

**SleepData - Sleep Disorders Clinical Platform**  
Delayed Sleep Phase Disorder

**Bárbara Luís Sequeira de Pinho**

Thesis to obtain the Master of Science Degree in

**Biomedical Engineering**

Supervisors: Prof. Doutor Mário Jorge Costa Gaspar da Silva

Prof. Doutor Bruno Emanuel da Graça Martins

Prof. Doutora Maria Teresa Aguiar Santos Paiva

**Examination Committee**

Chairperson: Prof. Doutora Ana Luísa Nobre Fred

Supervisor: Prof. Doutor Mário Jorge Costa Gaspar da Silva

Members of the Committee: Prof. Doutor Alberto Abad Gareta

**OCTOBER 2017**

# Agradecimentos

Quero agradecer a todos que, de alguma forma, foram essenciais ao desenvolvimento desta dissertação e que me acompanharam neste percurso.

Aos meus orientadores, o Professor Mário Silva, o Professor Bruno Martins e a Professora Teresa Paiva, por toda a disponibilidade, orientação e apoio fundamentais. Ao Instituto Superior Técnico, pela formação contínua nestes cinco anos, que me permitiu desenvolver os conhecimentos e capacidades necessárias a completar esta etapa do meu percurso académico.

Ao meu colega Tiago Castanheira, com quem desenvolvi o SleepData, por toda a dedicação a este projecto, discussão de ideias e pelo companheirismo diário, que tornou este percurso muito mais agradável.

Ao Centro de Electroencefalografia e Neurofisiologia Clínica Lda., por nos ceder os dados clínicos, e aos seus funcionários, que nos receberam e ajudaram a perceber o potencial do nosso trabalho e como o SleepData poderia ser moldado para se ajustar às necessidades diárias da clínica.

À Cátia Reis, pela constante disponibilidade, companheirismo, apoio e paciência, por me ter ajudado a perceber os contornos da Síndrome do atraso das fases do sono e esclarecido todas as dúvidas que foram surgindo, relativamente aos dados clínicos e à caracterização da síndrome.

A todos os meus amigos e colegas, pelo apoio durante esta etapa e pela partilha de todos os momentos que fazem dos anos de estudante únicos. Um agradecimento especial à Cláudia Silva, por ser um suporte constante, pela amizade e companheirismo que tornaram estes anos inesquecíveis, e por me motivar todos os dias a chegar mais longe. Ao Steve W., por me mostrar que todos os dias são uma aventura.

Aos meus pais, por todo o apoio, por garantirem que tenho sempre as melhores oportunidades para crescer, por serem o melhor exemplo de esforço e dedicação e por serem sempre os primeiros a acreditarem no meu sucesso. Ao meu irmão João, pelo apoio diário, pela amizade, pela paciência e por cuidar de mim todos os dias. À minha irmã Alexandra, por me dar motivos para querer ser o melhor exemplo. Devo a vós ter chegado até aqui.

# Resumo

Esta dissertação apresenta o design e desenvolvimento do SleepData, uma nova plataforma de informação para a gestão de dados clínicos de distúrbios do sono. A plataforma integra dados de múltiplas fontes produzidas por uma diversidade de dispositivos de monitorização e testes laboratoriais. O SleepData fornece ferramentas para análise estatística e ferramentas de suporte ao diagnóstico, especialmente concebidas para estudar pacientes com síndrome do atraso das fases do sono (SAFS) e insónia. A plataforma foi projetada para cumprir os princípios de dados FAIR para gestão de dados científicos, adotando o FHIR, um *standard* para a troca de informações de saúde, e o SNOMED CT e o LOINC, duas nomenclaturas médicas. Os dados estão protegidos em conformidade com as recomendações da Comissão Nacional de Protecção de Dados. Um estudo inicial sobre doentes com SAFS, com base nos dados de uma clínica de ponta integrados no SleepData, mostra que esta população apresenta um atraso das fases do sono de 1-4 horas, diversas comorbidades e fatores de *stress* comuns. Os resultados apontam para a hipótese das causas da SAFS neste conjunto de pacientes serem comportamentais e não fisiológicas, uma vez que o atraso das fases do sono é reduzido quando os doentes se encontram em ambientes com estímulos reduzidos. Além disso, os resultados do início da produção de melatonina em luz difusa, um marcador conhecido para SAFS, não são consistentes com a síndrome.

# Abstract

This dissertation presents the design and development of SleepData, a new information platform for managing clinical information on sleep disorders. The platform can integrate data from multiple sources produced by a diversity of monitoring devices and lab tests. SleepData provides tools for statistical analysis and diagnostic support tools specially designed to study patients with delayed sleep phase disorder (DSPD) and insomnia. The platform is designed to comply with the FAIR data principles for scientific data management, adopting FHIR, a standard for healthcare information exchange, and SNOMED CT and LOINC, two medical nomenclatures. Information is secured in compliance with the Portuguese Data Protection Authority recommendations. An initial study of patients with DSPD, based on data integrated into SleepData from a state of the art clinic, shows that this population presents a delay in phase of 1-4 hours, several comorbidities and common stress factors. The results point to the causes for DSPD in this set of patients being behavioural rather than physiological, as the delay is reduced when the surroundings have reduced stimuli. Additionally, the results of dim-light melatonin onset, a known marker for DSPD, are not consistent with the disorder.

# Palavras Chave

## Keywords

### **Palavras Chave**

Base de dados clínicos do sono

Sistemas de informação clínica

Síndrome do atraso das fases do sono

### **Keywords**

Clinical sleep database

Clinical information systems

Delayed sleep phase disorder

# Acronyms

<b>AASM</b>	American Academy of Sleep Medicine
<b>API</b>	Application Programming Interface
<b>CENC</b>	Centro de Electroencefalografia e Neurofisiologia Clínica
<b>CNPD</b>	Comissão Nacional de Protecção de Dados
<b>DBMS</b>	Database Management System
<b>DLMO</b>	Dim-light melatonin onset
<b>DOCX</b>	Microsoft Word Open XML Format Document file
<b>DSPD</b>	Delayed Sleep Phase Disorder
<b>EEG</b>	Electroencephalogram
<b>EMG</b>	Electromyogram
<b>EOG</b>	Electrooculogram
<b>ESS</b>	Epworth Sleepiness Scale
<b>FAIR</b>	Findable, accessible, interoperable and reusable
<b>FHIR</b>	Fast Healthcare Interoperability Resources
<b>HL7</b>	Health Level Seven
<b>HTML</b>	HyperText Markup Language
<b>HTTPS</b>	Hypertext Transfer Protocol Secure
<b>ICSD</b>	International Classification of Sleep Disorders
<b>IP</b>	Internet Protocol
<b>ISCED</b>	International Standard Classification of Education

**ISI** Insomnia Severity Index

**ISCO** International Standard Classification of Occupations

**JSON** JavaScript Object Notation

**LOINC** Logical Observation Identifiers Names and Codes

**MCTQ** Munich Chronotype Questionnaire

**MEQ** Morningness-Eveningness Questionnaire

**NIAAA** National Institute on Alcohol Abuse and Alcoholism

**PDF** Portable Document Format

**PSG** Polysomnography

**PSQI** Pittsburgh Sleep Quality Index

**REST** Representational State Transfer

**SCL90-R** Symptom-Checklist 90 Revised

**SSL** Secure Sockets Layer

**TXT** Text file

**UNESCO** United Nations Educational, Scientific and Cultural Organization

**XLSX** Excel Microsoft Office Open XML Format Spreadsheet file

**XML** Extensible Markup Language

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Thesis proposal . . . . .	2
1.2	Methodology . . . . .	2
1.3	Contributions . . . . .	3
1.4	Structure of the document . . . . .	4
<b>2</b>	<b>Sleep disorders</b>	<b>5</b>
2.1	Sleep and circadian rhythm . . . . .	5
2.1.1	Sleep functions . . . . .	7
2.1.2	Circadian rhythm and the sleep-wake cycle . . . . .	9
2.2	Diagnostic methods for sleep disorders . . . . .	12
2.2.1	Anamnesis . . . . .	12
2.2.2	Questionnaires . . . . .	14
2.2.3	Actigraphy and sleep logs . . . . .	15
2.2.4	Polysomnography . . . . .	17
2.2.5	Dim light melatonin onset measurement . . . . .	18
2.3	Delayed sleep phase disorder . . . . .	19
2.3.1	Prevalence, causes and triggers . . . . .	19
2.3.2	Diagnosis and treatment . . . . .	20
2.4	Overview . . . . .	23



<b>3</b>	<b>Supporting technologies and tools</b>	<b>25</b>
3.1	Data requirements for clinical applications . . . . .	25
3.1.1	Standard frameworks for health information systems . . . . .	26
3.1.2	Standardisation of medical nomenclatures . . . . .	28
3.1.3	Confidentiality, privacy and data security . . . . .	29
3.2	Software requirements for the SleepData platform . . . . .	31
3.2.1	Database management system . . . . .	31
3.2.2	Web server . . . . .	34
3.2.3	Web framework . . . . .	36
3.2.4	MEAN architecture . . . . .	37
3.2.5	Security and encryption . . . . .	37
3.3	Overview . . . . .	38
<b>4</b>	<b>SleepData</b>	<b>41</b>
4.1	Platform infrastructure . . . . .	41
4.2	Data modelling . . . . .	43
4.2.1	Clinical database . . . . .	43
4.2.2	Nomenclatures and coding systems . . . . .	45
4.2.3	User database and access control system . . . . .	46
4.3	User interface and user experience . . . . .	48
4.3.1	External dependencies . . . . .	48
4.3.2	User pages . . . . .	49
4.3.3	Uploading and managing exams . . . . .	54
4.3.4	Questionnaires and clinical forms . . . . .	54
4.3.5	Clinical dashboard . . . . .	55
4.3.6	Population statistics . . . . .	57
4.4	Overview on SleepData features . . . . .	60

<b>5</b>	<b>Characterisation of the DSPD population</b>	<b>63</b>
5.1	Characterisation . . . . .	63
5.1.1	General features . . . . .	63
5.1.2	Sleep pattern . . . . .	65
5.1.3	Questionnaires . . . . .	70
5.2	The role of DLMO in the diagnosis of DSPD . . . . .	71
5.3	Overview . . . . .	72
<b>6</b>	<b>Conclusions and future work</b>	<b>74</b>
6.1	Conclusions . . . . .	74
6.2	Future work . . . . .	76
	<b>Bibliography</b>	<b>80</b>
	<b>Appendices</b>	<b>81</b>
<b>A</b>	<b>Clinical database parameters</b>	<b>82</b>
<b>B</b>	<b>Dashboard parameters</b>	<b>89</b>
B.1	Delayed sleep phase disorder dashboard . . . . .	89
B.2	Insomnia dashboard . . . . .	91
<b>C</b>	<b>Population statistics chart list</b>	<b>94</b>
<b>D</b>	<b>CENC information</b>	<b>96</b>
D.1	Basic information . . . . .	96
D.2	Clinical record . . . . .	97
D.3	Actigraphy and related files . . . . .	97
D.4	Polysomnography and related files . . . . .	98
D.5	Night Diary . . . . .	99

# List of Figures

2.1	EEG of the second stage of NREM sleep . . . . .	6
2.2	EOG and EMG of the REM sleep phase . . . . .	6
2.3	Model of the two processes involved in the sleep-wake cycle . . . . .	10
2.4	Actigraphy device . . . . .	16
2.5	Core body temperature and melatonin levels over time . . . . .	21
3.1	MongoDB structure . . . . .	34
3.2	REST API Design . . . . .	36
3.3	MEAN stack architecture . . . . .	37
3.4	SSL verification protocol . . . . .	38
4.1	SleepData's software architecture . . . . .	42
4.2	SSL Certificate symbol: secure connection for SleepData platform . . . . .	42
4.3	Actigraphy report sample . . . . .	45
4.4	Login flow . . . . .	47
4.5	PSG cycle input . . . . .	49
4.6	Handlebars work flow . . . . .	49
4.7	Landing page . . . . .	50
4.8	Regular user homepage . . . . .	51
4.9	Homepages for different account types . . . . .	52
4.10	Web page schematics . . . . .	53
4.11	Query patients page . . . . .	55
4.12	List of possible equipment . . . . .	55

4.13	Glasgow sleep effort scale . . . . .	56
4.14	Partial health section of the Clinical Notes form . . . . .	57
4.15	Patient dashboard . . . . .	58
4.16	Insomnia patient dashboard . . . . .	58
4.17	Insomnia population statistics interface sample . . . . .	60
5.1	Age distribution (%) by age groups . . . . .	64
5.2	Reported features reported by the patient's, possibly related with DSPD . . . . .	65
5.3	Comorbidities distribution . . . . .	66
5.4	Distribution of perceived actigraphy schedules . . . . .	67
5.5	Distribution of total sleep time . . . . .	67
5.6	Sleep efficiency of DSPD patients computed by actigraphy . . . . .	68
5.7	Distribution of sleep latency . . . . .	68
5.8	Average bed time hour reported in a sleep diary, actigraphy and PSG . . . . .	69
5.9	Computed difference between average bed times . . . . .	69
5.10	PSG parameters of DSPD patients . . . . .	70
5.11	Distribution of questionnaires results . . . . .	71
5.12	Distribution of phase angle between time at sleep onset and time at DLMO . . . . .	72

# List of Tables

3.1	Feature comparison between relational and document-oriented DBMS . . . . .	32
4.1	SleepData's clinical database collections . . . . .	44
4.2	Types of accounts and their permissions . . . . .	47
4.3	DSPD dashboard parameters, variables and the respective colour codes . . . . .	59



# Introduction

The diagnosis and analysis of patients with sleep disorders depend on a vast ensemble of documents and exams, which are very heterogeneous not only in terms of sources, but also in used file formats. Furthermore, besides the reports provided by those machines, all other reports and clinical notes are often non-standardised, making it even more difficult to study and analyse the existent data, not to mention the handicap it represents in terms of data mobility and reproduction. Diagnostic machines often generate data files with up to 1 GB per exam, making it harder to process all the data. To complicate matters even more, there is no software system to manage in an integrated way all the data used in sleep medicine. This causes the process of diagnosis and study of sleep disorders to be often slow and puts the burden of diagnosis on the doctor alone. In addition, comparing large amounts of data or a large set of patients becomes a very limited task, leaving behind important information in each analysis. Therefore, there is a clear need for a platform composed of:

1. A database capable of integrating diagnostic files from every source;
2. Software for the automatic processing of these data, in order to implement statistical analysis and diagnostic supporting tools.

To understand the requirements and ensure that the platform created is useful for clinical practice and to verify the usability of the tools created, a group of IST students and faculty are working together with the *Centro de Eletroencefalografia e Neurofisiologia Clínica* (CENC), a state of the art clinic focused on sleep medicine with more than 5000 patients with sleep disorders. The lack of an integrated information system to support the handling and analysis of CENC's data highly limits the possibility for joint analysis of patients' clinical records.

As a proof of concept, a sleep disorder was studied in higher detail. Delayed Sleep Phase Disorder (DSPD) is a type of circadian rhythm disorder with a prevalence of about 16% in teenagers, with a reduction to 0.15% in the adult population, as indicated by Sheldon et al. [2014]. DSPD is often often misdiagnosed as other comorbidities, especially psychiatric disorders, according to Stores [2007]. The mistreatment of DSPD leads to serious psychiatric complications,

as reported by Okawa and Uchiyama [2007]. As mentioned by the American Academy of Sleep Medicine [2014], one of the most useful methods of diagnostic for circadian rhythm disorders is the measurement of dim-light melatonin onset (DLMO). Recent studies raise doubt as to the importance of DLMO as a marker. Taking all this into consideration, it seems relevant to characterise the DSPD population with DSPD and to study the importance of DLMO as a marker for this disorder.

## 1.1 Thesis proposal

The work presented in this dissertation had the following main objectives:

- Conceptualisation and development of a new information platform for managing sleep disorders, called SleepData, which is able to integrate clinical information of patients with sleep disorders, from several different sources and domains. SleepData enables the automatic update and processing of the data, to implement statistical analysis and tools specially designed with diagnostic purposes, as they give the physicians an overview of the patients' clinical records, as well as an overview of the set of population in study. This platform is scalable and adaptable for any clinic focused on sleep medicine. SleepData also aims to become a universal sleep data repository, being able to integrate the data of several clinics;
- Characterisation of CENC patients previously diagnosed with Delayed Sleep Phase Disorder (DSPD) using the SleepData platform. This characterisation highlights not only diagnostic parameters, but also general information about the patients. In particular, we intend to conclude whether dim-light melatonin onset measure is a reliable and effective marker, thus diagnostic method, for DSPD patients or not, taking in account the set of patients from CENC.

## 1.2 Methodology

To develop this thesis, I worked together with my colleague Tiago Castanheira, whose dissertation can be found in Castanheira [2017]. While the focus of my thesis was DSPD, my colleague's work was centred on insomnia. The development of this thesis followed the methodology described below:

1. Understand CENC's current state, including the different sources of information, how they are collected and treated;
2. Understand sleep disorders: what type of characterisation can be useful (giving special attention to DSPD), what are the current diagnostic tools and what are the specification of the files generated in the diagnostic process, in terms of file format, size and source;
3. Understand the requirements for the type of platform intended: existing standards for clinical applications, such as medical nomenclatures and protocols for information exchange, and data security requirements imposed by Portugal's legislation;
4. Choose the appropriate supporting information technologies and tools, such as database management system, web framework, and system architecture and security;
5. Conceptualisation and development of SleepData based on the collected information: configuration of the base software, modelling of the data, development of user interfaces for file insertion and diagnostic and analytic tools;
6. Transformation of CENC patients' clinical information to posteriorly load it to SleepData;
7. Conduct a study to characterise patient's diagnosed with DSPD;
8. Study the relevance of DLMO in the diagnosis of DSPD, considering the existent sample;
9. Evaluate the platform as a tool for clinical practice.

### 1.3 Contributions

In brief, the main contributions of this thesis are as follows:

- SleepData, a platform for sleep medicine which is capable of integrating clinical information from four different sources: medical records with the doctor's notes, standard questionnaires answered by the patients, the results of the exams performed, such as polysomnographies (PSG), actigraphies, DLMO measurement, as well as the technician's reports concerning each exam. SleepData provides tools for statistical analysis and diagnostic tools specially designed to study patients with DSPD or insomnia.
- The characterisation of a set of patients with delayed sleep phase disorder:
  1. 52% of patients manifest the first symptom between the ages of 6 to 20 years old;



2. Most patients experience family conflicts, traumatic experiences and stress or depression.
3. The most common comorbidities are anxiety, insomnia and depression;
4. More than 73% of the patients suffer from 2 to 4 comorbidities;
5. Most patients present a sleep phase delay of 1 to 4 hours;
6. Patients are able to fall asleep earlier and have better sleep efficiency when certain stimuli, like bright light and screen time, are reduced near bed time hour;
7. DLMO is not a good marker for DSPD for this set of patients.

SleepData meets the following requirements for clinical applications:

1. The data are findable, accessible, interoperable and re-usable due to the implementation of appropriate standards and nomenclatures;
2. Confidentiality, privacy, data security and encryption;
3. Scalable platform, capable of handling multiple users and simultaneous request.

## 1.4 Structure of the document

The rest of this document is organised as follows. Chapter 2 presents fundamental concepts related with circadian rhythm sleep-wake disorders and the diagnostic tools used, with focus on delayed sleep phase disorder, a view on the practical details of data collection methods, with CENC as a case study. Chapter 3 addresses the main data and software requirements for clinical applications and web applications, as well as a comparison of the chosen technologies and tools with other possible choices. Chapter 4 describes the platform infrastructure, data modelling, user experience and interface, analytic and diagnostic tools. Chapter 5 describes the characterisation of CENC's DSPD population using SleepData's tools. Finally, Chapter 6 concludes this dissertation by summarising the main findings of this work, and highlighting possible directions for future work.

# 2 Sleep disorders

This chapter presents fundamental concepts related with sleep and its disorders, with focus on the circadian rhythm and sleep-wake cycle (Section 2.1). Section 2.2 presents the main diagnostic tools and the practical aspects of data collection, with CENC as an example. Section 2.3 is dedicated to delayed sleep phase disorder, a main focus of this thesis. An overview of the topics approached in this chapter can be found in Section 2.4.

## 2.1 Sleep and circadian rhythm

To understand and study sleep disorders, it is first necessary to have a clear understanding of what normal sleep is and its functions, in human beings. As simply put by Carskadon and Dement [2011], sleep is a *reversible behavioural state of perceptual disengagement from and unresponsiveness to the environment*, where the individual often presents *postural recumbence, behavioural quiescence, closed eyes* and other indicators associated with sleeping. Although this is a vast and seemingly subjective definition, sleep may also be recognised through the variations in neural activity, usually detected by electroencephalogram (EEG), electrooculogram (EOG) and electromyogram (EMG). Using the physiologic parameters given by those tests, it is possible to recognise two different sleep states: non-rapid eye movement (NREM) and rapid eye movement (REM). NREM typically has minimal or fragmentary mental activity, being characterised by an EEG synchronous pattern of high voltage slow waves, with sleep spindles and K-complexes, as show in Figure 2.1. NREM sleep has three stages (N1, N2 and N3), distinguishable in the EEG. In contrast, REM sleep is an *activated brain in a paralysed body*, with high neural activity, muscle atony and rapid eye movement, often accompanied by muscle twitches and cardiorespiratory irregularities. REM stage is also characterised by the occurrence of dreams. Figure 2.2 represents the physiologic parameters characteristic of REM sleep phase. There is some disagreement regarding the definition of sleep onset, since a change in pattern in EEG is not always corresponding with a person's perception of sleep. Although it is said that the sleep state starts at the N1 stage, some prefer to only acknowledge sleep onset when an individual reaches N2. However, the process of *falling asleep* is not a discrete event, it is rather a continuous

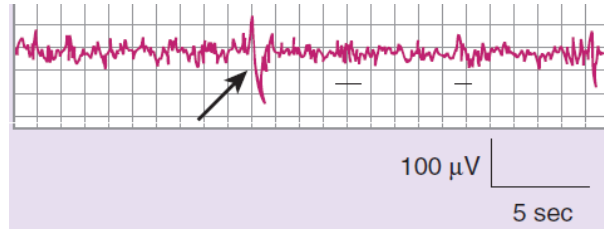


Figure 2.1: EEG of the second stage of NREM sleep. The arrow indicates a K-complex and the underlining shows two sleep spindles. Adapted from Carskadon and Dement [2011].

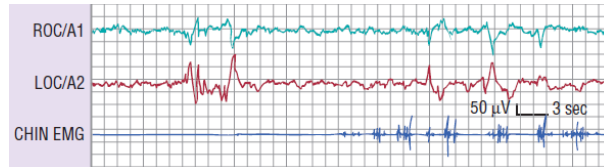


Figure 2.2: EOG and EMG of the REM sleep phase. On the left side is a burst of several rapid eye movements (out-of-phase deflections in right outer canthus [ROC]/A1 and left outer canthus [LOC]/A2). On the right side, there are additional rapid eye movements as well as twitches on the electromyographic (EMG) lead. Adapted from Carskadon and Dement [2011].

and progressive entry into a less vigilant state. For that reason there is consensus as to that the transition to sleep happens during N1 stage. After that, according to Billiard [2008], follows N2 and N3 stages (for about 10-25 minutes and 20-40 minutes, respectively), with a growing arousal threshold, which is the arousal needed to transit from a sleep state to a wakefulness state. REM sleep occurs next, usually after 80 to 100 minutes of sleep. The first cycle lasts for about 90 minutes. This NREM/REM sleep cycle repeats throughout the night, with duration of 100 to 120 minutes per cycle where the concentration of N3 is higher in the beginning of the sleep state and diminishes, giving space for longer REM sleep episodes. Typically, the first cycles are abundant in N3 sleep, but may be nonexistent in the following cycles. Also, the last cycle REM is the longest one. Normal sleep is considered to last around eight hours, giving room for about four complete cycles. Variations in this process are expected, depending on the age of the subject. For example, babies enter REM sleep before NREM and have shorter sleep cycles, with REM sleep corresponding to 50% of total sleep, value that decreases to 20% to 25%, values that are maintained throughout the years. Teenagers suffer a decrease of about 40% in N3 stage, decline that continues in adulthood, but with much lower steep. As every function in the human body, sleep is also affected by a number of internal and external factors. As said by Carskadon and Dement [2011], sleep-wake history, circadian phase, temperature, drugs and sleep disorders have a major impact on sleep.

### 2.1.1 Sleep functions

Sleep is one of the main bodily functions. Its importance to the maintenance of physiologic balance can easily be shown by studying the consequences of sleep deprivation. According to Silva [2014], the most noticeable consequences observed in most individual deprived of sleep are the increase in somnolence, reduction in psychomotor performance, lack of attention and reduced reaction time, reduced ability to form new memories, bad mood, fatigue and irritability. Looking into these observation more carefully, it is possible to discern the main sleep functions. In the same study by Silva [2014], six functions are emphasised, listed and briefly explained below:

1. Preservation of energy and promotion of anabolic processes: during sleep hours, the body temperature decreases, diminishing the metabolic processes and prioritising anabolic processes; there is a clear increase in the secretion of the growth hormone, known for stimulating protein synthesis and growth, bone mineralisation and increase in muscle mass, promoting lipolysis, reducing concentration liver glucose and stimulating the immunitary system; the secretion of prolactin increases as well, affecting osmoregulation, growth, reproduction and immunological modulation.
2. Thermoregulation mechanisms: it has been verified that NREM sleep is induced when the brain temperature reaches a certain threshold and since sleep reduces the body temperature, it is possible to infer that sleep contributes to the homeostatic control of brain temperature; the sleep induction by high temperature is done through hypothalamic neural networks, which inhibit neural groups responsible for waking up when exposed to a higher temperatures in the brain.
3. Brain *detoxification*: several experiences show that the concentration of several substances that promote sleep onset increases during wakefulness and decrease with sleep, leading to the hypothesis that sleep acts as a method of brain detoxification; the specificities of the mechanisms involved are yet to be established.
4. Promotion of immunologic responses: during NREM sleep, there is an increase in production of interferon gamma, translating into a prevalence of cellular immunity; during REM sleep, humoral immunity is prevalent, due to the increase of hormones like the growth hormone and prolactin, that stimulate the production of cytokines, like interleukin IL-2, responsible for the destruction of infected cells, and to the decrease in cortisol, that usually inhibits the production of that type of cytokine.

5. Brain development and maturation: REM sleep has a major role in the development and maturation of neural networks in primordial stages, by activating the central nervous system, which is consistent with the fact that newborns experience a higher percentage of REM sleep than adults.
6. Memory consolidation and brain plasticity: the slow waves present during NREM sleep have the ability to reactivate the hippocampus-neocortical circuits formed during recent learning processes in the vigilant period, having a major role in the process of memory consolidation; both REM and NREM sleep promote synaptic plasticity, which facilitates the consolidation of new memories.

Sleep is thus of utmost importance to a well functioning body, in terms of health, well-being and productivity. In a study by Dongen et al. [2003], the consequences of sleep deprivation are exposed highlighting its cumulative consequences, having as example the deterioration of reaction times. The study compared reaction times of three groups of subjects that experienced different total sleep times (8 hours, 6 hours and 4 hours), maintained for 14 consecutive days. An extra experiment of total sleep deprivation was made, in 3 nights without sleep. The main conclusions from this study showed that the restrictions highly compromised the subjects' reaction times and that accumulating vigilant hours worsened the results. For example, subjects that experienced 6 hours of total sleep time for 10 days, had the same chance of showing lapses in performance that subjects completely deprived of sleep for one night (0 hours of total sleep time). It was observed that either higher sleep deprivation for a short period of days or lower sleep deprivation for a longer period of days, had deteriorating effects in terms of performance in the subjects. In another study about sleep deprivation, Suchecki and Tufik [2014] point out to its relationship with stress, as the secretion of stress hormones increase after twelve hours of sleep deprivation. This highlights the fact that sleep is essential to survival and the lack of it represents a threat to the individual, as much as starvation or dehydration. In fact, Paiva [2014] mentions that animals, like flies and mice, die after 15 days of total sleep deprivation, alerting to the extreme importance of sleep.

Sleep disorders can be classified in the following way, according to the latest edition (third edition) of the International Classification of Sleep Disorders [American Academy of Sleep Medicine, 2014]:

- Insomnia;
- Sleep related breathing disorders;

- Central disorders of hypersomnolence;
- Circadian rhythm sleep-wake disorders;
- Parasomnias;
- Sleep related movement disorders;
- Other sleep disorders;

For the purpose of this thesis, this chapter will focus on circadian rhythm sleep-wake disorders, in particular delayed sleep phase disorder, without disregarding the importance of understanding and studying the other mentioned sleep disorders. Diagnostic methods common to most disorders will be discussed in Section 2.2.

### 2.1.2 Circadian rhythm and the sleep-wake cycle

The circadian rhythm is a type of biological rhythm present in all species, which regulates many physiological, neuroendocrinological and behavioural processes, characterised by repetition periods of about 24 hours. According to Geraldès and Paiva [2014], the circadian rhythm is generated endogenously, independent of external stimuli, meaning that it is generated even in the absence of environmental cues. The effect of this endogenous systems can be observed on a molecular level, as each cell has molecular oscillations, induced by activation/inactivation cycles of the so called *clock* genes. It is currently accepted that the main structures responsible for the circadian rhythm, or as it is often called, the main internal clock, are the suprachiasmatic nucleus (SCN) located in the hypothalamus. It is clear that environmental stimuli have the power to synchronise the circadian system. In fact, light is the main SCN synchroniser, also called *zeitgeber*, from the German, meaning *time giver*. This makes the retinohypothalamic tract the main synchronisation route. One of the examples mentioned by Geraldès and Paiva to verify this is that the production of melatonin drops when an individual is exposed to light. The SCN has a direct effect in the pineal gland, which is responsible for the production of melatonin. This structure also acts on several structures in the hypothalamus, the brain stem, the basal forebrain and the midbrain, thus regulating the production of cortisol, body temperature, the sleep-wake cycle and several other biological parameters.

The sleep-wake cycle cannot be explained solely by the effect of the main internal clock in the SCN, as explained by Geraldès and Paiva [2014]. Apart from the circadian process, a homeostatic process should be taken into consideration. In fact, this model explains that what determines

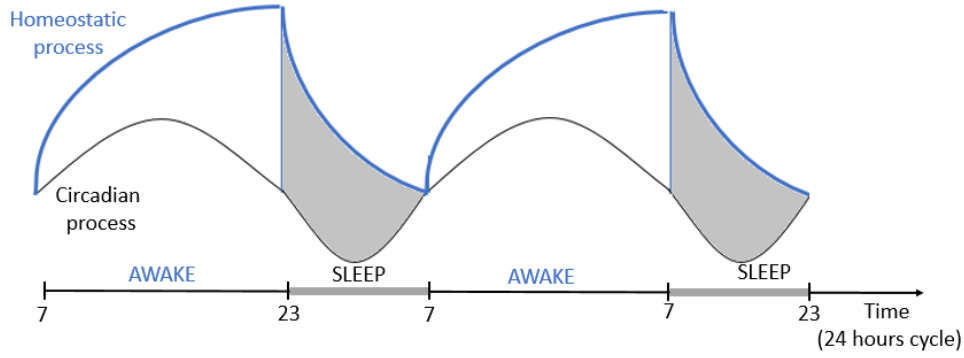


Figure 2.3: Model of the two processes involved in the sleep-wake cycle: circadian and homeostatic, in phase. Adapted from Geraldles and Paiva [2014].

the sleep and wake up schedule is the interaction between the circadian and the homeostatic process, and that a consolidated sleep can only be obtained when the two processes are in phase. To understand this concept, the two processes are represented in Figure 2.3. In practical terms, the SCN generates an alertness or arousal signal that increases its intensity during the day and declines during the night, increasing the production of sleep-inducing melatonin. The homeostatic process works as a regulator for the sleep *debt*, which increases during waking hours and gradually reduces during sleeping time. The markers for this process are slow sleep and slow waves, measured in EEG during NREM sleep. As the authors mention, the duration of these markers can be predicted by the amount of time a subject is awake, independently of the circadian process and are independent of the circadian phase in which sleep occurs. Koyama et al. [2014] describe the homeostatic variable has corresponding to fatigue and add that it may represent a neurochemical factor that accumulates during waking time, being responsible for the need of sleep feeling one has. This factor would be eliminated during sleep, much like the sleep *debt* mentioned before. One possible neurochemical factor is adenosine, product of adenosine triphosphate (the so called *energy molecule*), which increases with the time spent awake. The sleep-wake cycle usually includes 16 hours of wakefulness and 8 hours of sleep.

The influence of melatonin in the sleep cycle control should not be forgotten. As mentioned before, the production of this hormone is regulated by the circadian system, making it sensitive to light variations. According to Koyama et al. [2014], its production occurs during the night, specifically in the period between the twilight and the dawn, reaching its peak around 4 AM, when the exposure to light is nearly nonexistent. However, if there is exposure of the retina to bright light during that period, melatonin synthesis will cease, possibly completely, changing the sleep-wake circadian cycle.

Despite the countless bibliography available on this topic and the consensus there is among

several authors, it is important to keep in mind that the sleep parameters present high variability, inter and intraindividually. Variations are excepted in different age groups, genders, lifestyles and habits. According to Koyama et al. [2014], the sleep schedules are as variable as the distribution of schedules for certain activities. One can name the individual according to their preferences regarding these schedules, or better said, their chronotype. People that prefer to wake up early in the morning, perform physically and mentally demanding activities during the first hours of the day and go back to sleep in the first hours of the night are called the morning chronotype. The evening chronotype corresponds to individuals that prefer to wake up later and go to sleep later, concentrating the most demanding activities later in the day. Most people present an intermediate type. When individual preferences are too extreme and are the source of many undesirable symptoms, the subject may be suffering from a circadian rhythm sleep-wake disorder. These preferences are often associated to psychosocial factors that vary throughout the course of life. Czeisler and Buxton [2011] highlight the relevance of these factors, by comparing the sleep-wake cycle of humans to other animals. While the latter present schedules adapted to the time of day, the former overrides the natural environmental signals that synchronise the circadian sleep-wake rhythm. The appearance of the alarm-clock and artificial lighting can be considered the main guilty parts of this phenomenon, as people have the option to alter their waking hours due to work, social or recreational demands. Paiva [2014] refers the main social changes that this industrialised society brought, like the generalisation of work for shifts and intercontinental trips, demand of higher working rhythms and competition and, ultimately, the pressure to be online 24 over 24 hours. Due to all this, the respect for the circadian rhythm, or our biological clock, has fallen affecting not only sleep, but also meal schedules. The long-term consequences of this lack of respect for the internal clock may not be well established yet, but as the author explains it is impossible not to be aware of the increase of fatigue and stress complaints, giving as example the growing number of accidents related to fatigue, in which the sleep-wake cycle has an enormous effect.

According to the International Classification of Sleep Disorders (ICSD) by the American Academy of Sleep Medicine [2014], optimal sleep can only be obtain when there is a match of the sleep time with the sleep phase of the circadian sleep-wake cycle. As cited by Academy, a circadian rhythm sleep-wake disorder is caused by *alterations of the circadian time-keeping system, its entrainment mechanisms, or a misalignment of the endogenous circadian rhythm and the external environment*. The different types of disorders identified by the international classification are:

- Delayed sleep-wake phase disorder;



- Advanced sleep-wake phase disorder;
- Irregular sleep-wake rhythm disorder;
- Non-24-hour sleep-wake rhythm disorder;
- Shift work disorder;
- Jet lag disorder;
- Circadian sleep-wake disorder not otherwise specified.

Although they have different specificities, there is a set of symptoms shared by all the above disorders, such as difficulty maintaining sleep, excessive sleepiness during waking hours and impairments in social, occupational and educational performance. As one of the purposes of this thesis is to characterise a set of patients diagnosed with delayed sleep phase disorder, the specificities of this disorder will be approached in Section 2.3.

## 2.2 Diagnostic methods for sleep disorders

As this thesis aims to develop a platform that integrates clinical information from different sources, which includes the diagnostic reports and tools used to assess the state of each patient, it is important to survey the diagnostic methods that are used in clinical practice for sleep medicine.

### 2.2.1 Anamnesis

Anamnesis, or medical history, is the collection of information by a physician about a patient using specific questions. Although this method may not be as objective as other diagnostic tools, it is one of the most important in most medical specialities, as the information collected is highly susceptible to the patient's perception and to the patient's desire to give the doctor correct and exact information. However, it is the doctor's job to analyse and interpret the given information. As seen before, there are high social, behavioural and psychological components associated with sleep disorders, which highlight the importance of this diagnostic tool. Paiva and Pinto [2014] list the most important parameters one should collect when interviewing a patient with sleep complaints (although this list was written with the purpose of diagnosing insomnia, the parameters are common for most sleep disorders):

- The circumstances in which the symptoms started, such as age, triggers, evolution and duration;
- Sleep parameters: preferred sleep schedule and duration, actual duration of sleep habits before falling asleep, latency, sleep interruptions (frequency, duration and motive), behaviours during nocturnal awakenings, if the patient sleeps alone, nocturnal symptoms and how the patient feels in the morning;
- Severity and frequency of the symptoms;
- Consequences felt during day time, p.e fatigue, headaches and performance;
- Aggravating factors;
- Improvement factors;
- Sleep hygiene: habits and behaviours before going to bed and while in bed and use of stimulants (p.e nicotine, cocaine, caffeine and alcohol);
- Circadian rhythms: sleep schedules during work/school days, free days and during vacation days, shift work;
- Family and personal medical history;
- Treatments used (p.e. medication, psychotherapy) and their results;
- Stress factors: work, unemployment, multiemployment, personal conflicts, family conflicts, frequent travels, having children with sleep problems and any other stressor.

With a precise and extensive medical history it is possible to understand the outlines of a patient's complaints, allowing the doctor to do a pre-diagnosis and to prescribe further exams or other diagnostic tools if deemed necessary. For some clinical disorders, a physical examination is also needed, with results completing the medical record.

In practical terms, it is important to understand how a clinic manages this information in the daily activities. The doctor responsible for CENC, Dr. Teresa Paiva, collects the medical history in a handwritten clinical record, in the form of free text that covers all the points listed above.

### 2.2.2 Questionnaires

Another form of anamnesis is the use of standardised questionnaires that focus on specific aspects of sleep. The questionnaires can be answered by the patient or someone responsible for the patient, if the patient is a child or is unable to answer the questionnaires. Paiva and Penzel [2011] list the most used questionnaires in sleep medicine in the following categories <sup>1</sup>:

1. Sleep quality:

- (a) Pittsburgh Sleep Quality Index (PSQI): to assess sleep quality, latency, duration, efficiency, common sleep disturbances, medication used to sleep and daily impairments (available at <http://www.psychiatry.pitt.edu/node/8240>);

2. Sleepiness and vigilance:

- (a) Epworth Sleepiness Scale (ESS): to assess probability of falling asleep during different everyday tasks (available at <http://epworthsleepinessscale.com/>);
- (b) Glasgow Sleep Effort Scale: assess effort to fall asleep (available at <http://uofthenet.org/alliant/Garrison-Sleep/10-GSES.pdf>);

3. Insomnia:

- (a) Insomnia Severity Index (ISI): to assess insomnia's severity (available at [https://www.ons.org/sites/default/files/InsomniaSeverityIndex\\_ISI.pdf](https://www.ons.org/sites/default/files/InsomniaSeverityIndex_ISI.pdf));

4. Chronotype:

- (a) Morningness-Eveningness Questionnaire (MEQ): to assess schedules preferences (available at <http://www.cet-surveys.com/index.php?sid=61524>);
- (b) Munich Chronotype Questionnaire (MCTQ): assess sleep schedules, light exposure and other sleep-wake habits (available at <https://www.thewep.org/documentations/mctq>);

5. Symptoms:

- (a) Symptom-Checklist 90 Revised (SCL90-R): to assess nine symptoms of psychopathology and provides three global distress indices (available at <http://accesscm.org/wp-content/uploads/2013/03/SCL-90.pdf>).

---

<sup>1</sup>Only the relevant questionnaires for the purpose of this thesis are listed and described, as the list of possible questionnaires is highly vast.

CENC's patients fill these questionnaires, as needed, in paper.

### 2.2.3 Actigraphy and sleep logs

Sleep logs are filled in by the patient to document their sleep habits and schedules. The patient is instructed to complete the log every day, ideally for several consecutive weeks. The required information may depend on the doctor's instructions, which will request what information he or she deems more relevant to the patient's case. According to Bae and Avidan [2008], the usual requested information is:

1. The time the patient thinks he/she fell asleep;
2. The time the patient wanted to fall asleep;
3. Nocturnal awakening: quantity, duration and time;
4. The time the patient woke up and the time he/she got out of bed;
5. The time the patient wanted to wake up;
6. Comment on how the patient felt during the day;
7. Daytime naps: start and end time;
8. Medication use, if any.

Other relevant information could be the meal and exercise schedule. The author claims that this diagnostic method is not only helpful for the physician, but also for the patient. Sleep logs help patients have a better perception of their sleep problems and can also help patients see the evolution of their disorder. After completion of the sleep log, it should be taken for clinical sleep evaluation.

An actigraphy is a *measurement of physical activity (...) employed to estimate sleep and wakefulness based on relative levels of physical inactivity and activity* [American Academy of Sleep Medicine, 2014]. The device used for this exam is called an actigraph, a small wristwatch-like accelerometer (see Figure 2.4), which records the activity for several weeks with an adaptable sampling interval. The sampling interval can go from a few seconds to several minutes. With the help of dedicated software, the actigraphy is able to distinguish sleep and awake periods, measure the activity/inactivity levels, total sleep time, sleep latency, sleep percentage, total awake time,



Figure 2.4: Actigraphy device, ActTrust, sold by Condor Instruments (<http://www.condorinst.com.br/en/acttrust/>)

awake percentage, number of awakenings. The most recent devices have temperature and light sensors, allowing a more precise diagnosis.

Paiva and Penzel [2011] highlight the usefulness of the actigraphy, as it allows the determination of a robust circadian pattern. The precision of this device has been assessed by simultaneously running a polysomnography (method addressed ahead) and an actigraphy. Another point in favour of this method is the higher reliability when compared to sleep logs, where parameters like sleep latency, number of awakenings and even number of naps lack precision. The authors claim that the biggest advantage of the actigraphy is that it enables the assessment of the patients in their natural environment, without disrupting any habits. The sleep log poses some questions of veracity and precision, since it is completed by the patient without any medical supervision until it is completed. The combined use of sleep log with actigraphy, covering the same period of utilisation, gives a better understanding of the sleep-wake pattern of the patient and it enables the prediction of the dim-light melatonin onset (approached ahead). These diagnostic tools are often used in the evaluation of circadian rhythm disorders, excessive daytime sleepiness and insomnia, according to Bae and Avidan [2008].

CENC currently uses two distinct brands of actigraphs: Philips and Condor. The data is retrieved using the respective pod and software, which uploads the information to the computer. The raw data from Philips' actigraph, called Actiwatch2, is exported in XLSX file format, consisting of a spreadsheet with the information the device gathered for each epoch, where each row corresponds to a specific epoch and each column to a specific value measured. Actiwatch2 also registers the subject's and device's information, as well as the device setup. Philip's software allows the technicians and doctors to mark the epochs they find relevant for the study, which the software later uses to produce a report, with the information mentioned above. Condor's actigraph, ActTrust, exports its raw data in TXT file format, where each line corresponds to an epoch and the information in each line is separated by a semicolon. The text file also has information about the subject, device and device's setup, where the parameter and value are

separated by a colon. Like the Philip’s software mentioned previously, Condor’s software also allows the technician or doctor to mark the epochs deemed relevant, so that the software can produce a report. This report is exported in PDF file format, which includes the same parameters as the report provided by Actiwatch2. The size of the raw files exported by both actigraphs is highly dependable on the sampling interval, but usually has several megabytes.

CENC’s patients fill the sleep log during the recording period of the actigraphy, manually in a table with the required information (sleep, meals, work and exercise schedules).

#### **2.2.4 Polysomnography**

Like the name suggests, a polysomnography (PSG) is a set of tests used for sleep studies. It consists on the simultaneous and continuous recording of several electrograms and other variables through the course of one night, starting before sleep onset and finishing when the patient wakes up in the morning. According to Paiva and Penzel [2011], the PSG can be done in a sleep lab or on an outpatient basis, with or without supervision. Although the recordings of a PSG can vary according to its purpose, the pattern or classical PSG records:

1. Electroencephalogram (EEG);
2. Electrooculogram (EOG);
3. Electromyogram (EMG);
4. Electrocardiogram (ECG);
5. Airflow;
6. Oxygen saturation;
7. Body positioning and limb movement monitoring.

This method is appropriate to assess patients that have sleep episodes in inappropriate times, trouble sleeping or insomnia, difficulty staying awake or hypersomnia and atypical behaviours during sleep. It is also useful to assess how a therapy is progressing. It is considered to be a golden standard tool to assess sleep apnea, as it allows a precise diagnostic, allowing to assess the severity of the condition and the existence of other related problems.

CENC uses five different PSG systems Philips’ Alice 5, SOMNOmedics’SOMNOscreen, Natatus’ Embla, Embletta (portable version of Embla) and Nicolet. For each of these systems, the

data is exported using different software that register and treat data in different ways, which constitutes one of the major problems concerning the PSG information. The different software used export the data in different ways: one single file with all the data, one file for each channel acquired or one file for a number of sampling periods. Another problem is the dimension of the raw data of each exam exported, as each exam can take up more than 1 gigabyte of memory. From these files, it is possible to produce reports with different conclusions related with the recordings mentioned above. The clinic also records a Night Diary of the night corresponding to the PSG, if it is conducted in the sleep lab, where the patient's behaviour and perception is recorded.

### **2.2.5 Dim light melatonin onset measurement**

At the start of this chapter, I introduced the important role of melatonin in the synchronisation of the circadian rhythm. This hormone is produced between twilight and dawn. It is, therefore, not surprising that assessing the variation in time of the melatonin levels is considered to be a useful tool in the diagnosis of sleep disorders. According to Pandi-Perumal et al. [2007], the dim light melatonin onset (DLMO) is the most accurate marker for assessing the circadian phase and has been often used to evaluate problems related with the onset and offset of sleep. The DLMO test can be done through blood, urine or salivary samples. As the purpose of this test is to assess when the production of melatonin is above a certain threshold and to study its variation in time, several samples are needed, making salivary samples the most easy and practical to use. The most used method is the partial melatonin curve, which depicts the variation of melatonin through only a portion of the night, but if such is inconclusive, a 24h melatonin curve is required. In any case, nighttime saliva collection must be done under dim light conditions. The onset of melatonin in dim light conditions is defined at the time when a salivary concentration of melatonin of 4 pg/ml is reached. In adults, DLMO is considered normal when it occurs between 19:30 and 22:00 h, and between 19:00 and 21:00 h in children. Keep in mind that these schedules take into account the socially acceptable bed time. As mentioned by Fahey and Zee [2008], DLMO typically occurs around two to three hours before habitual bedtime.

Although these values are accepted by the medical community and DLMO is considered a gold standard for diagnosing circadian rhythm disorders, recent studies disagree with this claim. Quian et al. [2017] and Burgess et al. [2017] point out that the reference values do not account individual differences and that patients that suffer from circadian disorders tend to present higher variations (DLMO does not occur two to three hours before bedtime). Further investigation in

this topic is necessary.

CENC relies on an external laboratory to conduct the analysis, which later sends the results in paper.

## 2.3 Delayed sleep phase disorder

According to the ICSD, the delayed sleep phase disorder (DSPD) is characterised by a delay of two or more hours of the sleep-wake cycle, in comparison to the socially acceptable timing. Patients that suffer from this disorder have difficulty falling asleep at a socially acceptable hour, which results in shorter and insufficient sleep duration if they have to wake up early, as required by normal school or work schedules. Waking up at the socially acceptable time is also a major struggle for these patients, as they could not complete the sleep duration needed. However, if they are allowed to choose their preferred sleep-wake schedules, which represent a delay in timing, the sleep pattern and duration is normal. A common symptom of this disorder is low mental and physical performance during the first hours of the morning due to excessive sleepiness caused by sleep of short duration and by the interruption of sleep during the circadian phase of high sleep propensity. Many patients that suffer from DSPD present mental disturbances like mood disorders or depressive symptoms and may develop insomnia disorder [American Academy of Sleep Medicine, 2014]. Patients that suffer from this sleep disorder and that are not correctly diagnosed, tend to force themselves into a regular 9-5 workday schedule (or which ever schedule their environment forces), without clinical help to do so. This adaptation to a schedule outside the demands of their biological circadian rhythm is mostly unsuccessful, leading to several physical and psychological complaints, such as daytime sleepiness, fatigue, headaches, decreased appetite and depressed mood, as reported by Okawa and Uchiyama [2007].

### 2.3.1 Prevalence, causes and triggers

This disorder can be found in any age group, but it is more common in teenagers, showing a prevalence of 7 to 16%, as most show a *biological endogenous shift toward later bedtimes beginning around puberty* [American Academy of Sleep Medicine, 2014]. The higher prevalence in adolescents can also be explained by typical behavioural and social factors, such as school avoidance, social maladjustment and family dysfunction. Genetic factors are also thought to be associated with DSPD, including polymorphism in hPer3, the circadian clock gene, which explains why 40% of patients diagnosed with this disorder have a positive family history. This



genetic factor may be the source of DSPD onset in early childhood. Children with attention deficit hyperactivity disorder or autism spectrum disorder have an increased propensity for DSPD. Diagnosing and treating these two age groups is of utmost importance, as both teenagers and children with the evening chronotype often experience severe consequences later in life, like *decreased health-related quality of life, higher rates of (...) depression and suicidality, decreased sleep duration and increased daytime sleepiness, (...) impairments in academic functioning, and increased likelihood of substance use*. The common causes to all age groups are environmental factors. As mentioned previously, light is a main synchroniser of the circadian rhythm: increased exposure to bright light during the evening or decreased exposure to light during the morning may provoke a delay in the circadian phase. Other environmental factors responsible for this disorder may be inadequate adjustment to changes in work or social schedules, time zone changes and shift work. Stimulants abuse, like caffeine abuse is also a known trigger, delaying sleep onset and, thus, delaying the sleep-wake cycle. Other triggers include psychological, medical or environmental stressors. These stressors can take diversified forms, such as physical abuse, financial instability, death of a relative, chronic disease, or any other cause for concern.

### **2.3.2 Diagnosis and treatment**

DSPD is often often misdiagnosed as other comorbidities, especially psychiatric disorders, like psycho-physiological insomnia, depression or schizophrenia, leading to the inappropriate prescription of psychoactive drugs. This misdiagnosis is not only due to the lack of awareness of DSPD, but also due to the not so well defined limits that characterise this disorder, as mentioned by Stores [2007] and Shneerson [2005]. However, the ICSD provides a set diagnostic criteria for DSPD that must be met. The mandatory criteria for diagnosing a subject with DSPD are the following:

1. Significant sleep phase delay comparing with the socially accepted or desired sleep-wake timing, displayed by a chronic or recurrent inability to fall asleep and difficulty waking up at a required time;
2. Symptoms present for at least three months;
3. When patients are allowed to choose their sleeping schedule according to their preferences, they show improvements in sleep quality and duration, maintaining a delay in phase of the 24-hour sleep-wake cycle;

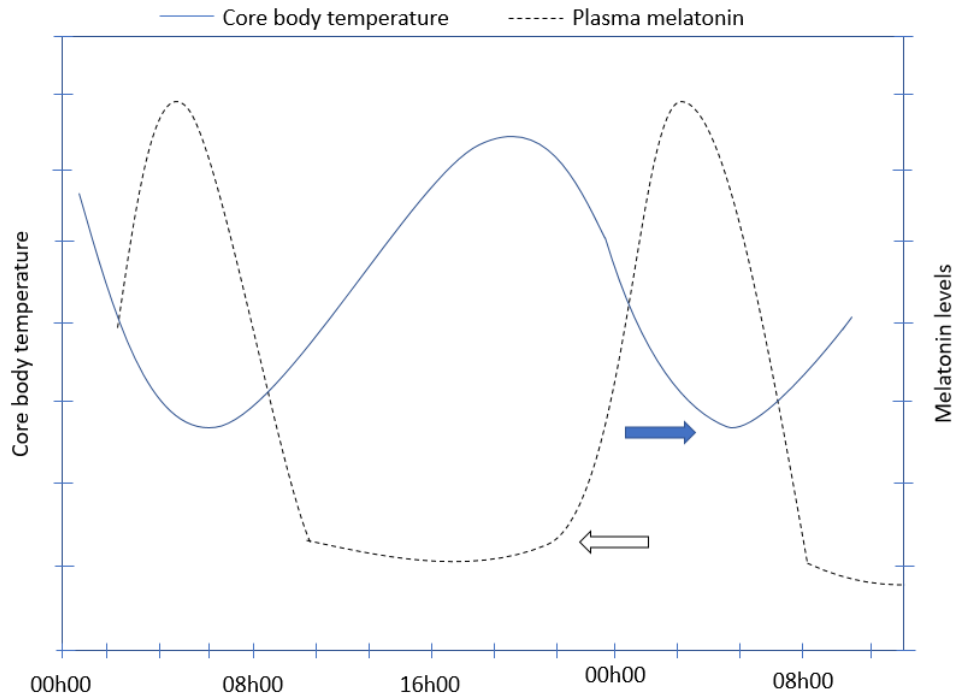


Figure 2.5: Representation of the phase relationship of core body temperature and melatonin levels over time. The white arrow indicates the dim light melatonin onset (DLMO), which typically occurs approximately 2–3 hours prior to habitual bedtime. The blue arrow indicates core body temperature minimum, which typically occurs 2 hours before habitual wake time. Adapted from Fahey and Zee [2008].

4. Minimum seven days sleep-log and, if possible, actigraphy monitoring, with both school/work and free days included, show a delay in the timing of the sleep period;
5. No other current medical condition, medication use or substance abuse explains the sleep disturbance better than DSPD.

The American Academy of Sleep Medicine (AASM) gives further specifications for some of the criteria. Sleep onset is usually delayed until 1 AM to 6 AM, both in work/school and free days. Wake time occurs in late morning or afternoon (when patients don't have social constraints that require them waking up earlier). Cultural and age variances should be taken into consideration. When a polysomnography exam is conducted during the schedule preferences of the patient, it shows a normal sleep pattern (if conducted at the socially acceptable schedule, increased sleep latency and decreased total sleep time are shown). Phase delay in melatonin production and temperature circadian cycles are observed as well. Figure 2.5 serves as reference for the normal circadian cycles of temperature and melatonin, for a subject whose bed time is usually around mid-night and wake up time around 8 AM.

The AASM adds that tools like chronotype questionnaires are also useful for diagnosis,

since patients suffering from DSPD usually score as evening types. Daytime sleepiness scale questionnaires are useful as well. A delay in the timing of other circadian rhythms like melatonin production, measured by DLMO, which results in the maintenance of the phase angle between rhythms, is also pointed out as desirable to confirm the disorder.

Although the measurement of DLMO is often used in clinical practice, the relevance of this tool is a topic of discord. DLMO models were constructed using group-averaged data and fail to reflect significant individual differences, as shown in a recent study by Quian et al. [2017]. According to another experiment by Burgess et al. [2017], those individual differences are even more notable in patients with DSPD. These studies raise doubts as to the importance of DLMO as a marker for DSPD.

Treatment for DSPD is extremely advised, given the high health risk previously discussed. Although in some cases increasing age may lead to an advance in the circadian sleep-wake cycle, re-positioning it to the desirable schedules, in many cases DSPD is a chronic condition that lasts into life. Repetition of the treatment may be needed, as the chronotype preference for later schedules is usually unaltered, as mentioned by the AASM.

Fahey and Zee [2008] explain the treatment options for DSPD. Usual treatments include phototherapy, behavioural and pharmacological treatments. Phototherapy, or light therapy, consists on the exposure to light. To treat DSPD, a range of 2500 to 9500 lux<sup>2</sup> is needed, provided by a light box, for one to three hours upon awakening. Light within the blue spectrum (wavelengths around 460 nanometers) is more effective and there are already many blue-green enriched light emitters available in the market. There may be some side-effects to this therapy, light headaches, nausea, visual problems and hypomania. It is important to highlight that while a higher exposure to light in the morning is useful to advance the sleep-wake cycle in DSPD patient, the same importance must be given to the avoidance of bright light exposure in the evening. Behavioural adjustments like avoiding bright screens, such as television, computers and smartphones, should be considered. Another way of avoiding bright light is through the use of dark glasses. The most common pharmacological treatment for DSPD is the use of melatonin, with doses of 0.3 to 3 milligrams given five hours before the habitual bedtime. Side effects include soporific effects, vasoconstriction and may aggravate asthma, so it should be prescribed with care. The combination of phototherapy and melatonin is advised for better results. Another treatment method used, although not as often, is chronotherapy, which consists in a progressive further delay of the sleep schedule by three hours every two days, until it reaches the desirable

---

<sup>2</sup>SI unit: measures luminous flux per unit area

schedule. The downside of this therapy is that the patients cannot have any work or social demands that conflict with the schedules of this therapy and longer length of time are needed to reach a successful result. Following this demanding therapy is often difficult for patients, as well as maintaining the new schedule. Finally, but not less important, behavioural strategies include a strict maintenance of the desired sleep schedules, avoidance of daytime naps, avoidance of stimuli in the evening (e.g. bright light, stimulants, physical exercise), regular exercise program and the maximisation of light exposure in the morning.

## 2.4 Overview

Sleep has different states, NREM and REM, distinguishable by different neural activity. NREM sleep has minimal mental activity, while REM sleep has high mental activity. A normal sleep pattern is composed of about four cycles of NREM/REM sleep, with REM corresponding to about 25% of total sleep time. Variations in different age groups are expected.

Sleep is one of the most important functions of the body, having a major role in six parameters:

1. Preservation of energy and promotion of anabolic processes;
2. Thermoregulation mechanisms;
3. Brain *detoxification*;
4. Promotion of immunologic responses;
5. Brain development and maturation;
6. Memory consolidation and brain plasticity.

Sleep deprivation is considered to be a stressful event with severe consequences to the organism, such as fatigue, reduced psychomotor performance, lack of attention, reduced reaction time and ability to form new memories, bad mood and irritability.

The most used diagnostic tools for sleep disorders are anamnesis, actigraphy and sleep logs, polysomnography, dim-light melatonin onset measurement and sleep related questionnaires. These tools present a high variability of format and parameters depending on the patient, the studied disorder and the brand or model of the equipment. The files generated by diagnostic

equipment can take up to 1 gigabyte of memory. The combined use of these tools, chosen appropriately depending on the clinical case, provides a more complete and a better understanding of the possible diagnosis.

The sleep-wake cycle can be explained by the interaction of the circadian rhythm with a homeostatic process. The circadian rhythm is generated endogenously, having the suprachiasmatic nucleus as main responsible structure. The most powerful *zeitgeber* is light, being the main circadian synchroniser, with the ability to delay or advance the circadian phase. The homeostatic process related with the sleep-wake cycle works as regulator of the sleep *debt*, correspondent to the feeling of need of sleep. A consolidated sleep can only be obtained if the circadian and homeostatic processes are in phase, according to this hypothesis. One's preference of sleep schedules can be distinguished in chronotype: morning, intermediate and evening chronotypes. If these preferences are too extreme, the person may suffer from circadian rhythm sleep-wake disorders, as they may not be aligned with the social or professional demands. This misalignment between the two has consequences like increased fatigue, stress and decrease of sleep quality.

Delayed sleep phase disorder is a type of circadian rhythm sleep-wake disorder, characterised by a significant delay in the sleep-wake cycle. These patients present normal sleep parameters when they are allowed to sleep in their preferred schedule (a delayed one, compared to the socially acceptable standard). Although DLMO is considered to be a good marker for DSPD, the results of recent studies raise doubts as to its importance. The most common treatments for this disorder include phototherapy, melatonin and behavioural adaptation, preferably with combined use.

# 3

## Supporting technologies and tools

This chapter addresses the requirements for clinical databases and web applications. Section 3.1 approaches data requirements for clinical applications, including the guidelines demanded by the Portuguese data protection authority. Section 3.2 covers the software requirements for clinical applications, with critical comparisons of the possible used tools. Finally, an overview of the chapter is available in Section 3.3.

### 3.1 Data requirements for clinical applications

Before designing and building a clinical platform, it is essential to understand its data requirements. The FAIR Guiding Principles, developed by FORCE11 - The Future of Research Communications and e-Scholarship, presents a set of principles to facilitate the discovery, access, integration and analysis of scientific data. FAIR stands for Findable, Accessible, Interoperable and Reusable, which are the characteristics scientific data should have to support both machines and humans.<sup>1</sup>

The FAIR data principles correspond to a set of rules data should follow in order to have the characteristics mentioned :

1. To be Findable:
  - (a) (meta)data are assigned a globally unique and eternally persistent identifier;
  - (b) data are described with rich metadata;
  - (c) (meta)data are registered or indexed in a searchable resource;
  - (d) metadata specify the data identifier.
2. To be Accessible:
  - (a) (meta)data are retrievable by their identifier using a standardised communications protocol;

---

<sup>1</sup>FORCE11: Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing version b1.0 - <https://www.force11.org/fairprinciples>

- (b) metadata are accessible, even when the data are no longer available.
3. To be Interoperable:
- (a) (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation;
  - (b) (meta)data use vocabularies that follow FAIR principles;
  - (c) (meta)data include qualified references to other (meta)data;
4. To be Re-usable:
- (a) meta(data) have a plurality of accurate and relevant attributes;
  - (b) (meta)data are released with a clear and accessible data usage license;
  - (c) (meta)data are associated with their provenance;
  - (d) (meta)data meet domain-relevant community standards.

According to Wilkinson et al. [2016], following these principles ensures a better data management, that depends not just on the proper collection, annotation and archival of data, but also the long-term care of that data. The author adds that one can choose which principles to follow, if not all, giving data different degrees of *FAIRness*. Another feature of the FAIR guideline is the distinction between metadata and data, which supports special circumstances. One example of these special circumstances given by Wilkinson et al. is highly sensitive or personally-identified data, where *publication of rich metadata to facilitate discovery (...) provides a high degree of ‘FAIRness’ even in the absence of FAIR publication of the data itself*. Clinical data is considered to be sensitive and measures to ensure its safety are extremely important.

As SleepData aims to be a platform that integrates and handles clinical data, as well as serve as means to study and characterise disorders, ensuring that the data is FAIR seems, almost, a demand. By following the FAIR data principles, Sleep Data can be built in a more robust, safe and useful way. Thus, the choices and decisions made to build this platform had these principles in consideration, as it will be shown in this chapter.

### 3.1.1 Standard frameworks for health information systems

Taking a closer look at the FAIR principles, we see that one of the requirements for data regarding reusability is that it follows domain-relevant community standards. In fact, by analysing the principles, one can see that the use of standards can easily help accomplish many other

requirements, like ensuring accessibility of the data. According to Hammond et al. [2014], standards are crucial to biomedical applications, due to the enormous diversity of data. If one chooses to ignore the use of standards, health information systems can easily become a set of unorganised, unconnected and independent units, making it far more difficult to assess a patient or a population as a whole.

Health Level Seven International (HL7), a standards development organisation, was founded in 1987 with the purpose of developing standard protocols for electronic health information exchange that support clinical practice and management. HL7 counts with the support of members from over 50 countries and representatives from several health related stakeholders. One of the standards created by HL7 for Clinical and Administrative Domains is HL7 Fast Healthcare Interoperability Resources Specification (FHIR)<sup>2</sup>. The FHIR standard for medical data structure and transfer uses modular components, called *Resources*, that can be assembled into working systems. Each resource presents a set of elements that compose it, how that resource is structured and the type and description of each variable. FHIR organises the Resources in different categories. The most relevant for SleepData are Clinical and Base resources. Clinical resources include diagnostic reports, observations, questionnaire responses, procedures, medication, while Base resources include patient, practitioner and appointments. A list of FHIR's resources is available on their website. FHIR structures each resource in a way that it is connected to other related resources. FHIR has other important and relevant features that make it a good choice for a standard protocol, such as:

- Fast and easy implementation;
- Multiple implementation libraries;
- Adaptable for local requirements;
- Strong foundation in Web standards (XML, JSON, HTTP);
- Support for RESTful architectures;
- Human-readable serialisation format;
- Common metadata;
- Extensible framework to support variation in healthcare.<sup>3</sup>

---

<sup>2</sup>Health Level Seven International (Consulted: 29/08/2017)- <http://www.hl7.org>

<sup>3</sup>HL7 FHIR Release 3 (Consulted: 29/08/2017) - <https://www.hl7.org/fhir/index.html>



Other organisations, like openEHR<sup>4</sup> develop standards for integration and analysis systems of clinical data. However, openEHR does not address protocols for information exchange.

### 3.1.2 Standardisation of medical nomenclatures

Wilkinson et al. [2016] propose the adoption of structured standards for clinical terms and codes to ensure that data interoperability, one of the FAIR principles for clinical data. The authors claim these standards serve two purposes:

- Saving developers from having to create their own nomenclatures
- Facilitating the exchange of clinical data among systems

One of the most used standards is SNOMED CT<sup>5</sup>, used in more than fifty countries. SNOMED CT is a multilingual clinical healthcare terminology, mapped to other international standards, like LOINC (which will be address later) and the International Classification of Disease. Its content is scientifically validated and facilitates the representation of clinical content in electronic health records. Each of SNOMED CT's concepts have an unique Fully Specified Name (FSN) and an unique numeric identifier, as well as formal definitions and synonyms. There is an hierarchical relationship between the concepts, identified by a semantic tag. One of the main features of this terminology is the vast set of classes of clinical terms it has, like body structures, clinical findings, geographical locations, event, organisms, pharmaceutical products, procedures, social context, substances. A comprehensive detailed list is available in SNOMED CT's web browser<sup>6</sup>. Adopting SNOMED CT as a nomenclature standard for clinical terms ensures that clinical information is recorded unambiguously, clearly and consistently, allowing accurate and comprehensive searches and removing language barriers. SNOMED CT contains a broad coverage of medical terms. To cover laboratory medicine and result reporting, SNOMED CT has recently incorporated LOINC<sup>7</sup>, or Logical Observation Identifiers Names and Codes.

In LOINC, each test, measurement, survey question, clinical document or observation, that have different clinical meaning, has an unique code. LOINC distinguishes each codable concept across six dimensions:

- Component, which is the substance or entity in question;

---

<sup>4</sup>openEHR (Consulted: 10/09/2017) - <http://www.openehr.org/>

<sup>5</sup>SNOMED International: SNOMED CT (Consulted: 30/08/2017) - <http://www.snomed.org/snomed-ct/what-is-snomed-ct>

<sup>6</sup>SNOMED CT: Web browser - <http://browser.ihtsdotools.org>

<sup>7</sup>LOINC: Get Started (Consulted: 31/08/2017) - <https://loinc.org/get-started/>

- Property, relative to the characteristic of the component;
- Time over which the observation was made;
- System upon which the observation was made;
- Scale of the component;
- Method used, when applicable;

This way, an unique code identifies with accuracy each test or result.

Other nomenclatures like the International Classification of Diseases<sup>8</sup> or ICSD are not as vast as SNOMED CT and LOINC. Even though ICSD is intended for sleep disorders, there are many terms used in clinical practice that this nomenclature does not include. Additionally, SNOMED CT already includes many terms present in other nomenclatures.

### 3.1.3 Confidentiality, privacy and data security

Goodman et al. [2014] advocate that developers of biomedical applications have the obligation to take into consideration the moral and ethical aspects related to they project. Among these aspects, the following deserve attention:

- Confidentiality;
- Privacy;
- Appropriate determination of who has access to the data;
- Appropriate selection of tools and how they can be used.

The author adds that balancing access to information and protection of patients' privacy is one of the main challenges, as making data easily accessible poses a threat to the appropriate use of clinical data. For example, there is information in an electronic health record that is not necessary for administrative purposes, like billing, or other health associated tasks. Unnecessary access to information should be restricted, to avoid harming the patients, or even to feed someone's curiosity. Goodman et al. [2014] go on reminding that people do not need to provide a motive for limiting access to their health data, as their privacy and data confidentiality are regarded as

---

<sup>8</sup>WHO - International Classification of Diseases (Consulted: 07/10/2017) - <http://www.who.int/classifications/icd/en/>

rights. If one wants to point out further reasons to protect privacy and confidentiality, we can look at how it benefits both individual and society. On the one hand, individuals that trust that their data are secure are more likely to be comfortable sharing clinical details with the clinicians, hence contributing to a better informed care, helping both practitioners and patients. On the other hand, if people do not trust their data to be safe, they will likely not seek professional assistance, posing a major threat to public health by increasing the risks of contagion. Data security not only avoids discrimination, bias and stigma, but also the misuse of data by third parties, like insurers. Measures to ensure data security can be divided in technological methods, like encryption and authentication (discussed in the next session), and policy approaches, which will be addressed next.

To make sure SleepData protects the interests and rights of the patients, the guidelines given by Portuguese Data Protection Authority called Comissão Nacional de Protecção de Dados (CNPd), regarding the use and treatment of personal data for Clinical Investigation were carefully analysed. The complete document, Deliberation number 1704/2015, was last reviewed in 2015 and is available online <sup>9</sup>. The security policies imposed by CNPD are the following:

1. The information system should:
  - (a) Ensure logical separation between health related data and the remaining personal information;
  - (b) Allow different data access levels, distinguishing clearance and privilege levels according to the users profile;
  - (c) Require passwords for user authentication;
  - (d) Update and delete user's profiles, when the privileges of the user change;
  - (e) Enforce frequent password changes.
2. Measures to prohibit access to information to unauthorised people should be taken;
3. The access to clinic records for clinical, epidemiological and genetic research requires the pre-existing treatment of personal data properly authorised;
4. Transmission of health information should always be encrypted;
5. Physical and logic access to the system's servers should be restricted and a logbook of the access should be maintained;

---

<sup>9</sup>CNPd - Deliration number 1704/2015 (Consulted: 31/08/2017) - <https://www.cnpd.pt/bin/orientacoes/orientacoes.htm>

6. Backups of the data should be made frequently and should be kept in places only accessible to the system’s administrator, or authorised personnel, bound by professional secrecy;
7. When the collection of information is not made by the health professional, measures to ensure the safe circulation of the information should be taken, to avoid data visualisation by unauthorised individuals;

CNPD policies add that everyone who, while performing their duties, gets access to the personal data treated is compelled to keep professional secrecy.

SleepData should follow the policies recommended by the data protection authority, in order to ensure the data inserted into the platform is safe and that the rights of the patient’s involved are protected.

## 3.2 Software requirements for the SleepData platform

### 3.2.1 Database management system

To create a clinical platform, it is necessary to give attention to the way the data will be organised in a database. Hence, a database management system (DBMS) must be chosen according to the requirements of such platform, in order to administer the database, create new entries, query the existent information, update and delete information.

The first step to choose the appropriate DBMS, is to opt between two main categories: relational or document-oriented. Relational DBMS are structured, using a *collection of tables to represent both data and the relationships among those data*, as described by Silberschatz et al. [2011], while document-based DBMS have no defined structure and are object-oriented. In particular, a relational database has a very rigid structure – a table can only store the information and data type that the design predicted – in contrast with document-oriented type which is much more flexible – the information and data type can be stored in any document freely, according to Silberschatz et al. [2011] and Vaish [2013]. Table 3.1 shows other significant and important features that distinguish these types of DBMS.

SleepData will process large amounts of data, with varying data formats depending not only on the patients’ cases but also the doctors’ assessments. SleepData aims to be scalable and

Table 3.1: Feature comparison between relational and document-oriented DBMS. Adapted from Mohamed et al. [2014]. and Vaish [2013]

Feature	Relational DBMS	Document-oriented DBMS
Data model	Schema must be defined before inserting data	No predefined schema. Has the ability to store unstructured, semi-structured and structured data
Data structure	Data clearly defined by their relationships	Both structured and unstructured data
Scalability	Very difficult to scale: limited by hardware resources (very costly and impractical)	Easily scalable: runs well on distributed systems (cloud)
Transactional reliability	Very high	Low
Complexity	High complexity when dealing with big amounts of data	Low complexity
Speed (performance)	Slower due to the number of joins needed for a query or update	Faster (typically, the information about an entity is in one single record)
Development time	Large time investment to design model and complex queries	Little to no time investment needed to design model or queries

the platform will not require complex transactions. Additionally, SleepData will store both unstructured (like clinical notes manual reports) and structured data (like automatically generated reports and laboratory results). Hence, a document-based DBMS is the type that best suits the requirements. In fact, Peek et al. [2014] claim that relational DBMS are not chosen for platforms dealing with big amounts of data because the volume, heterogeneity and complexity of the data make it inappropriate to fit into a predefined schema.

Looking into the many options available, MongoDB came to our attention not only as the most popular open source document-based DBMS, but also as a system that fits the requirements of SleepData. MongoDB’s characteristics are as follows:

- Flexibility of the data storage, in JSON-like documents (see Figure 3.1): each document can vary in terms of fields, according to the need;
- Easy access to data, provided by ad-hoc queries, indexing and real time aggregation;
- Distributed database, allowing high availability and horizontal scaling;
- Document model simple to learn and use;
- Has drivers for more than 10 programming languages.<sup>10</sup>

---

<sup>10</sup>MongoDB: Architecture - <https://www.mongodb.com/mongodb-architecture>

Although popularity is not a criterion itself, it often means that there is a big amount of documentation, tutorials and help-guides available online, which was verified to exist. This type of technical support, not only produced by MongoDB but also by current and former users, make it easier to have a deep understanding of the system and, thus, enables a consistent construction and use of the database. Another point in favour of MongoDB is the fact that many big companies are using it, like CERN, Cisco, Facebook, eBay, KPMG, MetLife and Barclay's, and including health care and medical companies, like the medical equipment manufacturer Medtronic and the pharmaceutical giant AstraZeneca, among others.<sup>11</sup> This fact can attest MongoDB's trustworthiness. Additionally, FHIR, the chosen standard protocol, has direct integration with MongoDB.

To implement the architecture of MongoDB's data model, it is important to understand just a few basic concepts:

- Document: JSON file with a set of key-value pairs, each key is called a field;
- Collection: group of documents with similar or related purpose. Documents in the same collection can have different fields;
- Database: container of collections.

Figure 3.1(top) represents the structure of a database in MongoDB, with the elements mentioned above. When a document is generated, if an unique ID is not set by the developer, MongoDB automatically generates an *object ID* that unequivocally identifies the document. In case there is related data, one can choose to add that data in the same document, as an embedded sub-document, instead of creating another document in another collection, with a reference to connect the two documents.<sup>12</sup> In Figure 3.1(bottom) , which represents two different samples of documents, the field 'contact' is an example of an embedded sub-document. In this case, both documents (left and right) represent similar information, for example, basic information of a patient, so they should be included in the same collection, even if they have different fields or structure. Relating with Figure 3.1(top), the sample on the left could be Document DB.A.1 and the one on the right could be Document DB.A.2.

---

<sup>11</sup>MongoDB: Industries - <https://www.mongodb.com/industries>

<sup>12</sup>MongoDB: Manual - <https://docs.mongodb.com/manual/>

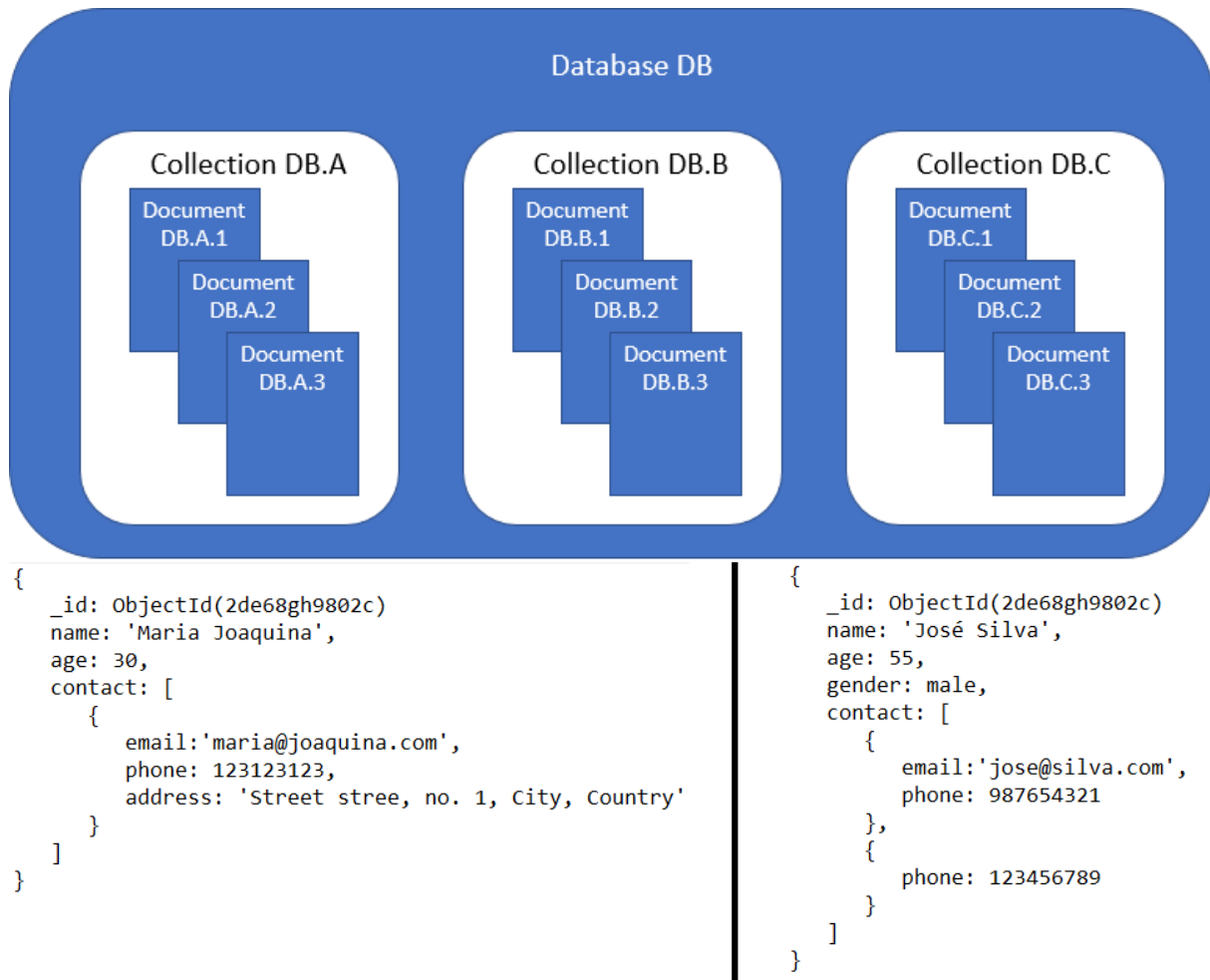


Figure 3.1: MongoDB structure: sample of a MongoDB database (top) and two samples of similar documents in MongoDB (bottom)

### 3.2.2 Web server

The SleepData database will be accessible to the end user through a website, which needs to be hosted by a computer that is connected to the internet at all times. The computer has a few essential requirements that depend on the extent of the project itself, like enough processing speed, large storage and memory capacity. Independently of the specifications of the computer, it must always provide a network interface with an IP address so that the website can be found. The web server is the software that handles the requests of client applications.<sup>13</sup>

Once again, a web server program had to be chosen from a set of open source solutions. From this set, the most used web server is Apache. Another alternative is building a web server with Node.js, a server-side JavaScript runtime built on Chrome's V8 JavaScript engine, since it

<sup>13</sup>WebDevelopersNote: What is a web server? (Consulted: 01/07/2017) - <http://www.webdevelopersnotes.com/what-is-web-server>

was mentioned several times for applications that use MongoDB, the database server selected previously. The main difference between the two is that Apache has a synchronous architecture, while a web server built with Node.js has an asynchronous one. Apache handles each request by a separate thread and blocks that thread until its I/O processing is done. In contrast, a Node.js web server runs every request in a single main thread with an event loop which waits on the completion of I/O subthreads. It cannot be said that one architecture is better than the other, as both have advantages and disadvantages. According to Lei et al. [2014], a Node.js web server can handle far more simultaneous requests due to its asynchronous, non blocking, event-based behaviour, making it the best choice for handling dynamic content and multiple users. The downside when handling all the requests in a single thread is that if something goes wrong the whole instance will crash. In contrast, as Apache handles each request in a separate thread, in case of a problem only that part will crash<sup>14</sup>. Despite this, Node.js is still a better solution for SleepData, since its design optimizes scalability in web applications. In fact, the performance study by Chaniotis et al. [2015], concludes that *Node.js offers client-server development integration, aiding code re-usability in web applications, and is the perfect tool for developing fast, scalable network applications*. Another detail in favor of using Node.js to build a web server is that some of Node.js's top members include Microsoft, IBM, Paypal, Google and Intel, which take an active voice in the maintenance of the community<sup>15</sup>.

Although Node.js is not as popular as Apache, and thus, does not have as much documentation, it was found that it is easy to use and that there are more than enough tools available online to understand and use this technology at its full potential. Node.js uses JavaScript as its main language, which is fairly easy to learn, and it has one of the biggest ecosystems of open source libraries, named *npm*<sup>16</sup>. *npm* consists of a free and shareable repository of JavaScript packages. These packages (or modules) are pieces of code developed to solve specific problems. *npm* makes it possible to download, reuse on any project and check for updates on any of these packages. This type of structure also makes it easy to manage different parts of the code. *npm* includes both packages that can be used on the server-side and packages that can be used on the front end.<sup>17</sup>

---

<sup>14</sup>Quora: What are the pros and cons of Node.js versus Apache web server? (Consulted: 01/07/2017) - <https://www.quora.com/What-are-the-pros-and-cons-of-Node-js-versus-Apache-web-server>

<sup>15</sup>Node.js: Members - <https://nodejs.org/en/foundation/members/>

<sup>16</sup>Node.js - <https://nodejs.org/en/>

<sup>17</sup>npm: documentation - <https://docs.npmjs.com/>



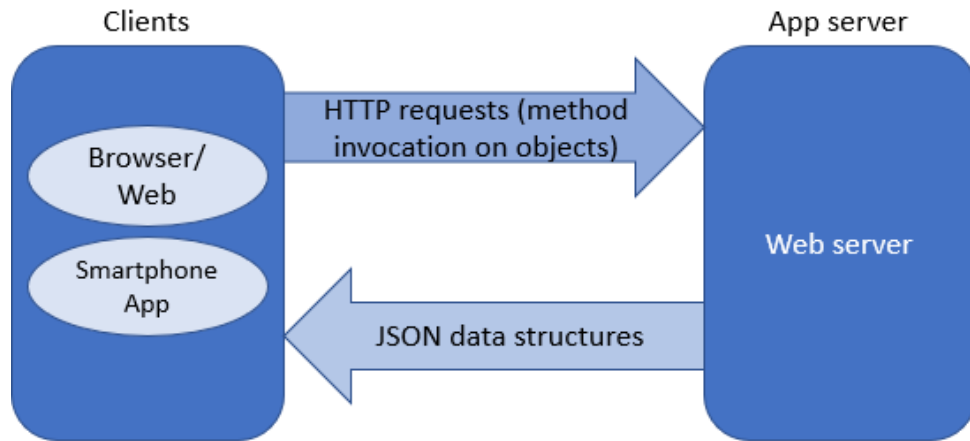


Figure 3.2: REST API Design

### 3.2.3 Web framework

A web framework supports the development of web applications, by providing a set of features commonly required to build any web application programming interface (API), like URL routing, database manipulation, templating and security.<sup>18</sup>

A web API is a source code interface with a set of functions that allow the access to features or data of a web application using HTTP protocol. Nowadays, REST is the most used API style in the world wide web. REST stands for Representational State Transfer and it is an architectural style that works by taking advantage of HTTP methodologies. The API sends GET, PUT, POST or DELETE requests, to retrieve, change or update, create and remove data objects on the server, respectively. Figure 3.2 shows how a RESTful API allows the communication between client and server.<sup>19</sup>

*Express*<sup>20</sup>, available from *npm*, is a light-weight, server-side web framework for Node.js, that simply extends the server's abilities without covering any of its previous features. One of the main feature this framework offers is routing, which defines how the server responds to a request from the client. Express has several HTTP utility methods that allow the design of a RESTful API<sup>21</sup>. Other middlewares (which are functions that deal with requests) that *Express* adds to Node.js are session, fundamental for authentication processes, file upload and static file server middlewares.

<sup>18</sup>WebArchive: Web application framework (Consulted: 04/07/2017) - [https://web.archive.org/web/20150723163302/http://docforge.com/wiki/Web\\_application\\_framework](https://web.archive.org/web/20150723163302/http://docforge.com/wiki/Web_application_framework)

<sup>19</sup>TechTarget: RESTful API (Consulted: 01/09/2017) - <http://searchcloudstorage.techtarget.com/definition/RESTful-API>

<sup>20</sup>Express - <https://expressjs.com/>

<sup>21</sup>Express: Routing - <https://expressjs.com/en/guide/routing.html>

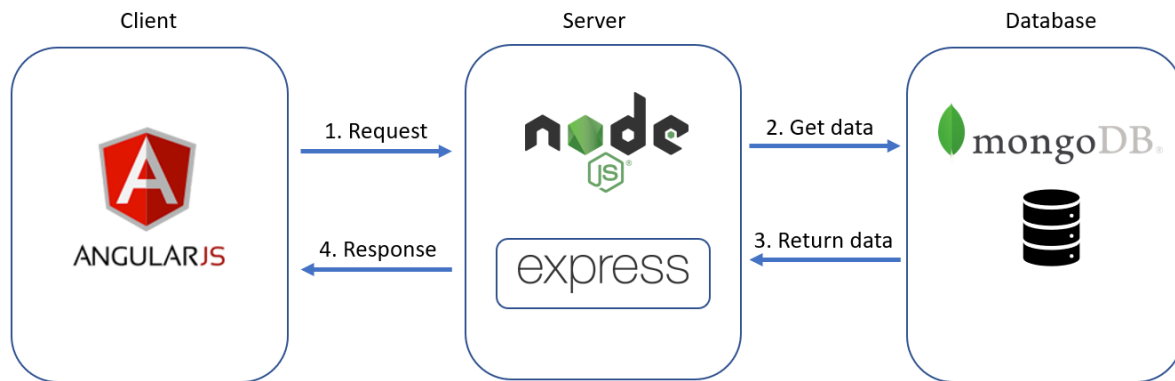


Figure 3.3: MEAN stack architecture

### 3.2.4 MEAN architecture

The elements above fit together and interact in an integrated platform. The combined use of Node.js, Express and MongoDB represents what is called the MEAN<sup>22</sup> stack architecture (see Figure 3.3). The MEAN stack is a full-stack JavaScript framework that stands for **M**ongoDB, **E**xpress, **A**ngularJS and **N**ode.js, and has all the components needed to build a web application. AngularJS is a front-end framework that runs in browser JavaScript engines, with the purpose to extend traditional HTML to present dynamic content. However, there are a few downsides to this technology, like a steep learning curve, its complexity and complicated debugging<sup>23</sup>.

### 3.2.5 Security and encryption

As required by the CNPD, the transmission of data should always be encrypted. In a system like the one we are creating, that means that the connection between client and server must be secured.

To ensure this security, SleepData should work under a HTTPS protocol. This means that the communication over HTTP is done under a connection encrypted by a secure sockets layer (SSL). The SSL connection uses a pair of keys: one public and one private, with the particularity that one can only be decrypted by the other. When a browser tries to access a website secured by SSL, it establishes a connection with the server to verify if the two keys match or not. In case they do, a third key is generated, called a session key, used to encrypt the transmitted data. This way the SSL connection provides encryption of the exchanged data between client and server, ensuring that no other connection is *listening* to the data and that the data is not

---

<sup>22</sup>MEAN - <http://mean.io/>

<sup>23</sup>DZone: The Cons of AngularJS (Consulted: 05/07/2017) - <https://dzone.com/articles/cons-angularjs>

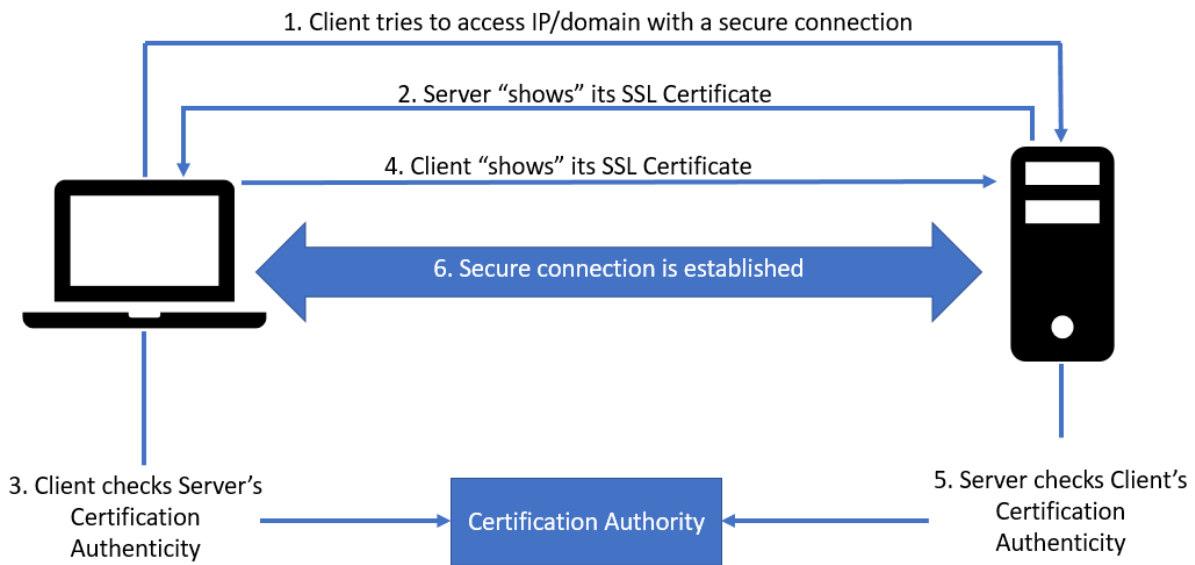


Figure 3.4: SSL verification protocol

modified or corrupted in the connection. To create the keys, a SSL certificate is needed. Anyone can create one, but browsers only trust the certificates created by a Certificate Authority. Using a certificate from a Certificate Authority provides authentication of the website, so that users know they are accessing the authentic, secure website they intend to communicate with.<sup>24</sup> Figure 3.4 represents the SSL protocol process.

### 3.3 Overview

Clinical data follow the FAIR principles in order to facilitate the discovery, access, integration and analysis of scientific information, as well as to ensure proper data retrieval, annotation, archival and long-term care of the data. To ensure the data has a high degree of *FAIRness*, the use of standards is necessary. HL7 has a complete standard protocol for information exchange, called FHIR, presenting a large set of modular components. Each component presents a data structure, including the type and description of each element. The use of FHIR ensure the data is more accessible and re-usable. Standards for medical nomenclatures are also important to make that FAIR, contributing to the interoperability and findability of the clinical data. SNOMED CT and LOINC are two of the most used nomenclatures, the first with focus on all types of clinical terms, the other with focus on laboratory medicine and result reporting. Both nomenclatures provide unique codes to each different term, ensuring data is unambiguously and clearly recorded. LOINC is incorporated in SNOMED CT, facilitating the combined used of the two.

<sup>24</sup><https://www.digicert.com/ssl/>

Ensuring the confidentiality, privacy and data security aspects of the data is a moral obligation of the developers of any clinical platform. Protecting a patient's right to privacy and confidentiality brings benefits both for the individual and the society, by enabling better care for the patient and better public health for the society. The main guidelines regarding the use and treatment of personal data for Clinical Investigations, provided by CNPD to ensure data security include:

1. Ensuring logical separation between health and personal data;
2. Restricting access to different levels of data privileges;
3. Requiring passwords for authentication;
4. Transmission of data should be encrypted;
5. Physical and logical access to the servers should be restricted;
6. Regular backups should be made;
7. Measures to ensure safe circulation of data should be taken.

MongoDB, an open source document-based DBMS, has the following properties:

1. Appropriate to store both structured and unstructured data;
2. Easily scalable;
3. Appropriate to deal with big amounts and different formats of data;
4. Low complexity;
5. Short development time needed.

Further advantages of using MongoDB are the highly available documentation and tutorials, that facilitate its use. Node.js, a server-side JavaScript runtime environment, has one of the biggest free and shareable repository of JavaScript packages that facilitate the resolution of specific development problems, both on the server and client sides, called *npm*. One of those packages is *Express*, a web framework. Together with AngularJS, these three technologies combined compose the MEAN stack architecture.

To ensure the security of the data, communication between server and client over HTTP should be done under an encrypted connection using the SSL protocol, which is the same as

saying it works under a HTTPS protocol. By acquiring a SSL certificate from a Certificate Authority, the users/browsers know that no other connection is *listening*, that the data is not modified or corrupted and that the website is authentic.

# 4 SleepData

This chapter describes the different steps involved in the conceptualisation and development of the clinical platform for sleep disorders, starting with the platform infrastructure, including the steps to ensure data security and encryption (Section 4.1). Section 4.2 details the data modelling and what type of information is stored in SleepData, followed by the specifications of user interface and user experience in Section 4.3, with examples of SleepData’s interface. Finally, Section 4.4 summarises the key aspects of the design of SleepData.

## 4.1 Platform infrastructure

SleepData has an architecture similar to the MEAN stack (see Figure 4.1): MongoDB was chosen as the database management system, Node.js was used to build the web server and *Express* was chosen as the web framework. Considering the time available to build SleepData, it was decided not to use AngularJS and to stick to simple JavaScript and HTML.

SleepData is currently at one of FCCN/FCT’s<sup>1</sup> data centers through the following domain: <https://sleepdata.inesc-id.pt>. This allows SleepData to be accessible to any device with an internet connection, whether it is a laptop, tablet or smartphone.

We have obtained a signed SSL certificate from DigiCert<sup>2</sup>, a Certificate Authority recognised by the most popular browsers. DigiCert Certificates are standard x509 certificates, a widely accepted international standard. This certificate ensures that the server a user connects to is indeed SleepData. This certificate can be verified when accessing the platform. When someone accesses SleepData, the browser informs the user on the security status with a lock icon or a green bar, like the one in Figure 4.2.

In this project, many *npm* packages were used to ensure that SleepData had the desired features, but the most important ones are the following:

- *mongodb*, which is the official MongoDB driver for Node.js;

---

<sup>1</sup>FCCN - <https://www.fccn.pt/>

<sup>2</sup>DigiCert - <https://www.digicert.com/>

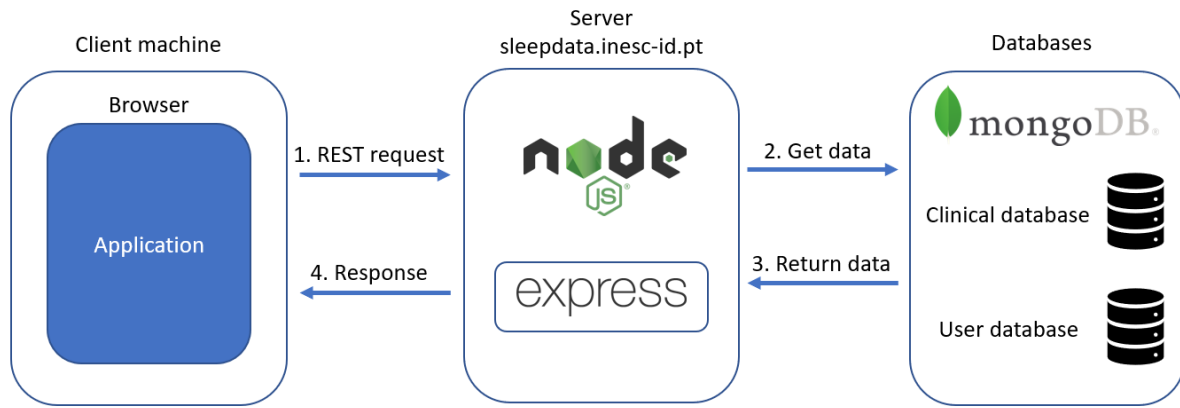


Figure 4.1: SleepData's software architecture

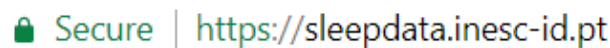


Figure 4.2: SSL Certificate symbol: secure connection for SleepData platform

- *express*, which is a minimalist web framework;
- *passport* to handle user authentication;
- *https* to handle HTTPS requests;
- *python-shell* to run Python scripts.

As one of the purposes of this platform is to integrate the information collected from sleep medicine clinics along with data directly input by users, it is necessary to enable uploading exams and medical reports to the platform. For that reason, Python scripts were written with the objective of handling the data received and parse it into the JSON format needed in the database. The package *python-shell* was crucial to run those scripts on the server side. Any other language could have been used. Python was chosen due to previous knowledge of the language and due to the amount of documentation available to aid the development of the scripts. SleepData currently has parsing methods for the reports exported by two models of actigraphs: Philips' Actiwatch2 and Condor's ActTrust. After parsing, the original files still kept in the database for future reference.

FHIR Resources were used as a base to structure SleepData information entries, giving data re-usable and accessible attributes, thus ensuring data is more FAIR. Regarding standard nomenclatures, we chose to use SNOMED CT and LOINC. The combined use of SNOMED CT and LOINC, by encoding the clinical terms used in SleepData whenever possible and deemed necessary, ensures that the language used in the platform is clear, unambiguous and consistent,

allowing data to be not only interoperable, but also findable, contributing to a higher degree of *FAIRness* of the data.

## 4.2 Data modelling

As a database in MongoDB is comprised of collections of documents, the data modelling of SleepData consisted in specifying the JSON structures for each main concept. Two databases were designed: a clinical database and a user database. The clinical database designed to store data about the patients (personal and clinical information, including exams and reports) and the clinical staff. The design of this database took into account not just the requirements mentioned previously and in the literature concerning sleep disorders, but also CENC's organisation. Another database was designed to manage user authentication. This way, user information and personal and clinical information is kept logically separated.

### 4.2.1 Clinical database

The database for clinical information was designed taking into consideration the several types of files that sleep medicine clinics produce for each patient and the recommendations of HL7's FHIR for modelling clinical resources, such as patients and observations. For each concept, a collection was created to assemble the documents for each instance of that concept. Several document types were structured as FHIR Resources<sup>3</sup>, with their content adapted to fit SleepData needs. Table 4.1 shows the collections designed for this database and the correspondent FHIR resource, as well as the type of information each collection represents. Besides the collections mentioned in Table 4.1, a collection for each type of questionnaire was also designed, based on the resource QuestionnaireResponse.

Every document, independently on the Resource it was based on, includes the following information:

- Type and category of document;
- Unique identification code, generated by MongoDB;
- Status (final, corrected, preliminary);

---

<sup>3</sup>FHIR: Resource Index (Consulted between March and June 2017) - <https://www.hl7.org/fhir/resourcelist.html>



Table 4.1: SleepData’s clinical database collections and their mapping to FHIR resources

Collection	FHIR’s resource	Description
Patient	Patient	Patient’s data and contact information
Professional	Practitioner	Professional’s data and contact information
ProfessionalRole	PractitionerRole	Information on the role of each professional and their schedule
ClinicalNotes	Observation	Parameters collected in clinical notes, with the exception of medication
Medication	Medication	Medication a patient takes, doses and active ingredients
Actigraphy	Observation	Actigraphy parameters from the automatically generated report
ActigraphyReport	Observation	Parameters of the report produced by the clinical staff concerning actigraphy results
PSG	Observation	Parameters of PSG observed by the clinical staff concerning PSG results
PSGReport	Observation	Conclusions taken out of the analysis of PSG and night diary
NightDiary	Observation	The diary produced by the clinical staff during a PSG
DLMO	Observation	Parameters of the DLMO exam

- Date and time: identifying when it was inserted in SleepData and when it was effectively produced;
- Unique identification number of the professional responsible for the document;
- Unique identification number of any related document (for example, an actigraphy report document, must be related with an actigraphy document);
- Unique identification number of the patient the document corresponds to, in case this is justified;
- Laboratory responsible for the exam, in case it was conducted by an external entity;
- Coding system for every field or exam, when possible.

An example of part of an actigraphy report document is shown in Figure 4.3. The requirements for clinical databases mentioned in the previous chapter were taken in consideration. For example, personal information and clinical information are kept on separate collections, providing a logical separation between the two, to protect the best interests of the patient, as required by CNPD’s guidelines.

The exact information registered in each collection is specified in Appendix A.

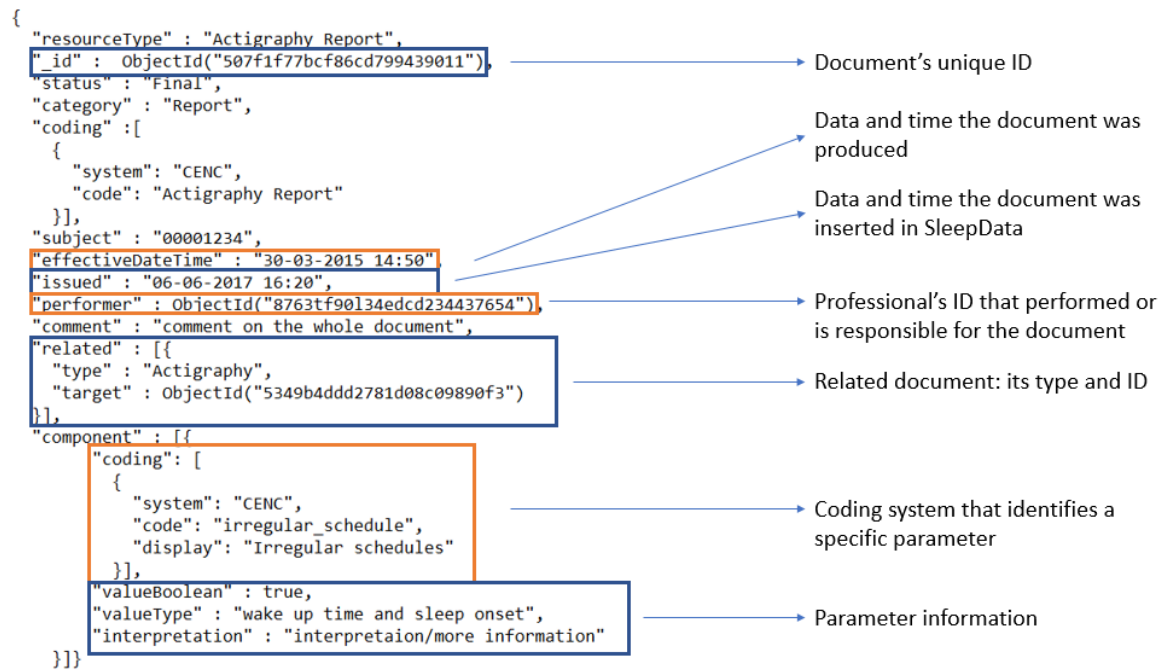


Figure 4.3: Actigraphy report sample with some explained parameters (all other are self explanatory)

## 4.2.2 Nomenclatures and coding systems

SNOMED CT and LOINC cover all the clinical related nomenclature. However, it is common for SleepData clinical records to include terms outside the clinical domain. For that reason other coding systems were used as well:

- International Standard Classification of Education<sup>4</sup> (ISCED2011, by UNESCO): to register the education level of a person;
- International Standard Classification of Occupations<sup>5</sup> (ISCO, by the International Labour Organization): to register the occupation (job) area
- Drinking levels by the National Institute on Alcohol Abuse and Alcoholism (NIAAA)<sup>6</sup>

Some parameters do not fit into any of the existent coding system. For that reason, a new coding system for CENC was created with those parameters, each with a different code and specific value, when justified (see parameter *irregular schedules* in Figure 4.3 as example).

<sup>4</sup>ISCED2011 (Consulted: 23/09/2017) - <http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf>

<sup>5</sup>ISCO (Consulted: 23/09/2017) - <http://www.ilo.org/public/english/bureau/stat/isco/>

<sup>6</sup>NIAAA: Drinking levels (Consulted: 23/09/2017) - <https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/moderate-binge-drinking>

### 4.2.3 User database and access control system

A database separate from the clinical database was created to store the information of each user account. This database is composed of one collection where each document represents a user account. The user database is essential for data privacy and security, as it is important to ensure access control in SleepData. *npm*'s package *passport*<sup>7</sup> was essential to the creation of different types of accounts with usernames and passwords. The *passport* package stores the encrypted password of each account, so that it cannot be accessed directly from the database. This way, each document on the user database stores the username, encrypted password and account/permission type. When logging in to SleepData, the server checks if the account exists and its permission type, in order to redirect the user to the correct webpages. The level of information given depends on the user's permission type. The platform distinguishes its users by giving them accounts with different types of permissions. There are three types of user accounts: Regular User, Professional User and Administrator. The login flow is shown in Figure 4.4. Table 4.2 specifies what actions each type of account can take. As shown, Regular User account is intended for patients or anyone who wants to contribute to the growth of SleepData as a repository of sleep medicine data by answering the available questionnaires. Professional User account is intended for professionals of the clinics that want to use SleepData, to manage clinical information about their patients. Administrator account is intended for administrative purposes only.

Since the purpose of SleepData is to hold data from multiple clinics, it is possible to distinguish Professional User accounts from different clinics. That way, every Professional User can perform the same actions, but only related to patients associated to that clinic (i.e. Professional User from Clinic A can only manage patients from Clinic A). Another important aspect is that Regular User accounts can be associated to a clinic by adding a patient identification field to the Regular User document in the database. This is only possible due to MongoDB's semi-structured information model. If the Regular User account document has a "ClinicA\_id" field, then the user is a patient from Clinic A.

An alternative to this access control system would be using MongoDB's Role-Based Access Control feature, which determines the user's access to database resources. However, we chose to control the access to the data using the *passport* package mentioned above.

---

<sup>7</sup>Passport - <http://passportjs.org/docs>

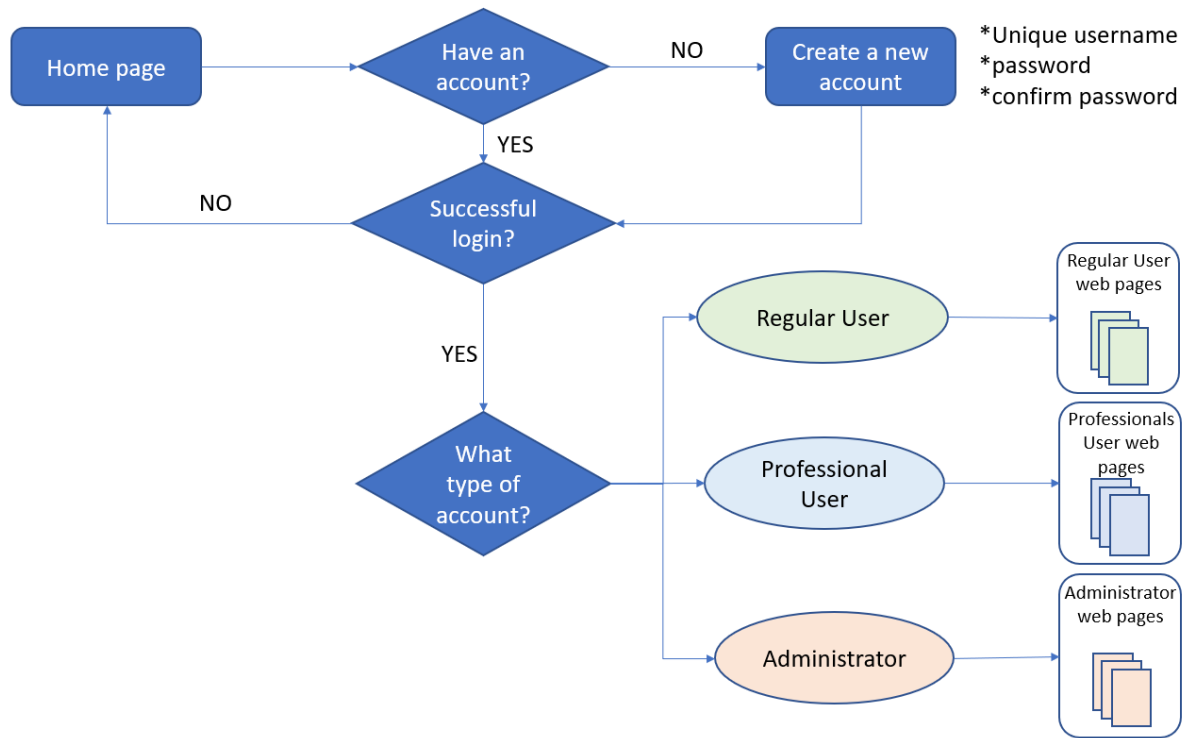


Figure 4.4: Login flow

Table 4.2: Types of accounts and their permissions.

Permission	Regular User	Professional User	Administrator
Manage personal information	✓	✓	✓
Answer questionnaires	✓	✗	✗
Manage clinical information of associated clinic	✗	✓	✗
View patients' statistics of associated clinic	✗	✓	✗
View patients' dashboard information of associated clinic	✗	✓	✗
Associate a Regular User account to a patient of associated clinic	✗	✓	✗
Manage permissions of any account	✗	✗	✓
Manage professionals and their roles	✗	✗	✓
Manage clinics	✗	✗	✓

## 4.3 User interface and user experience

The design of the user interface must take in to consideration practical clinical aspects, such as the way the information collected during a consultation is registered, both for the physicians and the patients, since the goal is to aid and support, and not to interfere or overburden the both parties. Therefore, the interface was built as intuitive as possible, taking into consideration CENC's personnel advices.

### 4.3.1 External dependencies

Several front-end external dependencies were used to provide the intended features for the user interface. jQuery and other JavaScript libraries were used for event handling, which means they allow responses to be activated upon the selection of an element.<sup>8</sup> For example, when filling in information about the duration of a patient's sleep cycles it is possible to add more cycles, but this option is only available if the user clicks a certain button, as shown in Figure 4.5. This provides a simpler interaction with the website and a more organised structure. Another important feature provided is the submission of information upon click a button with that purpose.

Although HTML was used as the default view engine, it was substituted by Handlebars<sup>9</sup> when needed. Handlebars is used to dynamically generate HTML pages, which simplifies the task of manually updating the data we want to show the user, in a clean and simple way, as shown in Figure 4.6. A Handlebars expression is any content between double curly braces, inserted into a HTML template, for example: `{{ variable }}`. The engine takes the variable, compiles it into a function which is executed by passing a JSON object as an argument, for example: `{'variable': 'Maria'}`. Finally, the function returns the required HTML after replacing the variables with the corresponding value. The JSON object is provided by the server so it is dependent on the response of the request made by the user. This is particularly useful for visualisation of the results of queries. Another advantage of Handlebars is that it provides scalability and maintainability, since it separates logic-less templates from server-side logic. The *npm* package that allows the use of Handlebars with Express is called *hbs*<sup>10</sup>.

For plotting on the browser, Google Charts<sup>11</sup>, an open source tool was used. This tool provides a vast selection of responsive charts, which are easily customisable and can be enabled

---

<sup>8</sup>jQuery - <https://jquery.com/>

<sup>9</sup>Handlebars.js - <http://handlebarsjs.com/>

<sup>10</sup>npm: hbs (Consulted: 11/07/2017) - <https://www.npmjs.com/package/hbs>

<sup>11</sup>Google Charts - <https://developers.google.com/chart/>

**Cycle duration (minutes)**

Cycle 1 Duration

**Add another cycle**

(a)

**Cycle duration (minutes)**

Cycle 1 Duration

Cycle 2 duration

**Add another cycle**

(b)

Figure 4.5: PSG cycle input: (a) only one cycle input available (b) upon clicking the button, another field appears

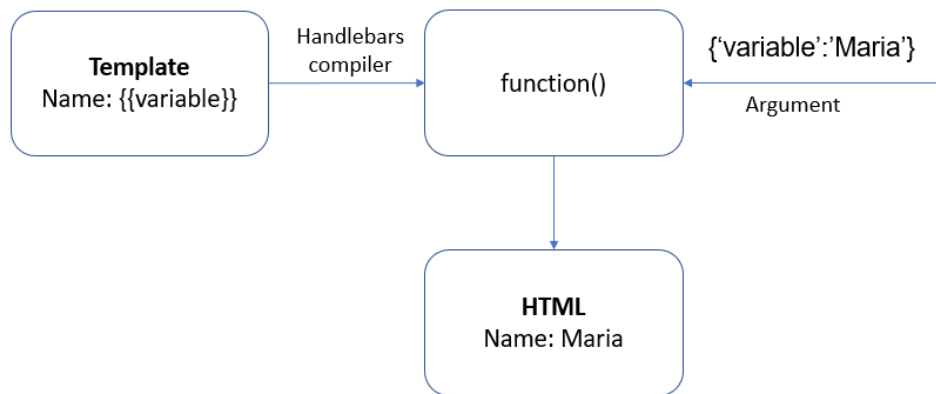


Figure 4.6: Handlebars work flow. Adapted from: <https://www.sitepoint.com/a-beginners-guide-to-handlebars/>

to work in an interactive dashboard. The charts are cross-browser and cross-platform compatible and do not require any plugins, making them a good choice for scalable solutions. Another possible tool was D3.js<sup>12</sup>. On the one hand, D3.js is more customisable than Google Charts and allows the design of more complex charts. On the other hand, Google Charts is simpler to use and easier to learn. Considering the time available and the low complexity of the intended plots, Google Charts was deemed more suitable for this project.

### 4.3.2 User pages

Upon accessing SleepData, the landing page allows the user to login or register, in case the user has no account (see Figure 4.7). By clicking the "Register" button, the login form is replaced by a registration form for new users. The landing page also doubles as the homepage for non-authenticated users, so it has some of the platform statistics, such as number of users, patients, clinical notes, exams and questionnaires.

---

<sup>12</sup>D3.js - <https://d3js.org/>

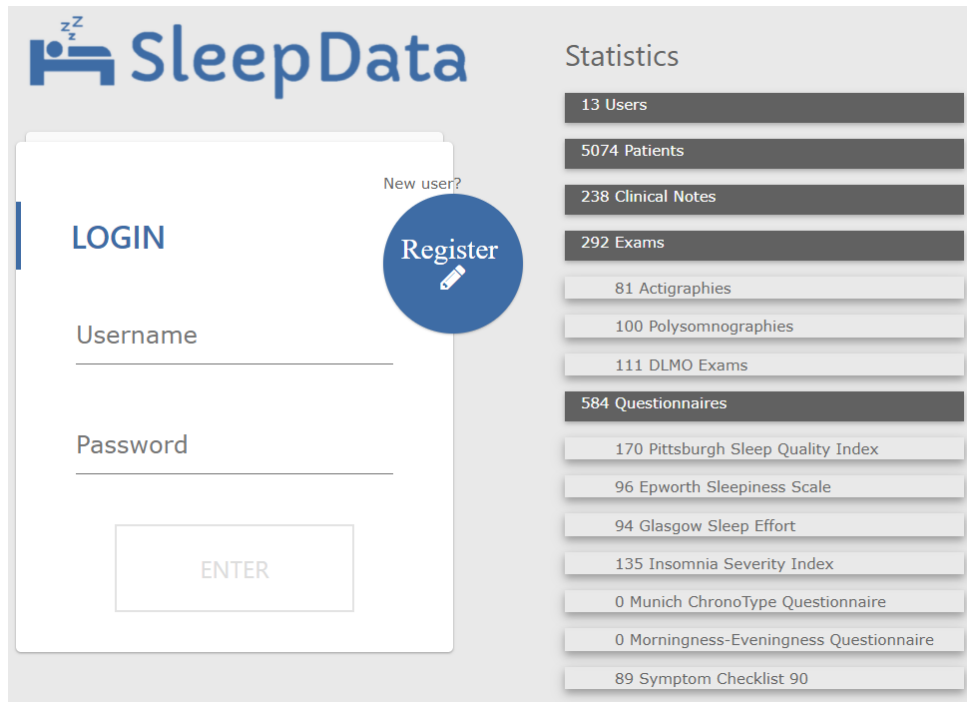


Figure 4.7: Landing page

After successful login, the user is redirected to the user homepage. As shown in Figure 4.8(a), the only two actions a user with Regular User permissions can take is filling in questionnaires or updating personal information, like the name, birthday and gender. The latter is common to all user types. The available questionnaires can be accessed by clicking the "Questionnaires" button, as shown in Figure 4.8. The user is then redirected to the chosen questionnaire. Users with Professional User permission are redirected to a homepage like the one in Figure 4.9(a). As shown, this type of user can:

- Perform administrative actions related to his/her clinic, like creating new patients files or map an existing patient record to a Regular User account;
- Fill in forms like Clinical Notes, Polysomnography Exam (important parameters from the raw files) and Night Diaries;
- Fill in Actigraphy and PSG reports;
- Add laboratory exams, like DLMO and Vitamin D;
- Add raw Actigraphy and PSG files;
- View patient data and populations statistics.

Finally, if the user has Administrator permissions, the user will be redirected to a page where it is possible to manage the permission of the existing accounts, manage professionals and their roles

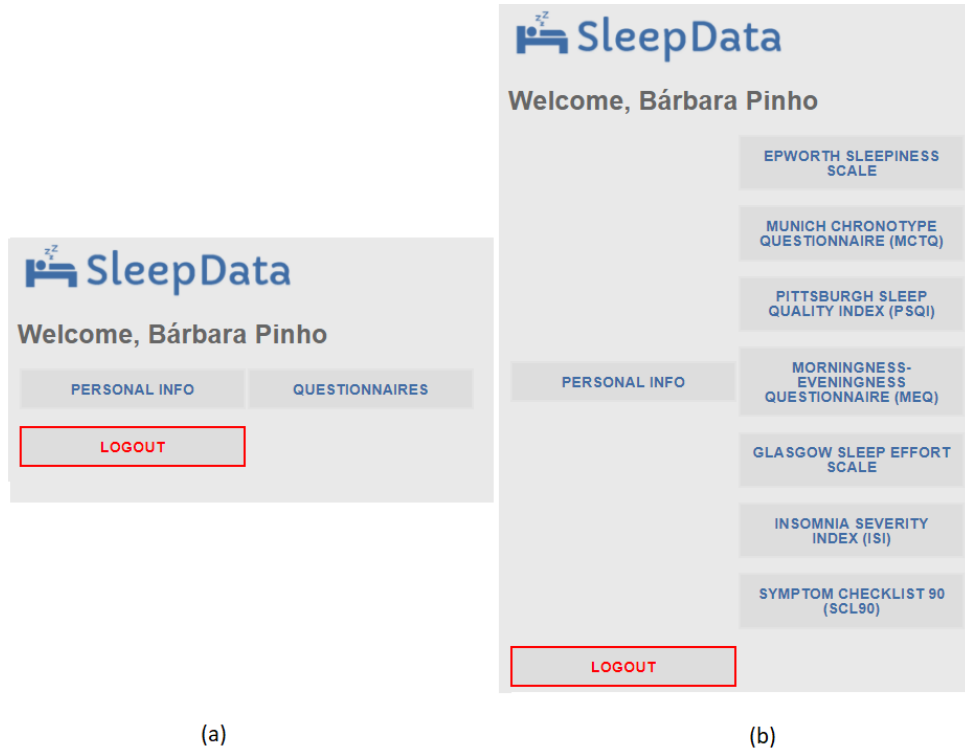


Figure 4.8: Regular user homepage: (a) before clicking the "Questionnaires" button and (b) after clicking the button, the available sleep questionnaires appear)

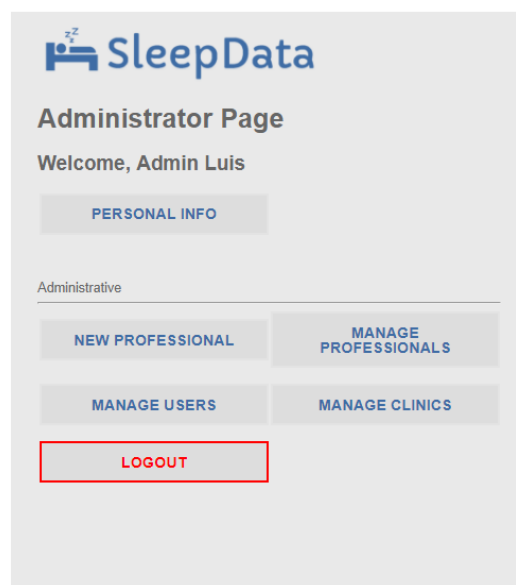
and manage clinics (see Figure 4.9(b)). By querying the existing users, the Administrator user can give the user Administrator, Regular User or Professional User (immediately assigning the Professional User to its clinic) permissions, delete users or map users to a professional's profile.

A schematic representation of the web pages that compose SleepData and how each user can navigate through the platform is available in Figure 4.10.





(a)



(b)

Figure 4.9: Homepages for different account types: (a) CENC-Professional and (b) Administrator

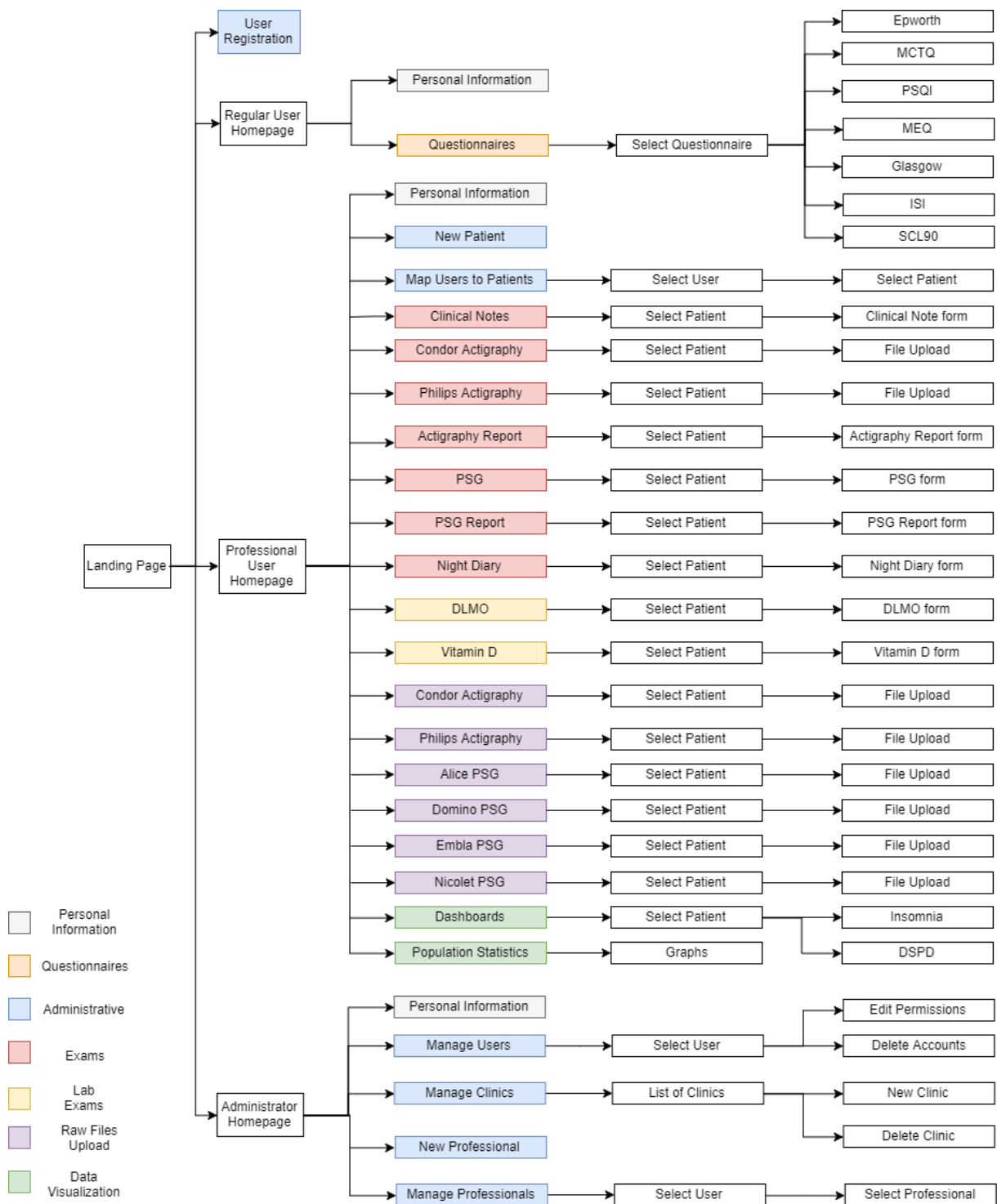


Figure 4.10: Web page schematics

### 4.3.3 Uploading and managing exams

Each file can be linked to other related files. For example, a Polysomnography Report, Night Diary and other raw files can all be linked to the same Polysomnography exam. SleepData does not allow to add or fill in documents related with patients before specifying the patient in advance. When performing an action related to patient data, the user is redirected to a page to pick the patient. Figure 4.11 is a snapshot of the query page, where the complete name, Patient ID at the clinic and birthday of the patient are shown, to avoid user mistakes. When adding a raw file, it is also mandatory to pick of the equipment that generates the file, as the parsing methods differ. This way, after clicking the "Actigraphy" or "PSG" buttons, a list with the recognised equipment appears, as shown in Figure 4.12. Afterwards, the user is redirected to the query page, where he/she can choose the patient the files refer to and upload the files. As different models export the data in different formats, the SleepData file upload window suggests the file format to let the user know what data type the platform is expecting to receive.

In case there is a need to update any previously uploaded file, when the user submits the updated version this does not delete the older version. Instead, when it comes to data visualisation, only the newest version of the file will be shown. This feature is important, since it can be relevant to assess previous versions of clinical files.

### 4.3.4 Questionnaires and clinical forms

All questionnaires and clinical forms display the name of the patient they relate to. Since most questionnaires are designed to compute at least one metric to be analysed by physicians later, it is not possible to submit a questionnaire without filling in all the required information (see an example in Figure 4.13).

Clinical forms do not have a restricted number of parameters nor mandatory fields, as variability is most often present. For that reason, the forms were designed in such a way that the user can choose to add extra fields, such as extra contact information or extra measurement fields (see Figure 4.5). The forms on the clinicians' side are the most complex, specially clinical notes, due to their subjectivity. For that reason these forms have the following characteristics:

1. Organised in sections by type of information (e.g. personal, health, work and schedules);
2. Allow adding comment fields when selected;
3. Show further options when selected.

Name	Patient ID	Bday	
JOSE [REDACTED]	[REDACTED]	[REDACTED]	ADD NOTE
DIOGO [REDACTED]	[REDACTED]	[REDACTED]	ADD NOTE

Figure 4.11: Query patients page (Note: The information is censored to protect patients' confidentiality)

Raw Files Upload

ALICE

CONDOR DOMINO (SOMNOSCREEN)

PHILIPS EMBLA

NICOLET

Figure 4.12: List of possible equipment

These features are shown in Figure 4.14, which is a partial snapshot of the Clinical Notes form. The information below the selected fields does not appear in the form until the field is selected. This allows a better overview of the form without overloading the user with unnecessary information.

### 4.3.5 Clinical dashboard

One of SleepData's main purposes is to facilitate data visualisation, in a way that clinical diagnosis can be performed in a more organised and direct way. That can be obtained by providing users with Professional User permission dashboards with the clinical information of each patient in their associated clinic. As the list of sleep disorders is vast and the clinical parameters that define and characterise each disorder differ, SleepData enables the visualisation of specific parameters and variables of a patient, according to the disorder the professional wants to study in detail. After choosing to view a patient's data, the Professional User is then redirected to a page with the patient's general information, where it is possible to choose the type of dashboard required, as shown in Figure 4.15. Two dashboards were designed: one to study DSPD and another one to study insomnia. Both dashboards follow the same principle: allow visualisation of specific variables and parameters, with colour cues to provide to the professional a quick understanding of the patient's state. Those colour cues work in the following ways: normal

## Escala de Sono de Glasgow



Nome: Bárbara Pinho
CENC ID: 234
Data: 11/10/2017

Para cada item escolha a opção que melhor corresponde à sua opinião.

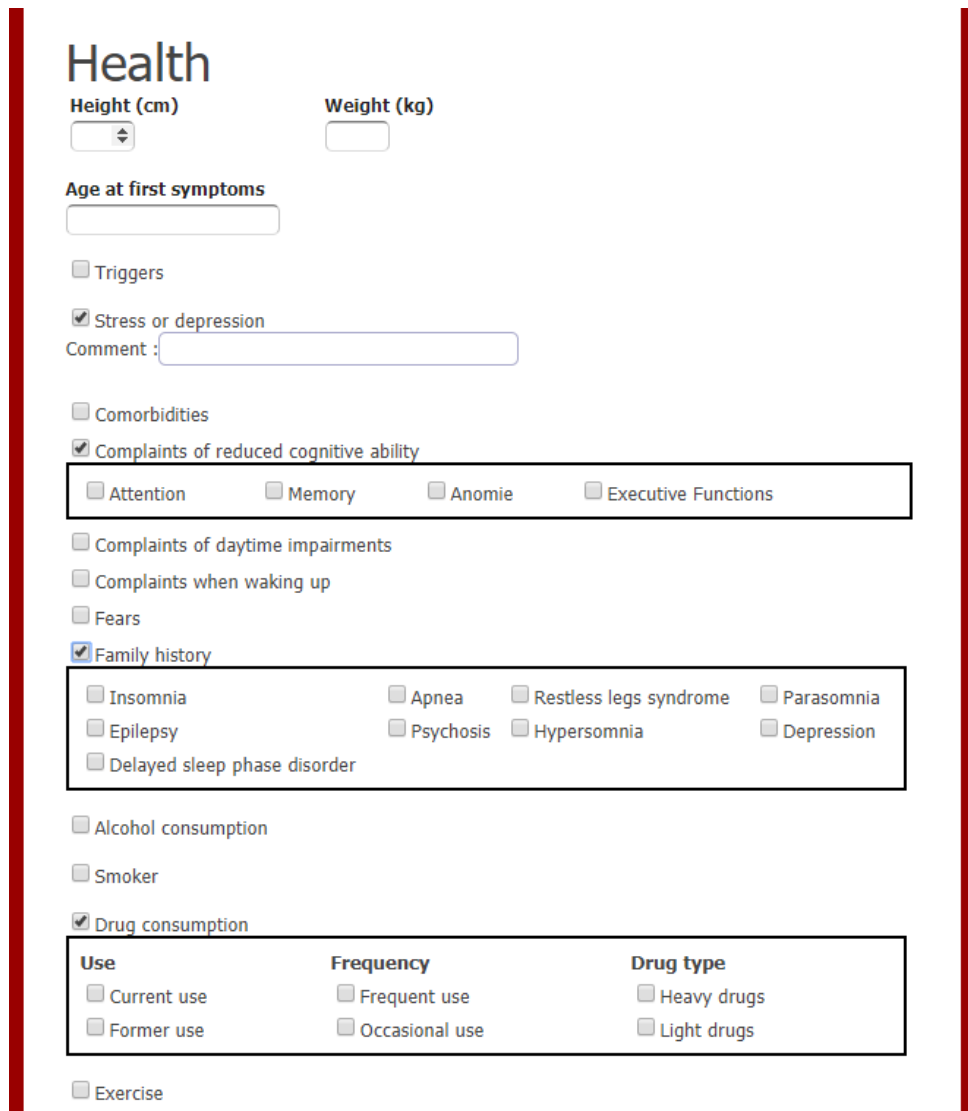
	Verdade	Neutro	Falso
À noite, esforço-me demasiado por dormir, quando deveria ser mais natural.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sinto que devia controlar o meu sono durante a noite.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
À noite, adio a minha ida para a cama com medo de não conseguir dormir.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Quando estou na cama preocupo-me por não dormir e não consigo dormir.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Não sou bom a dormir à noite.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fico ansioso por dormir, à noite, antes de me deitar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Preocupo-me com as consequências a longo prazo por não dormir durante à noite.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 4.13: Glasgow sleep effort scale

parameters are highlighted in green, abnormal parameters are highlighted in yellow, orange or red, depending on the scale and severity of the variable displayed. A snapshot of part of the insomnia dashboard of a random patient is shown in Figure 4.16. Although insomnia is not the main focus of my thesis, my colleague's thesis, Castanheira [2017], explores it in detail. Table 4.3 shows the parameters and colour codes related to actigraphy and PSG used in the DSPD dashboard. All other parameters of the DSPD dashboard and their colour code are described in Appendix B, as well as the ones used in the Insomnia dashboard. The parameters and colour codes used were adapted from the guidelines in the literature concerning normal sleep and the characterisation of DSPD and insomnia, detailed in Chapter 2. These were also reviewed and accepted by Dr. Teresa Paiva, the physician responsible for CENC and advisor of this dissertation.

Concerning DSPD, a few aspects of the dashboard should be highlighted:

1. Actigraphy often computes minimum, average and maximum values related with the days it was used. All values are shown in the dashboard. However, to avoid visual confusion in the dashboards, only the average values are coloured.



**Health**

Height (cm)  Weight (kg)

Age at first symptoms

☐ Triggers

☒ Stress or depression

Comment :

☐ Comorbidities

☒ Complaints of reduced cognitive ability

<input type="checkbox"/> Attention	<input type="checkbox"/> Memory	<input type="checkbox"/> Anomie	<input type="checkbox"/> Executive Functions
------------------------------------	---------------------------------	---------------------------------	--

☐ Complaints of daytime impairments

☐ Complaints when waking up

☐ Fears

☒ Family history

<input type="checkbox"/> Insomnia	<input type="checkbox"/> Apnea	<input type="checkbox"/> Restless legs syndrome	<input type="checkbox"/> Parasomnia
<input type="checkbox"/> Epilepsy	<input type="checkbox"/> Psychosis	<input type="checkbox"/> Hypersomnia	<input type="checkbox"/> Depression
<input type="checkbox"/> Delayed sleep phase disorder			

☐ Alcohol consumption

☐ Smoker

☒ Drug consumption

Use	Frequency	Drug type
<input type="checkbox"/> Current use	<input type="checkbox"/> Frequent use	<input type="checkbox"/> Heavy drugs
<input type="checkbox"/> Former use	<input type="checkbox"/> Occasional use	<input type="checkbox"/> Light drugs

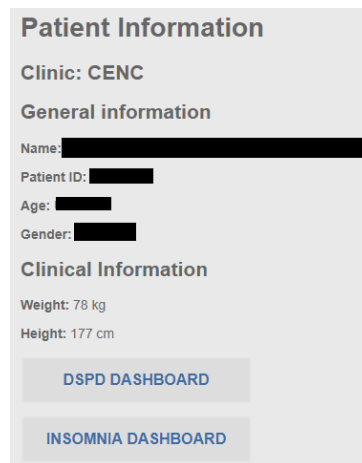
☐ Exercise

Figure 4.14: Partial health section of the Clinical Notes form

2. Some values do not have a colour code as their values cannot be seen as "good" or "bad". These parameters are often lifestyle related and may or may not be related to the disorder. Nonetheless, they are important to the process of diagnosis.
3. The PSQI (sleep quality) and MCTQ (chronotype) results presented are separated into work and free days. In patients with DSPD, these results often differ, as work/school schedules affect the desired sleep schedule, and thus sleep quality and computed chronotype.

#### 4.3.6 Population statistics

Another feature of SleepData is the ability to study and characterise the patient population in the platform with a certain disorder. For that purpose, several charts were designed, based on GoogleCharts. SleepData allows the visualisation of the charts for the general population of



**Patient Information**

**Clinic: CENC**

**General information**

Name: [REDACTED]

Patient ID: [REDACTED]

Age: [REDACTED]

Gender: [REDACTED]

**Clinical Information**

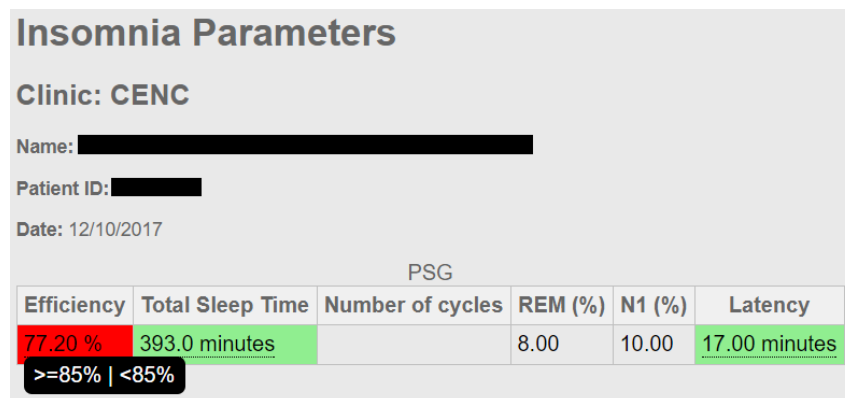
Weight: 78 kg

Height: 177 cm

[DSPD DASHBOARD](#)

[INSOMNIA DASHBOARD](#)

Figure 4.15: Patient dashboard. Note: sensitive information was censored to protect patient confidentiality



**Insomnia Parameters**

**Clinic: CENC**

Name: [REDACTED]

Patient ID: [REDACTED]

Date: 12/10/2017

PSG					
Efficiency	Total Sleep Time	Number of cycles	REM (%)	N1 (%)	Latency
77.20 %	393.0 minutes		8.00	10.00	17.00 minutes
>=85%   <85%					

Figure 4.16: Insomnia patient dashboard (partial snapshot, only PSG parameters). See as the efficiency colour division appears when hovering over the parameter. Note: sensitive information was censored to protect patient confidentiality.

patients on the data base or restricted to a particular disorder, such as DSPD and insomnia. The user can define which population to study, by selecting one of the buttons, as shown in Figure 4.17. All the charts are shown in the same page, organised in different sections.

SleepData Population Statistics charts are divided in different sections, as follows:

1. General information;
2. Clinical Notes;
3. Actigraphy;
4. DLMO;
5. Polysomnography;

Table 4.3: DSPD dashboard parameters, variables and the respective colour codes: only actigraphy and PSG parameters shown.

Source	Colour	Green	Yellow or Orange	Red
	Parameter			
Actigraphy and PSG	Sleep onset hour	Before 2 AM	-	After 2 AM
	Wake up hour	Before 11h00 (AM)	11h00 to 14h59	After 15h00
	Total sleep time	Before 11h00 (AM)	11h00 to 14h59	After 15h00
	Latency	$\leq 30$ minutes	-	$>30$ minutes
	Efficiency	$\geq 85\%$	-	$<85\%$

6. Comparison between bed time hour computed by actigraphy, computed by PSG and reported by patients;
7. PSQI;
8. ISI;
9. Epworth Sleepiness Scale;
10. Glasgow Sleep Effort Scale;
11. Symptom-checklist-90 revised.

An extensive list of the charts that compose each section in SleepData Population Statistics can be found in Appendix C.

Like in the previous subsection, the scales and slots of the charts were chosen according to the parameters approached in Chapter 2 and approved by Dr. Teresa Paiva.



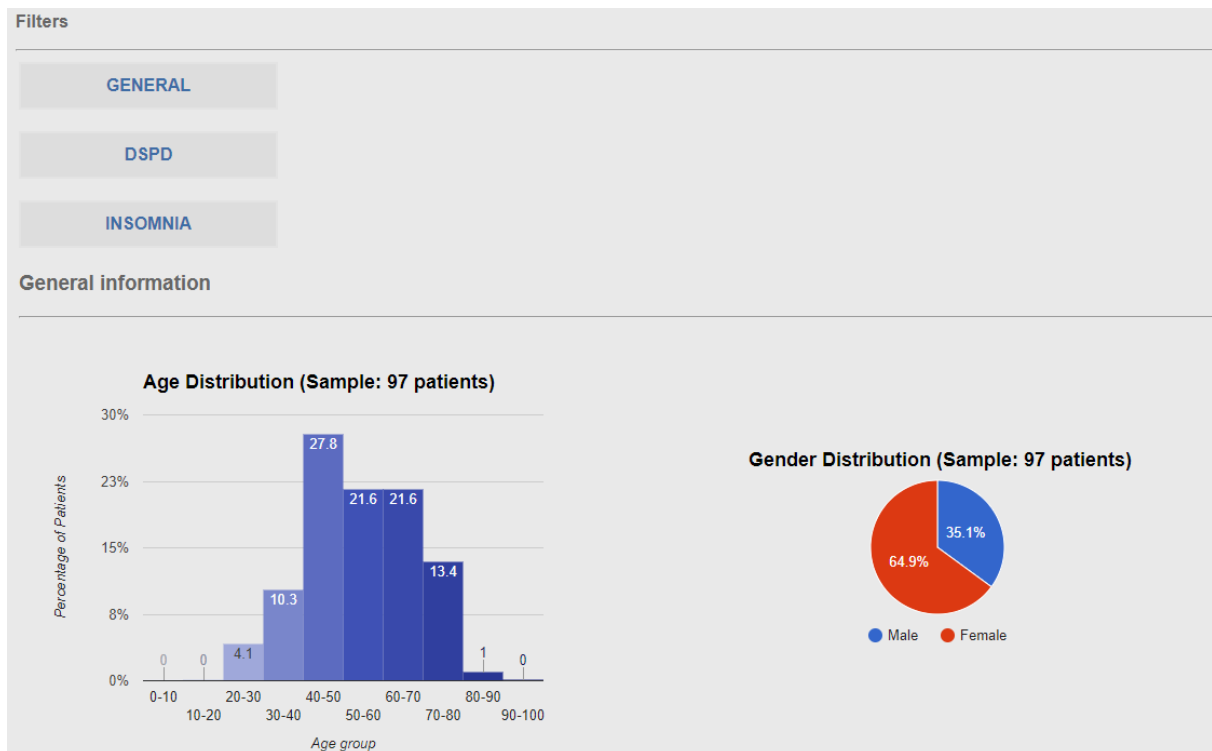


Figure 4.17: Insomnia population statistics interface sample. The dashboard section shown is related with general information.

## 4.4 Overview on SleepData features

To ensure data security and encryption, the platform runs under a HTTPS protocol and can be accessed in the following link: <https://sleepdata.inesc-id.pt>. SleepData is composed of two databases functioning together: one for clinical data, another one for user accounts and permissions. There are three types of accounts: Regular User, Professional User and Administrator.

To have a good understanding of the status of SleepData platform, it's useful to look at each requirement and how they were met:

**FAIR data** Use of unique identifiers, standardised FHIR Resources to structure clinical documents, medical nomenclatures (SNOMED CT and LOINC) and international standards (ISCED2011, ISCO and NIAAA). Use of CENC's coding system, which can be used by any other clinic.

**Logical separation between health and personal data** Health documents do not have personal information and vice-versa. By accessing a health document, it is not possible to know which patient it belongs to.

**Different access levels and controlled access** Different account permissions. Access to

documents in different collections granted based on account type. Password authentication implemented, but stronger authentication methods in library possible.

**Update and delete users** Administrator is able to delete accounts and change permissions.

**Prohibit access to unauthorised people** Authentication of platform by x509 certificate in SSL connections. Validation of users.

**Encrypted transmission** HTTPS connections enforced.

**Access to servers restricted** Platform hosted in secure data centre (FCCN).

**Frequent backups** SleepData host managed by data centre professionals.

**Everyone with access to personal data bound by professional secrecy** IST students and faculty signed confidentiality agreements before having access to the CENC data.

**Scalable DBMS** MongoDB: DBMS capable of managing big amounts of heterogeneous data with structured and unstructured information.

**Scalable web server** Node.js: capable of handling simultaneous requests, dynamic content and multiple users.

**Web framework that handles RESTful API** *express*

**Multi-clinic platform** SleepData is able to handle information of several clinics. It is possible to create new clinics.

**Integrate clinical information from different sources** Clinical notes, sleep questionnaires, actigraphy reports, PSG reports and DLMO exams.

**Automatic update and processing of the data** Clinical forms, questionnaires and parsing methods prepared to receive information and update database.

**Tools for statistical analysis** Several charts with data from all the diagnostic tools, that give an overview of the set of population in study.

**Diagnostic tools** Colourful dashboards with the most relevant information for two different disorders: DPSD and insomnia.

Considering the topics above, SleepData is a fairly consistent prototype for a more advanced and better platform for integrating sleep medicine data. SleepData could be easily be customised to meet the demands of different clinics. It is possible to add clinics and upload their information.

In fact, SleepData already has information about CENC patients. The personal information of over 5000 patients was uploaded, together with the clinical information of two selected sets:

- 145 patients previously diagnosed with DSPD;
- 100 patients previously diagnosed with insomnia.

Appendix D gives an extensive list of CENC's documents, with the information it has described thoroughly and how each file is collected. These two sets were used to test SleepData's capacity to deliver the intended information of population statistics and to work as a support diagnostic tool for DSPD (and for insomnia, in Castanheira [2017]).

Any clinic added by the Administrator can use SleepData to support the clinic's daily activities. Additionally, SleepData can serve as a universal sleep data repository. However, it cannot be said that the platform is completely ready for clinical practice. There are still many improvements to be made, both on the clinical and the infrastructural sides. For example, the platform does not yet allow the visualisation of every document inserted, such as different versions of clinical notes of the same patient, and can still be improved in terms of user authentication process and security.

MongoDB offers the possibility to encrypt the data stored in the database (also known as Encryption at Rest)<sup>13</sup>. With this feature, someone that would gain access to the system would not be able to see the data, unless they have the key that was previously configured to decrypt the data. However, this feature is only available in the paid MongoDB Enterprise version and not in the free version used for SleepData, so this feature could not be used.

---

<sup>13</sup>Encryption at Rest - <https://docs.mongodb.com/manual/core/security-encryption-at-rest/>

# 5

## Characterisation of the DSPD population

This chapter describes my characterisation of the DSPD population (Section 5.1), followed by an analysis of the role of DLMO in the diagnosis of this type of patients (Section 5.2). An overview of the topics approached in this chapter is available in Section 5.3.

### 5.1 Characterisation

The characterisation of the DSPD population was conducted using the data set provided by CENC. This data set is comprised of patients that were or are being treated for DSPD. All charts and results used in this characterisation were provided by SleepData's statistical analysis tools.

#### 5.1.1 General features

The clinical notes of the data set provided allow the analysis of the distribution of:

1. Age;
2. Age at the start of symptoms;
3. Bad consumption habits;
4. Comorbidities;
5. Possible related features.

Starting with age distribution, in Figure 5.1(a), it is possible to see that, out of 123 patients being treated with DSPD, more than 25% are young adults, with between 20 and 30 years old, and that more than 65% of patients have between 20 and 50 years old. In Figure 5.1(b), which represents the distribution of the age at first symptom reported by 75 patients, it is possible to see that 52% report having the first symptom between the ages of 6 and 20 years old. There is a clear decrease in the percentage of patients that report the start of symptoms taking place

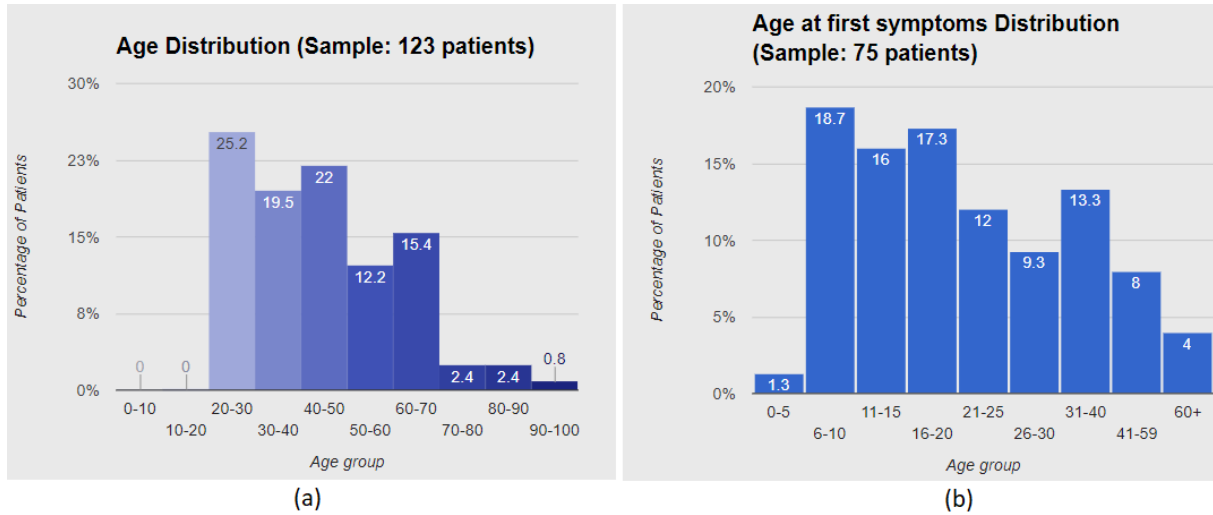


Figure 5.1: Age distribution (%) by age groups: (a) Age of the patients when they seek treatment; (b) Reported age at first symptom of DSPD

later in life, with the exception of the age group 31-40 years old. This is consistent with previous studies that claim that DSPD affects all age groups and that there is a higher prevalence of DSPD in children and teenagers. The two charts show that, although children and teenagers suffer from DSPD the most, they don't seek treatment until they reach adulthood. There are three possible explanations behind these results:

1. Increased difficulty in dealing with the disorder when confronted with certain responsibilities associated with adulthood – like work schedules – leads to increased search for medical treatment;
2. Parents believe the symptoms in children and teenagers are age related (or puberty related) and do not think medical advice is needed;
3. Cumulative effects of an untreated disorder leads to higher demand for medical treatment, in older age groups.

Figure 5.2 represents the distribution of features related with DSPD reported by patients: traumatic experience, car accidents, family conflicts, academic underachievement, stress at work place<sup>1</sup> and stress or depression. Each patient could report more than one complaint. More than a third of the examined patients complain about family conflicts, while at least 10% report traumatic experiences, academic underachievement or feeling stress or depressed. Only 28 out of 145 patients (19%) do not report one of the mentioned characteristics. While these features may

<sup>1</sup>The chart shows zero reports of stress at work place, as it was not reported objectively in the data set given and, therefore, not interpreted by SleepData.

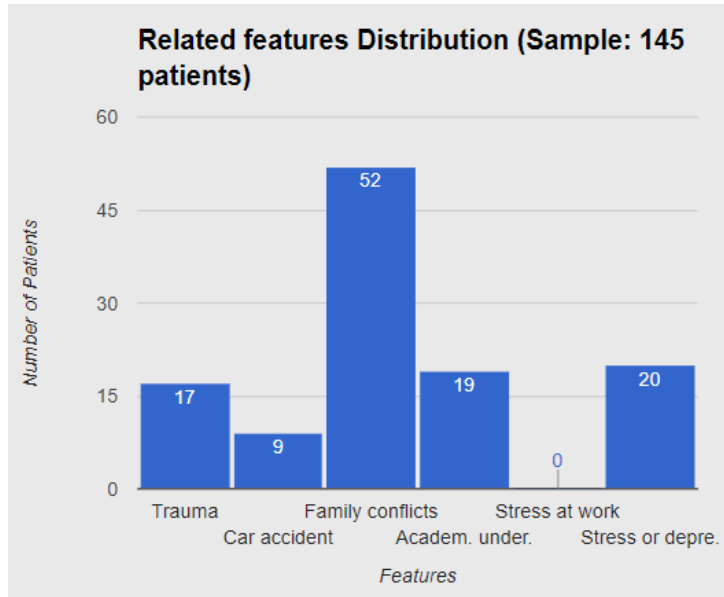


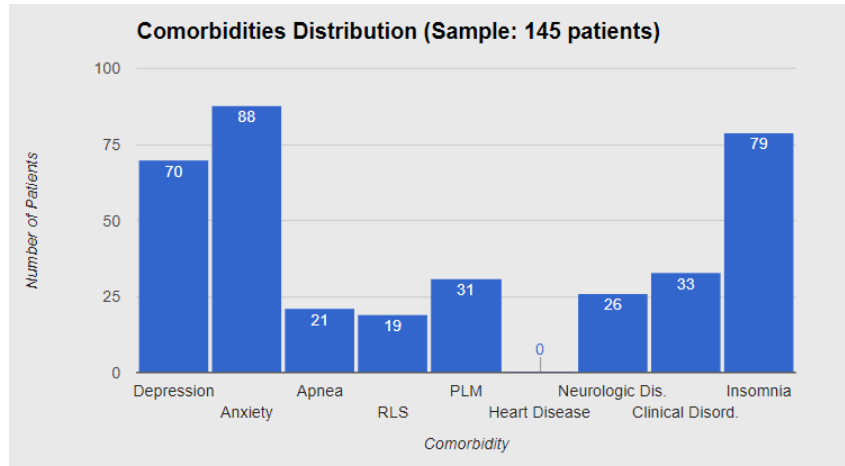
Figure 5.2: Reported features reported by the patient's, possibly related with DSPD

not be the reason behind DSPD, they might be some of the main triggers. The most common ones provoke stress and worry in the patients, leading to difficulty falling asleep, delaying the sleep phase of the patients.

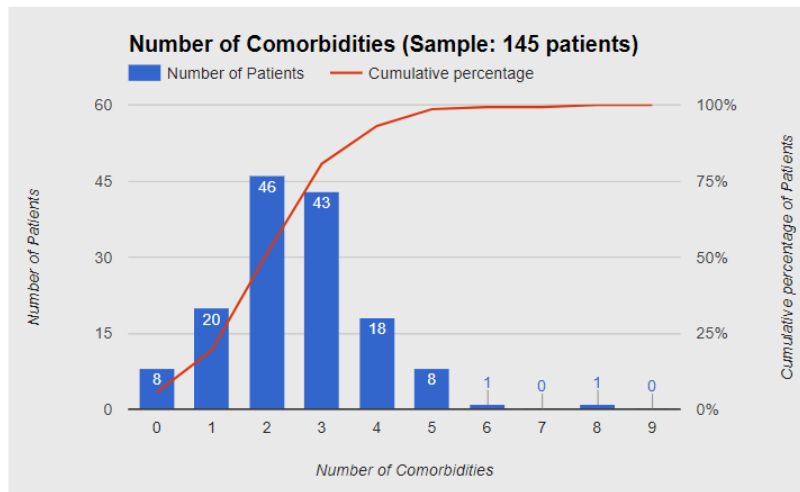
The distribution of comorbidities of patients diagnosed with DSPD can be seen in Figure 5.3(a). Among the studied comorbidities, the most common is anxiety, affecting about 60% of the population, followed by insomnia and depression, with 54% and 48% respectively. This is of no surprise, as DSPD is often misdiagnosed as depression or insomnia. Although the results refer to a sample of 145 patients, the combined number of comorbidities is much higher, as each patient often reports several comorbidities, as shown in Figure 5.3(b). It is possible to see that more than 73% of the patients have between two and four comorbidities. Given the common triggers to these disorders include the stressors mentioned above, it is expected that each patient develops more than one disorder. This supports the fact that DSPD is not easily diagnosed, as it is associated with many other comorbidities, thus leading to its misdiagnosis.

### 5.1.2 Sleep pattern

The usual sleep schedule of DSPD patients can be studied using actigraphy results. The actigraphy is often conducted for several weeks and there is no distinction between work and free days. Figure 5.4 shows the distribution of the usual bed time (a) and wake up time (b) of DSPD patients. As it is possible to see, about 80% of the patients go to bed between 1 and 4 AM and wake up between 9 AM and 12 PM, showing a phase delay of one to four hours, which



(a)



(b)

Figure 5.3: Comorbidities distribution: (a) distribution of main comorbidities and (b) distribution of number of comorbidities per patient with cumulative percentage

is consistent with previous studies. In this data set, phase delay can go as further as nine hours. Figure 5.5 represents the distribution of total sleep time, with the variance of the values. Out of 75 patients, 27 have an average total sleep time of 420 minutes (7 hours). The maximum values of total sleep time of those patients often reaches values of eight to nine hours. This indicates that these patients are getting enough sleep. However, about 30 patients are getting six or less hours of sleep. Comparing these results with the bed time and wake up time schedules, even though a delay in phase can be perceived, it is safe to assume one of three scenarios:

1. Patients have an even more delayed sleep phase than the one perceived in Figure 5.4, meaning they spend enough time in bed but are not getting enough sleep;
2. Patients suffer from insomnia, taking too long to fall asleep;
3. Patients go to bed at a time that agrees with their phase, but have to wake up before

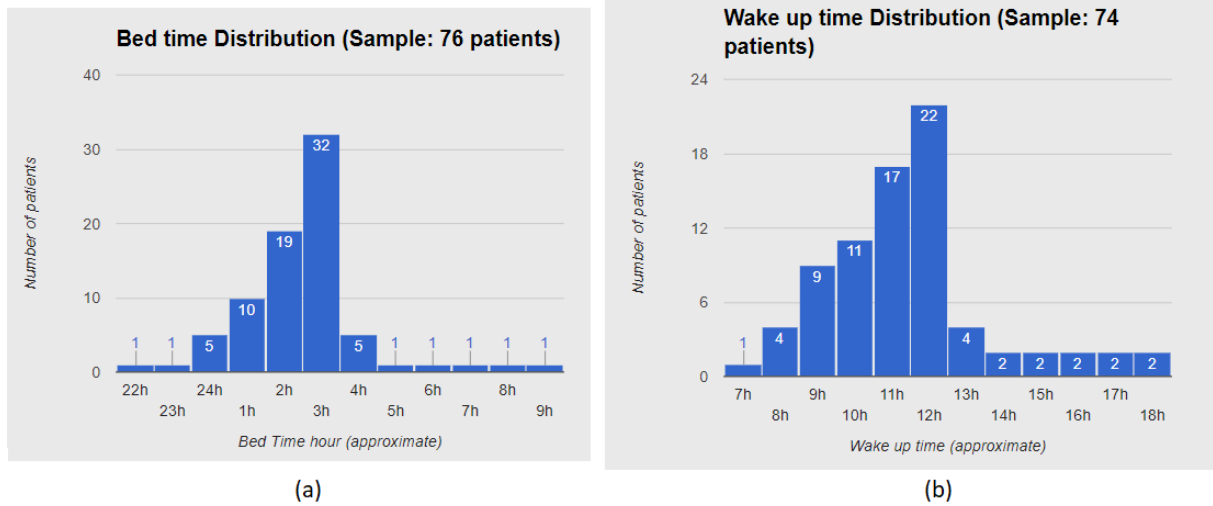


Figure 5.4: Distribution of perceived actigraphy schedules: (a) Bed time and (b) Wake up time

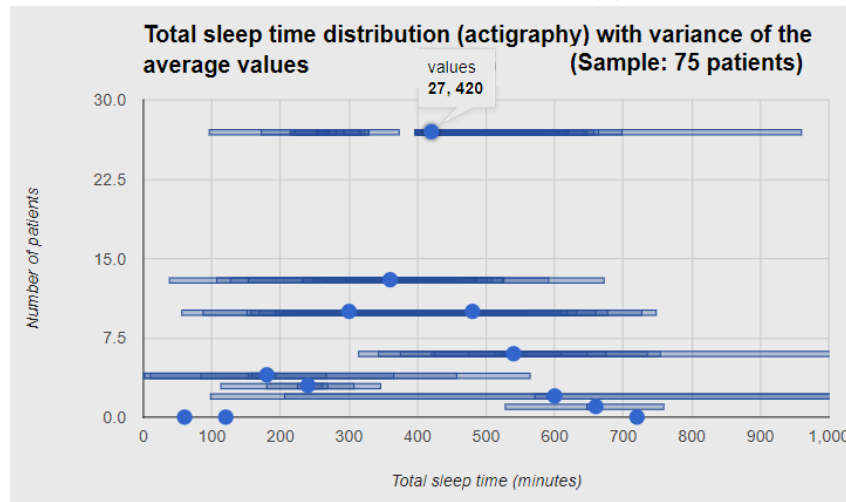


Figure 5.5: Distribution of total sleep time. The blue bars indicate the variance of the medium values and the density of the bar indicate which values are more frequent (denser areas indicate higher frequency).

completing the total sleep time advised.

The first scenario can be further analysed by computing sleep efficiency. The formula to compute sleep efficiency is the following:

$$\text{Sleep efficiency} = \frac{\text{Total sleep time}}{\text{Total time in bed}} \times 100$$

Figure 5.6 shows that 65.8% of the patients have reduced sleep efficiency (efficiency lower than 85%). This supports the first and second scenarios: the patients are either trying to follow a sleep schedule that is not on their sleep phase, probably sooner than that, or patients are taking a long time to fall asleep, a probable indicator of insomnia. Figure 5.7, which represents the



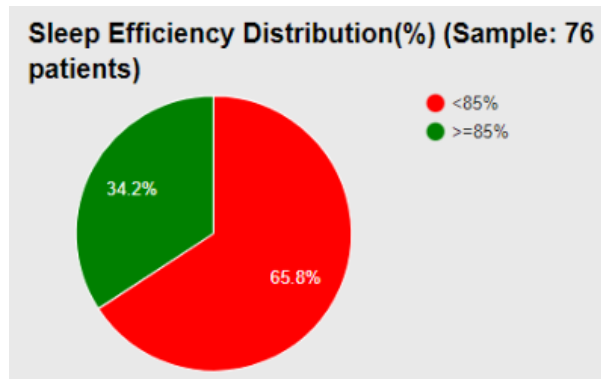


Figure 5.6: Sleep efficiency of DSPD patients computed by actigraphy

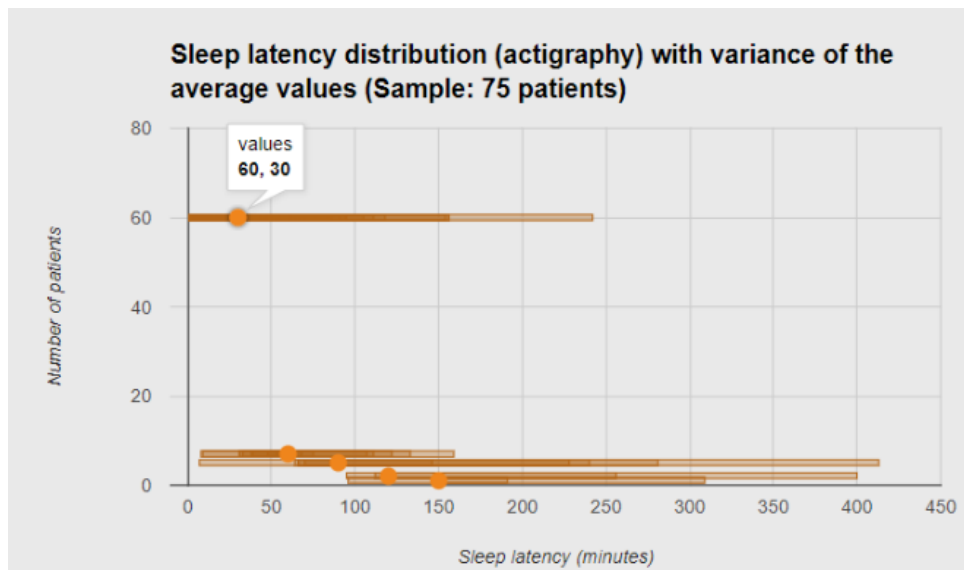


Figure 5.7: Distribution of sleep latency. The orange bars indicate the variance of the medium values and the density of the bar indicate which values are more frequent (denser areas indicate higher frequency).

distribution of sleep latency perceived by actigraphy, helps explore in more detail the second scenario. It is possible to see that the average latency for 60 out of 75 patients is 30 minutes or less. However, the variation bar shows that most of those patients experience maximum latency of up to two hours (120 minutes). These values do not exclude the second scenario, but show that the first one has higher probability.

Combining the results of actigraphy with PSG and sleep diary could give a fairly consistent overview on the sleep patterns of a patient. However, DSPD is a disorder mostly characterised by a delay in the sleep schedule, meaning it is extremely important to understand if the exams conducted are in the patient's sleep phase or out of it. Before looking at the results from other sources, it is important to compare them. Figure 5.8 compares the average bed time hour as reported by the patient in a sleep diary and as perceived by the actigraphy, and the bed time hour when the patient did a PSG. Looking at the chart it is possible to verify that the bed time

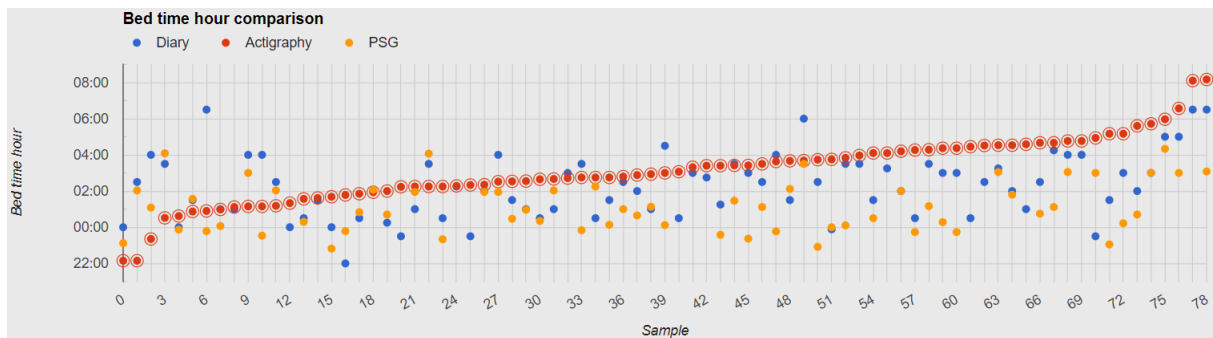


Figure 5.8: Average bed time hour reported in a sleep diary (blue), actigraphy (red) and PSG (yellow). Each vertical line (sample) represents the average values for the same patient. The samples are sorted in ascending order by bed time hour (actigraphy).



Figure 5.9: Computed difference between average bed time perceived by actigraphy and by PSG (top). Computed difference between average bed time perceived by actigraphy and reported by a sleep diary (bottom).

hour perceived by the actigraphy is often much later than the one reported by the patient or even later than the bed time hour of the PSG exam of the patient. This tendency is more notable for later hours. In fact, Figure 5.9 shows the results of computing the difference between the average bed time hour of the actigraphy and the other two sources:

- The bed time in actigraphy is nearly one hour later than the one reported by the patient;
- The bed time in actigraphy is nearly two hours later than the one in the PSG.

Bear in mind that the actigraphy is conducted for several days or weeks and is able to detect bed time hour with high precision. Looking at these results it is possible to point out two things:

1. Patient's often report their bed time hour to be sooner than it actually is: either it is misjudged by the patient or misreported on purpose;
2. PSG is often conducted earlier than the usual bed time hour of the patient.

The second situation needs further explanation. Polysomnography is most often conducted in a clinic's sleep lab. Patient's arrive early in the night at the clinic and have a lot less stimuli than at home: calm environment, no access to television or bright screens and dim-light conditions. The downside of the PSG exam is that patients often wake up earlier than expected, due to the clinic's higher activity in the morning. This affects the results of the exams in terms of wake up

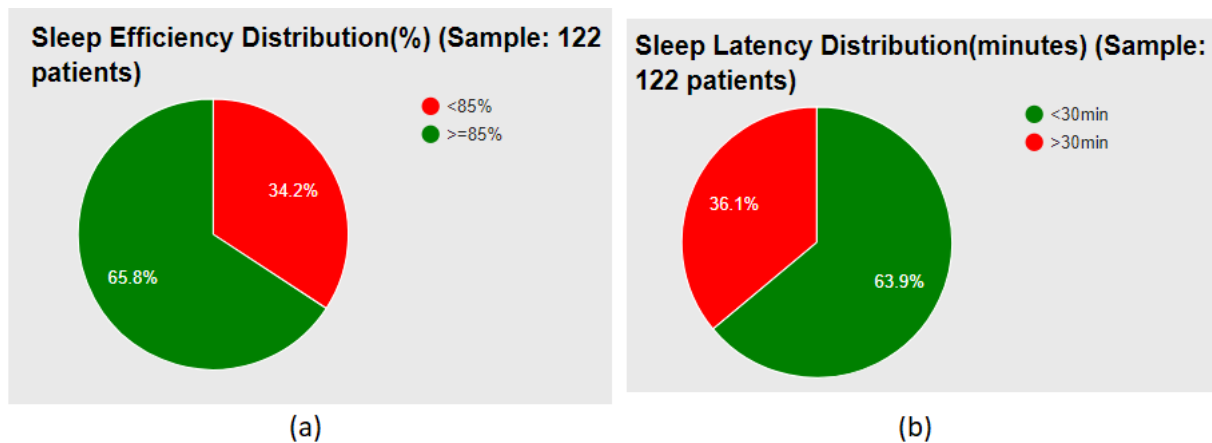


Figure 5.10: PSG parameters of DSPD patients: (a) sleep efficiency and (b) sleep latency

time, total sleep time and, thus, number of cycles. It would be far more useful to ensure that this exam was conducted during the whole sleep phase of the patient. Additionally, PSG only documents one night, unlike actigraphy. However, looking at the sleep efficiency and latency computed by the PSG, shown in Figure 5.10, it is possible note two aspects:

1. The results of sleep efficiency are completely opposite to the ones computed by actigraphy: out of 122 patients, more than 65% present sleep efficiency equal to or higher than 85%;
2. Although bed time is earlier in PSG than the patients' usual bed time, more than 63% fall asleep in 30 minutes or less.

These values lead to an interesting hypothesis concerning this set of patients with DSPD: the disorder may be mainly caused by behavioural aspects rather than physiological, since patients present a good and earlier sleep pattern when the stimuli are reduced.

### 5.1.3 Questionnaires

The Munich Chronotype Questionnaire and the Morningness-Eveningness Questionnaire would be very useful to assess the habits of the patients and their chronotype. However, the set of patients with DSPD provided by CENC did not have information on either. Figure 5.11 represents the distribution of results of ESS and PSQI. The results of ESS show that 50% of the patients have a score of 17 or higher, which indicates severe excessive daytime sleepiness. More than 40% present moderate excessive daytime sleepiness (score of 13 to 16), while the rest present at least mild excessive daytime sleepiness. None of these values indicate normal sleepiness. Nearly 95% of the patients that answered PSQI, scored higher than 5, which is associated

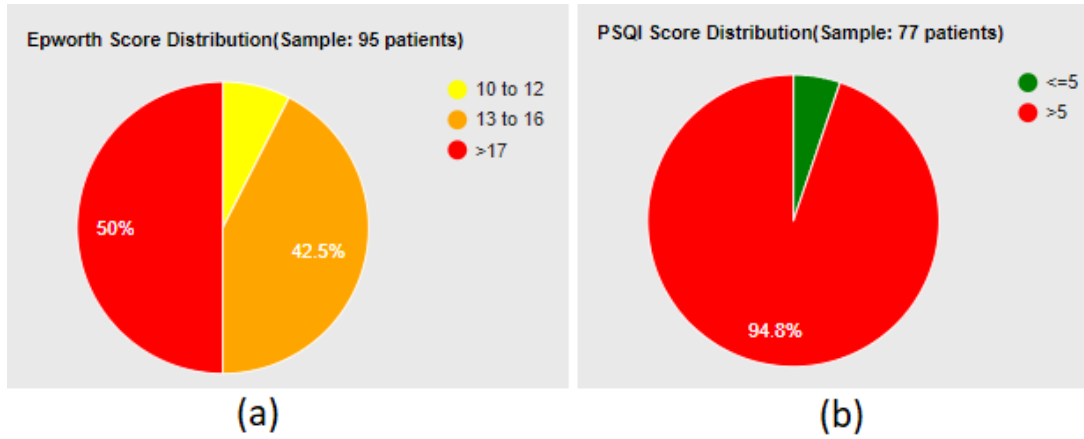


Figure 5.11: Distribution of questionnaires results: (a) Epworth Sleepiness Scale and (b) Pittsburgh Sleep Quality Index

with poor sleep quality. Both results are consistent with each other, as it is expected that patients with poor sleep quality suffer from higher sleepiness during the day, and consistent with the results of sleep efficiency computed by the actigraphy.

## 5.2 The role of DLMO in the diagnosis of DSPD

One way to determine if the melatonin circadian cycle is in phase with the sleep cycle, is to compute the phase angle between the time at sleep onset and the time at DLMO:

$$\text{Phase angle} = \text{Time at sleep onset} - \text{Time at DLMO}$$

A normal phase angle presents values between two to three hours (see Chapter 2). Only 8 out of 55 patients present a phase angle of 120 to 180 minutes (see Figure 5.12). Considering the set of patients who have values nearly normal (1-2 hours and 3-4 hours of phase angle), this represents only 40% of patients with the production of melatonin in phase with the sleep cycle. This does not mean that the other 60% were mistakenly diagnosed with DSPD, as the production of melatonin in phase is not a mandatory criterion to diagnose this disorder. About 45% of these patients present a delay of the production of melatonin higher than the delay in sleep phase (phase angle between -4 to 1 hours). These results are not consistent with the studied bibliography [Fahey and Zee, 2008]. By comparing these results with the ones presented in the literature, we can point out two aspects:

1. There is a high variation of phase angle, which can be explained by individual differences;

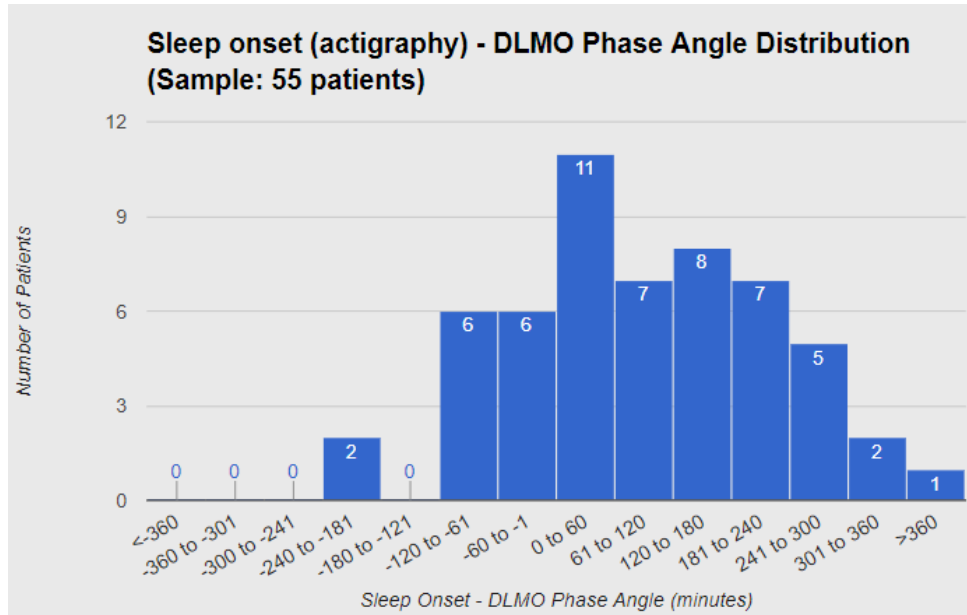


Figure 5.12: Distribution of phase angle between time at sleep onset (as computed by actigraphy) and time at DLMO, in minutes

2. DSPD can be often caused by behavioural aspects and not physiological.

Both these aspects point to same conclusion: DLMO may not be a good marker for DSPD, considering the data set available.

## 5.3 Overview

SleepData is loaded with clinical information from CENC patients. An analysis of the population with DSPD using the SleepData tools is possible. That population can be characterised as:

- First symptoms often appear between the ages of 6 to 20 years old;
- Family conflicts, traumatic experiences and stress or depression are common in most patients;
- Most patients suffer from anxiety, insomnia or depression;
- It is common for patients to suffer from two to four comorbidities;
- Most patients present a sleep phase delay of one to four hours;
- Patients bed time hour is earlier when the surroundings have reduced stimuli.

The results of DLMO are not consistent with the disorder, possibly due to individual differences. Additionally, DSPD can have behavioural and not physiological causes, leading to the conclusion that DLMO is not a good marker of DSPD for this set of patients.

# 6 Conclusions and future work

## 6.1 Conclusions

The main contribution of this thesis was SleepData, a clinical platform oriented for sleep medicine, developed in partnership with my colleague Tiago Castanheira.

In the development of SleepData some aspects concerning the technical requirements for clinical platforms were taken in consideration. To support the technical requirements, several technologies and tools were chosen. The use of standard protocol for information exchange, FHIR, for modelling the most common information resources, such as patient and clinical notes, together with nomenclatures like SNOMED CT and LOINC, make the integrated clinical data more FAIR. The guidelines provided by CNPD to ensure data protection were also taken in consideration. SleepData software architecture resembles the MEAN stack: MongoDB was chosen as a DBMS due to its scalability, its capacity to store structured and unstructured data and to deal with big amount of heterogeneous data. On the server-side, Node.js, a JavaScript runtime environment, was chosen to build the server, together with *express*, a web framework. The transmission of data between server and client is done under an encrypted connection by a SSL. The chosen tools proved to be appropriate for the type of platform we intended to build.

SleepData can be used to support the the daily activities of any sleep medicine clinic and could be easily customised to meet the demands of each clinic. Clinics can manage their patients and fill in clinical forms regarding clinical notes, actigraphies, PSGs and DLMO exams. Patients can answer sleep medicine questionnaires, which their physicians can later assess. SleepData has analysis tools that support the diagnosis of DSPD and insomnia and statistical tools to assess sets of patients. The platform was loaded with DSPD and insomnia patients from CENC, a state of the art sleep medicine clinic. It was possible to use SleepData's tools to assess these sets of patients.

To ensure the usability of SleepData's tools, the clinical outlines of sleep disorders were taken into consideration. There was a detailed study about the diagnostic tools used in clinical practice and about DSPD. The most predominant characteristic of DSPD is the delay in preferred sleep

schedules. In this analysis we also saw that the diagnostic tools present high variability of format and size and that the relevance of DLMO to the disorder's diagnosis needs further investigation. My colleague, Tiago Castanheira, studied insomnia in higher detail.

SleepData is a consistent basis for a more developed platform for integrating sleep medicine data. It should not be considered as a final product for clinics, as many features still need to be implemented. For example, for SleepData to be truly considered a universal sleep data repository, a new account type with access to the anonymised clinical data is needed. Additionally, it was decided not to use AngularJS, the front-end framework used in the MEAN stack. Although this did not affect the goals intended for SleepData, AngularJS could have provided a more consistent and dynamic interface. Regarding security, Encryption at Rest in MongoDB could not be implemented since this feature was only available in the paid version of MongoDB. However, this could be easily implemented and SleepData data would benefit from this extra security measure.

Another main contribution of this thesis was the characterisation of the set of CENC patients with DSPD, using SleepData tools for statistical analysis. Actigraphy results were very useful to assess the sleep schedules of the patients. The comparison of actigraphy results with other diagnostic tools, such as sleep quality questionnaires, PSG and clinical notes provided a better understanding of DSPD and of the characteristics of this population. Most patients present a delay in sleep phase of 1-4 hours, poor sleep quality and comorbidities associated with psychological disorders. One of the most relevant features that CENC patients with DSPD present was the ability to fall asleep earlier in environments with reduced stimuli. This may show that DSPD, in this set of patients, is caused by behavioural aspects, which is not surprising, since light is one of the main synchronisers of the circadian rhythms. Regarding DLMO, it was found that the results were not consistent with the disorder. In other words, contrary to the literature, the melatonin production cycle did not present the same phase delay as the sleep cycle. This supports studies that show that DLMO should not be used as a marker for DSPD.

In this analysis of DSPD, it was not possible to study several important parameters, such as, light exposure and schedules of physical activity. It would have been relevant to cross these parameters with the results mentioned above. However, this was not possible due to the structure of the data set provided.



## 6.2 Future work

SleepData is a consistent prototype and has the potential to become a more developed and advanced product. It is possible to foresee what improvements on interface, infrastructure and security would make SleepData a better platform for integrating sleep medicine data.

The following features could improve user interface and make SleepData more user friendly and useful for sleep medicine clinics:

**Managing exams** Allow visualisation of files inserted, such as clinical notes and reports, and allow update of each field in every file. Allow Professional User to compare different versions of the same file, to assess evolution of the patient.

**Dynamic content** Make dashboards and statistics pages more dynamic, to allow the user to choose what fields to see, what data set to study and what type of charts.

**Filters for other sleep disorders** Design dashboards and statistics pages to analyse specific data considering other sleep disorders than DSPD and insomnia.

**Multilingual interface** Provide translation in other languages.

**Sleep diary** Allow patients to fill in their sleep diaries in SleepData and integrate that information with their clinical data.

**Bulk inserts** Allow clinics to import data to sleep data in bulk. Parsing methods could be needed to parse the data into the correct JSON format.

**Add option to associate patient with diagnosis** Allow physicians to associate their patients with a diagnosis so that their patients clinical information is part of the statistical analysis of a disorder.

**Universal account** Create a universal account type that allows access to anonymised clinical information. The purpose of this feature would be to conduct clinical studies.

**Machine learning algorithms** The use of machine learning algorithms for the prediction of diagnosis would make SleepData even more useful for physicians. An integrated analysis of a patient's clinical data and other universal data available could ease the physicians burden of diagnosis by notifying or alerting them to certain parameters.

The following features would increase users' trustworthiness in SleepData, making the platform's security more robust:

**Email authentication** Allow users to authenticate with their email address.

**Two-factor authentication** Confirm a user's claimed identity by using two different components combined, for example, username/password (already implemented) and mobile phone verification (random code sent by text message).

**Allow password change** In case the user needs to recover the password or simply wants to change it, the user should be able to do so, after going through a thorough verification of identity.

Other aspects could be improved, depending on the demands of each clinic. Every clinic that intends to use SleepData can have their own needs and requirements, depending on the organisation of each clinic and on the diagnostic tools used. Those requirements should be taken into consideration if one wants to customise SleepData.

Regarding DSPD and the use of DLMO as a marker, further studies with larger and more detailed data sets are still needed. It is important to understand how relevant light exposure, physical exercise and other stimuli are, as they clearly affect the circadian sleep cycle. It is also relevant to study how these factors may have no effect on other circadian cycles, such as the production of melatonin.

# Bibliography

- American Academy of Sleep Medicine (2014). *International Classification of Sleep Disorders*. American Academy of Sleep Medicine, 3rd edition.
- Bae, C. and Avidan, A. (2008). Evaluation and testing of the sleepy patient. In SMITH, H. R., COMELLA, C. L., and HÖGL, B., editors, *Sleep Medicine*, chapter 2, pages 25–46. Cambridge University Press.
- Billiard, M. (2008). Normal sleep. In Smith, H. R., Comella, C. L., and Högl, B., editors, *Sleep Medicine*, chapter 1, pages 9–24. Cambridge University Press.
- Burgess, H. J., Park, M., Wyatt, J. K., Rizvydeen, M., and Fogg, L. F. (2017). Sleep and circadian variability in people with delayed sleep-wake phase disorder versus healthy controls. *Sleep Medicine*, 34:33–39.
- Carskadon, M. A. and Dement, W. C. (2011). Normal Human Sleep: An Overview. In Kryger, M. H., Roth, T., and Dement, W. C., editors, *Principles and Practice of Sleep Medicine*, pages 16–26. Elsevier Inc., 5th edition.
- Castanheira, T. (2017). *SleepData - Sleep disorders clinical platform: Insomnia population characterization*. Master thesis, Instituto Superior Técnico - Universidade de Lisboa.
- Chaniotis, I. K., Kyriakou, K.-I. D., and Tselikas, N. D. (2015). Is Node.js a viable option for building modern web applications? A performance evaluation study. *Computing*, 97(10):1023–1044.
- Czeisler, C. A. and Buxton, O. M. (2011). The Human Circadian Timing System and Sleep–Wake Regulation. In Kryger, M. H., Roth, T., and Dement, W. C., editors, *Principles and Practice of Sleep Medicine*, chapter 35, pages 402–419. Elsevier Inc., 5th edition.
- Dongen, H. P. A. V., Maislin, G., Mullington, J. M., and Dinges, D. F. (2003). The Cumulative Cost of Additional Wakefulness : Dose-Response Effects on Neurobehavioral Functions and Sleep Physiology From Chronic Sleep Restriction and Total Sleep Deprivation. *SLEEP*, 26(2):117–126.

- Fahey, C. D. and Zee, P. C. (2008). Circadian rhythm disorders. In SMITH, H. R., COMELLA, C. L., and HÖGL, B., editors, *Sleep Medicine*, chapter 4, pages 56–77. Cambridge University Press.
- Geraldes, R. and Paiva, T. (2014). Mecanismos circadianos de regulação do sono. In Paiva, T., Andersen, M. L., and Tufik, S., editors, *O Sono e a medicina do sono*, chapter 1, pages 35–44. Manole.
- Goodman, K. W., Cushman, R., and Miller, R. A. (2014). Ethics in Biomedical and Health Informatics: Users, Standards and Outcomes. In Shortlife, E. H. and Cimino, J. J., editors, *Biomedical Informatics*, chapter 10, pages 329–355. Springer.
- Hammond, W. E., Jaffe, C., Cimino, J. J., and Huff, S. M. (2014). Standards in Biomedical Informatics. In Shortlife, E. H. and Cimino, J. J., editors, *Biomedical Informatics*, chapter 7, pages 211–254. Springer, 4th edition.
- Koyama, R. G., Jr., S. d. A. F., and de Mello, M. T. (2014). Mecanismos circadianos reguladores do sono II. In Paiva, T., Andersen, M. L., and Tufik, S., editors, *O Sono e a medicina do sono*, chapter 1, pages 45–52. Manole.
- Lei, K., Ma, Y., and Tan, Z. (2014). Performance Comparison and Evaluation of Web Development Technologies in PHP, Python and Node.js. *International Conference on Computational Science and Engineering*, pages 662–668.
- Mohamed, M. A., Altrafi, O. G., and Mohammed O. Ismail (2014). Relational vs . NoSQL Databases : A Survey. *International Journal of Computer and Information Technology*, 03(03):598–601.
- Okawa, M. and Uchiyama, M. (2007). Circadian rhythm sleep disorders: Characteristics and entrainment pathology in delayed sleep phase and non-24 sleep-wake syndrome. *Sleep Medicine Reviews*, 11(6):485–496.
- Paiva, T. (2014). História, mitos e impacto. In Paiva, T., Andersen, M. L., and Tufik, S., editors, *O Sono e a medicina do sono*, chapter 5, pages 755–764. Manole.
- Paiva, T. and Penzel, T. (2011). *Centro de Medicina do Sono - Manual Prático*.
- Paiva, T. and Pinto, H. R. (2014). Clínica e terapêutica das insónias. In Paiva, T., Andersen, M. L., and Tufik, S., editors, *O Sono e a medicina do sono*, chapter 2, pages 230–262. Manole.

- Pandi-Perumal, S. R., Smits, M., Spence, W., Srinivasan, V., Cardinali, D. P., Lowe, A. D., and Kayumov, L. (2007). Dim light melatonin onset (DLMO): A tool for the analysis of circadian phase in human sleep and chronobiological disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 31(1):1–11.
- Peek, N., Holmes, J. H., and Sun, J. (2014). Technical challenges for big data in biomedicine and health: data sources, infrastructure, and analytics. *Yearbook of medical informatics*, 9(1):42–47.
- Quian, J., Moris, C., Phillips, A., Czeisler, C., and Scheer, F. (2017). Unexpected Increase in Melatonin Concentrations During Daytime Sleep in Simulated Night Work Protocol. *Journal of Sleep and Sleep Disorders Research*, 40.
- Sheldon, S. H., Ferber, R., and Kryger, M. H. (2014). *Principles and practice of pediatric sleep medicine*. Elsevier Inc., 2nd edition.
- Shneerson, J. M. (2005). *Sleep Medicine: A Guide to Sleep and its Disorders*. Blackwell Publishing Ltd, 2nd edition.
- Silberschatz, A., Korth, H. F., and Sudarshan, S. (2011). *Database System Concepts*. McGraw-Hill, 6th edition.
- Silva, F. H. L. (2014). Funções do Sono. In Paiva, T., Andersen, M. L., and Tufik, S., editors, *O Sono e a medicina do sono*, chapter 1, pages 5–19. Manole.
- Stores, G. (2007). Clinical diagnosis and misdiagnosis of sleep disorders. *Journal of Neurology, Neurosurgery & Psychiatry*, 78(12):1293–1297.
- Suchecky, D. and Tufik, S. (2014). Privação do sono e estresse. In Paiva, T., Andersen, M. L., and Tufik, S., editors, *O Sono e a medicina do sono*, chapter 1, pages 20–32. Manole.
- Vaish, G. (2013). *Getting Started with NoSQL*. Packt Publishing Ltd.
- Wilkinson, M. D. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, pages 1–9.

# Appendices

# A Clinical database parameters

The parameters of each collection are shown in the table below. Common parameters to every collection include:

- Type and category of document;
- Unique identification code, generated by MongoDB;
- Status (final, corrected, preliminary);
- Date and time: identifying when it was inserted in SleepData and when it was effectively produced;
- Unique identification number of the professional responsible for the document;
- Unique identification number of any related document (for example, an actigraphy report document, must be related with an actigraphy document);
- Unique identification number of the patient the document corresponds to, in case this is justified;
- Laboratory responsible for the exam, in case it was conducted by an external entity;
- Coding system for every field or exam, when possible.

Patient
Name
Marital status
Contact information (phone, address and email)
Gender
Nationality
Birthday
Deceased (boolean)

Language preference
Responsible person in case of minors and patients unable to take care of themselves (name, contact and relationship with the patient)
General practitioner (name, identification and institution)
<b>Professional</b>
Name
Contact information (phone, address and email)
Gender
Birthday
Deceased (boolean)
Language preference
Qualification
<b>ProfessionalRole</b>
Period of employment
Name and identification number
Organisation
Role
Speciality
Location
Contact information (of workplace)
Schedule
<b>ClinicalNotes</b>
Number of separations
Family conflicts (boolean)
Car accidents (boolean)
Education level (ISCED2011 levels and codes)
Academic underachievement (boolean)
Occupation (ISCO groups and codes)
Works in shifts (boolean)
Absenteeism (boolean)



Wake Up Time during the work days (minimum, average and maximum values)
Wake Up Time during the weekends (minimum, average and maximum values)
Getting up time during the work days(minimum, average and maximum values)
Getting up time during the weekends(minimum, average and maximum values)
Bed time during the workdays(minimum, average and maximum values)
Bed time during the weekends(minimum, average and maximum values)
Breakfast time(minimum, average and maximum values)
Lunch time(minimum, average and maximum values)
Snack time(minimum, average and maximum values)
Dinner time(minimum, average and maximum values)
Supper time(minimum, average and maximum values)
Skips meals (boolean)
Bad meal schedules (boolean)
Age at first symptoms
Traumatic event (boolean)
Stress and/or depression (boolean)
Triggering factor of the symptoms
Family history
Alcohol consumption (NIAAA drinking levels)
Narcotics usage (boolean and type of usage)
Smoker (boolean, type, quantity and number of years)
Comorbidities
Body weight
Body height
Irregular Schedules (boolean)
Work Schedule (type, beginning and end time, number of hours per day and week, number of days per week)
Long work shifts (boolean)
Stress at the workplace (type)
Physical exercise (boolean, number of days per week, usual time of exercise)

Excessive working hours (boolean)
Travels frequently(boolean)
Light exposure (type if abnormal)
Sedentary lifestyle (boolean)
Too many responsibilities (boolean)
Procrastination (boolean)
Takes homeopathic medicines (boolean)
Sexual orientation
Ethnic group
Religious affiliation
Complaints of reduced cognitive ability (type)
Complaints of daytime impairments (type)
Medication
Amount, type, usage, manufacturer, form, active ingredient and its amount for each drug
Actigraphy
Equipment used, including identification number and brand
Number of days used
Total sleep time, in minutes (minimum, average and maximum values)
Bet time (minimum, average and maximum values)
Getting up time (minimum, average and maximum values)
Sleep onset time (minimum, average and maximum values)
Sleep efficiency (minimum, average and maximum values)
Wake-time after sleep onset (minimum, average and maximum values)
Number of awakenings (minimum, average and maximum values)
ActigraphyReport
Irregular schedules (boolean and type)
Circadian rhythm profile
Reduced motor activity during the day (boolean)
Increased motor activity during the evening (boolean)
Increased motor activity during sleep (boolean)

Sleepiness episodes during the day (boolean)
Light exposure (type if abnormal)
Matching diary and actigraphy schedules (boolean)
Altered total sleep time (type)
PSG
Setup (EEG configuration)
Total sleep time
Time in bed
Sleep onset time
Wake up time
Sleep efficiency
REM latency
N1 sleep phase (%TST)
N2 sleep phase (%TST)
N3 sleep phase (%TST)
N3 temporal profile
REM sleep phase (%TST)
REM temporal profile
Number of sleep cycles
Duration per cycle
Last sleep cycle is REM (boolean)
Apnea hypopnea index
Oxygen desaturation index
Minimum oxygen saturation
Snoring (%TST)
Periodic sleep movements
Micro-awakening index
Fragmented sleep
Suggested immobilization test
PSGReport

Perception of sleep duration
Insomnia criteria
Reduction of sleep duration
Normal deep sleep
Normal REM sleep
Reduced REM sleep latency
Abnormal deep sleep temporal profile
Abnormal REM temporal profile
REM atonia
Alpha wave intrusion
Beta wave intrusion
Cardiac anomalies
Sleep apnea
Periodic limb movement disorder
Abnormal behaviours during the night
NightDiary
Time at Lights off
Time at Lights on
Patient's perspective of his/hers sleep (perceived sleep time and quality)
Trips to the bathroom
Meals during the night
Vocalizations
Crisis
Urination in bed (boolean)
Slept with someone
use of Continuous Positive Airway Pressure (CPAP)
Malfunctions of the equipment
Suggested immobilisation test result
DLMO
DLMO is not computable (boolean and reason)

Melatonin Level (time and quantity for each measure)
Time at DLMO

# B

## Dashboard parameters

### B.1 Delayed sleep phase disorder dashboard

DSPD dashboard parameters, variables and the respective colour codes are shown in the table below.

Source	Colour			
	Parameter	Green	Yellow or Orange	Red
Actigraphy and PSG	Sleep onset hour	Before 2 AM	-	After 2 AM
	Wake up hour	Before 11h00 (AM)	11h00 to 14h59	After 15h00
	Total sleep time	Before 11h00 (AM)	11h00 to 14h59	After 15h00
	Latency	<= 30 minutes	-	>30 minutes
	Efficiency	>=85%	-	<85%
Actigraphy Report	Light exposure	-	-	Too much exposure during night and/or not much during day
	Matching diary and actigraphy schedules	no colour code		
Clinical Notes	Age at first symptoms	no colour code		
	Works in shifts	no colour code		
	Traumatic experience	no colour code		

	Traumatic car accident	no colour code		
	Family history	no colour code		
	Family conflicts	no colour code		
	Stress at workplace	no colour code		
	Procrastination	no colour code		
	Stress or depression	no colour code		
	Trigger	no colour code		
	Comorbidities	False	-	True (with specifications)
	Fears	False	-	True (with specifications)
	Complaints when waking up	False	-	True (with specifications)
	Daytime impairments	False	-	True (with specifications)
	Cognitive complaints	False	-	True (with specifications)
	Alcohol consumption	no colour code		
	Drug consumption	no colour code		
	Smoking habits	no colour code		
DLMO	Time at DLMO	no colour code		
Epworth Sleepiness Scale	Score	$\leq 9$	Yellow: 10-12; Orange: 13-16	$\geq 17$
PSQI	Score	-	-	$> 5$
	Bed time	no colour code		
	Wake up time	no colour code		
MCTQ	All parameters	no colour code		

MEQ	Score	23-32 (Neutral: 16-22)	Orange: 11- 15	6-10
SCL-90-R	All score	no colour code		

## B.2 Insomnia dashboard

Insomnia dashboard parameters, variables and the respective colour codes are shown in the table below.

Source	Colour Parameter	Green	Yellow or Orange	Red
PSG	Sleep Efficiency	$\geq 85\%$	-	$< \%$
	Total Sleep Time (TST)	$> 6h$	5 to 6h	$< 6h$
	Latency	$\leq 30$ minutes	-	$> 30$ minutes
	Number of Cycles	$> 4$	2 to 4	$< 2$
	REM Sleep (%TST)	no colour code		
	N1 Phase (%TST)	no colour code		
PSG Report	Patient's Perception of TST	Correct (30min range)	Optimistic	Pessimistic
	Abnormal Deep Sleep Temporal Profile	False	-	True
	Alpha waves intrusion	False	-	True
	Beta waves intrusion	False	-	True
	Sleep Apnea	No	Light or moderate	Severe
	Restless Legs Syndrome	False	-	True
	Periodic Limb Movement Disorder	False	-	True



	Cardiac Anomalies	False	-	True
Actigraphy Report	Exposure to solar light	Good	-	Irregularities
	Frequent napping	False	-	True
Vitamin D Exam	Measures Vit D	>30ng/mL	10 to 30 ng/mL	>10mL
Clinical Notes	Comorbidities	False	-	True (with specifications)
	Daytime impairments	False	-	True (with specifications)
	Cognitive complaints	False	-	True (with specifications)
	Age at first symptoms	no colour code		
	Stress or Depression	False	-	True
	Works in shifts	no colour code		
	Traumatic experience	no colour code		
	Traumatic car accident	no colour code		
	Family history	no colour code		
	Family conflicts	no colour code		
	Stress at workplace	no colour code		
	Procrastination	no colour code		
	Stress or depression	no colour code		
	Trigger	no colour code		
	Fears	False	-	True (with specifications)
	Complaints when waking up	False	-	True (with specifications)
	Alcohol consumption	no colour code		
	Drug consumption	no colour code		

	Smoking habits	no colour code		
ISI	Score	$\leq 7$	Yellow: 8-14; Orange:15- 21	$\geq 22$
Epworth Sleepi- ness Scale	Score	$\leq 9$	Yellow: 10-12; Orange:13- 16	$\geq 17$
PSQI	Score	$< 5$	-	$> 5$
	Bed time	no colour code		
	Wake up time	no colour code		
Glasgow Sleep Effort Scale	All score	no colour code		
SCL-90-R	All score	no colour code		



# Population statistics chart list

SleepData Population Statistics charts are divided in different sections, as follows:

1. General information:

- (a) Age distribution by age groups;
- (b) Gender distribution;

2. Clinical Notes:

- (a) Age at first symptoms distribution by age groups;
- (b) Related features distribution;
- (c) Triggers distribution;
- (d) Light exposure distribution;
- (e) Bad habits distribution;
- (f) Main comorbidities distribution;
- (g) Distribution of the total number of comorbidities per patient;
- (h) Main day-time complaints distribution;
- (i) Distribution of the total number of complaints per patient;

3. Actigraphy:

- (a) Total sleep time distribution with variance of the average values (from 0 to 12 hours: 1 hour steps for average values; continuous scale for variation in minutes);
- (b) Sleep latency distribution with variance of the average values (from 0 to 150 minutes: 30 minutes steps for average values; continuous scale for variation in minutes);
- (c) Average bed time distribution (from 22h to 9h, in 1 hour steps);
- (d) Average wake up time distribution (from 7h to 18h, in 1 hour steps);
- (e) Sleep efficiency distribution (slots of  $\geq 85\%$ ,  $> 85\%$ )

4. DLMO:

- (a) Sleep onset (actigraphy) - DLMO phase angle distribution (from -6 to +6 hours, in steps of 1 hour);
  - (b) Sleep onset (sleep diary) - DLMO phase angle distribution (from -6 to +6 hours, in steps of 1 hour);
- 5. Polysomnography:
  - (a) Total sleep time variance (minutes);
  - (b) Total sleep time distribution (slots of <5 hours, 5-6hours, >6 hours);
  - (c) Sleep onset variance (minutes);
  - (d) Sleep onset distribution (slots of  $\leq 30$  minutes,  $> 30$  minutes);
  - (e) Sleep efficiency distribution (slots of  $\geq 85\%$ ,  $< 85\%$ )
- 6. Comparison between bed time hour computed by actigraphy, computed by PSG and reported by patients;
- 7. PSQI:
  - (a) Score variance;
  - (b) Score distribution (slots of  $\leq 5$ ,  $> 5$ );
- 8. ISI:
  - (a) Score variance;
  - (b) Score distribution (slots of 8-14, 15-21,  $> 21$ );
- 9. Epworth Sleepiness Scale:
  - (a) Score variance;
  - (b) Score distribution (slots of 10-12, 13-16,  $> 17$ );
- 10. Glasgow Sleep Effort Scale:
  - (a) Score variance;
- 11. Symptom-checklist-90 revised;;
  - (a) Comparison of score by symptom dimension, between targeted population and healthy population.

# D

## CENC information

The clinic's information on the patients can be distinguished by its source, type and format, as follows:

- Basic information, recorded in the management software *Clínicas*;
- Clinical record, handwritten by the doctor in paper;
- Actigraphy's raw data and automatically generated reports;
- Daily log related with actigraphy, written by the patient;
- Polysomnography's raw data and manually generated reports;
- Night diary, handwritten by the clinic's technicians in paper;
- Questionnaires answered by the patients, in paper;
- Results of DLMO and Vitamin D exams, sent in paper by an external laboratory.

Each source is described thoroughly in the subsections that follow, including how the data is recorded and treated.

### D.1 Basic information

The basic information of the patient (name, birthday, gender, address, phone number and nationality) is introduced in the management software *Clínicas*, which assigns an unique identification number to every new patient. This software is also used to schedule appointments, write medical prescriptions and invoices on the services provided at the clinic. Upon a consultation, clinical information could be added to the patient's file already created in *Clínicas*. However, this software is not adapted to sleep disorders, which is the main focus of CENC, so it is not used for this purpose.

## D.2 Clinical record

During the consultation, the doctor registers the information retrieved in a handwritten clinical record (or medical history), in the form of free text. This record has the following information:

- Symptoms the patient describes;
- Time-line and possible trigger for those symptoms;
- Height and weight;
- Current medication;
- Personal daily schedule concerning the time they go to bed and fall asleep, wake up, have breakfast, lunch and dinner, and exercise;
- Work schedule;
- Lifestyle information relevant for each patient: sedentarism, smoking, drinking, drug consumption), stress and other habits or information the doctor finds significant;
- Comorbidities and family history;
- Personal information that the doctor deems relevant, such as work status, academic degree, stress related episodes and family conflicts;
- List of questionnaires to be answered by the patient (and/or guardian) and exams the doctor considers necessary to support the diagnosis;
- Any other piece of information the doctor regards as relevant to diagnose the patient.

This clinical record does not obey any standardisation pattern, other than the one created by the doctor, who usually follows the same line of narrative.

## D.3 Actigraphy and related files

As previously mentioned, CENC currently uses two distinct brands of actigraphs: Philips and Condor. Each model produces an automatic report, after allowing the professionals to mark the epochs deemed relevant. Philips' actigraph report is exported in DOCX file format and includes:

- The patient's ID, gender and age;
- The recording period;
- Free text for indications for use, in the format of a table;
- The summary statistics of the exam, including the minimum, average and maximum values for each parameter, in the format of a table;
- Free text for comments, in the format of a table;
- Graphics correspondent to the raw data, in the format of images;
- The daily statistics for each day in the recording period, including the minimum, average and maximum values for each parameter, in the format of a table.

Condor's actigraph report, exported in PDF file format, includes the same parameters as the report provided by Philips' actigraph, with following additional parameters:

- The patient's date of birth;
- Hardware and software status.

Attached to each actigraphy, is a daily log of the sleep, meals, work and exercise schedules, corresponding to the recording period of the actigraphy. This report is manually filled in by the patient.

## D.4 Polysomnography and related files

Polysomnography includes a vast ensemble of raw data of each exam exported. Each exam is thoroughly analysed by the clinic's technician's who use that data to write reports with the most relevant parameters. One report includes the general values, like total sleep time, time in bed and wake up time, sleep onset time, sleep efficiency and latency, REM latency, percentage of each sleep phase, temporal profiles, number and duration of cycles, indexes of apnea-hypopnea, micro-awakenings and oxygen desaturation, percentage of snoring and periodic sleep movements. Another report is the critical analysis of the previous values, with topics like sleep perception, insomnia criteria, sleep duration, alpha and beta wave intrusions and the existence of sleep disorders.

## D.5 Night Diary

For each PSG exam, a Night Diary is created by a CENC's technician, regarding the behaviour of the patient, in the form of free text. Although there are no standard rules for how to register that information, the usual and most relevant parameters registered are:

- The time at which the lights were turned off and on;
- The patient's perception of the duration and quality of sleep;
- Number of trips to the bathroom and the time they took place;
- Number of vocalisations;
- If the patient had any crises;
- If the patient had any meals during the night;
- If the patient urinated in bed;
- If the patient used Continuous Positive Airway Pressure ventilator;
- If the patient slept with someone;
- If the suggested immobilisation test was performed and, if so, the result;
- Patient's complaints;
- Other comments the technician finds relevant.