Qvida+: Development of a Clinical Decision Support System

Daniela João

U. PORTO

FEUP **FACULDADE DE ENGENHARIA**
UNIVERSIDADE DO PORTO

Mestrado Integrado em Engenharia Informática

Supervisor: Luis Paulo Reis, Eliana Silva

February 28, 2021

Qvida+: Development of a Clinical Decision Support System

Daniela João

Mestrado Integrado em Engenharia Informática

February 28, 2021

Abstract

In these last few decades, there has been a significant increase in the average life expectancy due to improved general living conditions as well as to several advances in the medical field. Contrary to what happened then, people with chronic conditions live more now, and so it is essential to ensure their quality of life. Furthermore, in addition to prolonging their life, another goal of medical treatment is to maintain or increase patients' quality of life. Health Related Quality of Life (HRQoL) can be defined as the individuals' perception of their own health status (physical, functional, emotional and social) and the impact of their condition or treatment in their daily life(job, family, friends). The QVida+ project, based on recent scientific and technological advances in the HRQoL fields and mobile devices, intends to create an innovative paradigm when it comes to the assessment and application of HRQoL. The following step to this project and the aim of the current dissertation was to develop a clinical support system. Such system would gather all the data collected in previous steps of this project, such as biometric data (e.g., sleep, heart rate variability)and physical activity(e.g., number of daily steps) collected through smartbands, responses to self-report questionnaires and clinical data from cancer patients and provide healthcare professionals with more and better information about their patients. This system goal is to analyse patients' evolution regarding their health status and HRQoL, as well as consequently assisting healthcare professionals on future decisions with greater quantity and quality of information. After developing this system, a validation process was carried among a convenience sample inside the university community to mainly understand the usability of the system and some other aspects such as utility and effectiveness, among others. The results have shown positive feedback of the system from the participants. However, further validation among the target audience must be done in the future.

Keywords: Health Related Quality of Life, Clinical Decision Support System, Biometric Data, Oncology

Areas:

- CCS -> Applied computing-> Health care information systems.

- CCS -> Information systems -> Decision support systems.

- CCS-> Information systems -> Computing platforms

ii

Acknowledgments

This dissertation wouldn't have been possible without the help of several people, and therefore this dissertation belongs a little bit to everyone of them.

First, I would like to thank my family, particularly my parents, for always making sure that I would get the best education possible and for supporting and believing in my ideas, my work, and my dreams even when myself would doubt them.

Secondly, I would like to thank my supervisors Eliana Silva and Luís Paulo Reis, for their help and guidance throughout this journey, this project could not have been done without them.

I would also like to stress here, the importance of all my teachers and educators from kindergarten to college, for believing in my potential and always pushing me, and making me realize that there's no limits for what one can achieve when we put our minds and hard work in everything we do.

Finally, but not less important, I would like to thank my brother Daniel João, my cousin Catarina Antão, and all of my friends, particularly Carolina Azevedo and Daniel Fernandes, for helping me with the proofreading of this document, but specially for keeping me a sane person during this insane times and always making sure I was in my best mind state to work and finish what is the most important project of my life, so far.

"Alone we can do so little; together we can do so much." – Helen Keller

"I'm a great believer in luck, and I find the harder I work, the more I have of it. "

Thomas Jefferson, former President of the United States

Contents

List of Figures

List of Tables

Abbreviations

QoL	Quality of Life
HRQoL	Health Related Quality of Life
ML	Machine Learning
WHO	World's Health Organization
SF-36	RAND Medical Outcomes Study 36-Item Health Survey
QWD	Quality of Well-Being
WHOQOL-BREF	World Health Organization Quality of Life - Brief
EUROQ0L	The European Quality of Life
EORTC QLQ-C30	European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire
FACT	Functional Assessment of Cancer Therapy
CARES	Cancer Rehabilitation Evaluation System
EDA	Electrodermal Activity
BP	Blood Pressure
BVP	Blood Volume Pulse
HR	Heart Rate
IBI	Inter-Beat Interval
HRV	Heart Rate Variability
CDS	Clinical Support System
LISA	Leukaemia Intervention Scheduling and Advice
BMI	body mass index
ECOG	Eastern Cooperative Oncology Group
PET	Positron Emission Tomography
CT	Computed Tomography
IPI	International Prognostic Index
LDH	Lactate Dehydrogenase
RMSSD	Root Mean Square of Successive Differences
ms	Milliseconds
SDNN	Standard Deviation of Normal to Normal
SVM	Support Vector Machine

Chapter 1

Introduction

The perspective of effectively measure Health-Related Quality of Life (HRQoL) is highly desired by health care professional, in order to accurately describe and assess the impacts of certain illnesses, conditions or treatments in patients daily life and routine [1]. HRQoL is one of the several variables commonly studied in the field of healthcare research and it includes a wide range of human experiences, including functioning and subjective response to illness. The most recent interpretations of HRQOL are based on the World Health Organization's (WHO) definition of health as a state of complete physical, mental, and social well-being and not merely the absence of disease [2]. With this concepts in mind, with this project, we aim to develop a decision support system who gathers data from several sources, such as self-report questionnaires, biometric data collected through smartbands and clinical data from cancer patients to help and support health care professionals in their medical decisions. The present chapter will give an overview of the context of this project, describe the motivation behind it, which goals it should achieve, and finally, it will give a summary of its structure.

1.1 Context

This dissertation is inserted in the area of health informatics, which aims to improve health care through the use of information systems. This dissertation will also be developed within the scope of the QVida+ research project. The QVida+ research project was developed in a consorcium formed by the company Optimizer, Centro Algoritmi, from University of Minho and Instituto de Ciências de Vida e da Saúde (ICVS) also from University of Minho, and the Artificial Intelligence and Computer Science Laboratory (LIACC) from University of Porto. This dissertation will be done in integration with LIACC, which is responsible for the Artificial Intelligence component of QVida+. LIACC is a laboratory that supports the collaboration of researchers from different Faculties that work in the fields of Computer Science and Artificial Intelligence. The first overall objective of the QVida+ project was to create a first prototype of an information system that allows continuous evaluation of the Quality of Life (QoL) of patients and supports decision making based on this data. Some steps were already carried out with this purpose, namely with some projects and

dissertations related with the reliability and efficacy of biometric data gathered through wearable devices [3].

1.2   Motivation and Objectives

With the extension of life expectancy due to the evolution of medicine along with the crescent rise of chronic illnesses, significant attention has been focused on exploring the impact of physical and mental illness on the overall quality of life. Although, there are much literature on this topic, and many approaches have already been taken in this regard, there is not any approach yet that enables an accurate and efficient way to assess HRQoL due to the subjective nature of the patient's perception of quality of life [4]. Most literature addresses the topic of HRQoL assessment with questionnaires [5], [6], [7], some merge the questionnaire approaches with the integration of clinical data from patients in an information system [8], and more recently a few pieces of literature have emerged on HRQoL assessment using biometric data collected with wearable devices [9], [10]. However, there is still a need for a more complete integration of all these data to better assess patients' HRQoL, and that was the aim of the current dissertation. Having this major objective, with this dissertation, we intended to:

- Integrate all different types of data (clinical, biometric and self-reported) in a clinical decision support system;

- Find the best approach to display this data so it can add knowledge to health care professionals;

- Develop the decision support system;

With this clinical decision support system, we aim to:

- Provide healthcare professionals with greater quality and quantity of information about their patients' quality of life;

- Assist them in future health care decisions;

1.3   Document Structure

Besides this introduction, so far, this dissertation contains 5 more chapters. In chapter 2, it is presented an overview of the state of the art and related work to HRQoL assessment. In chapter 3, the literature review and related work for decision support systems are explored. In chapter 4, the problem definition of this project is stated, and all topics related to the development of the clinical decision support system are referred. In chapter 5, the process of validation of our system is analysed. Finally, chapter 6 is the conclusions chapter, presenting a summary of the project as well as project limitations and future work.

Chapter 2

HRQOoL Assessment

2.1 Introduction

In order to clearly understand the following sections, it is important to approach and explore some key concepts that are referred. The first and central concept of this dissertation is Quality of Life (QoL). This concept encompasses how individuals measure and evaluate their own well-being in multiple aspects of their life. Besides physical aspects, these evaluations include one's emotional reactions to life events, disposition, sense of life fulfillment and satisfaction with work and personal relationships [4]. In a more technical way, QoL is often referred as the difference between one's expectations and it is current state, in a specific period of time [11]. Thus, and accordingly to some authors only the individual himself can make this assessment which is dependent of his current lifestyle, past experiences, expectations, dreams and ambitions for the future.

Another key factor influencing an individual's quality of life can be it is current health status, and in this context a more specific concept related to QoL arises which is HRQoL. HRQoL can be referred as a concept that includes both positive and negative aspects of health, since good health is not only the absence of disease but also a state of complete physical, mental, and social well-being [12]. The negative aspects include disease and dysfunctions limitations, while positive aspects are more related with mental and physical well-being, full functioning, physical fitness, adjustment, and efficiency of both mind and body. Consequently, we can perceive HRQoL as a dynamic concept that encompasses several factors such as individual's physical health, physiological state, independence levels, social relationships and the relationship with their environment. These aspects are health related when influenced by illness, injury and treatments [13]. Besides the obvious importance for the life of the patient, these concepts are of mandatory importance to medical practice, especially in the measurement of the impact of treatments and diseases, specifically the ones with a chronic nature. The usual medical assessment is to measure physical indicators to evaluate the capacity of patients, but these measurements are often poorly correlated with functional capacity in their daily life. Also, the same clinical status can affect different patients in different ways, sometimes with impressive contrasts in their QOL [9].

Accordingly to the World Health Organization (WHO), HRQoL is a measure that can be quanti-

Table 2.1: HRQoL Domains

Physical Health	Psychological	Social Relationships	Relationship with the Environment
Activities	Self-Esteem	Sexual Activity	Economy
Pain	Negative Feelings	Social Support	Information
Insomnia	Body Image	Family	Means of Transportation
Mobility	Thoughts	Personal Relationships	Security
Medication	Spirituality		Services
Dyspnea	Positive Feelings		Free Time
	Psychologial Well-being		

fied objectively in several domains according to the Table 2.1. In the last few years, an accurate HRQoL assessment has been considered a medical goal and much research and literature has risen within this scope.

2.2   Self-Assessment Questionnaires

Inside the scope of self-assessment questionnaires we can find two main categories: generic questionnaires that offer an overall summary of HRQoL and specific questionnaires that focus on certain diseases or patient groups.

General HRQoL assessment instruments have been extensively studied and validated in many types of patients. Typically, they address general health perceptions of overall well-being, such as function in the physical, emotional and social domains [2]. The most common examples of this types of questionnaires are:

- RAND Medical Outcomes Study 36-Item Health Survey (SF-36) [14] is a self-administrated instrument that takes less than 10 minutes to complete and quantifies HRQoL in multi-item scales that address different concepts such as: physical function, role limitation because of physical hurdles, bodily pain, general health perceptions, social function, emotional well-being, role limitation because of emotional problems, and energy/fatigue. This questionnaire has a portuguese version which besides being adopted to the language is also adopted to the culture [15]

- Quality of Well-being scale (QWB) [16] that in its original version had 18 item and would require a trained interviewer, even though newer versions can be self-administrated. This questionnaire summarizes three aspects of health status (mobility, physical activity, and social activity) in terms of quality-adjusted life years, quantifying HRQoL as a number that may range from complete well-being to death.

- Sickness Impact Profile [17] measures the impact of a certain sickness on an individual's daily activities and behaviour. This questionnaire is self-administrated and contain 136 items which can take 30 minutes or longer to complete.

- The Nottingham Health Profile [18] covers six topics that may be affected by a certain illness or condition: pain, physical mobility, sleep, emotional reactions, energy and social isolation by using a series of weighted yes or no items. It contains 38 self-administrated item and is relatively quick to complete. This questionnaire has a portuguese version highly used in Endometriosis [19].

- Profile of Mood States [20] in its original format has 65 items, however the most common version is the shortened version with 35 items [21]. It is a self-administrated questionnaire to measure mental health. This questionnaire measures six main categories: anxiety, depression, anger, vigor, fatigue and confusion. The short version of this questionnaire also presents a portuguese version [22].

- World Health Organization Quality of Life - Brief (WHOQOL-BREF) [23] covers four mains domains: physical, psychological, social and environment with a set of 26 items that can be self-administrated. The answers to this questionnaire use a 5-point Likert scale, inquiring 'how much', 'how satisfied', or 'how completely' the patient feels towards the domain being investigated. The WHOQOL-BREF was validated among people with type 2 diabetes and got good to excellent psychometric properties of reliability and validity. There is a WHOQOL version adapted to the portuguese universe validated among elderly population [24] .

- The European Quality of Life (EuroQoL) [25] is a self-administered scale used to assess 5 domains: mobility, self-care, role or activity, family and leisure activities, and pain and mood. The patient answers the questionnaire by ticking the statement that best describes his sense of current state of health. In the end, the patient also rates their health on the EuroQol thermometer which is a scale from 0-100 (worst imaginable health state to best imaginable health state). This questionnaire performed very well in this scope and was pointed out to be simple to use. A portuguese validated version was already developed [26] .

When it comes to disease specific HRQoL questionnaires, they focus on an area of interest, and it may be specific to a disease, to a specific group of patients or to a certain malfunction or problem. These disease-specific instruments assess a specif domain in order to measure responsiveness to certain treatments or other clinical actions [13]. Because of cancer's well-documented impact of their malignancies and treatments, quality of life specific to cancer patients has been largely investigated. There have been several questionnaires developed and tested that measure the impact of cancer and even specific types of cancer on patients' daily life. These instruments are particularly relevant to decide about medical practices, such as the choice of treatment or surgery, etc. Some examples of these questionnaires, validated and tested in patients with several types of cancer, are:

- European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30) [27] is a cancer specific questionnaire designed to measure cancer-specific HRQoL. It is composed of 30 items that address several domains that are

common to all types of cancer patients such as functional scales (physical, role, emotional, cognitive functioning, and social functioning), a global health scale, symptom scales (fatigue, nausea/vomiting, and pain) and some items concerning dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties related to the disease/treatment [2]. This instrument presents a validated portuguese version [5]. Apart from this questionnaire who assesses the impact of cancer in a more generic way, some disease-specific modules were developed as well, according to methodologically rigorous techniques. Some of these extra modules are: breast cancer [28], lung cancer [29], prostate cancer [30] and head and neck cancer [31].

- Functional Assessment of Cancer Therapy (FACT) [6] is a 28 item, self-administrated questionnaire with two-parts, that includes a general item set suitable to all types of cancer patients - FACT-G (Functional Assessment of Cancer Therapy - General) and one of several item sets regarding specific cancers. Each item is a statement in which the patient agrees or disagrees within a 5 point range. This questionnaires addresses areas such as: physical, family/social, emotional, and functional. Disease specific modules are available for colorectal [32], breast [33], prostate [34], ovary [35], lymphoma [36], among others.

- Cancer Rehabilitation Evaluation System (CARES)[7] is composed by 59 items. It is a self-administrated questionnaires and addresses cancer-related quality of life with 5 multi-item scales: physical, psychosocial, medical interaction, marital interaction and sexual function. CARES has been validated in patients with different tumors, including urological tumor.

- UCLA Prostate Cancer Index [37] is a 20-item questionnaire that addresses disease-specific HRQoL in domains such as urinary, sexual, bowel function, bother, etc. It has been validated to assess disease-targeted HRQoL in men treated for early-stage prostate cancer.

General HRQoL aspects and disease specific aspects are correlated and may impact each other. This interaction must be considered when interpreting the results. Despite the apparent need and benefits that this type of HRQoL assessment brings to chronic patients, there are still several limitations related to its use in healthcare, which translates into a still very low adoption of this method. The first limitation concerns the difficulty of health care professionals to understand the meaning of an individual's score or even how to interpret a change in the score [4]. Another limitation lies in the fact that several patients don't have the skills or physical capabilities to fill such questionnaires, for instance patients who have already lost some of their body or brain functions such as their long term memory, their ability to focus or to write, due to their conditions. Another problem with these methods is the preclusion of conducting such assessments on a daily basis or at least in a frequent base. These questionnaires are often only administrated when patients go to their health units, which is as few times as possible to avoid negative impact on patients caused by frequents trips to these places which can, sometimes, be far away from their residences. One way to overpass the latter limitation have been the online questionnaires that can be answered at home, but in that case another issue arises, which is the lack of feedback existent by the system

which leads to a loss of interest and motivation by the patient to interact with the system. The last limitation found, is the large extension of some of these questionnaires that can lead to patients saturation and a consequent loss of reliability in their answers [1].

2.3 Biometric Monitoring to Assess HRQoL

The most common way to assess HRQoL is through self-assessment questionnaires. However this assessment lacks the accurate assessment of objective components that can be done more accurately through devices in a non-intrusive way. These objective components are related to clinical indicators that measure key variables to the patients performance towards their daily tasks. Thus, to obtain a more complete and appropriate assessment of HRQoL, both objective functioning and subjective well being must be taken into consideration. [10]. With the rise of health and fitness conscious, some interesting technology has emerged in that scope, such as fitness watches, smart bands or smartphone-powered fitness applications. These instruments are becoming more robust and complex which makes them very popular among the fitness and health culture [38].Acknowledging these concepts, the idea of using such technological devices integrated in a system, that provided biofeedback in real-time in a non-obtrusive way, both to patients and to the health care professionals, emerged [10]. In addition, patients being under the notion that they are constantly being monitored might make them feel more secure and consequently less stressed about their condition, which can lead to a more positive outcome [38]. The physiological data gathered from these methods can be:

- **Electrodermal Activity** (EDA) is measured in the form of skin conductance. Some of the glands responsible to fluctuate the EDA respond to psychological changes [39] which translates into a close relationship between the EDA measured and one's emotional state, arousal, stress, among others.

- **Electrocardiograph** (ECG) measures that are related with heart rate (HR) and therefore heart rate variability (HRV) and inter-beat interval (IBI). From the HRV we can infer even more measures such as the root mean square of successive differences between normal heartbeats (RMSSD), which is obtained by first calculating each successive time difference between heartbeats in milliseconds (ms) , and the the standard deviation of the IBI of normal sinus beats (SDNN) also measured in ms. Then, each of the values is squared, and the result is averaged before the square root of the total is obtained. These measures can be intimately related to fatigue, stress, arousal, anxiety, among others. The before-mentioned variables, specially when combined with other cardiac metrics such as blood pressure (BP), can be very helpful to identify positive and negative emotional states and even some pathologies related to cardiac health [40].

- **Electromyography** (EMG) measures muscle activity and tension. It may be related with pain, anxiety disorders, muscle retraining after injuries, incontinence, among others [41].

- **Breathing patterns** which consist in the measurement of the volume of air contained in an individual's lungs [38]. This biofeedback technique is relevant given the correlation of this variable to wide variety of disorders, including asthma, cardiovascular disorders, hypertension, anxiety, duodenal ulcers, among others [42].

- **Electroencephalography** (EEG) also called neurofeedback, measures brain waves. These metrics may be used to treat conditions related with trauma, anxiety, insomnia, pain and several conditions related with mental disorders [41].

- **Thermal biofeedback** which consists in the measurement of body temperature with high sensitive sensors [43]. This biomaker may present specific values when an individual is in a situation of mental exertion, worrying, psychosocial stress, anxiety, among others. Temperature biofeedback has also shown correlation with sleeping patterns [44].

- **Number of Steps and distance** which are measured through step counters or pedometers and can be relevant to link the relationship between health and physical activity. This metric helps to categorize people into more or less active, can perform as very motivational and even induce better habits [45].

- **Pulse Oximetry** is a method for monitoring a person's oxygen saturation. This metric is particularly important to get biofeedback from patients suffering with conditions related to asthma, pneumonia, lung cancer, anemia, heart attack, heart failure among others [46].

- **Sleeping Patterns** which can be more accurately measured through polysomnography or actigraphy. These techniques rely in some physiologic parameters discussed before such as: brain dynamics with electroencephalography, muscle activity with Electromyography, heart physiology, respiratory function and even eye movement. Some of the most common sleep measures used to understand sleep patterns are sleep duration, sleep efficiency, number of times a person wakes up during sleep time, calories burnt during this period, etc. These metrics are particularly relevant in measuring HRQoL given that sleep disturbances are common in critically ill patients and they contribute to patient morbidity [47].

In order to achieve results that accurately translate the HRQoL of patients, the choice of the wearable device and consequently the variables to be measured must be taken into consideration. These wearable devices can be divided into two categories: Consumer Devices and Medical Devices. In the scope of this dissertation, both devices present features that would add value to the final project such as the focus on being everyday usable, the unobtrusiveness and battery life of the consumer devices and the high precision, research level quality readings and accuracy of the medical ones [38]. Medical Devices offer advanced systems for physiologic research with sensors that measure most biometric variables like EDA, ECG, EMG, breathing patterns from which its possible to infer a set of secondary variables. These devices present high quality materials, readings and software to present and analyse the data. However, in terms of weight and size there aren't many differences among them, being in general much heavier and larger then the consumer

devices. Their prices can be 40 times the time of non medical devices. The Consumer Devices, unlike medical devices, focus essentially on ludic activities, such as hiking, fitness tracking, calories counting, biofeedback videogames, etc. They are increasingly becoming smaller, lighter, more unobtrusive, more robust, more complex and more autonomous. Some of the most popular consumer devices in the market are [38]:

- Feel Emotional Wristband [48] from Feel is able to measure movement and activity, HRV, EDA and body temperature. After gathering the data, Feel uses Advanced Artificial Intelligence (AI) such as machine and deep learning methods, and signal processing algorithms to analyze vital signs and identify effective patterns that are presented to the users in the form of daily report through the Feel App. This service is especially related to the control and understanding of one's emotions, and is particularly used by people suffering from mental health disorders. The most basic version of this wristband is retailed from 130€.

- Zenta Wristband by Vinaya [49] is able to measure skin temperature, physical activity, EDA, HRV, blood oxygen saturation and even noise, from which when combined to other variables can also identify sleep patterns. This system uses a a set of machine learning methods that are able to identify patterns related to one's anxiety, well-being, movement, among others, and understand how those affect one's mood. The integration of this system to a smartphone is easy, and it keeps the user informed with alerts and notifications about its mental and physical state. The most basic version of this wristband is retailed from 132€.

- Microsoft Band 2 [50] measures HR and HRV, movement and activity through a GPS system, calories burn, sleep quality and quantity, O2 levels and body temperature. It also includes ambient light sensors, a gyroscope and accelerometer, a UV sensor, a barometer, a microphone, and it is waterproof. It is embedded with technology from Microsoft Health which gives a proper visualization, data processing methods and actionable insights based on one's data. The system can be easily integrated with any smartphone on the market and it can easily connect through them with Bluetooth mechanisms. Despite having been discontinued, it is still possible to find them for sale within a range price of 55€ to 85€ depending on their size. Inside the scope of the consumer devices, this solution is one of the most complete and research-focused solution .

- Jawbone UP3 [51] measures oxygen levels, EDA, body temperature, HR and HRV and it can infer respiration and sleeping patterns. It also includes fitness tracking, regarding caloric intake and physical activity as well as sleep patterns, sleep quality and physical activity. However, it does not provide an integration API to access the data in real-time. It retails from 22€.

- Apple Watch [52] which is already in its fifth series of watches, retails from 240€ to 460€ depending on the version and the type of customization required. The latest versions are become more health oriented and include several heart health tools such as HRV, HR, BP, among others, several ECG tools, a built-in compass, respiration patterns measurements

through photoplethysmographic measurement, GPS tracking, an accelerometer, and it has all international emergency numbers on speed dial, which the watch may use when detecting some irregularity such as a fall. Within an Apple ecosystem is probably on of the most promising devices. It enables an API called watchOS [53] for the development of apps to integrate in the watch and it supports the integration with some specific health applications. However the lack of integration with other operative systems and APIs outside the scope of Apple systems makes it very limited.

- Galaxy Watch [54] by Samsung has already two models and retails from 309€. This device can count steps, number of floors climbed, measure HR and HRV, identify sleeping patterns and identify stress levels according to the breathing sensors. It also provides tools to manage one's weight and it records calories, water and caffeine intakes. The Galaxy Watch already offers several Samsung Health applications with the possibility of customization. Similarly to what happened with Apple, Samsung provides an API [55] for developers to built their own watch applications and integrate them in their Galaxy Watch.

- WearOS [56] by Google is an API for developers, free of charge, to built their own watch application and integrate them in most Android based smartwatches in the market like the Fossil Sport [57] or Huawei Watches [58]. This is not the conventional type of system, however the possibility of choosing the watch that better suits the user, regarding its sensors, integrated with a customized application that suits the user's purpose makes this solution interesting for research purposes.

- BITalino [59] which retails from 130€ to 180€ and it is a set of hardware and software tools designed for developers, engineers and technology lovers to build their own gadgets or software applications. It offers mechanisms to perform ECG, EDA, EMG, EEG and movement and activity measurements. Thermal biofeedback can easily be integrated through the set of software tools BITalino offers. The platform also provides a complete software platform (Open Signals) to visualize, analyse and process data in real-time. This solution requires some extra work regarding the construction of the device and its software but its the most versatile and customized to users' needs.

In the scope of the Qvida+ project, a study to assess HRQoL through biometric monitoring was already conducted [10]. This study focused on cancer patients with lymphoma, from whom it was collected biometric data related to their sleep, hrv and physical activity. In order to understand which type of device would better suit the project, some variables were analysed such as features, signal processing, precision, sensor reliability, intrusiveness, practicality and cost. When considered all points, Microsoft Band 2, according to a review [38] of all devices done before the referred project, seemed the most complete solution regarding its cost-effectiveness analyses. These patients wore this device during one or two weeks at pretreatment, and during two consecutive weeks at first, third and sixth chemotherapy cycle. At the end of each week they answered

a HRQoL self-assessment questionnaire, the FACT-Lym. This study focused in finding and evaluating possible relationships between self-perceived HRQoL and in physiological data in terms of HRV metrics once HRV can be valuable in understanding pathological conditions in an individual's well-being. The results of this study have shown similar patterns in the FACT-G total score and in the parameters evaluated through the biometric data collected. In specific, the FACT domains and the HRV metrics have the lowest average levels on the first cycle and tend to increase along the following cycles. These results seem to be in concordance with the literature found in chemotherapy treatments on these conditions, which point out that individuals under such treatments are more severely affected in the first cycle. Given the good results obtained in this study, the biometric data variables used to assess HRQoL in the current dissertation were the same as the ones in this study.

Chapter 3

Clinical Decision Support Systems

3.1 Decision Support Systems in Healthcare

In the last decade, some institutions in the medicine field have been acknowledging some issues related to the quality of health care and advocate the use of clinical decision support (CDS) system to support most clinical actions, providing more efficiency and speed to all health care services [1]. A CDS system can be defined as any electronic system designed to directly aid clinical decision making by utilizing individual patients characteristics to generate patient-specific assessments or recommendations [60].

These systems provide healthcare professionals and patients, a set of tools to build knowledge and understand person-specific information. This information should be filtered, analysed and presented in an appropriate way that may help to enhance health and health care.

CDS systems may present a variety of tools to enhance decision-making in the clinical workflow. These tools include computerized alerts and reminders to care providers and patients, clinical protocols, patients clinical data, documentation templates, diagnostic support, relevant references to information within the context, among others [61].

Some particularly good examples that illustrate the good results obtained through decision support systems are:

- LISA (Leukaemia Intervention Scheduling and Advice) project [62] which is a web-based decision-support system for trial management of childhood acute lymphoblastic leukaemia. In this type of condition, continuous chemotherapy is a key component of the treatment. During this process, there are several decisions that must be taken weekly, regarding the doses administrated to the patients. The choice of theses dosages is not trivial, specially if the patient's data is not readily available, or the professional is not much experienced. LISA was designed to facilitate this process, whether by presenting the patients data in real-time or by helping the professionals making decisions, according to the protocol used in such processes. However, the system doesn't prevent healthcare professionals from not following what was advised by the system, if they decide so. After testing the system with several

professionals with different levels of expertise, results showed that LISA reduced significantly the number of erroneous prescriptions, while reducing the time taken by novices to reach a decision for each case.However, the system increased the time taken by experts to reach a decision. The majority of the professionals said they would use the system if it was available. Fig 3.1 presents Lisa's interface.

Figure 3.1: Lisa's interface

- A Multi-parametric Decision Support System for the Prediction of Oral Cancer Reoccurrence [63]. Oral squamous cell carcinoma is the most predominant neoplasm of the head

and neck region, which can produce very aggressive and unpleasant outcomes for the patients suffering from the condition. In this project, a CDS system was formulated to integrate several types of data such as clinical, imaging and genomic data from patients. The primary purpose was to identify the factors that influenced the progression of this condition and then predict its potential relapses. To test the system, data from 41 patients was analysed. This approach presented an accuracy of 100 per cent when predicting relapses for this disease, either between patients with and without disease reoccurrence. The identification of these factors, towards the proliferation of oral squamous cell carcinoma, facilitates its treatment and may be crucial to prevent its progression or future recurrences.

However, despite increasing emphasis on CDS systems and in its role in improving healthcare processes, there is still few evidence supporting widespread use. Most studies conducted in that regard, conclude that CDS systems are effective at improving healthcare workflow, however evidences on economic, workload, and efficiency outcomes remains limited [64]; [60] . Some of the reason of why CDS systems aren't yet disseminated through hospitals, clinics and health units are:

- lack of information and knowledge from health care professionals towards the benefits of CDS systems;

- lack of acceptance by health care professionals;

- high-priced systems, which is not in concordance with the level of priority acknowledged by healthcare providers towards the systems;

- lack of integration with other computer platforms, which implies a lot of work to introduce detailed information about patients in the system;

- lack of robust computer networks in most health units;

- lack of skilled human resources to create and maintain the databases for such systems;

- reduced ability to share knowledge and guidelines between systems.

3.2 Decision Support Systems and HRQoL

The aim of CDS systems to measure and manage HRQoL is to provide both to healthcare professionals and to patients themselves a set of tools to measure, on a daily basis, patients' quality of life without interfering with the normal functioning of the health units and patients' normal routine. These systems are intended to convert tacit knowledge into explicit knowledge, producing reports and guidelines for each patient's health problems. However, such task is not trivial, given the ambiguity and the variability of how each person feels towards a certain variable, making it hard to assess accurately certain factors that may influence HRQoL. This lack of linearity, adds considerable difficulty in computerizing such behaviours. In terms of functionalities, some of the tools that must be considered to integrate in these systems should be: accurate metrics to represent

the patient's quality of life and evolution in real-time, evaluation of health gains, assess the impact of the consequences of the disease, treatment and medication, use of data mining processes to categorize and compare patients within a certain type of disease or treatment, provide statistical analysis related with survival factors, quality of life, health status across time for a specific condition or treatment, etc, and finally the addition of health, survival and quality of life predictions.

There is not much literature in this regard, and very few projects were done in this scope, however there is one that deserves mentioning which is QoLiS-Quality of Life Platform [65] implemented in the Portuguese Institute of Oncology in Porto. This project recorded the answers of the patients to HRQoL questionnaires (QLQ-C30 and QLQ-HN35) and combined these results with the patients clinical data, which could be migrated from the existent platform in use in this institution. Then the system could predict, in real-time, the HRQoL of the patients. In technical ways, this process was achieved through a On-Line Transaction Processing approach which allows the consolidation of the information of different databases on different technologies from different isolated platforms. To perform this approach, several modules were created in order to process, extract and analyse the data, with a different set of tools of statistics and data mining algorithms. In terms of used technologies, the mobile modules were developed in ADF Mobile, the front-end model was developed in Oracle ADF and the database was developed with Oracle12C. The interface layer presented to the client was developed in HTML5. The statistical and clinical reports were achieved using Java and JasperReports. Regarding the server, two technologies were used: Weblogic for pages related to clinical management, and Glassfish for reports from modules that could be combined with other applications. The webpages were developed on JavaScript, with applets responsible for the communication with the client and CSS modules responsible for the graphic style of the application. The software was written in Java, and some SQL data manipulation was embedded in Java code in the business layer which contained the rules for accessing the server. The strategies considered for this project were the use of the Rasch Model which relies on the model of item response theory (IRT) used to perform analysis of questionnaires or other instruments used to measure variables. According to this model, the statistics are independent of the items of the test, the results of the respondents are independent of the difficulty of the test, the analysis of the item takes into account level of knowledge of respondents, the analysis of the questionnaires doesn't require additional test to verify the trust and the statistics of the items and the ability of the respondents are analysed on the same scale. The IRT models usually use a logistic function to link the observed variables to the actual characteristics. Apart from this relationship, Rasch created a mathematical expression that takes into consideration other parameters such as probability of success, the ability of the respondent and the difficulty of the item, assigning a probability of success on the answer to a question regarding efficiency of the respondent and difficulty of the item. The information gathered through the questionnaires and the results obtained through this model consisted in the dependent variable to be later used in the algorithms chosen. Regarding these dependent variables, 18 aspects related to the patients clinical data were pointed out such as education level, marital status, years of smoking, local metastasis, etc. Regarding the classification of the data, the oncology institution physicians chose a friendly way to present

data: bad quality of life, normal quality of life and good quality of life. After performing several data mining techniques to the data to check errors, multicollinearity, among others, several algorithms were tested to compare how the several variables previously mentioned can relate to the classifications pointed out by the physicians. The algorithms used were:

- **K-Nearest Neighbor** is a non-parametric method used for classification problems, and it classifies the input value according to the classification of the closest value found [66].

- **Naives Bayes algorithm** applies the Bayes theorem which describes the probability of an event, based on prior knowledge of conditions that might be related to the event [67].

- **Decision Tree** builds classification in the form of a tree structure. It turns a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision nodes and leaf nodes. The topmost decision node in a tree corresponds to the best predictor called root node. This structure provides an easier way to understand and process data [68].

- **W-SMO algorithm** is a method of decomposition, by which an optimization problem of multiple variables is decomposed into a series of sub problems each optimizing an objective function of a small number of variables, while all the other variables are treated as constants that remain unchanged in the sub problem [69].

In Fig 3.2 it is presented the interface of this CDS system, where it is visible the presence of modules such as patients information, appointments, questionnaires, clinical data, QOL reports and messages/notes.

Figure 3.2: QoLiS' interface

Chapter 4

Project Implementation

In this chapter will be presented all sections regarding the system implementation, starting by the definition of the problem and specification of the requirements of the system. After this, a section about the development technologies will be presented as well as its relational model. Finally, the use cases, functionalities and interface of our system will be explained.

4.1 Problem Definition

As it was mentioned in previous chapters, this project intends to be innovative in the field of CDS systems and in the assessment of HRQoL. The purpose of this CDS is to constitute an helpful tool for health care professionals, whether for them to be better informed about the health status of their patients, or to help them in decision making related to future treatment directions, or even treatment interruptions in the oncology panorama. This CDS system will integrate patients' data from 3 different sources: their clinical data, biometric data collected through a smartband and patients answers to HRQoL assessment questionnaires. The purpose of this would be to present healthcare professionals with an accurate HRQoL assessment regarding patients' self-percepted HRQoL, but also taking into account a more objective component such as their own physical and physiological functioning. The proposed architecture for this project is presented in Fig 4.1.

As already mentioned in chapter 2, the QVida+ project [10] had already gathered some data in previous steps of its development. This project accompanied lymphoma patients before, during and after their chemotherapy cycles, where it was gathered patients clinical data, biometric data through a smartband during two weeks for each participation moment, and answers from HRQoL questionnaires in different stages of their treatments. Although, it was not possible to access that data during the development of the current project due to bureaucratic constraints and to the current pandemic situation, our goal was always to integrate that data in our CDS, so all the data and data variables that our CDS requests and presents were created to host all these QVida+ data gathered before. The goal was to make our CDS completely ready to be populated by these data. That said, all of the data present in the system at the time of the delivery of this project is fictional and illustrative.
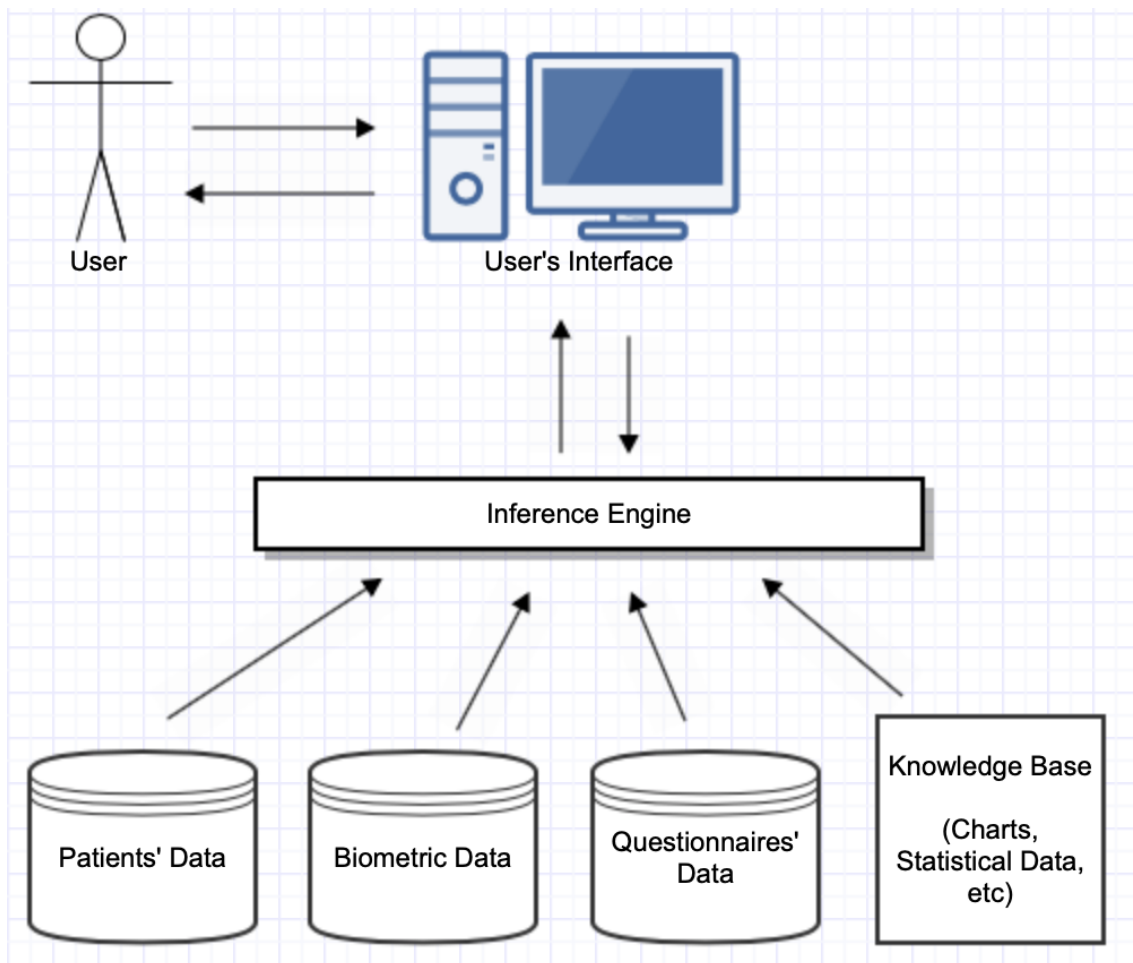
Figure 4.1: Proposed Architecture for CDS system

- The clinical data gathered in previous steps of the QVida+ project includes:

  - the chemotherapy plan for each patient, how many cycles were performed, and the initial and final dates of the cycles.

  - their weight, height and body mass index (BMI).

  - if the patient is a smoker or not and the amount of packs of cigarettes per year if so.

  - their usual medication.

  - their personal and family clinical history.

  - their ECOG [70] score.

  - biopsy and immunophenotyping results.

  - Positron Emission Tomography (PET) and Computed Tomography (CT) Scan results in a diagnose stage and after the third cycle of chemotherapy.

  - Ann Arbour staging, International Prognostic Index (IPI), hemoglobin levels, Lactate dehydrogenase (LDH) levels and microglobulin levels.

- – diagnostic information such as type of lymphoma.

    – radiotherapy information such as location, frequency, intensity and initial and final dates.

- The biomteric data gathered includes:

    – HRV from which was possible to infer the root mean square of successive differences (RMSSD) between normal heartbeats and standard deviation of the interbeat interval (SDNN).

    – sleeping measures such as duration of all sleep moments during the day, sleeping efficiency, number of times the patient wakes up, calories burned and the HRV while sleeping.

    – physical activity such as number of steps per day, distance, sedentary, moderate and light activity time, calories burned and HRV.

- The questionnaires used to assess HRQoL were:

    – SF-36 [14] questionnaire answered in the first moment of participation of the patient whether in a pre-treatment phase or in the first cycle of treatments, and in the final moment of participation.

    – FACIT [71] questionnaire gathered in a pre-treatment phase and in the first, third and sixth cycle.

4.2 Ethical Considerations

Healthcare privacy is a central ethical concern involving the use of patients data in healthcare. With vast amounts of personal information widely accessible electronically some problems arise such as the potential for abuse by employers, insurers, and the government. To prevent such occurrences, some ethical considerations must be taken in order to protect individual privacy as much as possible [72]. In our system, some ethical concerns were also taken into account, and we wanted to make sure the the patients data was only available to the healthcare professionals following them. To achieve that, we made sure that the administrator of the system was the only person responsible to certify users, so that only healthcare professionals of an health unit can log in in the system, after permission grated by the administrator. This administrator, which must be someone from the health unit that might not necessarily be an healthcare professional, has almost full privileges through out the system, however he/she can't access to any patients' clinical information, such as general clinical data, radiotherapy data, chemotherapy data, biometric data, questionnaires data and final HRQoL scores. Although, both administrator and healthcare professionals can add patients to the system, the administrator can only fill some general information (non-clinical) on them, and must always associate them with the healthcare professional following them. When the healthcare professional adds a patient to the system, this patient becomes automatically associated

with the same healthcare professional, and he/she can add, edit or access all clinical data, biometric data, questionnaires data, and HRQoL final scores of this patient. Besides that, even though the healthcare professional has access to the full list of patients in the system, whether to schedule appointments or to see some general information, they can only access personal and clinical data if followed by them and associated to them.

4.3 CDS Requirements

As previously mentioned, this CDS intends to aid healthcare professionals in their medical decisions providing them a wide range of information and knowledge, regarding lymphoma's patients HRQoL, in a very visual and easy to access type of way. To achieve so, it is imperative to find the best approach when it comes to organize and display the clinical data, the questionnaires' data and the biometric data. Given the amount of data, it is particularly easy to sort it out in a confusing or overwhelming way which, in turn, can have the opposite effect desired on healthcare professionals. With this in mind, the first main requirement would be to find the best way in which the different types of data would complement each other. Besides that, another main requirement, still in this regard, is to present them in a visual and intuitive way, making use of tables and charts when necessary, as well as additional information about specific variables when relevant, so that the professional accessing it won't waste to much time reading and understanding the information. When it comes to the management of the data, it is crucial to make sure that the health care professional can add or edit the data he/she finds relevant, such as adding or changing the patient's clinical data, add chemotherapy cycles information, update questionnaires' answers, etc. However, we believe that this should not be the only way to populate the database, given that there are already several patients with full data records that should already be present in the system, and creating one by one directly in the system would demand to much time and effort. To prevent this, one essential requirement must be the possibility of populating the database directly with data files from the patients. Another aspect we thought relevant, was to give the CDS a management component, so that it was possible to create doctors, patients and allow the scheduling of appointments in order to be easier for doctors to find the information about their patients when accessing their own appointments, without needing to search for a specific patient. Like mentioned in the previous section, it is mandatory that only certified healthcare professionals, from the health unit, would be able to register and access the patients data and it is also imperative that only the doctor following a specific patient has access to its data.

4.4 Development Technologies

Regarding the choice of development technologies, the first main task was to choose the framework in which our system would be developed. We thought relevant to make it a web application so it could be more easily accessible to anyone, from anywhere, as long as it is a certified healthcare professional who is already registered in the CDS. With that in mind and due to its

simplicity, flexibility, reliability and scalability, the Django framework [73] was chosen. Besides that, Django is an open-source framework for back-end web applications based on Python [74], which was an advantage given the simplicity and familiarity with that programming language. The Django framework has a Model-View-Template architecture (Fig 4.2) [75] which allows developers to change the visual part of an app and the business logic part separately, without their affecting one another. The three layers (Model, View, and Template) are responsible for different things and can be used independently. The model part is the source of information about the data, containing all the fields and behaviors of the data that has been stored. Django supports 4 different databases, but for this project we decided to use SQLite [76], which is the default one, and it also the one we were more familiar with. These models contain information about the data and are represented by attributes (fields). A model is a simple Python class, and since it knows nothing about other Django layers, the communication between layers is possible only via an application programming interface (API). Models hold business logic, custom methods, properties, and other things related to data manipulation. Additionally, models allow developers to create, read, update, and delete objects in the original database.

The view executes tree tasks: it accepts HTTP requests, applies business logic provided by Python classes and methods, and provides HTTP responses to clients' requests. This means that the view fetches data from a model and either gives each template access to specific data to be displayed or processes data beforehand.

The templates are files with HTML code that are used to render data. The contents of these files can be both static or dynamic. Since there mustn't be any business logic in a template, it is there only to present data.

Django is also a batteries-included framework. This means that it comes with several useful packages that allow the developer to worry just about importing them, instead of writing code for certain parts from scratch. Some of this packages are Authentication with auth package [77], Admin interfacing with admin package [78] and Session management with Sessions package [79]. the admin interface is particularly important for our project given that the admin must certify the users and associate them to specific instances of the model doctor, as it will be better explained in further sections. Django offers an administrative interface that is both professional and versatile, and the developer doesn't have to worry about that. Another relevant libraries from Django were django-adaptors [80] which is a tool that allows the developer to transform easily a CSV/XML file into a python object or a django model instance, and django-csvimport [81] which is a generic importer tool to allow the upload of CSV files for populating data. Both these libraries were particularly relevant to populate the database with the data records that had been already gathered. Still regarding Django, the library django-extensions [82], combined with python libraries such as pygraphviz [83] or pydot [84] allow the developer to easily get automatically the relational model of all the Django system implemented. Like mentioned, in previous sections, it was important for us to make sure that our CDS had a strong visual component, so that it will cost less effort and time to users to find and understand the information they're looking for. For that, we had to make sure that charts of the data were displayed along with the data tables. To do that we decided to use

Chart.js [85]. Chart.js is an open source JavaScript [86] library that helps render HTML5 charts. Chart.js is responsive and counts with 8 different chart types.
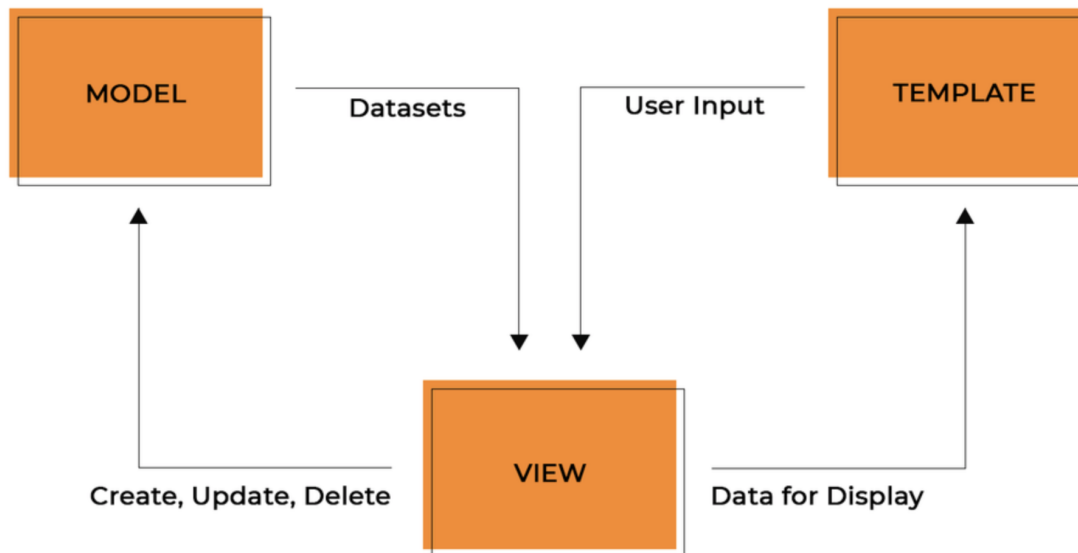

Figure 4.2: Django's Architecture

4.5 Relational Model

Regarding the database logic, there are 11 different models which map to a different single database table. The models presented in our system are: Patient, Doctor, Appointment, ClinicalData, ChemotherapyData, RadiotherapyData, SF_36, Fact_Lym, HrvData, SleepData, PhysicalActivityData(Fig 4.3).

The model Patient has as required attributes the name, age, address, condition, gender and contact of the patient. Apart from these, height, weight, IMC, medication and whether it is a smoker or not are also attributes of this model, although these ones can be null. The model Patient is the most important in the CDS given that it is the foreign key to every model in our system apart from the model Doctor. The model Doctor has as attributes the name, mobile, specialization and username which is the link between the doctor and the certified used. This model, along with the Patient model, are foreign keys to the model Appointment. The model Doctor it is also a foreign key for the model Patient. This model is particularly important given that the administrator must link each instance of this model to a user, so that each user can only access the data associated to their own patients and appointments. The attributes of the Appointment model are the date, time, the doctor and the patient. This model is mainly relevant for management reasons and to access the patient information faster through the appointments search, when a user is checking their own appointments, whether searching by the patient's name or the appointment date.
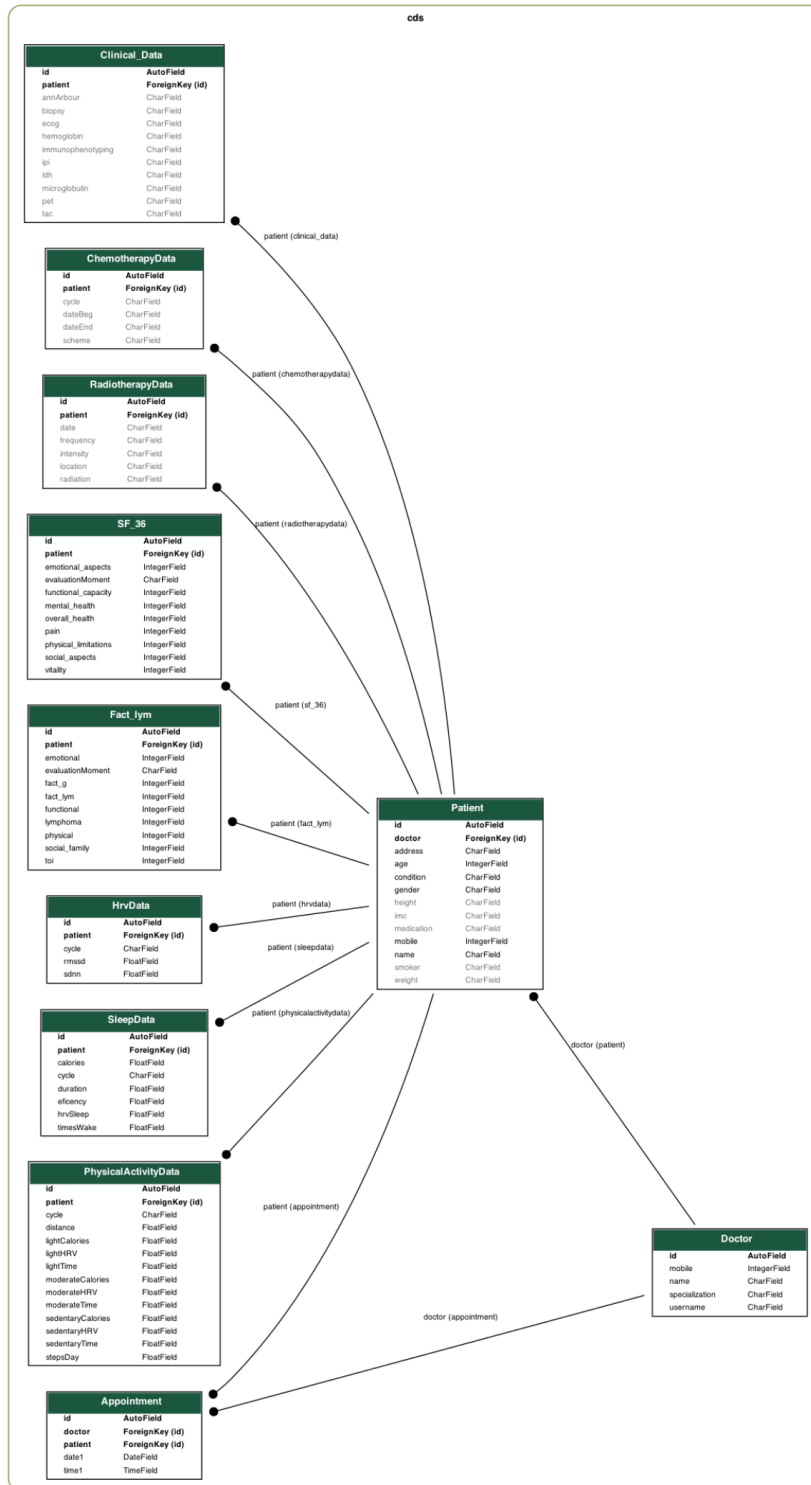
Figure 4.3: CDS relational model

The attributes on the ClinicalData model are the patient, the Ann Arbour staging classification, the biopsy result, the ECOG classification, the hemoglobin levels, the emunophenotyping result, the IPI classification, the LDH levels, microglobulin levels, the PET classification and the TAC classification. None of these attributes are mandatory. The foreign key to access this information is the patient attribute. The ChemotherapyData keeps the data related to every cycle of chemotherapy of each patient. This data consists in the patient itself, which is the foreign key to this table, the cycle of chemotherapy, the date of beginning and end of that cycle, and the chemotherapy scheme for that cycle. In the RadiotherapyData model it is where all data related to radiotherapy sessions are saved. This data consists, once again, in the patient itself as foreign key to the table, the date of the session and then the frequency, intensity, location and type of radiation. In the scope of questionnaires data, there are 2 models, the SF_36 model and the Fact_Lym model. These models keep the classification of each component of the questionnaire, the moment of evaluation and the patient associated to those classifications, which is once again the foreign key to each of these tables. The attributes that represent each SF_36 component are emotional_aspects, functional_capacity, mental_health, overall_health, pain, physical_limitations, social_aspects and vitality. In the case of the Fact_Lym model, its keeps the Fact G classification, the Fact Lym classification, TOI and lymphoma classification as well as emotional, physical, functional and social/family classifications. Regarding the Biometric Data, the data is kept in three different models, HrvData, SleepData and PhysicalActivityData. The HrvData keeps the HRV metrics for each patient in each chemotherapy cycle. So its attributes are obviously the chemotherapy cycle, the patient, and the RMSSD and SDNN values. The SleepData model is where all data related to sleep of each patient for each chemotherapy cycle is kept. This data consists on the calories burnt per night on average, sleep duration per night, on average, sleep efficiency, the hrv during sleeping times and the number of times the patient wakes up on average. The PhysicalActivityData models keeps the data related to all physical activity of each patient during a chemotherapy cycle, whether its sedentary, light or moderate activity. These data consists on the distance and steps per day. Besides these two attributes, it also keeps the calories burnt, the HRV values and the time spent on each type of physical activity (sedentary, light and moderate).

4.6   Use Cases

A UML with the use cases of this system is presented in (Fig 4.4). There are three types of users in this system, the administrator, the certified user, and the uncertified user. The uncertified user has very little privileges in this system, being only able to access the login page, the about page and the contacts page.

The certified users, intended to be doctors certified by the administrator of the system, have all the privileges the uncertified users have. Furthermore, they have full privileges when it comes to their patients' data. These users can add patients to the system as long as they're the doctor following them. They can add and access specific medical info that the administrator of the system can't, such as chemotherapy and radiotherapy cycles, biometric data and questionnaires data. They

are the only users with access to evaluation tables and charts regarding their patients HRQoL. These users can also schedule appointments of their own, and for this functionality they have access to the full list of patients in the system, and not only the ones they're following. Regarding other doctors, each certified user can access a list of doctors in the system, and access a limited amount of information on them, however they can't access which patients they're following, their appointments, or their patients' data.

The administrator of the system is responsible for creating the certified users and manage them. This is also the only user capable of adding doctors to the system and link them with the certified user, given that a certified user is intended to be a doctor. Apart from these, the administrator can also add patients to the system and associate them with any doctor in the system, which is something the certified users can't do, given that they can only add patients if they're the doctors following them. Although the administrator can create patients and fill some general information fields about them, this user can't add or access more specific information, such as radiotherapy information, chemotherapy information, questionnaires data, biometric data or evaluation tables and charts. Regarding appointments scheduling, the administrator is able to make any appointment with any doctor or patient.

Figure 4.4: UML Use Cases

4.7   System Functionalities and Interface

Some of the general functionalities of this system were already discussed in the previous section. However, we thought to be relevant to discuss some of them with greater detail. Regarding the functionalities available to all types of users, there are 3 different views available, the login page, the about page and the contacts page (Fig 4.5). After making the login, this system can be defined in two different components, the management component and the CDS to evaluate HRQoL.
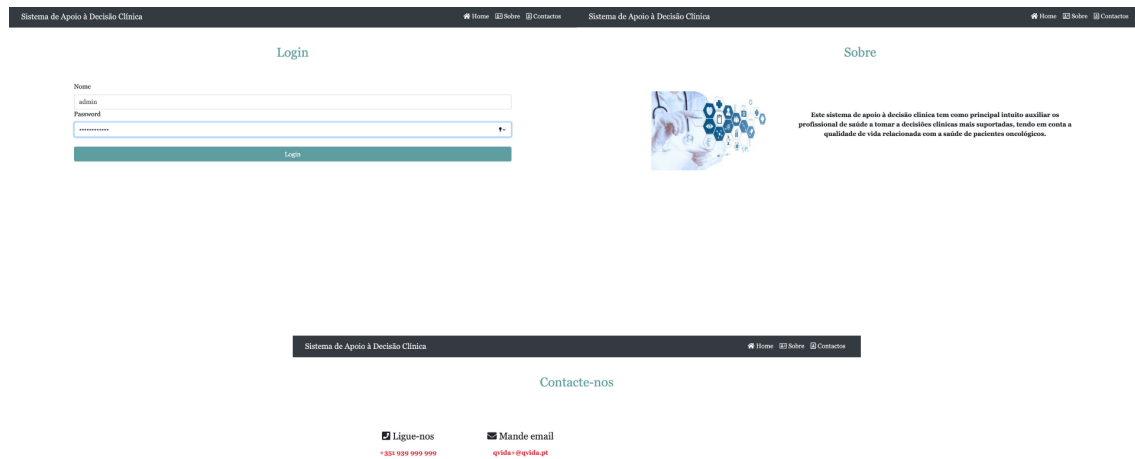
Figure 4.5: Login Page, About Page and Contact Page

4.7.1   System's Management Component

Regarding the system management component, there are some substantial differences between the views of the administrator of the system and the certified users. The first noticeable difference appears on the navigation bar, where the option of adding a doctor under the doctor tab is only available for the admin, whereas the certified user, under this tab, only has the option to view the list of doctors in the system. Besides the navigation bar, the biggest difference when logging in as an administrator, is the access to the admin panel control in the index page (Fig 4.6).

This admin panel control is the automatic Django admin interface, mentioned previously, which reads metadata from the models of the system to provide a quick interface where the administrator can manage content of the system. This panel can only be accessed by the administrator and its used as an organization's internal management tool (Fig 4.7).

In our system, this view is particularly helpful for the admin to certify the users and link them to the doctors instances via username (Fig 4.8).

Through the panel, the administrator can also add patients, add doctors and schedule appointments in the system, however the system presents specific views for these purposes. Another

Figure 4.6: Differences between the Administrator Index on the left and the Certified User Index on the right.
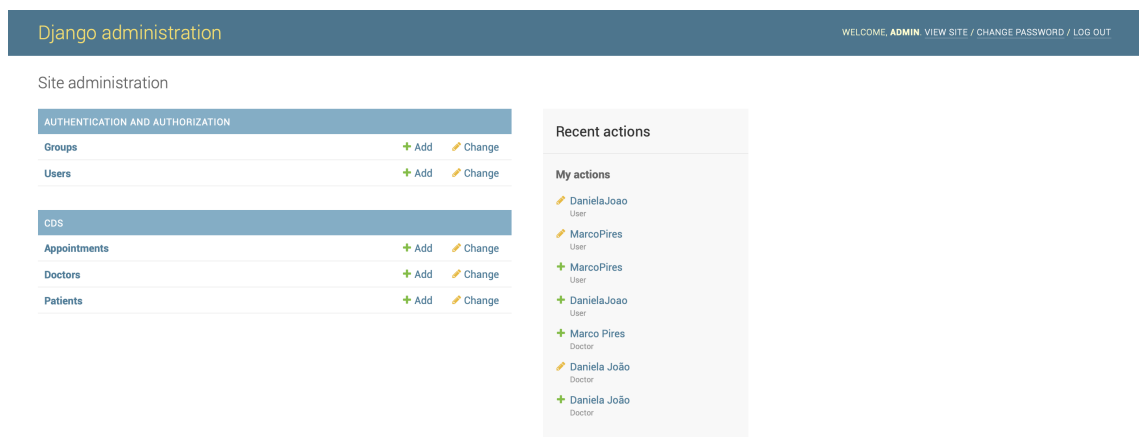

Figure 4.7: Django admin interface of our CDS

important difference between the interfaces of the administrator and the certified user is the access to patients data. Although, both users are able to create a patient, the difference in this functionality is that the administrator must associate them with a doctor while that association is automatically made for the certified user (Fig 4.9).

Besides that, the administrator is not able to add or access specific data patients or HRQoL evaluations of any patient (Fig 4.10).

In this view, we notice that the administrator doesn't have access to the links available for the certified user, which references the personal page of each patient. Another difference here, is that for the same instant in time, we realize that the list of patients presented for the administrator is larger than the one presented for the certified user. This happens because the administrator has access to the complete list of patients in the system, whereas the certified users only have access to their own patients. Similarly to what happens when creating a patient, the scheduling
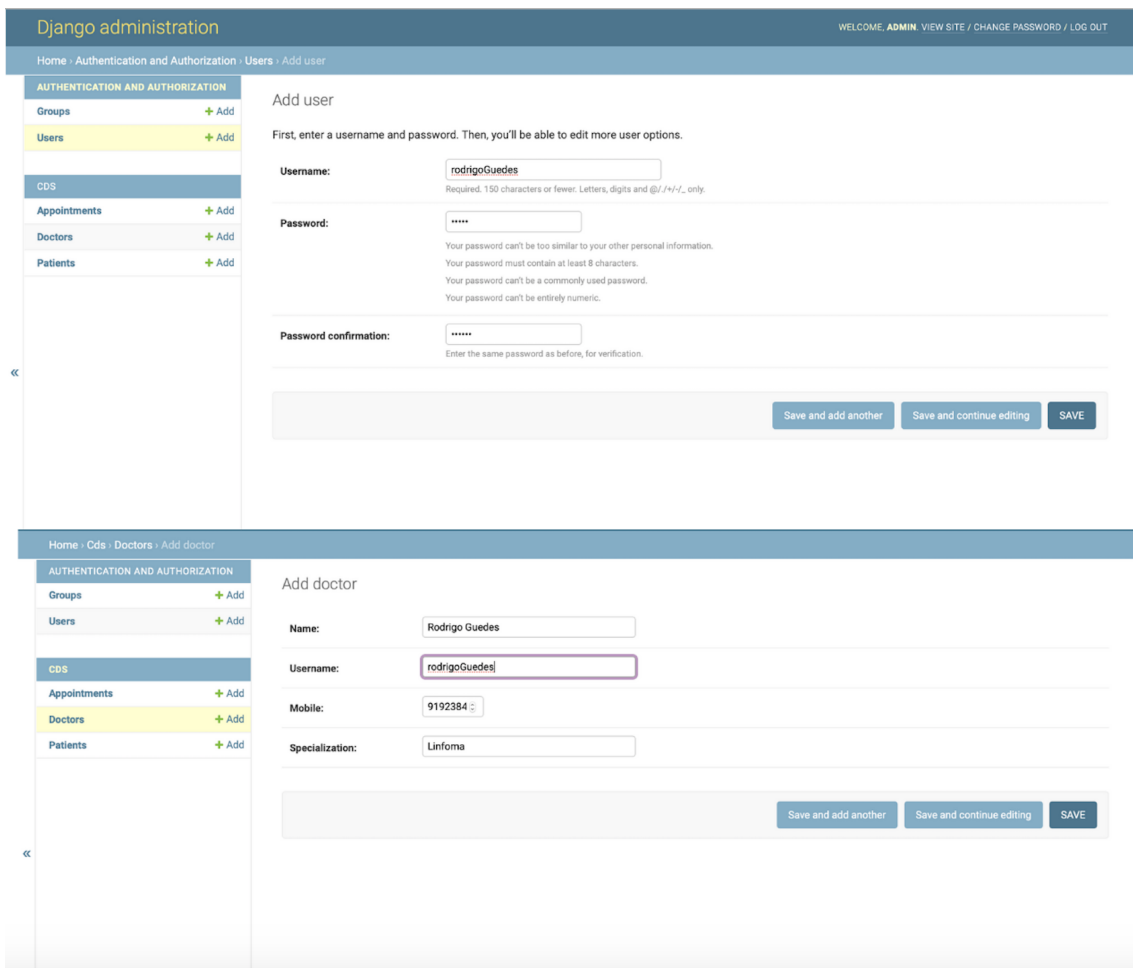
Figure 4.8: Creation of certified user by the admin in the first view. Creation of the doctor, associated to the certified user created in the first views via username, in the second view.
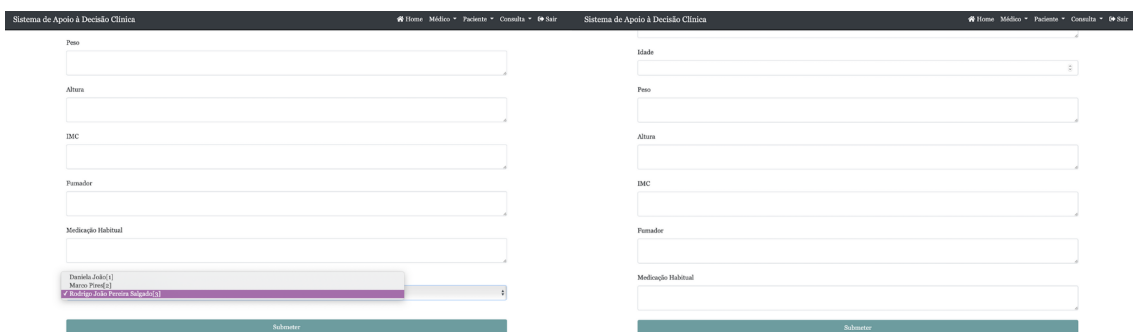

Figure 4.9: View of part of the form for creating a patient for the admin user in the right, and the certified user in the left.

of appointments also present some peculiarities for each user. The first difference, regarding appointments management is in the view of the form to schedule them. Like in the form to add a
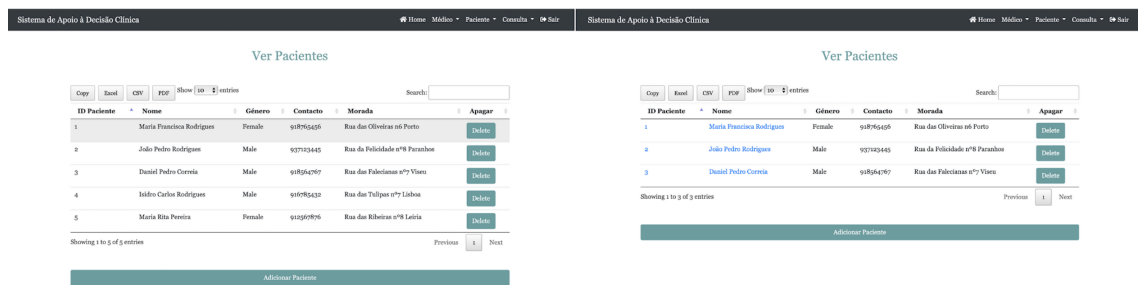
Figure 4.10: View of the list of the patients in the system for the administrator in the right and for the certified user in the left.

patient to the system, in this one, the certified user can only make appointments of they're own, so the doctor field for each appointment is filled automatically whereas in the admin view for scheduling an appointment, this user must chose which doctor is leading the appointment (Fig 4.11).
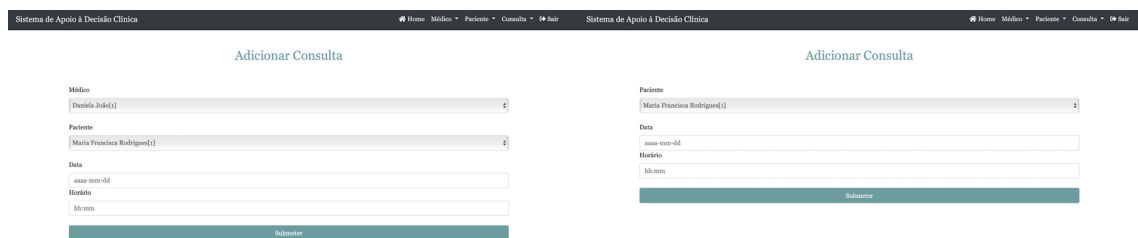
Figure 4.11: View of part of the form for creating a patient for the admin user in the right, and the certified user in the left.

Still in this view, and differently from what happened in the view where the patients where listed, the certified users, like the administrator, have access to all the patients in the system to make an appointment, and not only the patients they're following. When it comes to the view of listing all appointments, for the administrator, all upcoming appointments in the system are presented, while in the same view for the certified users, only the appointments they're leading are shown (Fig 4.12).

In this view, in the part where the names of the patients are displayed along with the rest of the information for their appointments, the certified user has the option to access the link over these names to access these patients' personal page whereas these links aren't available for the same view for the administrator.

4.7.2 System's CDS to evaluate HRQoL

Given that the patients' data is only available for certified users, in this component of the system, it no longer makes sense to talk about different users. The first view, when accessing the patient's

Figure 4.12: View of the list of the patients in the system for the administrator in the right and for the certified user in the left.

personal page, displays all the data inserted on the system for this patient when it was created (Fig).

Figure 4.13: View of the patient general information, inside the general vision tab on the left side bar

In this view, we soon notice two navigation bars, besides the main one on the top. The navigation bar on the left side presents the tab to access the general vision regarding patients data, clinical data and radiotherapy and chemotherapy data. The second tab references the page where all the information, answers and HRQoL evaluations regarding questionnaires related to HRQoL are displayed. The third tab presents information related to biometric data on the patient, and this

section is divided in three, the HRV metrics, the sleep metrics and the physical activity metrics. The final tab presents a general evaluation on the patients HRQoL regarding the patients' questionnaires' data and biometric data. The navigation bar on the top, below the main one, presents a mean to navigate inside any tab on the left side, so in this specific tab of the general vision of the patient, we notice 4 tabs on the top, the patients' data, the clinical data, the radiotherapy data and the chemotherapy data. Regarding the patients' data tab, like previously mentioned, it displays to the health care professional some general information inserted by the time this patient was added to the system, as well as the possibility to change this information. The second tab, presents the clinical data associated to this patient. If the patient was just added to the system, this view will ask for data on this regard to the user (Fig 4.14).

Figure 4.14: View of the patient's clinical data when there's still no information on this regard about a patient, on the left. On the right is the form to add this information to the system.

When this data is already present in the database, this view will display that data along with the possibility to edit this information (Fig 4.15). These data, like all data in this system, can be added by the user directly on the system or, if the amount of data is larger, it can be inserted directly in the database through JSON, XML or YAML files.

The following tab has the information related to the patient's chemotherapy cycles. This view presents a table where each line holds information on each cycle, such as the date of beginning and end of treatment as well as the chemotherapy scheme for that cycle (Fig 4.16). This table also allows the health care professional to search for a specific cycle, date or scheme, as well as download the content of this table in a pdf, excel or csv format, or copy the content to the clipboard. Besides that table, it also presents the option to add a chemotherapy cycle and the option to see the HRQoL evolution associated to these cycles. This option references the last tab on the left side navigation bar, where a general evaluation of the total HRQoL is made.

Figure 4.15: View of the patient's clinical data when this information is already present in the system.

Figure 4.16: On the left is the view of the patient's chemotherapy cycles. On the right is the form to add chemotherapy cycle information.

In the last tab of the general vision of the patient, there's the information on the radiotherapy session. Similarly to what happened in the clinical data tab, when there's still no information on this regard, the system request the insertion of this data along with a form to do so (Fig 4.17).

When this data is already present in the system, this view displays data such as the date of the session, the location, type of radiation, intensity and frequency as well as the option to edit this information (Fig 4.18).

The second tab on the left side navigation bar, is where we find the information on each patient's questionnaires. After getting on this tab, in the first view, the system requests the user to choose

Figure 4.17: On the left is the view of the patient's radiotherapy session when there's still no information on this regard about a patient. On the right is the form to add this information to the system.

Figure 4.18: View of the patient's radiotherapy session when this information is already present in the system.

which questionnaire's data (sf-36 or fact-lym) is to be analysed (Fig 4.19).

Inside the page of each questionnaire, the views and functionalities are the same for both questionnaires. The first view after choosing the questionnaire to be analysed will let the user select whether the data should be presented in a table or in a chart (Fig 4.20).

When the option selected is the table view, a table with the results for each component of the questionnaires will be displayed, organized chronologically. In this view is also possible to add more moments of evaluation or delete the ones available. As a way to read this data more efficiently and effectively, the range values for each component of the questionnaire is presented

Figure 4.19: First view inside the questionnaires tab on the left side navigation bar. The user must select which questionnaire results wants to analyse.

Figure 4.20: This is the first view after the user selects the questionnaire to be analyzed, and must choose now in which way wants this data to be displayed.

in the head of the table, and the results of these evaluations are displayed along with red downward arrows or green growing arrows, to better compare such values with previous results of the same components in other moments of evaluation. As it happens with several tables in our system, also in these tables, the option to search for certain values or moments of evaluation is available, as well

as the possibility to download this information in csv, pdf, or excel as well as copy this information to clipboard. Still in this view, the possibility to visualize this data in a chart is available without having to navigate to the previous view. In the chart view, a chart with all the components of the questionnaire previously selected is presented. This is a line chart where in the x axis are presented the moments in time in which the questionnaire was taken, usually coincidental with moments of pre or post chemotherapy cycles. Above this chart, the user is also able to change which dimensions of the questionnaire wants to be represented in the chart, making it easier to read and absorb the information, if the user is searching for specific parameters of the patients HRQoL. From this view, the user can choose to return to the table view, without having to navigate backwards. In Fig 4.21 are presented both views, table and chart, for a questionnaire's data, in this case the SF-36.
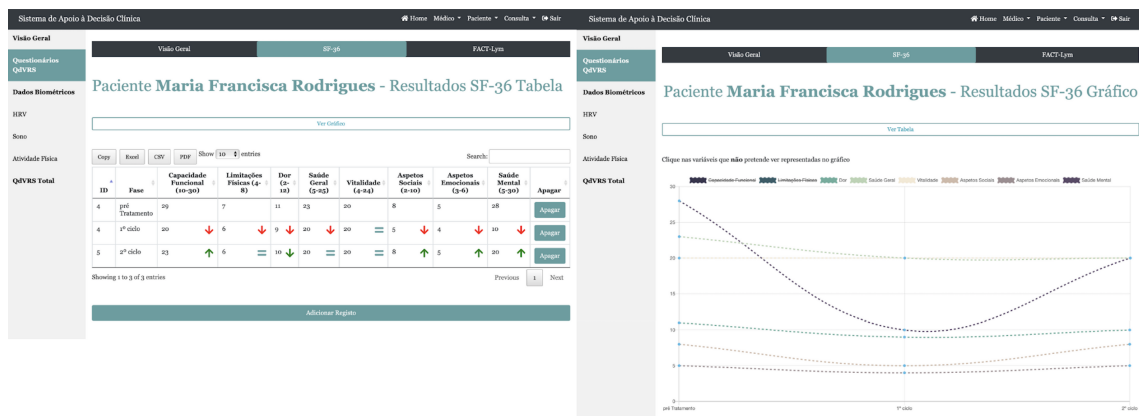
Figure 4.21: Table view with questionnaire's data on the left, and chart view for the same data on the right.

The next tab on the left side is the one that takes the user to the management of the biometric data. The first view in this tab lets the user decide which type of biometric data (HRV, sleep or physical activity data) wishes to visualize (Fig 4.22). These different types of biometric data can also be accessed through the navigation bar on the left side.

The first metrics, inside the biometric data, are the HRV metrics. Inside this view (Fig 4.23), the user has access to both the table and chart that hold the values for the patient's mean SDNN and RMSDD for each chemotherapy cycle.

As in previous tables, with this one it is possible to make searches on specific values or chemotherapy cycles, delete specific cycles, and download its content in pdf, csv or excel, as well as copy to clipboard. Besides this, this table also presents modals in both variables, which are represented through buttons for more information in the table head. These modals hold some explanation on the relevance of these variables and contain some reference values for them (Fig 4.24). The chart is presented below the table and the user can choose to visualize none, only one or both variables.

Another important type of biometric data to help assess the HRQoL are sleep metrics. The system view (Fig 4.25) for this data follows the same logic as the view for the HRV metrics.

Figure 4.22: First view inside the Biometric Data Tab

For this metric, a table and a chart are also displayed. This table presents all the functionalities previous discussed for the HRV table such as search mechanisms, download content, modals with relevant information on the 5 variables, option to delete and add cycles, as well as the possibility to add great amounts of data directly to the database. Regarding the chart, the user can choose as well which variables want to see displayed.

The last type of biometric data in our system are the physical activity metrics. Inside the scope of the physical activity data, the general variables are the number of steps per day on average and the distance walked in a day on average. Besides that, there are 3 types of physical activity approached in the system, sedentary activity, light activity and moderate activity. For each of these types of activity, there are some variables taking into account such as time spent in each activity, HRV and calories burnt. Given that, and in order to have a clearer vision of these metrics, the tab on top, inside the physical activity metrics references 4 different views. The first view (Fig 4.26), presents an overview of all variables in the table, and represents in the chart the variables concerning physical activity in general as well as the time spent in each specific type of activity.

The other 3 views (Fig 4.27) in this scope, represent the specific variables for each type of activity (sedentary, light and moderate). These representations are made through tables and charts as well. Regarding all the tables and charts on this scope, the logic is similar to the previous types of biometric data.

Finally, the last tab, on the left side navigation bar, it is the total HRQoL of the patient, where a first view is presented with two options for the user to choose between the Self-perceived HRQoL and HRQoL biometrically measured (Fig 4.28).
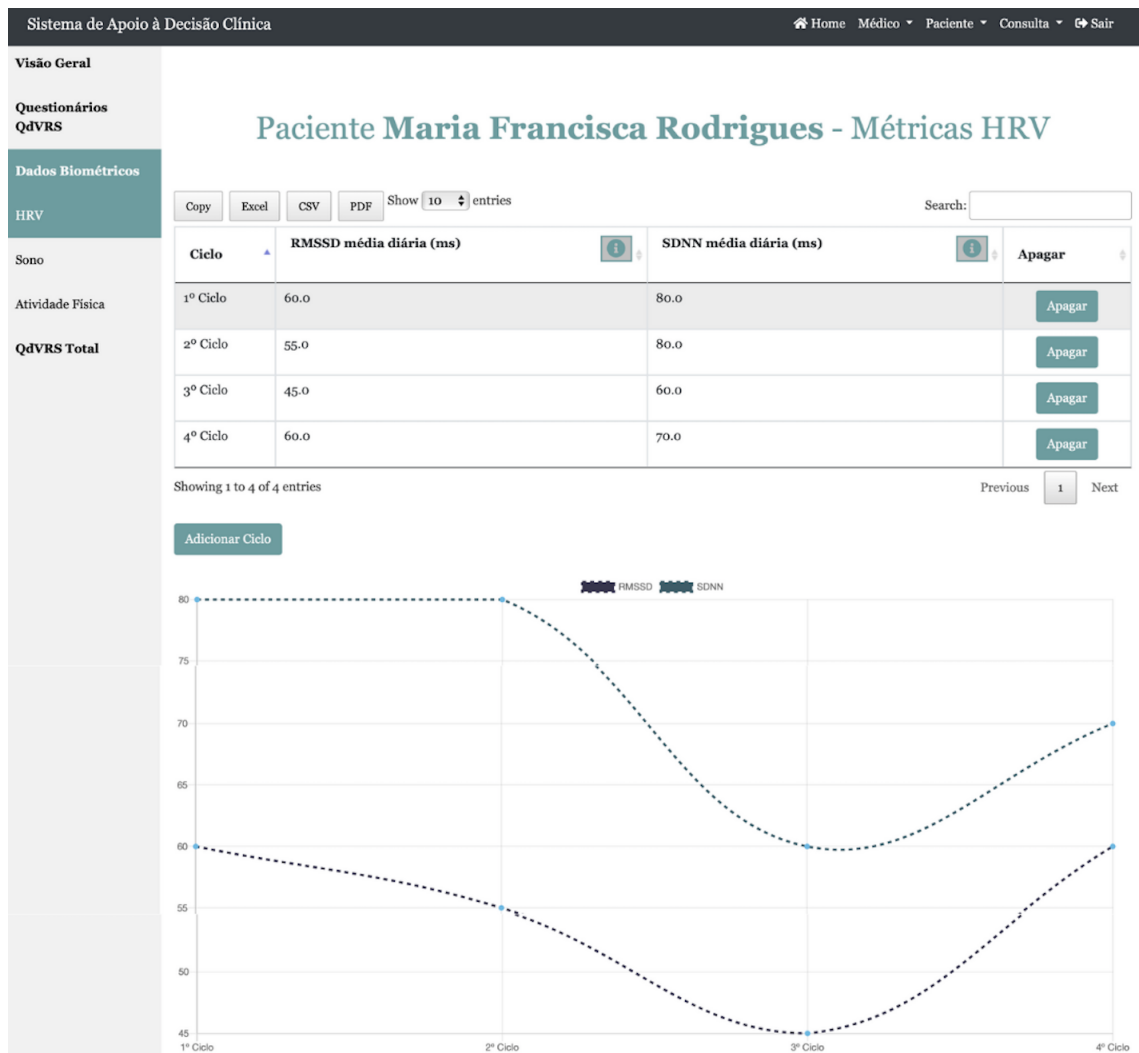
Figure 4.23: HRV metrics view inside the biometric data scope

Inside the Self-perceived HRQoL the final scores for the Fact-lym are displayed as shown in Fig.4.29.

These scores are the TOI (Trial Outcome Index) which is the sum of the physical, functional and lymphoma specific concerns, the Fact-G score which is the sum of the physical, social, emotional and functional well-being and the Fact-lym score which is the sum of the physical, social, emotional, functional and lymphoma specif concerns. For the Sf-36 questionnaire, no score is presented in this tab since the measure itself has no final scores, and each score is associated to only one dimension of the questionnaire, and those are already presented in the Sf-36 tab inside the questionnaires scope. Inside the HRQoL biometrically measured, the final score for each biometric measures in our system are presented and a chart is displayed with the HRV total score, the sleep total score and the physical activity total score (Fig. 4.30).

Each of these scores were calculated based on the variables inside each measure, for each we had an ideal value. For instance, in the case of the HRV metrics, both variables SDNN and

Figure 4.24: Modal for the SDNN variable

RMSSD have reference values. According to the literature, an individual with RMSSD and SDNN values superior to 100ms is considered an healthy individual [87]. Taking this into account, and giving a weight of 50% to each variable the total score of HRV is calculated as follows:

$$rmssd * 0,50 + sdnn * 0,50 = HRVTotalScore$$

Regarding the sleep metrics, only two variables were chosen to calculate the total score of this dimension, given that the number of times a person wakes up per night as well as their calories burnt per night can be very changeable regarding each individual, which makes it hard to set an idyllic value for such variables. Given that we decided to take into account for the total score of the sleep metrics only the duration of sleep, which has an ideal value, considering the 8 hours of sleep per night recommended by the literature, and the efficiency of sleep which is already given in percentage. This way, and also giving a weight of 50% to each variable, the total score of the sleep dimension is calculated as follows:

$$efficiency * 0,50 + \frac{duration * 100}{8} * 0,50 = SleepTotalScore$$

For the total physical activity score, the logic is similar to the previous measures, however, in this case, 3 variables were chosen: the number of steps per day, which has a recommended value by the literature, of 10 000 steps per day [88], the time of moderate activity which should be of at least half an hour per day, and the time of light activity which, taking into account the 30 minutes recommended of moderate activity and the maximum of 10 hours of sedentary activity recommended [89], should be of at least 13.5 hours. Thereby, and giving a weight of 33,3% to
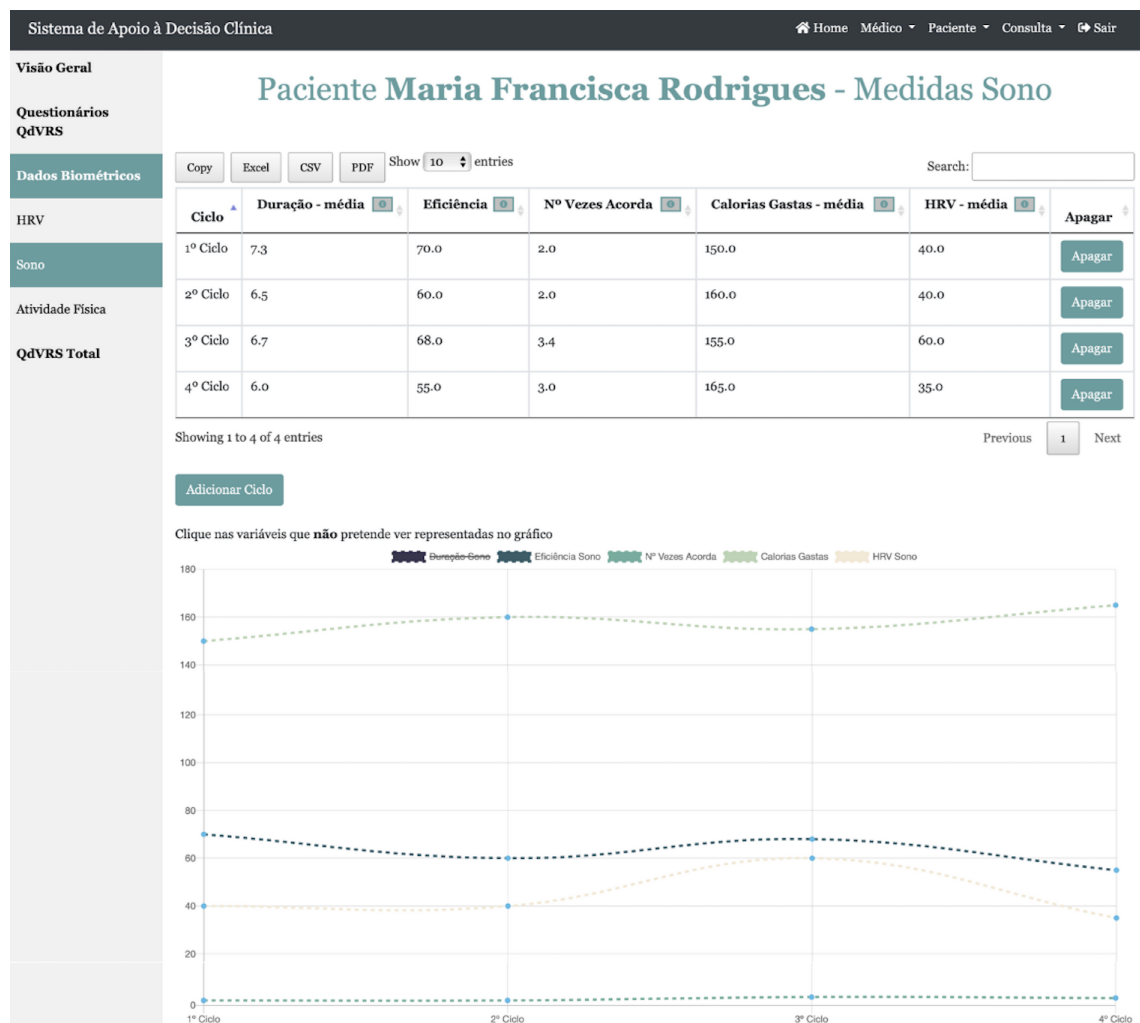
Figure 4.25: Sleep metrics view inside the biometric data scope

each variable, the total score of the physical activity dimension is calculated as follows:

$$\frac{stepsDay*100}{10000}*0,33+\frac{lightTime*100}{13.5}*0,33+\frac{moderateTime*100}{0.5}*0,33=PhysicalActivityTotalScore$$
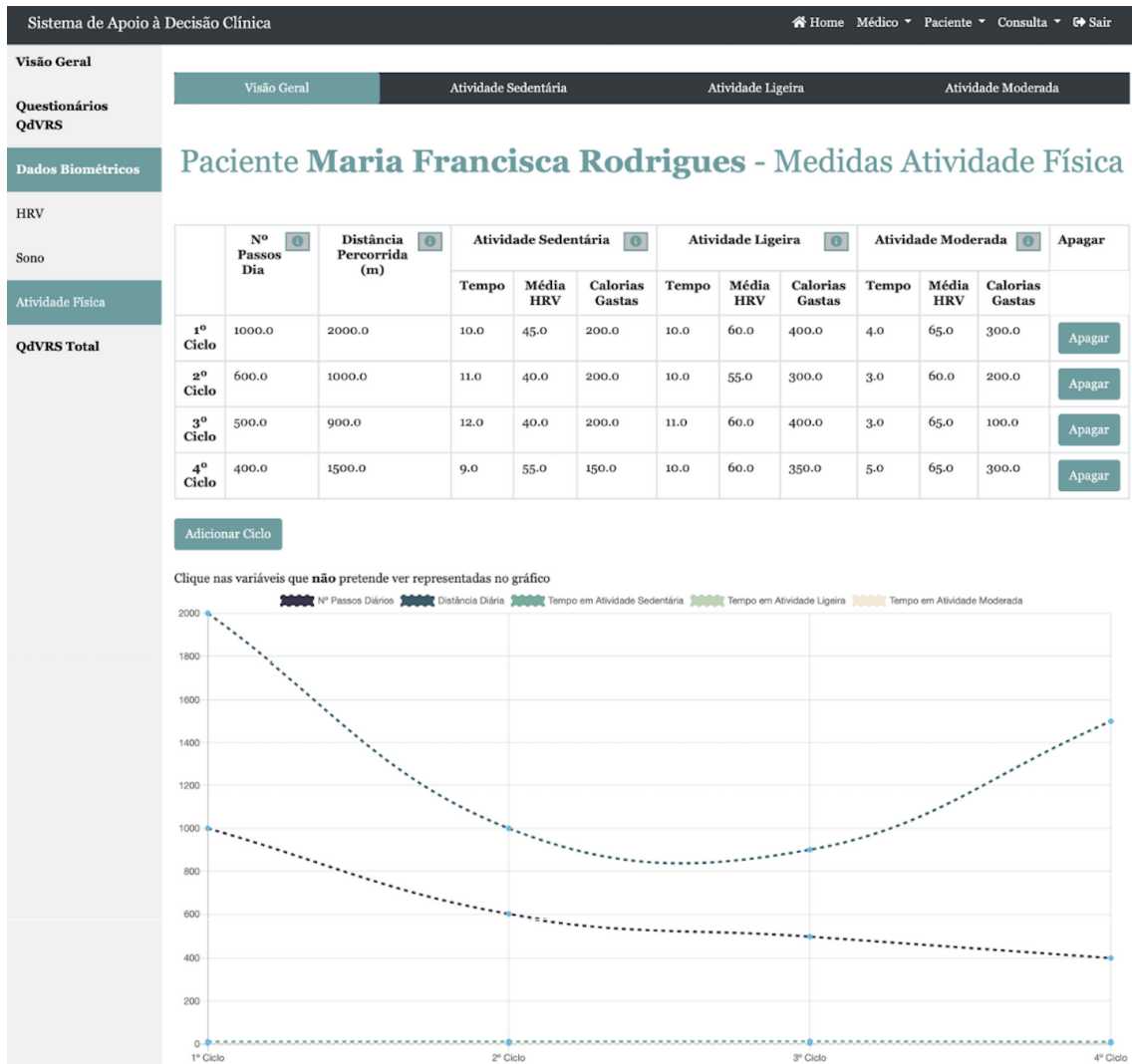
Figure 4.26: Overall physical activity metrics view inside the biometric data scope
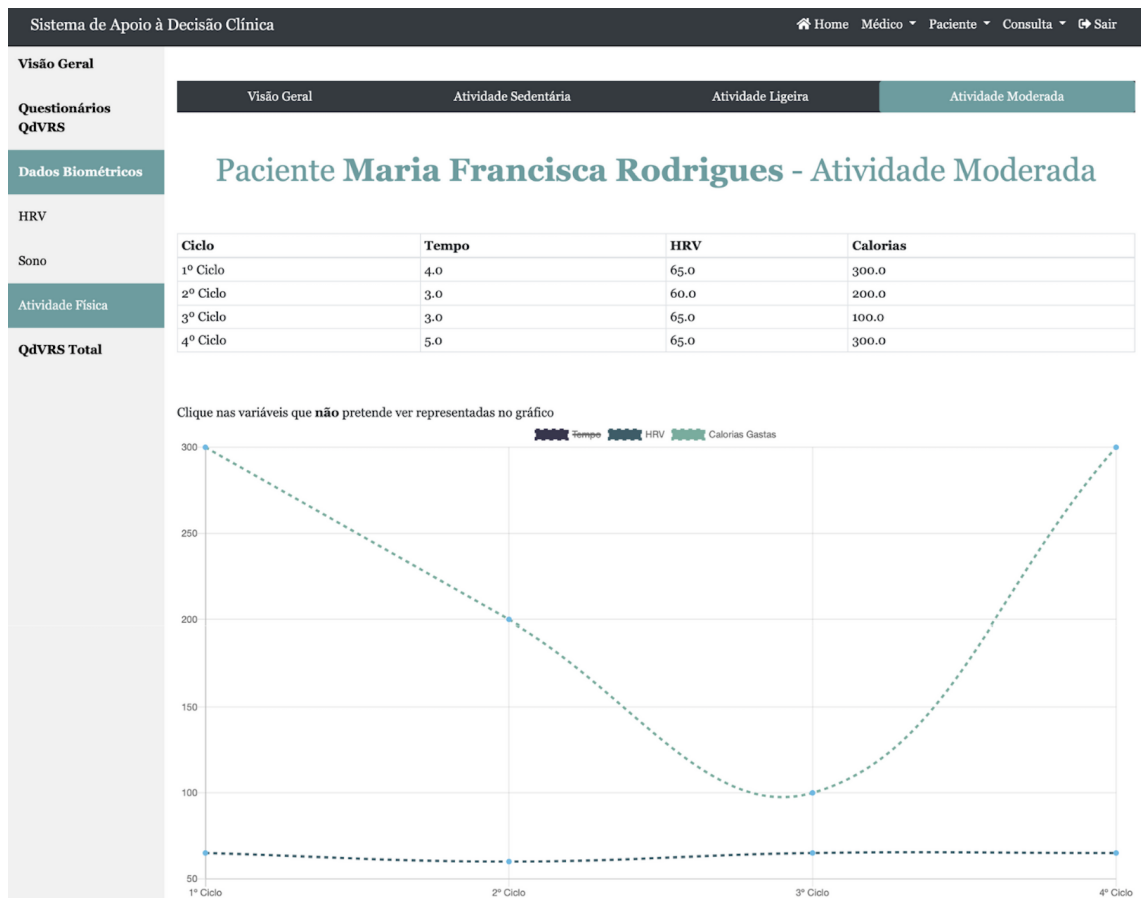
Figure 4.27: Specific type of physical activity metrics view inside the biometric data scope. In this case, moderate activity.

Figure 4.28: First view inside the Total HRQoL tab

Figure 4.29: View of the self-perceived HRQoL inside the Total HRQoL tab

Figure 4.30: View of the HRQoL biometrically measured inside the Total HRQoL tab

Chapter 5

Qvida+ CDS Validation

The Qvida+ CDS was originally intended to be tested and validated among healthcare professionals, specifically in the oncology field. However, due to the current situation of the pandemic we are facing, that was not possible to achieve. Hence, the system validation process had to be redefined. The most convenient option was to still make some functionality and usability tests among a convenience sample inside the university community, mainly with students of the Faculty of Engineering and the Faculty of Medicine of the University of Porto, even though there were also a small percentage of students from other educational institutions participating in this process. To perform this evaluation of the system, a survey was designed to mainly evaluate the usability of the system with 27 questions on that regard. Besides that, some questions were also elaborated in regard to system efficiency, effectiveness, data exposure and visual component of the CDS. [88].

5.1 Procedures

In order to able the interaction with the system, a fictional certified user was created. Some fictional patients associated to this certified user were also created, and the fields for their clinical data, biometric data and HRQoL questionnaires data were filled for illustrative purposes. The participants were mainly asked to participate in the survey by email or social media. They were asked to read a brief description of the system as well as understand its goal. Some information regarding the functionalities of each user (administrator and certified used) were given, and the credentials created for each user for this validation process were provided. At this point, the participants were invited to open the link for the CDS and interact with the system. After this, they were asked to fill the Google Form with the questionnaire mentioned previously, that can be found in Appendix A. This questionnaire is in portuguese as it was presented to all the participants. The form begins with an inform consent in order for the participants to agree with the usage of their data for this survey purpose. After this, a small set of socio-demographic questions regarding age, gender, education level and whether the participant studies in the healthcare field, are presented. Since the participants were not the target audience for this study, the following questions on this questionnaire are not so much about the system content itself, but mainly about the usability of

the system. In the last section, the participants are also asked about their personal opinion on dimensions of the system such as effectiveness, utility, data exposure, etc.

5.2   Measures

The questionnaire elaborated for the validation of our system can be divided in 3 sections as presented in table 5.1.

Table 5.1: Validation Questionnaire Dimensions

Section	Dimension	Items
Administrator Section	Usability and Functionalities	items 1 to 8
Certified User Section	Usability and Functionalities	items 9 to 27
General Opinion Section	General Usability	item 28
	General Utility	item 29
	General Effectiveness	item 30
	Data Exposure	item 31
	Visual Component	item 32

The first section concerns the usability of the system as an administrator, the second section is still about the usability of the system but as a certified user, and the third section is mainly about the general opinion of the participants regarding the general usability of the system as well as its utility, effectiveness, data privacy and visual component of the system. The first section of this questionnaire, as mentioned before, concerns the functionalities and usability of the system regarding the administrator. In this section, the participant must answer 8 questions about the functionalities and usability of the management component of the system. These questions intend to assess the participant's ease in performing certain tasks related to the administrator role. The participants answer these questions selecting the opinion that most relates to their experience. The options for each of these questions were "I have done it easily", "I have had some difficulties doing it" and "I wasn't able to do it".

The following section, concerns the functionalities and usability of the system from the certified user point of view. In this section, 19 questions are presented to the participant, 4 of them regarding the functionalities of the management component available for the certified user and 15 questions about the CDS component that only this user can access. The participant, similarly to the previous section selects the answer that resonates more with their experience, between the options "I have done it easily", "I have had some difficulties doing it" and "I wasn't able to do it".

In the last section, 5 opinion questions are presented to the participant. The first question regards their general opinion on the usability of the whole system. The options to be selected are "Easy to use", "I had some difficulties interacting with it", "I had a lot of difficulties interacting with it" and "I could not interact with it". The second question on this section aim to assess the patients' opinion on the utility of the system to the healthcare professionals. The options to be selected are "I think it is a useful resource for the health professional" and "I don't think it is a useful resource for the health professional". The following question concerns the effectiveness of the system

regarding the patients. The options to be selected in this question are "The system presents many benefits to the patients", "The system presents some benefits to the patients" and "The system doesn't present any benefits to the patients". The fourth opinion question in this section intends to comprehend how comfortable the participants would be, as patients, to expose their data on this system. The options presented for this one are "I would feel safe about my data being exposed in the system", "I have some concerns about the exposure of my data in the system" and "I would not feel safe about my data being exposed in the system". The last question on this system is about the visual component of the CDS and how was the participants' perception of the information displayed. The options to be selected in this question are "The information is presented clearly, and is easily understandable", "I had some difficulties understanding the way the information is presented" and " The information is not presented clearly and it is difficult to understand". All five questions in this opinion section, are presented along with an extra option with an open text box to express a different opinion than the ones in the options.

5.3 Results and Discussion

The sample for this survey consists of 23 participants as shown in Table 5.2.

Table 5.2: Survey Sociodemographic Data.

Sociodemographic Data		
Characteristic	**N**	**%**
Age		
21	1	4.3%
22	1	4.3%
23	13	56.6%
24	5	21,8%
25	2	8.7%
26	1	4.3%
Total	23	100%
Gender		
Female	12	52.3%
Male	10	43.4%
Other	1	4.3%
Total	23	100%
Education		
Basic Education	0	0%
Secondary Education	1	4.3%
Superior Education	22	95.7%
Other	0	0%
Total	23	100%
Works/Study in the Healthcare Field		
Yes	13	56.5%
No	10	43.5%
Total	23	100%

Their mean age is 23,4 years old with a 1,03 standard deviation, with the youngest participant being 21 years old and the oldest being 26 years old. From this sample, 52.3% are female, 43.4% are male and 4.3% from other gender. Regarding their education, 95.7% of the sample had superior education and only 4.3% had just secondary education. Finally, regarding their field of studies/work being in healthcare, 43.5% of the sample answered yes and 56.5% answered no.

Given that the first two sections of our questionnaire aim to evaluate how easily the participants can navigate and perform certain specific tasks on our system, we analysed our sample as whole. In the last section, given the opinion nature of the questions we believed it was relevant to differentiate the sample in healthcare students and other students.

5.3.1 Administrator Section Results

Regarding the administrator section the results can be found in the chart on Fig.5.1. The horizontal axis represents the percentage of participants who selected each of the 3 options presented for this section. The items of this section are represented in the vertical axis. From this analysis we can infer that, generally, the participants were very successful in the execution of these tasks, with only two items rising more concern. These concerning items are the first one, where 8,7% of the participants claimed that could not create patients at all, and the fourth item, where a relevant percentage of the participants (21,7%) claimed having some difficulties in certifying and associating each user to a doctor. This hurdle, in this last item, can be explained by the lack of knowledge/explanation in the introduction of the questionnaire regarding the association of users and doctors via username, as well as its purpose. In item 2 and 5 none of the participants claimed any difficulty, and the remaining items present minor percentages of participants having some sort of difficulty performing the corresponding task.

5.3.2 Certified User Section Results

The results regarding the certified user section can be found in the chart on Fig.5.2. The horizontal axis represents the percentage of participants who selected each of the 3 options presented for this section. The items of this section are represented in the vertical axis. Similarly to the previous section, we can infer that most participants were successful in performing these tasks with most items having none or minor percentages of participants claiming some difficulties. The item rising more concern in this section is item 10 where 8.7% of the participants weren't able to create patients and 8.7% claimed having difficulties performing the same task. After analysing the results of this questionnaire, and after realizing that both in the administrator section and in the certified user section, the creation of patients was one of the tasks where participants were having more difficulties, a problem in the system was found related to the field of the patient's medication. This field was originally created to receive until a max of 30 characters, which, in this process of validation, revealed not to be sufficient. The medication field was now updated to receive until a maximum of 100 characters.
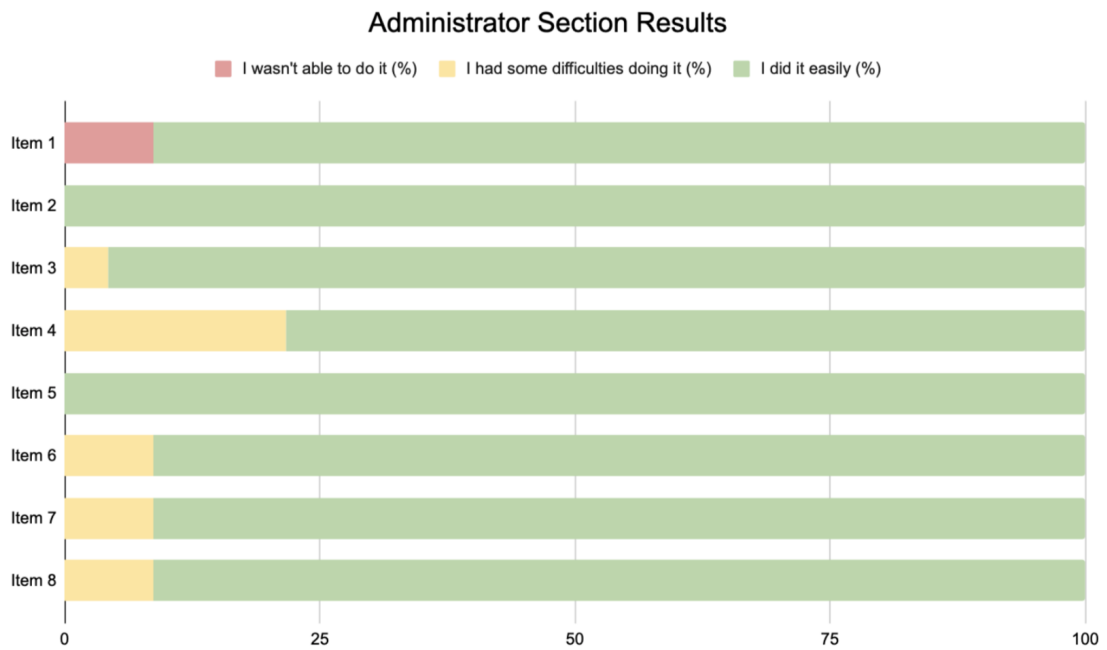
Figure 5.1: Administrator Section - Chart Results
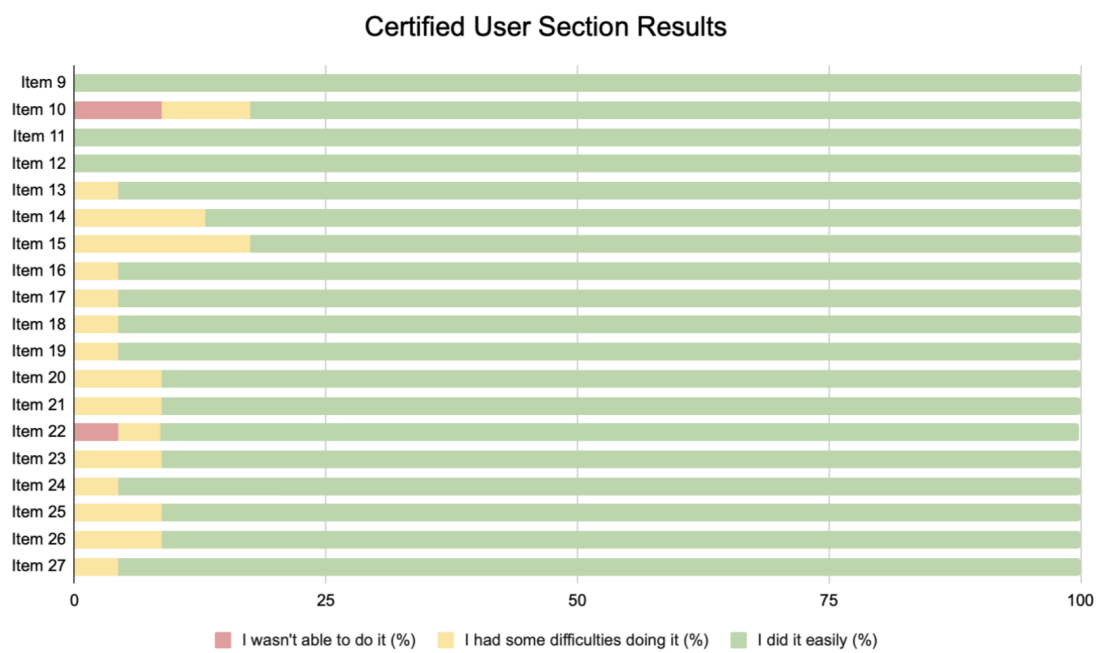
Figure 5.2: Certified User Section - Chart Results

5.3.3   Opinion Section Results

As mentioned previously, in this section, given the opinion nature of the items as well as some dimensions such as utility and effectiveness of the system, we thought relevant to analyse our

sample divided in healthcare students and others students. The first item in this section concerns the overall usability of the system and the results are presented in the chart on Fig.5.3. In the analysis performed for this item, we realized that 80% of the healthcare students participating in this survey found our system easy to use as well as 92.3% of other fields' students. Only 1 participant that weren't healthcare student had some difficulties using the system and only one healthcare student claimed having a lot of difficulties interacting with it. There was also 1 healthcare student selecting the option "Other" and in the open box for that option the participant wrote "Easy to use, I just couldn't create new patients". This problem could be related to the medication field mentioned in the previous subsection, which was solved after the analysis of this questionnaire. In general, the results have proven that our system is easy to manage and use.
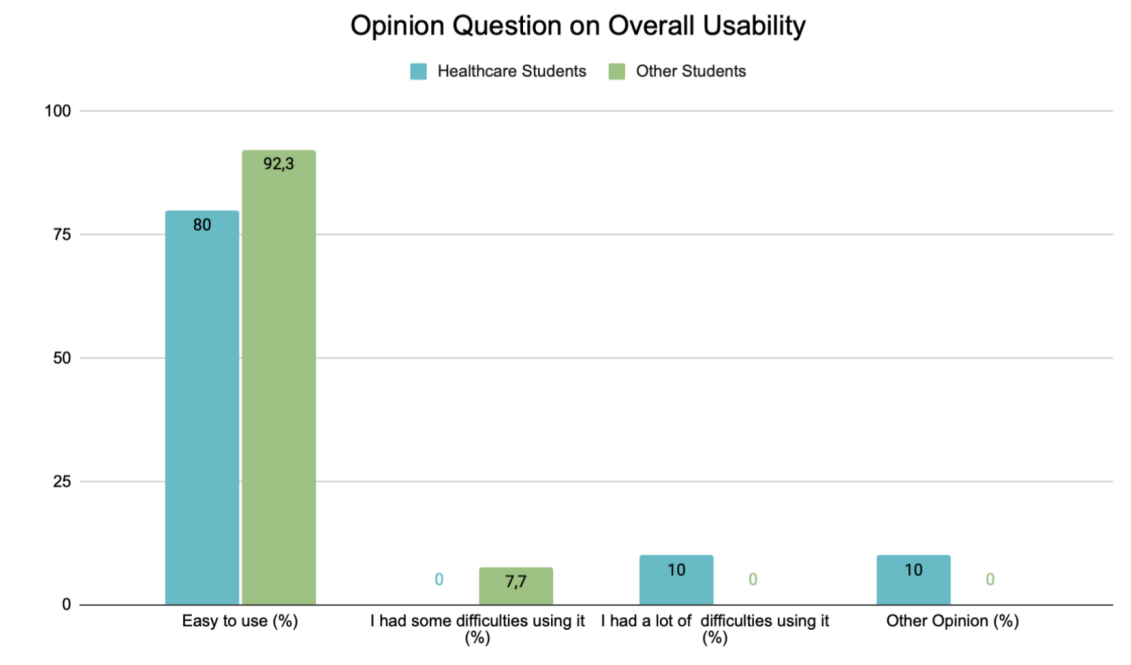
Figure 5.3: Opinion Question on Overall Usability - Chart Results

The second item in this section was about the overall utility regarding the healthcare professionals. The results for this question are presented in the chart on Fig.5.4. The opinions of the participants regarding this item were unanimous, and all participants of the survey, healthcare students and others, believed the system to be a useful resource for healthcare professionals. The results of this item analysis revealed an optimal general opinion, regarding one of our main goals which was creating an helpful resource to healthcare professionals.

The third item in this section was about the overall effectiveness regarding the benefits of the system for the patients. The results of this item are presented in the chart on Fig.5.5. The results have shown that 60% of the healthcare students, participating in this survey, believed the system presents many benefits to patients, and 40% of these students claimed the system presents just some benefits to patients. When it comes to other participants, that were not healthcare students,
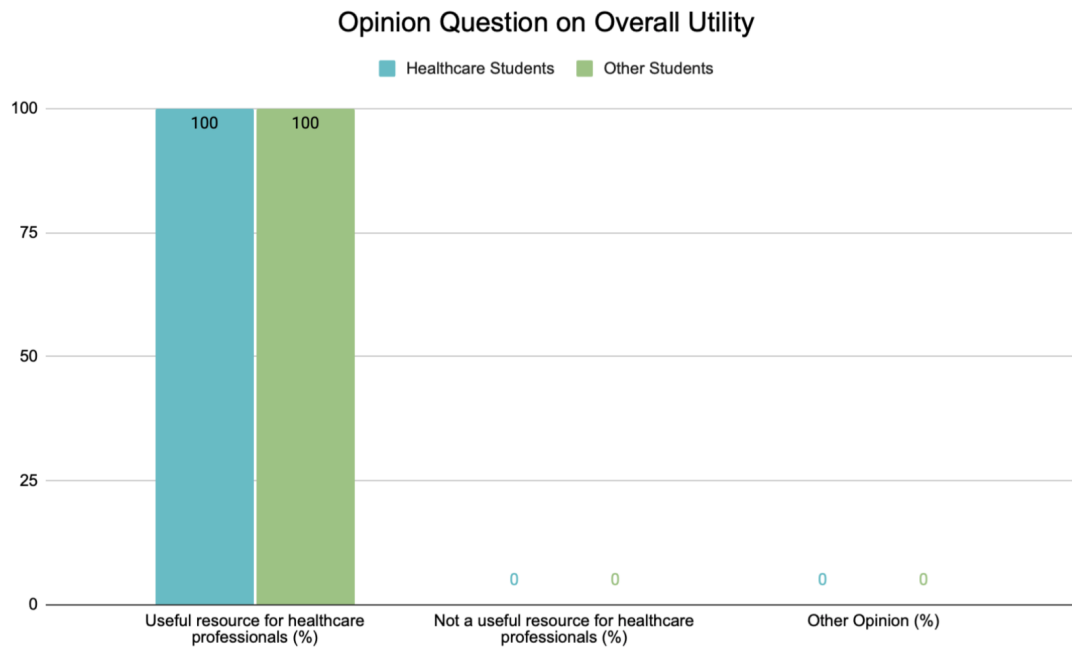
Figure 5.4: Opinion Question on Overall Utility - Chart Results

77% of them believed it presents many benefits to patients as well, and only 23% believed it presents just some benefits. None of the participants, healthcare students or not, believed that the system doesn't present benefits to patients. Through this results, although satisfactory, we believe that there is still room for improvement on this regard.

The fourth item in this section relates to the patients' data exposure in the system. The results gathered for this item are presented in the chart on Fig.5.6. In this subject, 60% of the healthcare students stated to feel comfortable to expose their own data in the system, as patients, as well as 62% of the other participants. The sample stating some concerns about such exposure was of 40% for healthcare students and 30.1% for other students. There was also a participant, not healthcare student, claiming that would not feel safe exposing his/her data on the system. These results, seem to point out that there's still room to improvements to be made regarding the privacy of the data.

Finally, the last item in this section regards the visual presentation of the CDS data. The results gathered for this last item are presented in the chart on Fig.5.7. Regarding this subject, all healthcare students claimed the information was clearly presented and easily understandable, as well as, 76.9% of the other participants, that were not healthcare students. The percentage of participants that were not healthcare students claiming some difficulties understanding the way the information was presented was 23.1%. None of the participants, healthcare students or not, selected the option stating the information is not clearly presented nor understandable. This results were pretty satisfactory given the minor percentage of participants claiming difficulties understanding the data, and taking into consideration that in this percentage, none of the participants were healthcare students. Given that, those difficulties in understanding the data might be related not necessarily to the way

Opinion Question on Overall Effectiveness
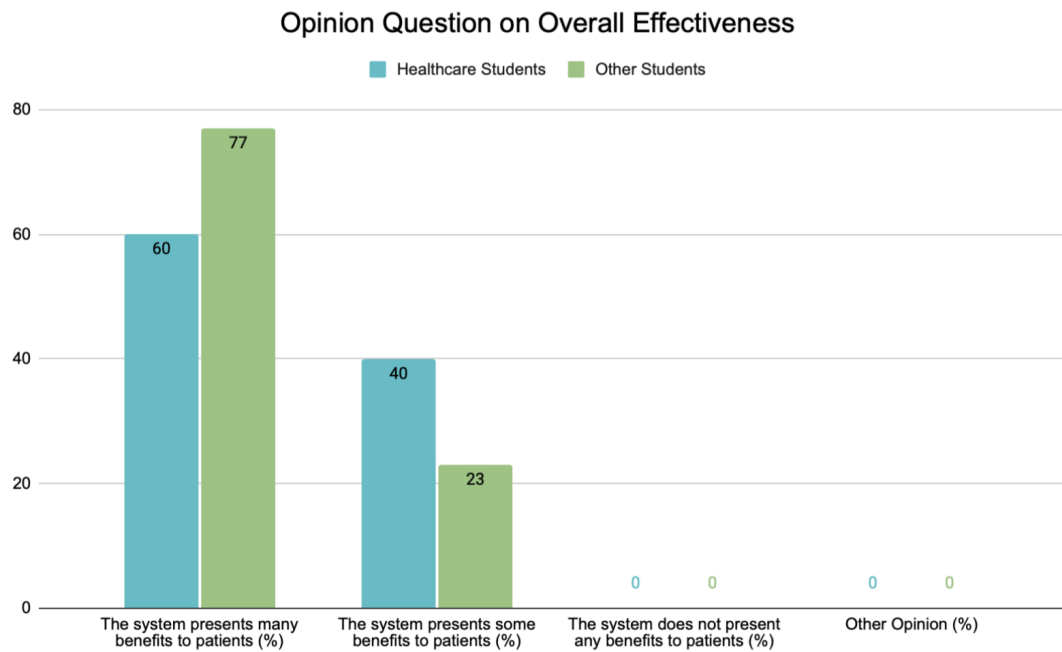
Figure 5.5: Opinion Question on Overall Effectiveness - Chart Results

Opinion Question on Data Exposure

Figure 5.6: Opinion Question on Data Exposure - Chart Results

the data is presented but maybe with the clinical character of the data content itself.
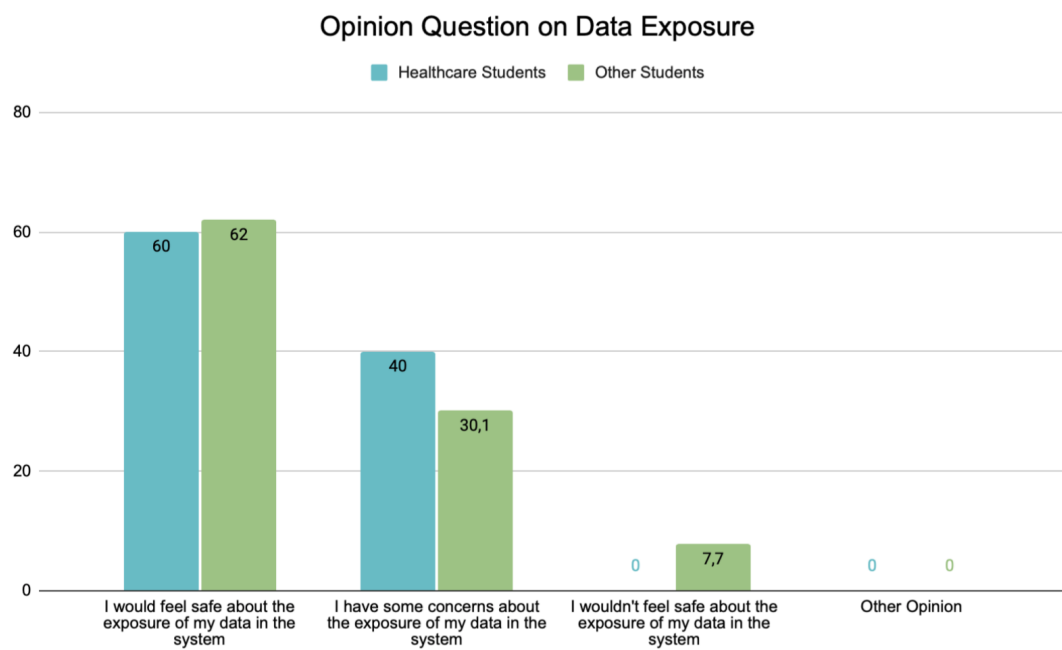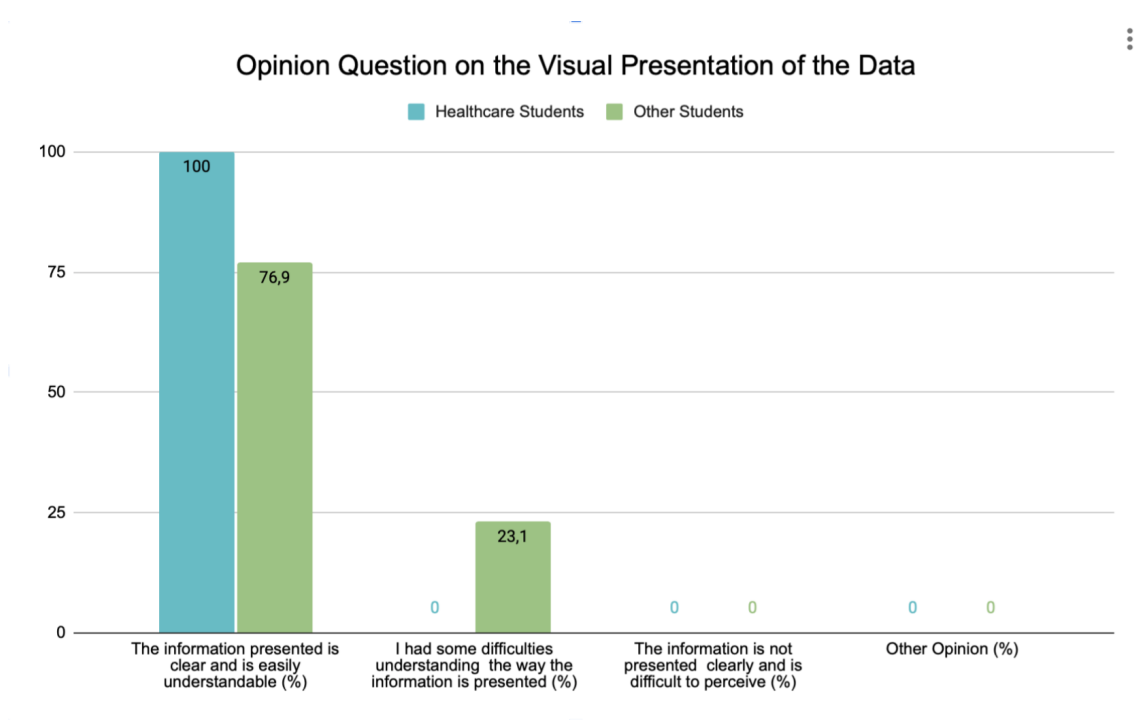
Figure 5.7: Opinion Question on the Visual Presentation of the data - Chart Results

Chapter 6

Conclusions

6.1 Summary

Even though some work have been done, in the last decades, to better assess HRQoL in patients with chronic illnesses, there is still a long way to go in this field, whether in terms of a less unobtrusive assessment, a more frequent assessment, a more objective assessment or even in the way this assessment is made available to patients and healthcare professionals. With this project, we intended to build a robust CDS system capable of doing this assessment accurately, taking into account as many variables as possible and capable of displaying such information to the healthcare professionals as a mean to better inform them about their patients' health status and help or support decisions regarding future steps on patients' treatments.

To achieve that, we tried to embody as many HRQoL variables regarding self-assessment HRQoL, with patients' answers to questionnaires, concrete physical variables with biometric data, as well as a richer set of clinical data. Given the importance of confidentiality and ethical considerations in healthcare, some measures were taken in that regard, such as, not giving access to patients' data to the administrator or to any doctor not following that specific patient, as well as making the administrator the only user capable of granting accesses to users in our system. Another important consideration taken into account was to make sure that our system would be very intuitive, visual and easy to read so that the health care professional could understand the data without much effort or without wasting to much time.

6.2 Limitations

Due to the Covid 19 pandemic, there were two main limitations to our system. At first, the idea was to integrate our CDS with Machine Learning algorithms, and some investigation was even made on that regard, which narrowed the options for that integration to some algorithms: a Bayesian network, Tree Based Model, a Support Vector Machine (SVM) and K-means clustering. Some of that research is available in Appendix B.

Although, we believed such integration would be very valuable to our system, whether to better

understand the data and its patterns, but also to make more accurate predictions and classifications regarding the patient's HRQoL and the history of similar patients in our system, the amount of data necessary to perform such tasks was never made available. These data, gathered in previous steps of the Qvida+ project, which encompassed clinical data, self-assessment data and biometric data, was difficult to get for two main reasons. The first reason was that the amount of patients with full records of the data, gathered before, might have not been enough to use Machine Learning techniques. However, the main reason to drop this integration was due to the impossibility of dealing with all the bureaucracies necessary to be able to access that specific healthcare data. In normal circumstances, this process could take several time, but in a pandemic panorama, where so many people were working from home and where other priorities came first, this goal became unrealistic. With that in mind, we thought reasonable not to get stuck in this part of the project due to the lack of data, and we felt the necessity to re-think our project to some extent and reset a new path to it, by turning it into a more analytical and visual system.

The second main limitation found during the development of this project, was its validation among healthcare professionals. The idea at the beginning was to validate our CDS, idyllically in a oncology center, with the target users of our system, the health care professionals following cancer patients. However, once again due to the pandemic panorama, such task was once again impossible to accomplish, and this validation was done with a convenience sample of college students.

6.3 Future Work

Regarding future work, we believe that the next step should be the integration of artificial intelligence in this system, so that the assessment of HRQoL of new patients could be the most accurate possible, given the data of other patients in the system. With that added knowledge, the CDS would be able to make predictions regarding the future of the patients and therefore be a more valuable tool to aid health care professionals in clinical decisions. This step would also improve the legibility of the data, once it would not only be displaying the data in charts and tables, but could also make suggestions based on the history of similar patients.

With this integration, the formulas used in the final dashboard of the total HRQoL, to calculate the final score of the HRV, sleep and physical activity dimensions, could be improved and made more accurate.

Another crucial step would be the interaction of health care professionals, specifically professionals dealing with cancer patients, so that their feedback could be taken into account, which might constitute a valuable input.

Concerning the results of the validation survey, we believe that there is still some thinking to be done regarding the patients' data privacy, in order to make them feel more comfortable about that exposure.

Appendix A

Qvida+ Decisio Support System - Validation Questionnaire

QVida+ Sistema de Apoio à Decisão Clínica - Questionário de validação

No âmbito da dissertação de mestrado "Qvida+ Desenvolvimento de um Sistema de Apoio à Decisão Clínica" do mestrado integrado em Engenharia Informática e Computação da Faculdade de Engenharia da Universidade do Porto foi desenvolvido um sistema de apoio à decisão clínica. Este sistema pretende reunir dados biométricos coletados a partir de smartbands, respostas a questionários de autorrelato de qualidade de vida relacionada com a saúde (QdVRS) e dados clínicos de pacientes oncológicos. Este sistema tem o intuito de fornecer aos profissionais de saúde uma melhor e maior quantidade de informação sobre a qualidade de vida dos seus pacientes a cada fase de tratamento, geralmente associadas a ciclos de quimioterapia, e por isso auxiliá-los em decisões clínicas futuras.

Este sistema tem dois usuários principais, o administrador e o usuário certificado, que serão os profissionais de saúde que seguem os pacientes inseridos no sistema.
O sistema pode ser definido por duas componentes principais, a componente administrativa e a componente do sistema de apoio à decisão (SAD). Relativamente à componente administrativa, o administrador tem privilégios totais sobre esta componente, podendo criar pacientes, médicos, certificar usuários e marcar consultas. O administrador não tem acesso à componente de SAD e por isso não tem acesso aos dados clínicos, dados de questionários ou dados biometricos de nenhum paciente.
O usuário certificado, relativamente à componente administrativa, pode marcar consultas, se ele for o profissional de saúde responsável pela mesma e pode criar pacientes desde que seja o profissional a seguir esse paciente. Relativamente à componente de SAD, este tem privilégios totais sobre esta componente, desde que relativamente aos seus próprios pacientes.

Link para Sistema de Apoio à decisão Qvida+:

https://deployment-qvida.herokuapp.com

Todos os dados introduzidos no sistema, neste momento, são fictícios e puramente ilustrativos.

Para utilizar o sistema como administrador, estas são as credenciais:
Username:admin
Password:admin

Para utilizar o sistema como usuário certificado (profissional de saúde), estas são as credenciais:
Username:Médico1
Password:qvida123

*Obrigatório

Embora não faça parte do público alvo, este questionário pretende avaliar a experiência de utilização do sistema. Nesse sentido, gostaria de solicitar a sua participação na presente investigação. Para isso, necessitará de aceder ao mesmo através do link referido anteriormente e interagir uns minutos com o mesmo. De seguida, gostaria de solicitar que preenchesse o seguinte questionário. A participação nesta investigação é voluntária e todos os dados recolhidos são anónimos e confidenciais, e serão usados apenas na presente investigação. Se pretender algum esclarecimento sobre este estudo, por favor contacte-me através do up201505892@fe.up.pt. Agradeço desde já a sua colaboração. *

☐   Compreendo e aceito.

A sua idade *

A sua resposta

O seu sexo

☐ Feminino

☐ Masculino

☐ Outro

☐ Outra:

Habilitações Literárias

☐ Ensino Básico

☐ Ensino Secundário

☐ Ensino Superior

☐ Outro

☐ Outra:

Trabalha na Área da Saúde? *

○ Sim

○ Não

○ Outra: _____

Após interagir com o sistema, como administrador, indique a sua experiência relativamente às funcionalidades do mesmo. *

	Não consegui	Tive algumas dificuldades	Consegui facilmente
Consegui criar pacientes	○	○	○
Conseguir criar médicos	○	○	○
Consegui criar usuários	○	○	○
Consegui certificar usuários associando-o a um médico via username	○	○	○
Consegui agendar consultas	○	○	○
Consegui visualizar e interagir com a lista de pacientes no sistema	○	○	○

	Não consegui	Tive algumas dificuldades	Consegui facilmente
Consegui visualizar e interagir com a lista de médicos no sistema	○	○	○
Consegui visualizar e interagir com a lista de consultas no sistema	○	○	○

Após interagir com o sistema, como usuário certificado, indique a sua opinião relativamente às funcionalidades do mesmo. *

	Não consegui	Tive algumas dificuldades	Consegui facilmente
Consegui agendar consultas	○	○	○
Consegui criar pacientes	○	○	○
Consegui visualizar e fazer pesquisas na lista de pacientes no sistema	○	○	○
Consegui visualizar e fazer pesquisas na lista de consultas no sistema	○	○	○
Consegui adicionar/editar e visualizar dados clínicos do paciente	○	○	○

Consegui adicionar/editar e visualizar dados relativamente a ciclos de quimioterapia do paciente	◯	◯	◯
Consegui adicionar/editar e visualizar dados relativamente a sessões de radioterapia do paciente	◯	◯	◯
Consegui adicionar/editar e visualizar dados relativos aos questionários de QdVRS	◯	◯	◯
Consegui visualizar as tabela de dados relativas a cada um dos questionários de QdVRS(Sf-36 e Fact_lym)	◯	◯	◯
Consegui visualizar e interagir com o gráfico apresentado para cada um dos questionários de QdVRS (sf-36 e fact-lym)	◯	◯	◯

Consegui visualizar a tabela e o gráfico relativos às medidas biométricas HRV	○	○	○
Consegui adicionar/editar e visualizar dados relativos às medidas biométricas HRV de cada paciente	○	○	○
Consegui visualizar informação a partir dos botões "+ info" na tabela de dados relativa às medidas biométricas HRV	○	○	○
Consegui visualizar a tabela e o gráfico relativos às medidas biométricas de Sono	○	○	○
Consegui adicionar/editar e visualizar dados relativos às medidas biométricas do sono de cada paciente	○	○	○
Consegui visualizar informação a partir dos botões "+ info" na tabela de dados das medidas biométricas do Sono	○	○	○

Consegui visualizar a tabela e o gráfico relativos às medidas biométricas da Atividade Física	○	○	○
Consegui adicionar/editar e visualizar dados relativos às medidas biométricas da atividade física de cada paciente	○	○	○
Consegui visualizar informação a partir dos botões "+ info" na tabela de dados das medidas biométricas da Atividade Física	○	○	○
Consegui visualizar e interagir com a dashboard final relativa a QdVRS total	○	○	○

Após interagir com o sistema durante algum tempo, indique a sua opinião geral relativamente à usabilidade do mesmo

○ Fácil de usar

○ Tive algumas dificuldades em interagir com o sistema

○ Tive muitas dificuldades em interagir com o sistema

○ Não consegui interagir com o sistema

○ Outra: _____

Indique a sua opinião relativamente à utilidade do sistema

○ Penso que é um recurso útil ao profissional de saúde

○ Não é um recurso útil ao profissional de saúde

○ Outra: _____

Indique a sua opinião relativamente à efetividade do sistema

○ O sistema traz muitos benefícios pacientes

○ O sistema traz alguns benefícios aos pacientes

○ O sistema não apresenta qualquer benefício aos pacientes

○ Outra: _____

Enquanto paciente, indique a sua opinião relativamente à exposição de dados no sistema

○ Sentir-me-ia seguro relativamente à exposição dos meus dados no sistema

○ Tenho alguns receios relativamente à exposição dos meus dados no sistema

○ Não me sentiria seguro relativamente à exposição dos meus dados no sistema

○ Outra: _____

Indique a sua opinião relativamente à visualização dos dados do paciente?

○ A informação é apresentada de forma clara, e é facilmente perceptível.

○ Tive algumas dificuldades em perceber a forma como a informação é apresentada

○ A informação não é apresentada de forma clara e é de difícil percepção

○ Outra: _____

Submeter

Appendix B

Machine Learning Algorithms research.

- **Bayesian Network:** this type of approach is ideal when trying to calculate the probability of a cause given some observed evidence. Most problems of these type are complex with many inter-related variables, and there might be several variables and even more potential causes. In more technical ways Bayesian networks aim to model conditional dependence, and therefore causation, by representing conditional dependence by edges in a directed graph. Through these relationships, one can efficiently conduct inference on the random variables in the graph through the use of factors. They can be used for a wide range of tasks including prediction, anomaly detection, diagnostics, automated insight, reasoning, time series prediction and decision making under uncertainty [90].

- **Tree Based Models:** Tree based algorithms empower predictive models with high accuracy, stability and ease of interpretation. Methods like decision trees, random forest, gradient boosting which are tree based models are becoming very popular among the data science community. These strategies are mostly in classification problems. The decision tree method, for instance, works for both categorical and continuous input and output variables. In this technique, the population or sample is split into two or more homogeneous sets based on the most significant difference in the input variables. These algorithms don't require any statistical knowledge to read and interpret them and its graphical representation is very intuitive. They are also very useful in data exploration since these models are one of the fastest ways to identify most significant variables and relation between two or more variables, which is also one of the goals in this stage of the project. These type of algorithms don't require as much data cleaning as some other modeling techniques since they're not influenced by outliers and missing values to a fair degree. Another big advantage, specially in this project, is that data type is not a constraint, so it can handle both numerical and categorical variables [91].

- **Support Vector Machine (SVM):** is another popular algorithm for classification challenges. In the SVM algorithm, each data item is plotted as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of

a particular coordinate. Then, the classification is made by finding the hyper-plane which is basically the main differentiator between two classes. In other words, it means that given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier [92].

- **K-means clustering:** K-means clustering tries to partition the dataset into K pre-defined distinct non-overlapping subsets which are the clusters, where each data point belongs to only one group. It tries to make the intra-cluster data points as similar as possible while also keeping the clusters as far as possible. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid (arithmetic mean of all the data points that belong to that cluster) is at the minimum. The less variation within the clusters, the more homogeneous the data points are within the same cluster [93].

References

[1] L. P. Reis, B. M. Faria, J. Gonçalves, Rocha, and V. Carvalho, "A survey on clinical decision support systems concerning quality of life," pp. 1–8, 2016.

[2] M. S. Litwin, "Health-related quality of life," pp. 237–251, 2007.

[3] C. Qvida+, "Estimação contínua da qualidade de vida para auxílio eficaz à decisão clínica." http://qvida.pt, 2020.

[4] P. Theofilou, "Quality of life: Definition and measurement," *Europe's Journal of Psychology*, vol. 9, pp. 150–162, 2013.

[5] J. Pais Ribeiro, C. Pinto, and C. Santos, "Validation study of the portuguese version of the qlc-c30-v.3," vol. 9, pp. 89 – 102, 00 2008.

[6] M. T. King, M. R. Stockler, D. F. Cella, D. Osoba, D. T. Eton, J. Thompson, and A. R. Eisenstein, "Meta-analysis provides evidence-based effect sizes for a cancer-specific quality-of-life questionnaire, the fact-g," *Journal of Clinical Epidemiology*, vol. 63, no. 3, pp. 270 – 281, 2010.

[7] B. Schouten, J. Hellings, E. Hoof, P. Vankrunkelsven, P. Bulens, F. Buntinx, J. Mebis, D. Vandijck, and W. Schrooten, "Validation of the flemish cares, a quality of life and needs assessment tool for cancer care," *BMC Cancer*, vol. 16, 2016.

[8] C.-H. Chang, D. Cella, G. A. Masters, N. Laliberte, P. O'Brien, A. Peterman, and D. Shevrin, "Real-time clinical application of quality-of-life assessment in advanced lung cancer," *Clinical Lung Cancer*, vol. 4, no. 2, pp. 104 – 109, 2002.

[9] A. S. Dias, "Modelos de previsão de qualidade de vida através de sensores biométricos," 2017.

[10] E. Silva, J. Aguiar, A. Oliveira, B. M. Faria, L. P. Reis, V. Carvalho, J. Gonçalves, and J. Oliveira e Sá, "An approach to assess quality of life through biometric monitoring in cancer patients," pp. 607–617, 2019.

[11] K. C. Calman, "Quality of life in cancer patients–an hypothesis.," *Journal of medical ethics*, vol. 10 3, pp. 124–7, 1984.

[12] World Health Organization, "Constitution." https://www.who.int/about/who-we-are/constitution, 2020.

[13] X.-J. Lin, I.-M. Lin, and S.-Y. Fan, "Methodological issues in measuring health-related quality of life," *Tzu Chi Medical Journal*, vol. 25, pp. 8–12, 2013.

[14] R. Hays, C. Sherbourne, and R. Mazel, "The rand 36-item health survey 1.0," *Health economics*, vol. 2, pp. 217–27, 10 1993.

[15] P. Ferreira, "Creation of portuguese version of the mos sf-36 part i, cultural and linguistic adaptation," *Acta médica portuguesa*, vol. 13, pp. 55–66, 01 2000.

[16] W. Sieber, E. Groessl, D. Ph, K. Kistler, T. Ganiats, and R. Kaplan, "Quality of well being self-administered (qwb-sa) scale," 01 2008.

[17] B. M. . G. Pollard, Bobbitt, "The sickness impact profile: Reliability of a health status measure," vol. 14, pp. 146–155, 1976.

[18] K. Demet, F. Guillemin, N. Martinet, and J. André, "Nottingham health profile: Reliability in a sample of 542 subjects with major amputation of one or several limbs," *Prosthetics and orthotics international*, vol. 26, pp. 120–3, 2002.

[19] C. Nogueira-Silva, P. Costa, C. Martins, S. Barata, C. Alho, C. Calhaz-Jorge, and F. Osório, "Validation of the portuguese version of ehp-30 (the endometriosis health profile-30)," *Acta Médica Portuguesa*, vol. 28, 2015.

[20] D. McNair, L. Droppleman, M. Lorr, Educational, and I. T. Service, "Profile of mood states (poms)," 1971.

[21] S. Shacham, "A shortened version of the profile of mood states," *Journal of Personality Assessment*, vol. 47, no. 3, pp. 305–306, 1983.

[22] M. F. Viana, P. L. Almeida, and R. C. Santos, "Adaptação portuguesa da versão reduzida do perfil de estados de humor - poms," 2012.

[23] W. Group, "The world health organization quality of life assessment (whoqol)," *Social Science Medicine*, vol. 46, no. 12, pp. 1569 – 1585, 1998.

[24] M. R. M. P. C. S. L. B. A. R. P. L. Vilar, Manuela Simões, "Adaptação e validação do whoqol-old para a população portuguesa: Resultados da implementação de grupos focais," *Revista Iberoamericana de Diagnóstico y Evaluación - e Avaliação Psicológica*, 2014.

[25] R. Brooks, "Euroqol: the current state of play," *Health Policy*, vol. 37, no. 1, pp. 53 – 72, 1996.

[26] P. Ferreira, L. Ferreira, and L. Pereira, "[contribution for the validation of the portuguese version of eq-5d].," *Acta médica portuguesa*, vol. 26, pp. 664–75, 11 2013.

[27] EORTC, "European organisation for research and treatment of cancer." https://www.eortc.org, 2020.

[28] W.-C. Chie, K.-J. Chang, C. F. Huang, and W.-H. Kuo, "Quality of life of breast cancer patients in taiwan: validation of the taiwan chinese version of the eortc qlq-c30 and eortc qlq-br23," *Psycho-oncology*, vol. 12, pp. 729–35, 2003.

[29] B. Bergman, N. Aaronson, S. Ahmedzai, S. Kaasa, and M. Sullivan, "The eortc qlq-lc13: a modular supplement to the eortc core quality of life questionnaire (qlq-c30) for use in lung cancer clinical trials," *European Journal of Cancer*, vol. 30, no. 5, pp. 635 – 642, 1994.

[30] G. Borghede and M. Sullivan, "Measurement of quality of life in localized prostatic cancer patients treated with radiotherapy. development of a prostate cancer-specific module supplementing the eortc qlq-c30," *Quality of Life Research*, vol. 5, no. 2, pp. 212–222, 1996.

[31] A. C. Sherman, S. Simonton, D. C. Adams, E. Vural, B. Owens, and E. Hanna, "Assessing Quality of Life in Patients With Head and Neck Cancer: Cross-validation of the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Head and Neck Module (QLQ-H
amp;N35)," *Archives of Otolaryngology–Head Neck Surgery*, vol. 126, pp. 459–467, 04 2000.

[32] H. Colwell, S. Mathias, M. Turner, J. Lu, N. Wright, M. Peeters, D. Cella, and G. Devercelli, "Psychometric evaluation of the fact colorectal cancer symptom index (fcsi-9): Reliability, validity, responsiveness, and clinical meaningfulness," *The oncologist*, vol. 15, pp. 308–16, 02 2010.

[33] M. J. Brady, D. F. Cella, F. Mo, A. E. Bonomi, D. S. Tulsky, S. R. Lloyd, S. Deasy, M. Cobleigh, and G. Shiomoto, "Reliability and validity of the functional assessment of cancer therapy-breast quality-of-life instrument.," *Journal of Clinical Oncology*, vol. 15, no. 3, pp. 974–986, 1997.

[34] J. Hong, S. Jeon, H. Lee, Y. Choi, S. Kim, and H. Choi, "The functional assessment of cancer therapy-prostate (fact-p) scales in men with prostate cancer: Reliability and validity of the korean version," *Journal of Korean medical science*, vol. 21, pp. 295–9, 05 2006.

[35] K. Basen-Engquist, D. Bodurka-Bevers, M. A. Fitzgerald, K. Webster, D. Cella, S. Hu, and D. M. Gershenson, "Reliability and validity of the functional assessment of cancer therapy–ovarian," *Journal of Clinical Oncology*, vol. 19, no. 6, pp. 1809–1817, 2001.

[36] D. Cella, K. Webster, J. Cashy, L. Kutikova, M. Burgess, B. Lin, L. Bowman, A. Liepa, J. Gauthier, S. Gregory, S. Johnson, and B. Cheson, "Development of a measure of health-related quality of life for non-hodgkin's lymphoma clinical research: The functional assessment of cancer therapy - lymphoma (fact-lym).," *Blood*, vol. 106, pp. 750–750, 11 2005.

[37] M. Litwin, R. Hays, A. Fink, P. Ganz, B. Leake, and R. Brook, "The ucla prostate cancer index: Development, reliability, and validity of a health-related quality of life measure," *Medical care*, vol. 36, pp. 1002–12, 08 1998.

[38] P. Nogueira, J. Urbano, L. Reis, H. Lopes Cardoso, D. Silva, A. Rocha, J. Gonçalves, and B. Faria, "A review of commercial and medical-grade physiological monitoring devices for biofeedback-assisted quality of life improvement studies," *Journal of Medical Systems*, vol. 42, 2018.

[39] C. Chow, "DeepBUFS: Deep Learned Biometric User Feedback System," pp. 150–154, 2017.

[40] M. Pelc, Y. Khoma, and V. Khoma, "ECG Signal as Robust and Reliable Biometric Marker: Datasets and Algorithms Comparison," *Sensors*, vol. 19, p. 2350, 2019.

[41] Web MD, "How is electromyogram (emg) used in biofeedback therapy?." https://www.webmd.com/a-to-z-guides/qa/how-is-electromyogram-emg-used-in-biofeedback-therapy, 2020.

[42] M. Conde Pastor, F. Javier Menéndez, M. T. Sanz, and E. Vila Abad, "The Influence of Respiration on Biofeedback Techniques," *Applied Psychophysiology and Biofeedback*, vol. 33, no. 1, pp. 49–54, 2008.

[43] F. J. Keefe and E. T. Gardner, "Learned control of skin temperature: Effects of short- and long-term biofeedback training," *Behavior Therapy*, vol. 10, no. 2, pp. 202–210, 1979.

[44] M. Conde Pastor, F. Javier Menéndez, M. T. Sanz, and E. Vila Abad, "The Influence of Respiration on Biofeedback Techniques," *Applied Psychophysiology and Biofeedback*, vol. 33, no. 1, pp. 49–54, 2008.

[45] D. Bassett, L. Toth, S. Lamunion, and S. Crouter, "Step counting: A review of measurement considerations and health-related applications," *Sports medicine (Auckland, N.Z.)*, vol. 47, 2016.

[46] Health Line Media, "Pulse oximetry, purpose and uses." `https://www.healthline.com/health/pulse-oximetry`, 2020.

[47] R. Bourne, C. Minelli, G. Mills, and R. Kandler, "Clinical review: Sleep measurement in critical care patients: Research and clinical implications," *Critical care (London, England)*, vol. 11, p. 226, 2007.

[48] Sentio Solutions Inc., "Biomarkers digital therapeutics for mental health." `https://www.myfeel.co`, 2020.

[49] Wearable LTD, "Zenta is a biometric bracelet that wants to track your emotions." `https://www.wareable.com/health-and-wellbeing/zenta-vinaya-specs-price-features-release-date`, 2020.

[50] TechRadar, Future Us Inc, "Microsoft band 2 review curves and sensors in all the right places." `https://www.techradar.com/reviews/wearables/microsoft-band-2-1306006/review`, 2016.

[51] TechRadar, Future Us Inc, "Jawbone up3 review future-proofing the fitness tracker?." `https://www.techradar.com/reviews/wearables/jawbone-up3-review-1297589/review`, 2018.

[52] Apple Inc, "The #1 smartwatch in the world. times two.." `https://www.apple.com/pt/watch/`, 2020.

[53] Apple Inc., "Create apps for watchos." `https://developer.apple.com/watchos/`, 2020.

[54] Samsung, "Galaxy watch bluetooth 46mm." `https://www.samsung.com/pt/wearables/galaxy-watch-r800/`, 2020.

[55] Samsung Developers, "Develop apps for galaxy watch." `https://developer.samsung.com/galaxy-watch-develop`, 2020.

[56] Google LLC, "Wear os by google." `https://developer.android.com/wear`, 2020.

[57] Fossil Group, Inc., "Sport smartwatches." `https://www.fossil.com/en-us/smartwatches/learn-more/sport/`, 2020.

[58] Huawei Technologies Co., "Huawei watch gt kv." https://consumer.huawei.com/pt/wearables/watch-gt/, 2020.

[59] plux | wireless biosignals, "Best industrial enabling tech 2017 in europe by the dg-connect." https://bitalino.com, 2020.

[60] P. Pawloski, G. Brooks, M. Nielsen, and B. Olson-Bullis, "A systematic review of clinical decision support systems for clinical oncology practice," *JNCCN Journal of the National Comprehensive Cancer Network*, vol. 17, pp. 331–338, 2019.

[61] HealthIT.gov, "Clinical decision support." https://www.healthit.gov/topic/safety/clinical-decision-support, 2020.

[62] J. Bury, C. Hurt, A. Roy, L. Cheesman, M. Bradburn, S. Cross, J. Fox, and V. Saha, "Lisa: A web-based decision-support system for trial management of childhood acute lymphoblastic leukaemia," *British journal of haematology*, vol. 129, pp. 746–54, 2005.

[63] K. Exarchos, Y. Goletsis, and D. Fotiadis, "Multiparametric decision support system for the prediction of oral cancer reoccurrence," *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 16, 08 2011.

[64] T. Bright, A. Wong, R. Dhurjati, E. Bristow, R. Coeytaux, G. Samsa, V. Hasselblad, J. Williams, M. Musty, L. Wing, A. Kendrick, G. Sanders, and D. Lobach, "Effect of clinical decision-support systems: A systematic review," *Annals of internal medicine*, vol. 157, 2012.

[65] B. Faria, J. Gonçalves, Rocha, and L. Reis, "A clinical support system based on quality of life estimation," *Journal of Medical Systems*, vol. 39, pp. 1–11, 08 2015.

[66] P. Cunningham and S. Delany, "K-Nearest Neighbour Classifiers," *Mult Classif Syst*, 2007.

[67] P. Kaviani and S. Dhotre, "Short Survey on Naive Bayes Algorithm," *International Journal of Advance Research in Computer Science and Management*, vol. 04, 2017.

[68] L. Rokach and O. Maimon, *Decision Trees*, vol. 6, pp. 165–192. 2005.

[69] X. Zhou, Y. Ma, Z. Cheng, L. Liu, and J. Wang, *A simplification on SMO algorithm and its application in solving -SVR with non-positive Kernels*, pp. 878–883. 2010.

[70] Ecog-Acrin, "Ecog performance status." https://ecog-acrin.org/resources/ecog-performance-status, 2020.

[71] K. Webster, D. Cella, and K. Yost, "The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation," *Health and quality of life outcomes*, vol. 1, p. 79, 2003.

[72] "Ethical sharing of health data in online platforms – which values should be considered?," 2017.

[73] Django Software Foundation, "Django makes it easier to build better web apps more quickly and with less code.." https://www.djangoproject.com, 2020.

[74] Python Software Foundation, "Python is a programming language that lets you work quickly and integrate systems more effectively." https://www.python.org, 2020.

[75] STEELKIWI INC, "Why is the mvt architecture important for any django web application?." https://steelkiwi.com/blog/why-django-best-web-framework-your-project/, 2021.

[76] SQLite Consortium, "What is sqlite?." https://www.sqlite.org/index.html, 2020.

[77] Django Software Foundation, "User authentication in django." https://docs.djangoproject.com/en/3.1/topics/auth/, 2021.

[78] Django Software Foundation, "The django admin site." https://docs.djangoproject.com/en/3.1/ref/contrib/admin/, 2021.

[79] Django Software Foundation, "How to use sessions." https://docs.djangoproject.com/en/3.1/topics/http/sessions/, 2021.

[80] Python Software Foundation, "django-adaptors 0.2.5." https://pypi.org/project/django-adaptors/, 2021.

[81] Python Software Foundation, "django-csvimport 2.16." https://pypi.org/project/django-csvimport/, 2021.

[82] Michael Trier, Bas van Oostveen and contributors, "Welcome to the django-extensions documentation!." https://django-extensions.readthedocs.io/en/latest/, 2021.

[83] Python Software Foundation, "pygraphviz 1.7." https://pypi.org/project/pygraphviz/, 2021.

[84] Python Software Foundation, "pydot 1.4.1." https://pypi.org/project/pydot/, 2021.

[85] Chart.js Org, "Chart.js." https:https://www.chartjs.org/docs/latest/, 2021.

[86] Pluralsight, "25 years of javascript." javascript.com, 2021.

[87] F. Shaffer and J. Ginsberg, "An Overview of Heart Rate Variability Metrics and Norms," *Frontiers in Public Health*, vol. 5, p. 258, 2017.

[88] C. M. C. M. Leandro, Gabrielle dos Santos; Moro, "MÉTODOS PARA AVALIAÇÃO DE SISTEMAS DE APOIO À DECISÃO," *XV Congresso Brasileiro de Informática em Saúde*, p. 9, 2016.

[89] Healthline Media, "How many steps do i need a day?." https://www.healthline.com/health/how-many-steps-a-day, 2021.

[90] K. Iqbal, Y. Xu, H.-W. Hao, Q. Ilyas, and H. Ali, "An overview of bayesian network applications in uncertain domains," *International Journal of Computer Theory and Engineering*, vol. 7, pp. 416–427, 2015.

[91] A. Cutler, D. Cutler, and J. Stevens, "Tree-based methods," pp. 1–19, 2008.

[92] S. Salcedo-Sanz, J. L. Rojo-Álvarez, M. Martínez-Ramón, and G. Camps-Valls, "Support vector machines in engineering: an overview," *WIREs Data Mining and Knowledge Discovery*, vol. 4, no. 3, pp. 234–267, 2014.

[93] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, 2002.