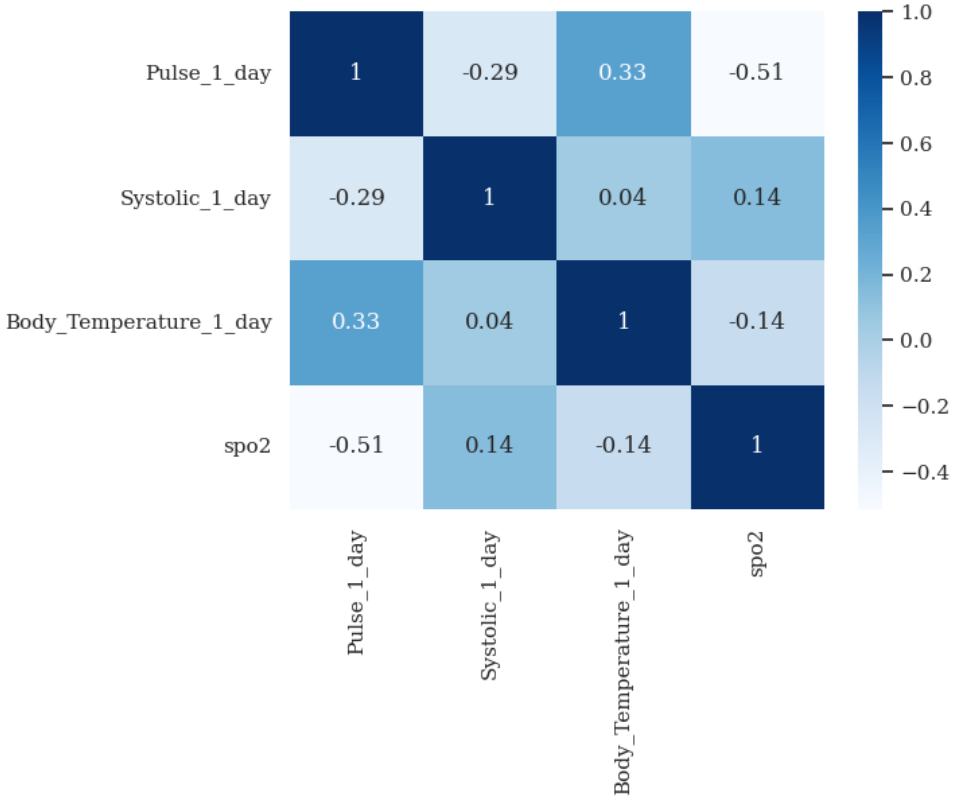

FIGURE 3.18: Correlation matrix of values of SpO2 parameter with the pulse rate, systolic blood pressure and body temperature values of the following day, using the dataset for the patient with ID no. 156.

To ensure the selection of the most optimal model architecture for predicting a specific vital sign, we trained and evaluated six distinct machine learning models. These models encompassed a diverse range of architectures, namely ARIMA (autoregressive integrated moving average), LSTM (long short-term memory), BILSTM (bidirectional long short-term memory), GRU (gated recurrent unit), LightGBM (light gradient boosting machine), and XGBoost (extreme gradient boosting).

The training process was preceded by essential hyperparameter tuning, which is a critical step in developing machine learning models. This tuning allowed us to optimize the models for the best possible performance. In our case, the models' performance was assessed using the root mean square error (RMSE), which measures the difference between prediction and the ground truth in the regression algorithm evaluation.

Table 3.5 presents an example of the RMSEs achieved for the fifth-day predictions via different machine learning model architectures for each vital sign prediction using the dataset for the patient with ID no. 156.

TABLE 3.5: Root mean square error values for the 5th-day predictions of different model architectures trained using the dataset for the patient with ID no. 156.

Model	SpO2	Heart Rate	Systolic Blood Pressure	Body Temperature
ARIMA	2.080718	7.089329	9.783878	0.247163
XGBoost	0.817778	0.96435	2.407083	0.302518
LightGBM	0.064668	0.380769	2.170715	0.058705
GRU	0.083168	0.110159	0.130179	0.131379
LSTM	0.092241	0.573169	0.135822	0.137075
BILSTM	0.084948	0.113384	0.132097	0.130094

As a result of our evaluation, we saved the models that demonstrated the lowest root mean square error (RMSE) for each vital sign. Consequently, we had 4 distinct models for each of the 14 patient-specific datasets, with each model specialized in predicting a specific vital sign.

Table 3.6 presents an example of the RMSEs for the 5th-day predictions achieved by the best machine learning model architectures for each vital sign prediction using the dataset for the patient with ID no. 156.

TABLE 3.6: Root mean square error values for the 5th-day predictions using the best model architectures trained on the dataset for the patient with ID no. 156

Vital Sign Predicted	Type	RMSE
SpO2	LightGBM	0.064668
Heart Rate	GRU	0.110159
Systolic Blood Pressure	GRU	0.130179
Body Temperature	LightGBM	0.058705

3.6.2 Production

In this section, we present the incorporation of the previously described predictive models into the ICDSS.

The vital signs prediction module presented in Figure 3.19 is composed of two sub-processes: a data pre-processing stage followed by the application of predictive models. The data pre-processing stage is essential to ensure that the data is in the correct format and that the vital sign measurements are appropriately integrated with the external measurements, as previously mentioned in Section 3.6.1.

The vital signs prediction process takes place daily, and the resulting predictions are stored in the database for future reference. Subsequently, the early warning module utilizes this data to assess and calculate the risk of a patient experiencing deterioration within the following five days.

When a new patient is integrated into the system, the prediction for each vital sign is calculated as the average of the predictions from all the models that predict the particular vital sign. After a period of 6 months, the error (root mean squared error—RMSE) of each predictive model is analyzed by measuring the distance between the values predicted by each model and the actual values of the vital

signs for each patient. The model with the lowest error is the one associated with the patient.

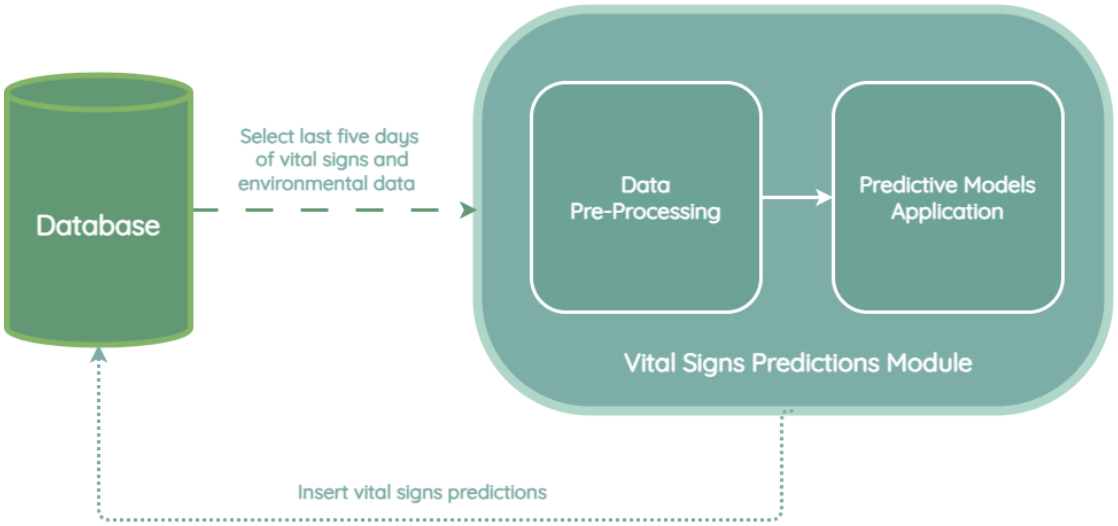

FIGURE 3.19: Vital signs predictions module architecture

3.7 Early Warning Score Calculation Module

In this module, the risk of a patient experiencing deterioration is assessed using the early warning score (EWS) clinical protocol. The EWS is utilized for monitoring and detecting the risk of health deterioration in patients and it is calculated by combining vital signs and clinical data, such as heart rate, blood pressure, respiration rate, body temperature, oxygen saturation (SpO2), and degree of consciousness. Individual scores for each vital sign are then totaled up, resulting in a total EWS score.

The higher the overall EWS score, the more likely a patient is suffering from a health deterioration. This clinical protocol presented in Table 3.7 is indicated by Hope Care SA's medical team.

TABLE 3.7: Early warning score clinical protocol suggested by Hope Care SA's medical team

	Description	0 Points	1 Point	2 Points	3 Points
SpO2	Difference between the predicted value for the day and the value from the previous day	<3%	3-5%	6-7%	>7%
Heart Rate	BPM Value	46-100	101-110	111-115	>115 or <46
Systolic Blood Pressure	Percentage difference between the predicted value for the day and the baseline value	<20%	>=20%	>=23%	>=25%
Body Temperature	Temperature value in Celsius	<37.5	37.5-37.9	38-38.4	>38.5

Similar to the vital signs prediction module, the early warning score calculation is performed daily, and the resulting scores are stored in the database.

Chapter 4

Demonstration and Evaluation

4.1 System Demonstration

To demonstrate how the ICDSS addresses the research question, we present a system trial with the incorporation of a new patient. We use the patient with ID no. 300. The patient health information used in this trial consists of historical information for a three-year period consisting of HRMS monitoring provided by Hope Care SA through the HCAlert platform.

The monitoring for the patient with ID no. 300 was initiated on 21th of April, 2022. The ICDSS received a notification from the HCAlert platform, regarding the need to incorporate this new patient, leading to the creation of a new record in the database. All vital signs monitored for the patient with ID no. 300 were transmitted to the HCAlert platform and subsequently extracted by the ICDSS, starting from 21th of April. These vital signs underwent analysis through the biometric sign error detection module. As no outliers were detected in the vital signs, they were seamlessly integrated into the database..

Table 4.1 presents the last five days of data extracted from the database for vital sign predictions on the 25th of April, 2022.

TABLE 4.1: Last 5 days of data extracted from the database for vital sign predictions on the 25th of April.

Date	Heart Rate (BPM)	Body Temperature (ºC)	SpO2 (%)	Systolic Blood Pressure (mmHg)	T MED (ºC)	HR MED (%)	PR QTD (mm)	pm25 (Count)
2022-04-21	61.0	36.2	95.0	95.0	9.60	63.25	1.86	1.66
2022-04-22	63.0	36.0	95.0	93.0	7.53	82.97	23.25	0.93
2022-04-23	59.0	36.5	96.0	96.0	8.95	69.24	1.91	0.58
2022-04-24	65.0	36.2	96.0	100.0	10.79	67.82	0.29	1.14
2022-04-25	57.0	35.9	96.0	102.0	12.35	65.43	0.01	2.63

By the 25th of April, a sufficient amount of vital sign data is available to provide insights into the patient's risk of health deterioration. The ICDSS proceeds with the prediction of vital signs and subsequently calculates the early warning score. Various models are employed to forecast the patient's vital signs for the initial 6 months of integration. The risk information regarding the patient's potential deterioration is provided to the HCAlert platform through a JSON file.

Table 4.2 presents the vital sign prediction values for the 26th of April. The predicted vital signs are then used to calculate the risk.

TABLE 4.2: Predicted Values of Vital Signs from 26th of April to 30th of April

Date	SpO2 (%)	Heart Rate (BPM)	Systolic Blood Pressure (mmHg)	Body Temperature (Celsius)
2022-04-26	95.028053	63.863962	98.327346	36.244274
2022-04-27	94.801013	64.027884	98.783749	36.162657
2022-04-28	94.948091	64.413307	99.589877	36.218256
2022-04-29	95.127560	64.438053	99.516291	36.246443
2022-04-30	95.054558	64.429125	99.496265	36.196343

Table 4.3 presents the values of the early warning score calculated on the 25th of April.

TABLE 4.3: Calculated values of the early warning score from 26th of April to 30th of April.

Date	SpO2 (%)	Heart Rate	Systolic Blood Pressure	Body Temperature
2022-04-26	0	1	0	0
2022-04-27	0	1	0	0
2022-04-28	0	1	0	0
2022-04-29	0	1	0	0
2022-04-30	0	1	0	0

Listing 4.1presents part of the structure of a part of the JSON file concerning the predicted vital signs and early warning score calculated from the 26th of April to the 30th of April.

Listing 4.1 Structure of the JSON file provided to HCAlert for patient risk information on the 25th of April.

```
{'predict_date': '2022-04-26',
'global_ews_score': 1,
'vitals':
 '{"spo2": {
  "predict_value": "95.02805293812013",
  "predict_score": "0", "units": "{\%}"},
 "pulse": {
  "predict_value": "63.86396198309728",
  "predict_score": "1", "units": "BPM"},
"systolic": {
  "predict_value": "98.32734618907372",
  "predict_score": "0", "units": "mmHg"},
"body_temperature": {
  "predict_value": "36.244273924492624",
  "predict_score": "0", "units": "ºC"}}}
```

After an evaluation spanning over 6 months, we focused on identifying the most suitable models to enhance the care of patient 300. Our selection process prioritized models with the lowest root mean square error (RMSE), as shown in Table 4.4.

TABLE 4.4: Root mean square error (RMSE) values of the top selected models for predicting the vital signs of patient 300.

Dataset used to train the model	Model	Parameter	Value (RMSE)
304	BILSTM	Spo2	0.285014
181	GRU	Heart Rate	1.520008
184	BILSTM	Systolic Blood Pressure	1.904305
181	GRU	Body Temperature	0.250580

We analyzed the patient's data from the previous 6 months; we provide a new basal value that reflects the patient's health condition, which is, consequently, used for the patient's clinical protocol adjustment, as shown in Listing4.2.

Listing 4.2 Suggested new basal values for patient 300 to the HCAlert platform.

```
{
'spo2': {
  'median_value': 96.0,
  'number_of_months': 6},
'body_temperature': {
  'median_value': 35.6,
  'number_of_months': 6},
'pulse': {
  'median_value': 73.0,
  'number_of_months': 6},
'systolic': {
  'median_value': 99.0,
  'number_of_months': 6}
}

```

During the course of 6 months, while closely monitoring patient 300's health, we detected an error involving one of the SpO2 measurements. Initially, this measurement seemed to comply with the clinical rules and was considered valid. However, upon atypical measurement validation, it became evident that the probability of this value ($p = 0.01599$) belonging to the distribution of SpO2 values for patient 300 was relatively low, falling below the threshold of 0.05. Due to this fact, the measurement was discarded from the dataset.

Figure 4.1 presents the distribution of SpO2 values of patient 300 analyzed for the error alert validation.

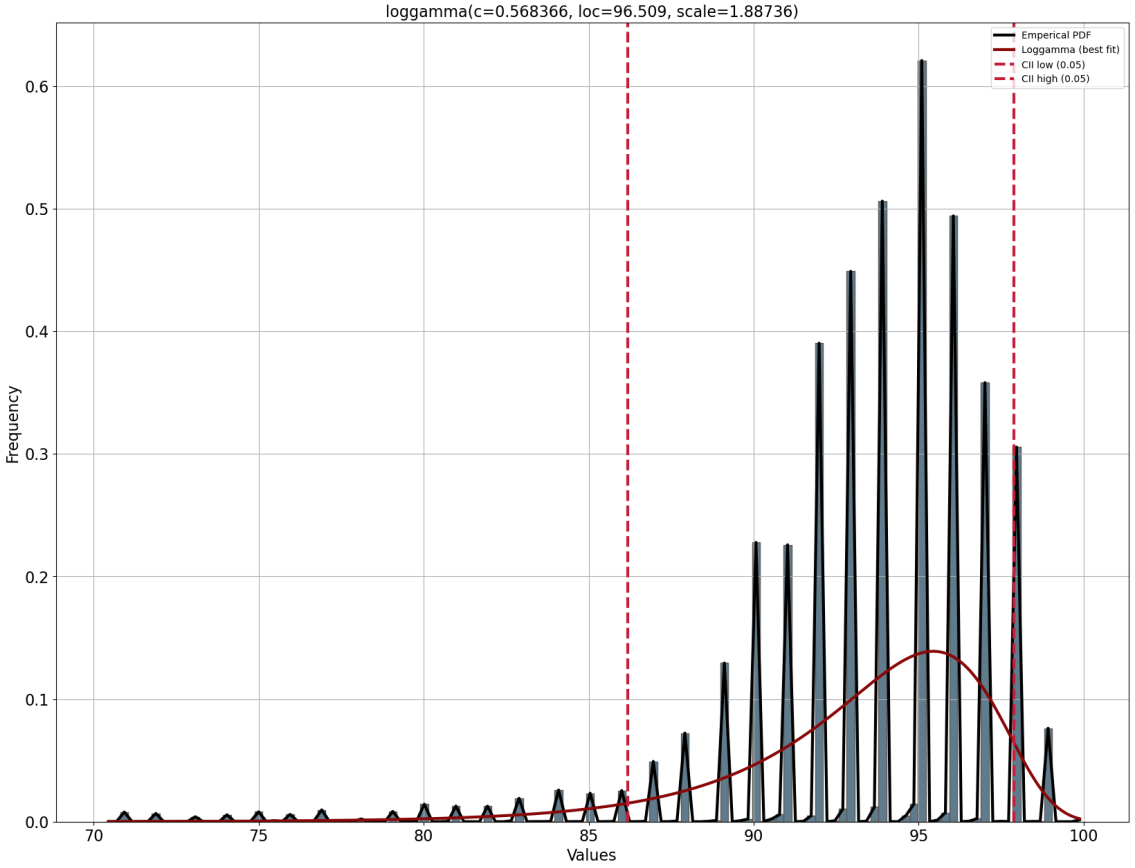

FIGURE 4.1: Distribution of SpO2 values analyzed of patient 300.

On the 25th of October, the ICDSS provided essential health information about the risk of patient deterioration. However, this risk was generated using predictions from the selected best models, as mentioned earlier.

Table 4.5 presents the last five days of extracted data from the database for vital sign predictions on the 25th of October.

TABLE 4.5: Last 5 days of data extracted from the database for vital sign predictions on the 25th of October.

Date	Heart Rate (BPM)	Body Temperature (ºC)	SpO2 (%)	Systolic Blood Pressure (mmHg)	T MED (ºC)	HR MED (%)	PR QTD (mm)	pm25 (Count)
2022-10-21	68.0	35.60	96.0	96.0	15.05	79.13	3.94	1.94
2022-10-22	74.0	35.80	96.0	96.0	14.91	74.00	18.72	1.20
2022-10-23	70.0	35.90	95.0	94.0	14.15	67.11	5.45	2.91
2022-10-24	72.0	35.80	97.0	93.0	14.32	72.58	1.47	1.93
2022-10-25	76.0	35.00	95.0	98.0	16.13	64.89	7.87	1.94

Table 4.6 presents the vital sign prediction values from the 25th of October. The predicted vital signs are then used to calculate the risk.

TABLE 4.6: Predicted vital sign values from the 26th of October to 30th of October.

Date	SpO2 (%)	Heart Rate (BPM)	Systolic Blood Pressure (mmHg)	Body Temperature (ºC)
2022-10-26	96.386055	70.779388	95.078346	35.292265
2022-10-27	96.228622	72.117355	94.664948	35.597720
2022-10-28	96.208916	72.186485	94.973228	35.796912
2022-10-29	96.297836	73.253487	95.260201	35.886715
2022-10-30	96.020462	72.828354	96.042572	35.796912

Table 4.7 presents the early warning score values calculated on the 25th of October.

TABLE 4.7: Calculated early warning score values from the 26th of October to the 30th of October.

Date	SpO2 (%)	Heart Rate (BPM)	Systolic Blood Pressure (mmHg)	Body Temperature (ºC)
2022-10-26	0	1	0	0
2022-10-27	0	1	0	0
2022-10-28	0	1	0	0
2022-10-29	0	1	0	0
2022-10-30	0	1	0	0

Listing 4.3 presents the structure of a JSON file concerning the predicted vital signs and early warning score calculated from the 26th of October to the 30th of October.

Listing 4.3 Structure of the JSON file provided to HCAlert for patient risk information on the 25th of October.

```
{'predict_date': '2022-10-26',
  'global_ews_score': 1,
  'vitals': '{
  "spo2":{
    "predict_value": "96.38605499267578",
    "predict_score": "0", "units": "\%"},
  "pulse": {
    "predict_value": "70.85945892333984",
    "predict_score": "1", "units": "BPM"},
  "systolic": {
    "predict_value": "94.98711395263672",
    "predict_score": "0", "units": "mmHg"},
  "body_temperature": {
    "predict_value": "36.07156866129014",
    "predict_score": "0", "units": "ºC"}}'},

```

4.2 System Evaluation

We performed a set of white-box tests, evaluating each module for its functionality (unit tests) and integration with the related modules of the system (integrated tests). Afterward, we conducted a survey to gather feedback from two medical professionals to evaluate the system based on a set of criteria inspired by Prat et al. [42]. Based on the positive feedback collected from the survey, it appears that the system was well-designed and valuable for managing the treatment of COPD patients.

Table 4.8 shows the evaluation given by two medical professionals specialized in COPD disease. The sample was slightly small, but highly significant since these medical professionals had experience in this disease. They were asked to answer questions, indicating a number between 1 and 5, where 1 corresponds to not relevant or not useful and 5 corresponds to very relevant or very useful [43].

TABLE 4.8: Results of the evaluation of the system by medical professionals.

Criteria	Questions	Objective Statement	Eval 1	Eval 2
Clinical Impact on Patients Treatment	Indicate the importance of an smart clinical decision support system capable of provide a 5-day Early Warning Scores for monitoring patients with COPD.	Importance of the intelligent clinical decision support system for monitoring patients with COPD	**5**	**5**
Patients Life Quality Impact	Indicate the impact of a smart clinical decision support system providing a 5-day Early Warning Scores on the quality of life of a patient with COPD.	Impact of a clinical intelligent decision support system on the quality of life of a patient with COPD	**5**	**5**
Utility	Indicate the usefulness of a system for healthcare professionals that generates information whenever there are changes in patients' baseline values.	Usefulness of a clinical intelligent decision support system that notifies patients baseline values modifications	**4**	**5**
	Indicate the importance of a system that provides short time horizon (in minutes) Early Warning Scores for the clinical follow-up of patients with COPD.	Importance of a clinical intelligent decision support system on the clinical follow-up of patients with COPD.	**5**	**5**
	Indicate the usefulness of a real-time alert system for healthcare professionals whenever an abnormal measurement occurs for a specific patient.	Usefulness of a clinical intelligent decision support system that notifies abnormal measurements detections	**5**	**5**
Consistency with the organization	Indicate the relevance of involving healthcare professionals in defining clinical intervals for abnormal measurements.	Clinical validation on the definition of intervals for abnormal measurements	**5**	**5**
	Indicate the relevance of involving healthcare professionals in defining the formula for calculating the basal value.	Clinical validation on the definition of the basal value calculation formula	**4**	**5**
	Indicate the relevance of involving healthcare professionals in selecting environmental and clinical parameters (e.g., vital signs) that most influence the clinical progression of patients with COPD.	Clinical validation on the selection of environmental and biometric signs that most influence the clinical progression of patients with COPD	**5**	**5**
Integration with clinical protocols	Indicate the relevance of the adopted Early Warning Score matrix for clinical decision-making and adjustment of therapeutic protocols for patients.	Relevance of the adoption Early Warning Score matrix for clinical decision-making and adjustment of therapeutic protocols for patients	**5**	**4**

Chapter 5

Conclusions & Future Work

5.1 Conclusions

In this dissertation, we developed a system prototype that answers our research question: "Is it possible to automatically monitor and analyse the risk of a potential health deterioration of COPD patients?". This system aims to provide early information concerning a patients health status evolution in order to support the treatment of patients with COPD.

As mentioned in Section 3, the ICDSS comprises two primary components: the vital signs prediction module and the early warning score calculation module. These components specifically address the research question.

The vital signs prediction module, as mentioned in Section 3.6, generates vital sign predictions using different types of model architectures. These predictive models are optimized using a fine-tuning process, with each model corresponding to a specific patient with a specific health profile. As demonstrated in Section 3.6.2, the integration of predictive models developed using data from fourteen different patients shows that the ICDSS has the flexibility to predict vital signs and, in turn, calculate the patient deterioration risk for various health profiles. This system has the ability to evolve and adapt to every patient condition since the first stage corresponds to using an ensemble of models to predict vital signs and the second stage corresponds to only using models with the lowest RMSE.

The early warning score calculation module uses vital sign records and determines the patient health deterioration based on a clinical protocol.

The ICDSS is also composed of three other modules: biometric sign error detection, basal value monitoring, and the communication manager.

The biometric sign error detection ensures the quality of all information concerning vital signs by validating, in a two-phase process, whether the vital sign values fall within the normal range for general COPD patients and subsequently, within the specific patient's normal range using a probability density function.

The basal value monitoring analyzes the vital signs and suggests recommendations for new basal values to the patient if they deviate from the baseline provided by the HCAlert platform. The communication manager deals with all connections between the ICDSS modules, the HCAlert platform, and weather information sources.

The ICDSS system completed the white-box tests, including unit tests and integration tests.

All of these tests validate its functionality and contribution to preventing and potentially improving patient treatment by offering an early indication of the patient's risk for deterioration.

Despite our ability to leverage real-time telemonitoring patient data, we employed clinical historical longitudinal data that was gathered over a substantial period of time (2–3 years) through a telemonitoring application. This extended time frame enabled us to formulate conclusions regarding the system's validity, supported by the early warning score implementation and the errors of the applied predictive models.

5.2 Limitations

The non-approval of the incorporation of new patients by the ethics committee associated with the HC PSI project made the testing and analysis of the ICDSS effectiveness in providing quality information regarding patient health deterioration risk difficult.

The scarcity of data was a limitation in our study, and two key aspects contributed to this challenge. Firstly, the measurements we had access to were not collected at hourly intervals, which restricted our ability to capture fine-grained variations in the data. The absence of hourly data points hindered our capacity to discern short-term patterns and trends, potentially hiding crucial insights that might have emerged with more frequent data collection.

Another significant data gap stemmed from the lack of information concerning home sensors, specifically data related to humidity levels. Humidity is a vital environmental factor that influences various aspects of indoor comfort, air quality, and overall well-being. All houses are different, with varying insulation and heating, leading to distinct risk profiles. Even two houses in the same location can exhibit varying humidity levels and significantly different temperatures (better insulated houses, air conditioning/heating, dehumidifiers, etc). The absence of the essential sensor data limited our ability to comprehensively assess the interplay between different environmental parameters, potentially leading to an incomplete understanding of the complex dynamics within the studied environment.

Despite the limitations, the system was validated, end-to-end, and clinically recognized as important for COPD monitoring, being adjustable enough to integrate these data sources if included in the project and handle a lower granularity of information to make predictions.

5.3 Communication

During this dissertation, we have contribute to the scientific community with a publication regarding the mentioned artifact. This article is named "Intelligent

Clinical Decision Support System for Managing COPD Patients" and is in editorial process for MDPI's special issue *Transforming Precision Medicine: The Intersection of Digital Health and AI*[44].

5.4 Future work

As part of our future work, we will aim to identify some potential advancements to pursue. Firstly, we will aim to validate the effectiveness of the ICDSS (clinical deterioration surveillance system) by obtaining real-time patient data through the HCAlert platform. Analyzing these data over an extended period will help us assess the accuracy and quality of early information provided by the ICDSS, particularly regarding a patient's risk of deterioration.

To enhance the robustness of our research, we will seek to access a more extensive and diverse dataset that includes patient data from different countries.

Expanding our data collection to the international stage will ensure that our findings are relevant to a broader population.

Adopting a more inclusive approach involves considering a broader range of age-related values. By including individuals across various age groups, we could reveal some patterns and trends that may be present within different life stages.

To achieve more precise and detailed analyses, we propose incorporating more daily frequent recordings. This higher data capture frequency will enable us to detect subtle fluctuations and temporal dynamics that might be missed in less frequent sampling, providing real-time insights into patients' vital signs.

Additionally, the integration of sensor technology to monitor indoor humidity and temperature levels would facilitate the extraction of valuable insights regarding the relationship between environmental factors and health deterioration.

By pursuing these advancements, we seek to increase the importance and reliability of our research, which could ultimately contribute to better patient treatment.

References

[1] World Health Organization (WHO), 2023. www.who.int/data/global-health-estimates, Accessed on 24.09.2023.

[2] Portuguese Society of Pneumonology, 2022. https://www.sppneumologia.pt/saudepublica/dpoc/, Accessed on 24.09.2023.

[3] Portuguese Lung Foundation, "Observatório nacional doenças respiratórias 2022," 2022. https://ondr2022.fundacaoportuguesadopulmao.org/, Accessed on 24.09.2023.

[4] Royal College of Physicians (2017) National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity7 in the NHS., 2023. https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2, Accessed on 24.09.2023.

[5] M. Al-rawashdeh, P. Keikhosrokiani, B. Belaton, M. Alawida, and A. Zwiri, "IoT adoption and application for smart healthcare: A systematic review," *Sensors*, vol. 22, p. 5377, July 2022.

[6] J. T. Kelly, K. L. Campbell, E. Gong, and P. Scuffham, "The internet of things: Impact and implications for health care delivery," *Journal of Medical Internet Research*, vol. 22, p. e20135, Nov. 2020.

[7] K. Y. Ngiam and I. W. Khor, "Big data and machine learning algorithms for health-care delivery," *The Lancet Oncology*, vol. 20, pp. e262–e273, May 2019.

[8] M. Hobensack, J. Song, D. Scharp, K. H. Bowles, and M. Topaz, "Machine learning applied to electronic health record data in home healthcare: A scoping review," *International Journal of Medical Informatics*, vol. 170, p. 104978, Feb. 2023.

[9] J. Gregório, L. Reis, M. Peyroteo, M. Maia, M. M. da Silva, and L. V. Lapão, "The role of design science research methodology in developing pharmacy ehealth services," *Research in Social and Administrative Pharmacy*, vol. 17, pp. 2089–2096, 12 2021.

[10] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *Journal of Management Information Systems*, vol. 24, pp. 45–77, 12 2007.

[11] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, and D. Moher, "The prisma 2020 statement: An updated guideline for reporting systematic reviews," *International Journal of Surgery*, vol. 88, p. 105906, 2021.

[12] G. Pare, M. Jaana, and C. Sicotte, "Systematic review of home telemonitoring for chronic diseases: The evidence base," *Journal of the American Medical Informatics Association*, vol. 14, pp. 269–277, May 2007.

[13] S. Jang, Y. Kim, and W.-K. Cho, "A systematic review and meta-analysis of telemonitoring interventions on severe COPD exacerbations," *International Journal of Environmental Research and Public Health*, vol. 18, p. 6757, jun 2021.

[14] C. Kruse, B. Pesek, M. Anderson, K. Brennan, and H. Comfort, "Telemonitoring to manage chronic obstructive pulmonary disease: Systematic literature review," *JMIR Medical Informatics*, vol. 7, p. e11496, Mar 2019.

[15] D. Achelrod, J. Schreyögg, and T. Stargardt, "Health-economic evaluation of home telemonitoring for COPD in germany: evidence from a large population-based cohort," *The European Journal of Health Economics*, vol. 18, pp. 869–882, oct 2016.

[16] N. Hong, C. Liu, J. Gao, L. Han, F. Chang, M. Gong, and L. Su, "State of the art of machine learning–enabled clinical decision support in intensive care units: Literature review," *JMIR Medical Informatics*, vol. 10, p. e28781, 3 2022.

[17] J.-L. Pépin, B. Degano, R. Tamisier, and D. Viglino, "Remote monitoring for prediction and management of acute exacerbations in chronic obstructive pulmonary disease (aecopd)," *Life*, vol. 12, p. 499, 3 2022.

[18] K. Exarchos, A. Aggelopoulou, A. Oikonomou, T. Biniskou, V. Beli, E. Antoniadou, and K. Kostikas, "Review of artificial intelligence techniques in chronic obstructive lung disease," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, pp. 2331–2338, 5 2022.

[19] J. wen Lu, Y. Wang, Y. Sun, Q. Zhang, L. ming Yan, Y. xi Wang, J. han Gao, Y. Yin, Q. yue Wang, X. lian Li, and G. Hou, "Effectiveness of telemonitoring for reducing exacerbation occurrence in copd patients with past exacerbation history: A systematic review and meta-analysis," *Frontiers in Medicine*, vol. 8, 9 2021.

[20] D. Sanchez-Morillo, M. A. Fernandez-Granero, and A. Leon-Jimenez, "Use of predictive algorithms in-home monitoring of chronic obstructive pulmonary disease and asthma," *Chronic Respiratory Disease*, vol. 13, pp. 264–283, 8 2016.

[21] C. Carlin, A. Taylor, I. van Loon, G. McDowell, S. Burns, P. McGinness, and D. J. Lowe, "Role for artificial intelligence in respiratory diseases—chronic obstructive pulmonary disease," *Journal of Hospital Management and Health Policy*, vol. 5, pp. 27–27, 9 2021.

[22] A. A. Rajeh and J. Hurst, "Monitoring of physiological parameters to predict exacerbations of chronic obstructive pulmonary disease (copd): A systematic review," *Journal of Clinical Medicine*, vol. 5, p. 108, 11 2016.

[23] I. Martín-Lesende, E. Orruño, A. Bilbao, I. Vergara, M. C. Cairo, J. C. Bayón, E. Reviriego, M. I. Romo, J. Larrañaga, J. Asua, R. Abad, and E. Recalde, "Impact of telemonitoring home care patients with heart failure or chronic lung disease from primary care on healthcare resource use (the telbil study randomised controlled trial)," *BMC Health Services Research*, vol. 13, p. 118, 12 2013.

[24] S. Mogo, V. E. Cachorro, and A. M. de Frutos, "Morphological, chemical and optical absorbing characterization of aerosols in the urban atmosphere of valladolid," *Atmospheric Chemistry and Physics*, vol. 5, pp. 2739–2748, Oct. 2005.

[25] M. C. Arranz, M. F. M. Moreno, A. A. Medina, M. A. Capitán, F. C. Vaquer, and A. A. Gómez, "Health impact assessment of air pollution in valladolid, spain," *BMJ Open*, vol. 4, p. e005999, Oct. 2014.

[26] H. Bayram, "Effect of global climate change-related factors on copd morbidity," *Tanaffos*, vol. 16, pp. S24–S24, 01 2017.

[27] World Health Organization (WHO), 2021. Review of Evidence on Health Aspects of Air Pollution: REVIHAAP Project: Technical Report. Available online : `https://apps.who.int/iris/handle/10665/341712`, Accessed on 24.09.2023.

[28] J. Lee, H. M. Jung, S. K. Kim, K. H. Yoo, K.-S. Jung, S. H. Lee, and C. K. Rhee, "Factors associated with chronic obstructive pulmonary disease exacerbation, based on big data analysis," *Scientific Reports*, vol. 9, p. 6679, 4 2019.

[29] World Health Organization (WHO), 2022. `https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health`, Accessed on 24.09.2023.

[30] Z. Liu, A. Alavi, M. Li, and X. Zhang, "Self-supervised contrastive learning for medical time series: A systematic review," *Sensors*, vol. 23, p. 4221, Apr. 2023.

[31] C. Bui, N. Pham, A. Vo, A. Tran, A. Nguyen, and T. Le, "Time series forecasting for healthcare diagnosis and prognostics with the focus on cardiovascular diseases," in *6th International Conference on the Development of Biomedical Engineering in Vietnam (BME6)*, pp. 809–818, Springer Singapore, Sept. 2017.

[32] N. Kaieski, C. A. da Costa, R. da Rosa Righi, P. S. Lora, and B. Eskofier, "Application of artificial intelligence methods in vital signs analysis of hospitalized patients: A systematic literature review," *Applied Soft Computing*, vol. 96, p. 106612, 11 2020.

[33] J. Xie, Z. Wang, Z. Yu, and B. Guo, "Enabling timely medical intervention by exploring health-related multivariate time series with a hybrid attentive model," *Sensors*, vol. 22, p. 6104, 8 2022.

[34] S. Sang, F. Qu, and P. Nie, "Ensembles of gradient boosting recurrent neural network for time series data prediction," *IEEE Access*, pp. 1–1, 2021.

[35] D. B. da Silva, D. Schmidt, C. A. da Costa, R. da Rosa Righi, and B. Eskofier, "Deepsigns: A predictive model based on deep learning for the early detection of patient health deterioration," *Expert Systems with Applications*, vol. 165, p. 113905, 3 2021.

[36] F. Haselbeck, J. Killinger, K. Menrad, T. Hannus, and D. G. Grimm, "Machine learning outperforms classical forecasting on horticultural sales predictions," *Machine Learning with Applications*, vol. 7, p. 100239, Mar. 2022.

[37] H. Chacón, V. Koppisetti, D. Hardage, K.-K. R. Choo, and P. Rad, "Forecasting call center arrivals using temporal memory networks and gradient boosting algorithm," *Expert Systems with Applications*, vol. 224, p. 119983, Aug. 2023.

[38] Global Initiative For Chronic Obstructive Lung Disease (GOLD), 2023. `https://goldcopd.org/2023-gold-report-2/`, Accessed on 24.09.2023.

[39] R. Draxler and G. Hess, "Hybrid single-particle Lagrangian integrated trajectories (HY-SPLIT): Version 4.0-Description of the Hysplit_4 Modeling System," vol. 12, Jan. 2010.

[40] G. Rolph, A. Stein, and B. Stunder, "Real-time environmental applications and display sYstem: READY," *Environmental Modelling &amp Software*, vol. 95, pp. 210–228, Sept. 2017.

[41] R. Chandra, S. Goyal, and R. Gupta, "Evaluation of deep learning models for multi-step ahead time series prediction," *IEEE Access*, vol. 9, pp. 83105–83123, 2021.

[42] N. Prat, I. Comyn-Wattiau, and J. Akoka, "Artifact evaluation in information systems design-science research–a holistic view," 2014.

[43] R. Likert, "A technique for the measurement of attitudes," *Archives of Psychology*, vol. 140, pp. 1–55, 1932.

[44] J. Pereira, N. Antunes, J. Rosa, J. C. Ferreira, S. Mogo, and M. Pereira, "Intelligent clinical decision support system for managing COPD patients," *Journal of Personalized Medicine*, vol. 13, p. 1359, Sept. 2023.