Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN
SCIENZE E TECNOLOGIE DELLA SALUTE
CICLO XXXV

Settore concorsuale: 09/G2 – BIOINGEGNERIA
Settore Scientifico Disciplinare: ING-INF/06 BIOINGEGNERIA ELETTRONICA E INFORMATICA

[IN SEARCH OF] DIGITAL BIOMARKERS FOR OBJECTIVE PAIN ASSESSMENT

Presentata da: Serena Moscato

Coordinatore Dottorato
Marco Viceconti

Supervisore
Lorenzo Chiari

Co-supervisori
Stefano Canestrari
Pietro Cortelli

Esame finale anno 2023

To the memory of those who fought
for the advancement of knowledge

Gutta cavat lapidem

ABSTRACT

Pain is one of the most physiologically complex phenomena a human being experiences in life, involving an intricate network of neural systems whose contribution and effect on other physiological mechanisms are not yet fully understood. Given its inherent complexity, pain is often challenging to describe comprehensively, with several factors playing a role in its under- or over-estimation. Standard pain assessment methods, based on explicit verbal communication, often fail to provide reliable and accurate information about the pain experience. This poses a critical challenge in the clinical context, where clinicians are called upon to treat pain promptly, efficiently, and personally. In the era of ubiquitous and inexpensive physiological monitoring, coupled with the advancement of artificial intelligence, these new tools appear as the natural candidates to be tested to address such a challenge.

This thesis aims to conduct experimental research to develop digital biomarkers to support the pain assessment process. The thesis is organized in the following five sections.

Section I presents state of the art regarding pain neurophysiology and assessment methods currently used in clinical settings.

Section II collects methods for appropriately conditioning physiological signals and analyses to evaluate the impact of possible confounding factors and how to control them.

The last three sections review three different pain conditions.

Section III is devoted to pain assessment in cancer patients. It includes a systematic review regarding the physiological reaction to pain in this population, and a method to automatically assess pain using physiological signals recorded in real-world contexts from cancer patients enrolled in a home palliative care program.

Section IV illustrates a study involving healthy subjects and chronic low back pain patients stimulated with noxious stimuli. Reactions to different experimental conditions in the two populations are presented, along with automatic classifiers that distinguish between healthy subjects and chronic low back pain patients.

Section V relies on pain in neurorehabilitation settings. The protocol of a clinical study involving multiple sclerosis patients undergoing neurorehabilitation treatment and a pilot study with preliminary results are presented.

In conclusion, implementing methods to support pain assessment using physiological signals and artificial intelligence algorithms has proven feasible. However, further studies on larger populations are needed to consolidate better the results presented here. A preliminary signal quality check is essential prior to any model development, along with including personal and health information in the models themselves, whenever possible, to limit their confounding effects on the physiological signals. A multimodal approach involving several physiological signals should be preferred for better performance, although also unimodal analysis revealed interesting aspects of the pain experience. Different pain sources can lead to different physiological reactions. Thus, specific models using different physiological parameters are needed in order to develop robust pain assessment.

Such an approach can enrich the routine clinical pain assessment procedure by giving the possibility of monitoring pain when and where it is actually experienced and without the involvement of explicit communication. This, in turn, would enable better characterization of the pain experience, improve antalgic therapy personalization, and bring timely relief, with the ultimate goal of ameliorating the quality of life of patients suffering from pain.

ACKNOWLEDGEMENTS

The PhD has been a long, exciting, tiring, and satisfying experience that would not have been so good if I had not had some people beside me.

First and foremost, I would like to express my sincere gratitude to Prof. Lorenzo Chiari, by whom I felt valued and who trusted me from day zero. Although it was not always easy, I am deeply grateful for the opportunity I have had to build my own career path.

I am also thankful to Prof. Pietro Cortelli, my co-supervisor, together with the Ospedale Bellaria researchers staff, for enriching my PhD project with a broader perspective. Thanks to Fabio La Porta, who enthusiastically accepted the challenge proposed in this thesis.

A huge thanks goes to my colleagues at Personal Health Systems Lab and Biolab Sabato, Luca, Elena, Pierpaolo, Silvia, Valerio, Ilaria, Jose, Addisu. Besides being precious sources of knowledge, they have been the perfect companions, always ready to lend me a hand. Special thanks to Igor, who has always been present and supportive. I am also grateful to my friend Giulia, with whom I shared happy and difficult moments of our PhD journey. Thanks to all the people I met at Unibo, students, researchers, and administrative staff, each of you has placed a building block in the person I am now.

This PhD project would never have come into being without the collaboration with Fondazione ANT. Thanks to Silvia, Andrea, Rita, and Vittoria for their unlimited support and valuable discussions we had along the road.

Many thanks to Prof. Yingzi Lin and the Intelligent Human Machine Systems Lab, for giving me the opportunity to work on an exciting project and spend a period of my life in Boston.

I would like to extend my sincere thanks to ADI Bologna (Associazione Dottorandi e Dottori di Ricerca in Italia – PhD students and PhDs in Italy Association) fellows, for their commitment and for getting me passionate about political and social issues of the academic career.

Lastly, I would be remiss in not mentioning my friends and my family, whatever their latitude or longitude. Thank you for always making me feel all your support, even hundreds or thousands of miles away. A special mention to Mom and Dad, my roots and safe haven, and to Silvio, who makes me experience the concept of “unconditional”.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

Acronym	Definition
A	Area under PPG pulse
A1	Area under PPG pulse between systolic foot and next systolic peak
A2	Area under PPG pulse between systolic peak and next systolic foot
Acc	Accuracy
ACC	Accelerometer
ACCor	Arterial Cingulate Cortex
AdaBoost	Adaptive Boosting Algorithm
AI	Artificial Intelligence
A_{ind}	Activity Index
ANS	Autonomic Nervous System
ApEn	Approximate Entropy
APSC	Normalized Average Power of EDA phasic component
AR	Activity Range
AUC	Area Under ROC Curve
BOLD fMRI	Blood-Oxygen Level-Dependent functional Magnetic Resonance Imaging
BP	Blood Pressure
BPI	Brief Pain Inventory
BQ	Basic Quality
CAN	Central Autonomic Network
Cfs	Correlation-redundancy feature set
CNS	Central Nervous System
Corr	Correlation
diaBP	Diastolicc Blood Pressure
DH	Dorsal Horn
DN4	Doleur Neuropathic - 4 questions
DR EDR	Dynamic Range of EDA phasic component
ECG	Electrocardiogram
EDA	Electrodermal Activity

EMG	Electromyography
EPs	Evoked Potentials
ESC	Electrochemical Skin Conductance
FLOP	Floating Point Operations
Fmax	Maximum Frequency of EDA signal
freqEDR	Frequency of EDA phasic component's peaks
HF	High Frequency Power
HQ	High Quality
HR	Heart Rate
HRV	Heart Rate Variability
IASP	International Association for the Study of Pain
IBI	Interbeat Interval
ICD-11	International Classification of Diseases, 11th revision
INSC	Integral EDA phasic component
KNN	k Nearest Neighbors
LANSS	Leeds Assessment of Neuropathic Symptoms and Signs Pain Scale
LEPs	Laser Evoked Potentials
LF	Low Frequency Power
LF/HF	Ratio between Low Frequency and High Frequency Power
Log	Logistic
LR	Low Threshold
MAP	Mean Arterial Pressure
maxampEDR	Maximum amplitude of EDA phasic component's peaks
MCC	Matthew's Correlation Coefficient
meanAI	Mean Activity Index
meanampEDR	Mean amplitude of EDA phasic component's peaks
meanEDA	Mean EDA
meanEDL	Mean EDA tonic component
meanHR	Mean heart rate
meanTEMP	Mean temperature
minampEDR	Minimum amplitude of EDA phasic component's peaks
ML	Machine Learning
MP	Multilayer Perceptron

MPQ	McGill Pain Questionnaire
MPS	Measurable Physiological Signals
NB	Naive Bayes
NCA	Neighborhood Component Analysis
NN	Neural Network
NPSI	Neuropathic Pain Symptom Inventory
NRS	Numerical Rating Scale
NS	Nociception-specific
PAG	Periaqueductal Gray Matter
PAT	Pain Assessment Tools
PD-Q	painDETECT questionnaire
PET	Positron Emission Tomography
pNN50	Percentage of successive NN intervals differening for more than 50%
PNS	Parasympathetic Nervous System
PPG	Photoplethysmography
PRISMA	Preferred Reporting Items for Systematic Review and Meta-Anlaysis
PTX3	Ionotropic Purinergic Receptors
QAT	Quality Assessment Tool
QST	Quantitative Sensory Testing
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies 2
Resp	Respiration
RF	Random Forest
RMSC	Normalized Root Mean Square of EDA phasic component
RMSSD	Root Mean Square of Successive Differences
RR	Respiration Rate
RVM	Rostral Ventral Medulla
S1	Secondary Somato-Sensory Cortex 1
S2	Secondary Somato-Sensory Cortex 2
SD1	Poincaré plot standard deviation perpendicular the line of identity
SD2	Poincaré plot standard deviation along the line of identity

SDANN	Standard deviation of the average NN intervals for each 5 min segment
SDNN	Standard Deviation of Normal Heartbeats
Sens	Sensitivity
SIE	Symbolic Information Entropy
slopeEDA	Slope of EDA
slopeEDL	Slope of EDA tonic component
slopeTEMP	Slope of Skin Temperature
SNS	Sympathetic Nervous System
Spec	Specificity
SPIDER	Sample Size, Population, Intervention, Study Design, Evaluation, Research Type
SpO2	Saturation of Hemoglobin with Oxygen As Measured by Pulse Oximetry
SQI	Signal Quality Index
stdAI	Standard deviation of Activity Index
stdampEDR	Standard deviation of EDA phasic component's peks
stdEDA	Standard deviation of EDA
stdEDL	Standard deviation of EDA tonic component
stdTEMP	Standard deviation of Skin Temperature
SVM	Support Vector Machine
sysBP	Systolic Blood Pressure
T1	Time interval between PPG systolic foot and successive systolic peak
T2	Time interval between PPG systolic peak and successive systolic foot
TEMP	Skin Temperature
TP	Total Power
TRP	Transient Receptor Potential
VAS	Visual Analogue Scale
VR	Virtual Reality
VRS	Verbal Rating Scale
WD	Wearable Device

WDR	Wide Dynamic Range
WHO	World Health Organization

INTRODUCTION

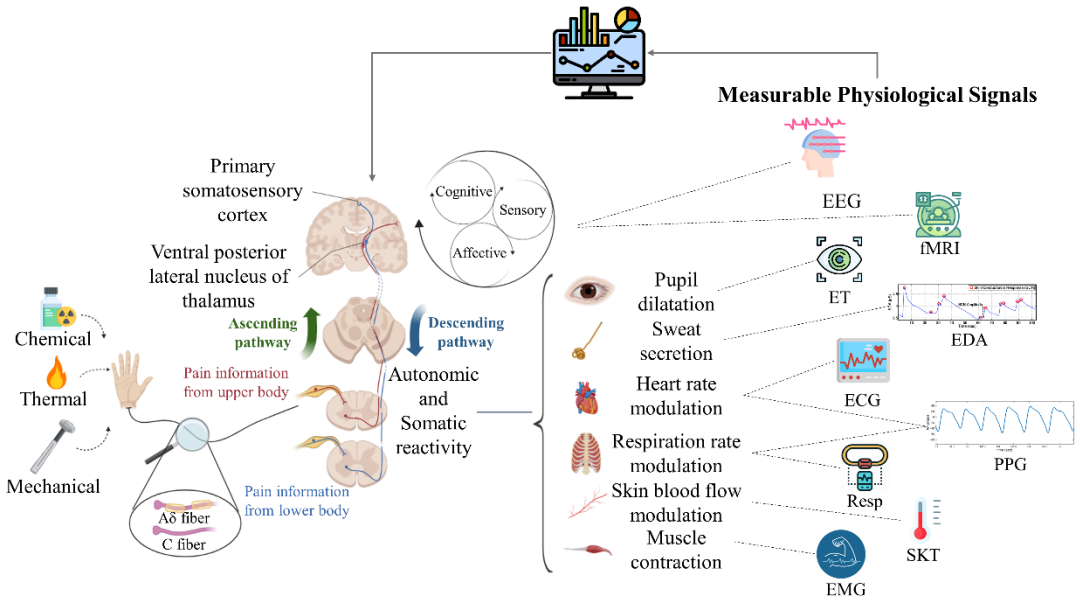
Pain is one of the most complex experiences a human being can feel, as is a mixture of sensory, cognitive, and emotional information. It works as an alarm bell to potential bodily threats, but it sometimes can arise because of a malfunctioning in the nerovus system, worsening the quality of life.

Pain is usually clinically assessed through subjective evaluations by asking the patient about the intensity and quality of the perceived pain. These pieces of information are then used to relieve pain by selecting the most appropriate antalgic therapy and during treatment to evaluate its effectiveness. Pain assessment is also often used as a diagnostic or prognostic factor for several pathologies. It follows that a reliable method to assess pain is paramount to solving the condition underlying pain itself.

Especially in the last years, the reliability of the current pain assessment tools has been questioned: they can indeed be used only when the patient is sufficiently wakeful and cooperative; such methods can be used only occasionally, not allowing continuous monitoring of the pain level; lastly, since they imply a verbal communication, they are highly dependent on the patient's willingness and way to represent his or her experienced pain.

Although some physiological processes contributing to the whole pain experience are not fully understood yet, it is well known that it alters several physiological functions as a reaction of mainly autonomic and somatic nervous systems, which are not voluntarily controlled. Thus theoretically, if the influenced physiological signals were properly

monitored (e.g., thanks to standard biomedical sensors) and analyzed, they could hint at the perceived pain. A depiction of this rationale is given in Figure 1.



EEG: electroencephalography. fMRI: functional Magnetic Resonance Imaging; ET: Eye Tracking. EDA: Electrodermal Activity. ECG: Electrocardiography. PPG: Photoplethysmography. Resp: Respiration. SKT: Skin Temperature. EMG: Electromyography.

Figure I Rationale of the PhD project

Such an approach could lead to several advantages when compared to the current pain assessment tools: it would not require the cooperation of the subject, having the dual effect of not being subject to the willingness to communicate pain, and being able to be used even in the extreme case of unresponsive patients, such as those with disorders of consciousness or coma; if the physiological signals are recorded through wearable devices, pain then could be continuously monitored, having the possibility of analyzing pain reaction when it is actually experienced.

It is worth specifying that such an approach is not that straightforward to be applied: physiological signals should be properly conditioned before being exploited to extract relevant features to pain assessment; several other factors besides pain can influence the same physiological signals, so the possible confounding factors should be properly controlled; the association between features extracted from the physiological signals and pain is not a direct one, rather a combination of changes in the physiological parameters can lead to an appropriate pain assessment.

This PhD project aims to develop automatic methods to assess pain by using physiological signals. To pursue this aim, the following were set as practical objectives:

- To select and/or develop preprocessing methods to appropriate condition physiological signals and control confounding factors
- To define an ideal pipeline for objective pain assessment
- To apply the designed pipeline to different health conditions.

The thesis is organized as follows, as also shown in Figure 2.

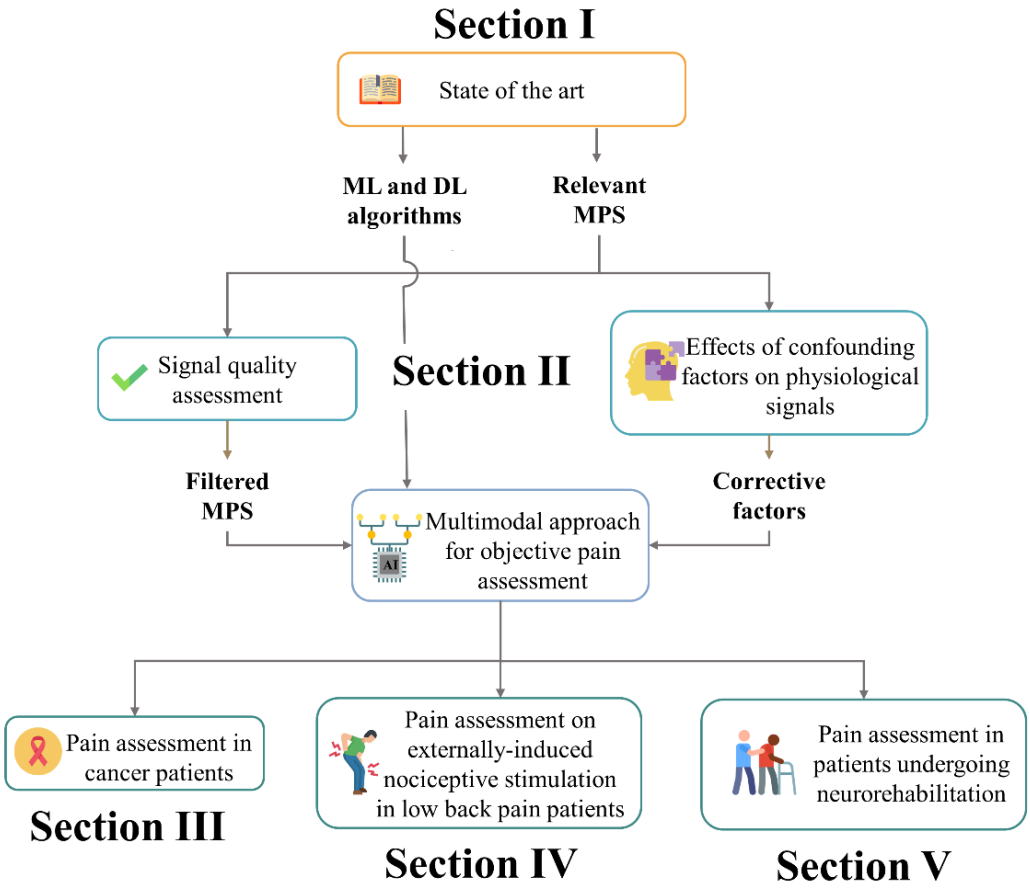


Figure II PhD project overview

Section I provides an overview of what is known about pain, describing the known neurophysiological mechanisms and its pathological conditions (Chapter 1), along with the current and innovative methods used for pain assessment (Chapter 2).

Section II is devoted to methods for appropriately preprocessing physiological signals before being used for pain assessment. A method will be presented to automatically check the quality of Photoplethysmography (PPG) signal, widely used in wearable devices to monitor the cardiovascular system (Chapter 3). PPG signal has also been assessed in its changes due to some sources of variability on it (Chapter 4). Lastly, the impact of several

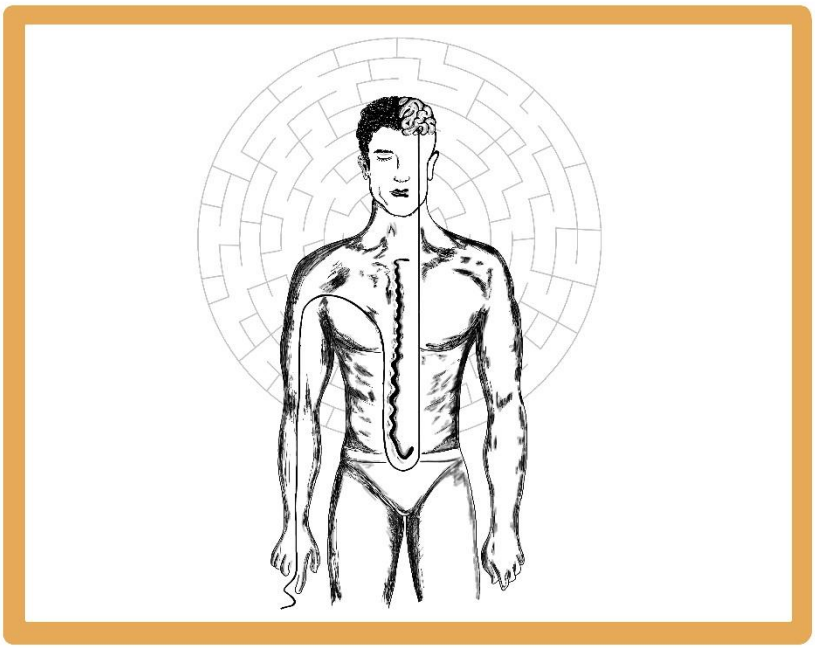
confounding factors on different physiological signals' circadian rhythm is also provided (Chapter 5).

Section III shows the work related to assessing pain on externally induced noxious stimuli in healthy subjects and chronic low back pain patients. Here a study to assess the physiological reaction to different noxious stimuli and appreciate possible differences between the two populations will be presented (Chapter 6), together with a machine learning approach to automatically classify healthy subjects and chronic low back pain patients based on physiological signals (Chapter 7).

Section IV relies on pain assessment in a cancer population. It consists of a systematic review aiming to assess the association between physiological signals and cancer pain (Chapter 8) and the design of an automatic pain assessment method that can distinguish between the absence and presence of pain based on physiological signals recorded in a real-world context (Chapter 9).

Section V investigates pain in patients undergoing neurorehabilitation. A study protocol to assess pain in this population through wearable sensors is presented (Chapter 10), together with a pilot study whose aiming to draft a practical pipeline to optimize the data collection process (Chapter 11).

SECTION I: STATE OF THE ART



1. PAIN

1.1 Definition of pain

The International Association for the Study of Pain (IASP) enacted the definition of pain for the first time in 1979, describing it as "*an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage*" [1].

The definition has been intended to provide a common understanding of pain, highlighting the main aspects of this complex experience:

- Subjectivity, as it relies on the personal experience rather than on the physiological processes;
- Multidimensionality, by expressing pain as both a sensory and emotional experience;
- Presence of actual or potential tissue damage as an inescapable condition for the pain to manifest itself [2].

Over the next four decades, several discoveries were made on pain and its related mechanisms, leading to criticism against the first definition of pain. Firstly, after the introduction of the biopsychosocial pain model by Loeser [3], it was clear that the definition lacked two components: cognition and psychosocial factors. Secondly, the definition emphasizes the verbal communication of pain, leaving apart cognitively impaired or non-verbal individuals who cannot express pain with their voice [4].

Several attempts have been made to propose a new definition of pain [5]–[7]. However, only in 2020 the IASP Presidential Task Force released the new definition of pain after a 2-year period of consultation. The new definition is "*an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage*" [8]. A comparison between the two definitions and the related notes is given in Table 1.1.

Table 1.1 Definition of pain, 1979 vs 2020. Adapted from [8]

IASP definition of pain (1979)	**Revised IASP definition of pain (2020)**
An unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage	An unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage.
Notes	Notes
Pain is always subjective. Each individual learns the application of the word through experiences related to injury in early life. Biologists recognize that those stimuli which cause pain are liable to damage tissue. Accordingly, pain is that experience which we associate with actual or potential tissue damage. It is unquestionably a sensation in a part or parts of the body, but it is also always unpleasant and therefore also an emotional experience. Experiences which resemble pain, e.g., pricking, but are not unpleasant, should not be called pain. Unpleasant abnormal experiences (dysaesthesiae) may also be pain but are not necessarily so because, subjectively, they may not have the usual sensory qualities of pain. Many people report pain in the absence of tissue damage or any likely pathophysiological cause; usually this happens for psychological reasons.	Pain is always a personal experience that is influenced to varying degrees by biological, psychological, and social factors. Pain and nociception are different phenomena. Pain cannot be inferred solely from activity in sensory neurons. Through their life experiences, individuals learn the concept of pain. A person's report of an experience as pain should be respected Although pain usually serves an adaptive role, it may have adverse effects on function and social and psychological well-being. Verbal description is only one of several behaviors to express pain; inability to communicate does not negate the possibility that a human or a nonhuman animal experiences pain.
There is no way to distinguish their experience from that due to tissue damage if we take the subjective report. If they regard their experience as pain and if they report it in the same ways as pain caused by tissue damage, it should be accepted as pain. This definition avoids tying pain to the stimulus. Activity induced in the nociceptor and nociceptive pathways by a noxious stimulus is not	Etymology Middle English, from Anglo-French peine (pain, suffering), from Latin *poena* (penalty, punishment), in turn from Greek *ποινε* (payment, penalty, recompense). *The Declaration of Montreal, a document developed during the First International Pain Summit on September 3, 2010, states that "Access

pain, which is always a psychological state, even though we may well appreciate that pain most often has a proximate physical cause.	to pain management is a fundamental human right."

1.2 Classification of pain

Although the definition of pain may provide a common framework, to achieve effective treatment of pain, it is necessary to gain other pieces of information about it. The guidelines to properly describe pain give a systematic approach to do that for pain classification, disseminated by the IASP Task Force of Taxonomy and also adopted by the World Health Organization (WHO) [9], [10].

Pain can be classified according to four different characteristics:

- Pathophysiological mechanism

  Nociceptive pain: it arises from the activation of nociceptors, sensory receptors deputed to detect damaged or potentially damaged tissue. This type of pain is further divided into somatic pain, arising from superficial tissues, and visceral pain, coming from internal organs and blood vessels

  Neuropathic pain: it is caused by structural damage and/or dysfunction in nerve cells in both the central and peripheral systems

  Mixed pain: in some cases, neuropathic pain may coexist with nociceptive pain. The different pathophysiological mechanisms involved in the two types of pain may work together to produce mixed pain

  Nociplastic pain: pain that arises from altered nociception despite no clear evidence of actual or threatened tissue damage causing the activation of peripheral nociceptors or evidence for disease or lesion of the somatosensory system causing the pain [11]

- Duration

  Acute: it is characterized by a sudden onset, it is felt immediately after injury and can be of severe intensity, but is usually of short duration. It is often linked to a injury or trauma, representing the alarm bell to the body to take action

  Chronic: it is continuous or recurrent pain that persists beyond the normal expected healing time, lasting for more than 3 months [12]

Breakthrough: it is characterized by a temporary increase in pain severity above the pre-existing baseline level. It is of sudden onset, severe and short duration, and may occur independently of any stimulus

Episodic: it occurs intermittently over a long period of time; episodes can vary in intensity, quality and frequency over time, being consequently unpredictable

End of dose: it occurs when drugs cannot reach the correct therapeutic level

- Etiology

  Malignant: it is a pain linked to cancer and related treatment

  Non-malignant: a pain not derividing from cancer or related treatment

- Anatomical district [13]

It is worth mentioning that there still exist some forms of pain that cannot be easily classified based on the aforementioned factors, such as fibromyalgia and primary headaches [12].

1.3 Pain Theories

Over the centuries there have been several theories about the mechanisms underlying pain, representative of the thinking and culture of the time [14]. Although more recent theories have replaced the older ones, each has helped to add a piece of knowledge to a phenomenon that is still not fully understood today. In this paragraph the main pain theories are presented.

1.3.1 Intensity Theory

The first attempt at describing pain perception is attributed to Plato, who, in his oeuvre Timaeus, in the fourth century B.C., described the pain as a more emotional rather than sensorial experience occurring when a stimulus is presented as more intense and lasting longer than usual [15]. More than a thousand years later, in 1874, Wilhelm Erb reintroduced the same notion, suggesting that pain arises when any strong sensory stimulus is imposed. The intensity theory has been further developed by Arthur Goldscheider, who carried out an experiment in which he demonstrated that the rapid and repeated presentation of a subthreshold tactile stimulus leads to experiencing what was called by the subject of the experiment “unbearable pain”. Goldscheider concluded that either repeated subthreshold stimulation by some form of summation or suprathreshold stimulation could cause pain [16]. Although this theory has been abandoned in favour of

the specificity theory, it contributed to the development of more recent theories that are still valid today.

1.3.2 Specificity Theory

René Descartes originally introduced Specificity Theory in Treatise of Man published in 1664, which basically relies on the assumption that different types of sensations belong to different pathways. A fundamental part in the development of this theory by Descartes is given by his description of nerves, called fibrils in his work, described as hollow tubules transmitting both sensory and motor information. Probably one of the most historically famous pictures related to pain is the illustration drawn by Louis La Forge based on Descartes' description of the pain pathway in Treatise of Man, presented in Figure 1.1: when the stimulus (A), in the form of fire in this specific case, reaches the body (B), particles from the stimulus move towards the skin and tug on the fibril, which makes open the pores of the other end of the fibril in the brain (d, e), allowing the particles to reach the fibril cavity (F). As a response, the "animal spirits" flow from the cavity into the fibril, making the subject look at the fire, move the foot away from the stimulus, and in general to act protective mechanisms [16].

Almost two centuries later, Charles Bell contributed to expanding the Specificity Theory, indicating a more complex organization of the brain with respect to the one given by Descartes and corroborating the principle for which there is a specific pathway for pain information [15], [16].

Specificity theory received several contributions in the following years. After substantial advancements in neurophysiology, Schiff and Woroschiloff between 1854 and 1859 established the location of a pain pathway in the spinal cord, separated from that of touch [16]. Later on, in 1894, Maximillian von Frey demonstrated that the somatosensory system is composed of specific receptor for tactile, hot, cold, and pain receptors, each with a specific transmission method [17], perfectly aligned with the Specificity theory.

Specificity theory has been the main paradigm to explain pain for the whole 19th and much of the 20th century. One of the main drawbacks of Specificity theory is that it has always focused on sensory functioning, without considering all the other factors that are now well-established to play a key role in the complete perception of pain, which will be formalized in later theories.

Figure 1.1 Drawing of pain pathway by Louis La Forge based on Descartes' theory

1.3.3 Pattern Theory

The Pattern theory is mainly ascribed to the psychologist John Nafe, who rejected the presence of specialized somatosensory pathways as an explanation for different types of perception. In the 1930s, he proposed that different somatosensory sensations come from different signal patterns in the brain that can be modulated in frequency and timing [16]. This theory helped explain complex pathological pain conditions, such as allodynia, despite its refusal the presence of different sensory pathways, which are now consolidated by several experiments [17].

1.3.4 Gate Control Theory

The Gate Control theory allowed bringing together the Specificity and Pattern theories, revolutionizing the field of pain research. It was theorized by Ronald Melzack and Patrick Wall in 1965, and it stated the presence of a “gate” within the dorsal horn, which is responsible for inhibiting or facilitating pain impulses [18]. Besides this kind of control, the two scientists also stated that another control mechanism is given by the cortical areas of the brain, thus introducing for the first time the psychological and cognitive aspects of pain [17].

1.3.5 Pain neuromatrix

Originally introduced by Melzack in the 1960s, the pain neuromatrix theory's publication enshrines the multidimensional characteristics of pain once and for all. Melzack owes this

insight to his work with amputees and their phantom limb pain: without any real sensorial stimulus, patients with amputated limbs still suffered from pain [19]. So pain is not only due to a mere somatic stimulus. Rather, it is the result of the interactions of different systems that modulate the final perception of pain, both at the peripheral and central level [17]. This brand new perspective on pain allows us to apply new strategies for pain relief, healing not only the physical aspect of pain, but also the cognitive and emotional aspects of it.

1.4 Neurophysiological mechanisms of pain

Pain perception involves the synergistic work of both Peripheral Nervous System and Central Nervous System (CNS). While the former comprises nerves and ganglia that gather the nociceptive signalling to be conveyed at the central level, the latter integrates the information into a complex net of cerebral and cortical regions. The main events occurring in pain experience are transduction, transmission, perception, and modulation [20].

1.4.1 Transduction

Several body districts, from the skin to subcutaneous structures, have specialized sensory receptors that are activated in response to nociceptive stimuli, called nociceptors. Most nociceptors are free nerve endings of primary sensory neurons. There are four main classes of nociceptors:

- Thermal nociceptors: activated by extreme temperatures (< 5°C or >45°C); they are peripheral endings of small diameter myelinated Aδ axons;
- Mechanical nociceptors: activated by pressure on the skin, they are peripheral endings of myelinated Aδ axons as well;
- Polymodal nociceptors: activated by mechanical, chemical or thermal stimuli, this class of nociceptors is found at the terminations of unmyelinated C axons of small diameter
- Silent nociceptors: they can be found in the viscera; several chemical agents lower their their firing threshold.

The transduction from nociceptive stimulus to nociceptor depolarization occurs by axon membrane receptors, including transient receptor potential (TRP) ion channels, tetrodotoxin-resistant Na+ channels, and ionotropic purinergic receptors (PTX3).

Depolarization activated by the receptor then causes an action potential to be triggered that propagates to central areas of the nervous system [21].

The different classes of nociceptors are widely distributed throughout the body and often act in a coordinated manner: when a nociceptive stimulus elicits the body, there is an initial fast, sharp pain (called “first pain”) transmitted by the fast myelinated Aδ axons, with a diameter of 2-5 μm and a conduction velocity of 5-15 m/s, and then a slow pain appears (called “second pain”), transmitted by the C fibers with a diameter smaller than 2 μm and a conduction velocity of 0.5 – 2 m/s [21], [22]. A representation of the action potential speeds of different axons is presented in Figure 1.2.

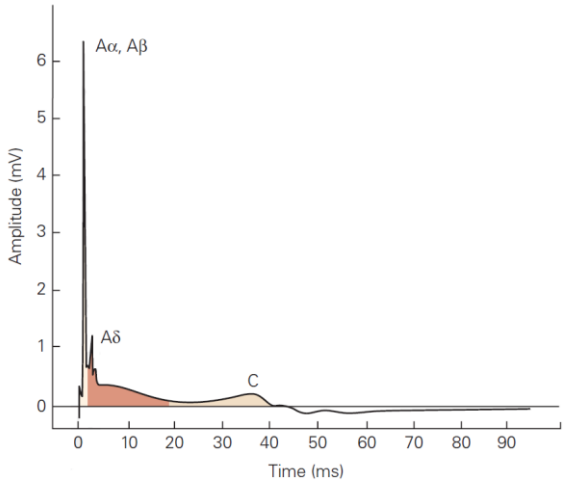


Figure 1.2 Axons' action potential speed. Adapted from [21]

1.4.2 Transmission

1.4.2.1 Peripheral Transmission

The nociceptive signal travels through axons to the central body of nociceptors, which are located in the dorsal root ganglia or trigeminal ganglia. The central branches of such neurons then terminate in the spinal cord, most precisely in the dorsal horn (DH). DH is divided into six layers, called laminae, which are the sites where the primary afferent fibers synapse with second-order neurons, which can be of three types:

- Nociception-specific (NS): respond selectively to high threshold nociceptive stimuli conveyed by Aδ and C fibers located in laminae I, II, and III
- Wide dynamic range (WDR): respond to a variety of sensory stimuli, receiving information from somatic and visceral nociceptors, found in laminae V and VI

- Low-threshold (LR): respond only to innocuous stimuli, and produce complex responses since the incoming signal is given by passing through several intermediate synapses located in laminae VII and VIII [23]

There is a close link between the anatomical organization of DH laminae and their functions in sensory processing [21]. For example, in lamina V nociceptive information from both somatic and visceral nociceptors is conveyed, which leads to the "referred pain" phenomenon, that is, the perception of a visceral pain rising on the surface of the body [24].

Activation of neurons located in the DH occurs mainly through the release of glutamate and neuropeptides, which perform coordinated actions to regulate the action potential [21].

1.4.2.2 Central Transmission

A subset of projection neurons in the DH transmits the nociceptive information to central nervous districts mainly via five ascending pathways:

- **Spinothalamic tract**: this is the main ascending pathway in which nociceptive information from lamina I, V, VI, and VII neurons travels, located in the white matter of the spinal cord. It is composed of two parts: the lateral spinothalamic tract, which transmits pain and temperature sensation, the anterior spinothalamic tract, which transmits information about touch and pressure [20].
- **Spino-reticular tract** contains axons from neurons in lamina VII and VIII, which projects to thalamus and hypothalamus via nuclei of brainstem reticular formation [22].
- **Spino-mesencephalic (or spino-parabrachial) tract**: it contains axons from lamina I and V. Information transmitted along this tract is thought to be responsible for the affective component of pain, as the axons project the information in part to the parabrachial nucleus, which in turn projects to the amygdala, the key nucleus of the limbic system that regulates emotional state.
- **Cervicothalamic tract**: it contains axons from neurons in the lateral cervical nucleus, which in turn receives input from neurons in lamina III and IV of the dorsal horn. Most axons in this tract terminate in the midbrain nuclei and the lateral ventroposterior and posteromedial nuclei of the thalamus.

- **Spinohypothalamic tract**: it contains the axons of neurons present in lamina I, V, and VIII. These axons project information to hypothalamic nuclei that act as autonomic control centers involved in the regulation of neuroendocrine and cardiovascular responses that accompany pain perception [21].

1.4.3 Perception

Conscious perception of pain is the result of processes occurring at the level of the brain, with the active involvement of several brain regions, representing the "Pain Neuromatrix" [25], whose teory has been presented earlier in this chapter. Several cortical and subcortical networks are involved in pain perception, activating sensory, limbic, and associative areas [26].

1.4.3.1 Thalamic projections

At the thalamic level, the different tracts have mainly two termination sites: ventrocaudal and medial. Neurons in the ventrocaudal thalamus project information directly to the primary and secondary somato-sensory cortex (S1 and S2). The medial thalamus receives, in addition to some indirect input, significant input from the brainstem, subsequently projecting the information to S1 and S2. Information received by the ventrocaudal thalamus is thought to be mainly responsible for acute, well-localized pain arising on the body surface, while information coming to the medial thalamus responds more to stimuli from deep somatic and visceral structures [27].

1.4.3.2 Thalamus-Encephalon connections

Third-order neurons, with cell bodies in the ventral posterior lateral nucleus of the thalamus, project the signal to S1 and S2, specifically to Broadman's areas 1-3, involved in sensory discrimination of pain (intensity, localization, and quality) and integration of pain information with visual, auditory, and gustatory inputs, respectively [27]. The signal is also projected to cortical limbic areas, such as the anterior cingulate cortex (ACCor) and insular cortex, involved in mediating affective/emotional components [28]. Simultaneously, nuclei adjacent to the thalamus receive projections from the spinothalamic tract and mediate some behaviours in response to pain, such as arousal level and emotion [27]. In particular, activation of the ACCor and insular cortex is associated with the subjective experience of pain [29].

The nociceptive signal is also transmitted through the spino-reticular and spino-mesencephalic tracts to brain areas involved in the emotional and affective aspects of pain, such as the amygdala, basal ganglia, hypothalamus, periaqueductal gray matter (PAG), nucelus accumbens, and rostral ventral medulla (RVM) [30].

1.4.3.3 Psychological factors

Attention

The emblematic example of the influence of attention on pain perception is the wounded soldiers who do not feel pain at all as long as they are on the battlefield [31]. The influence of attention and mind on pain has been well recognized in Eastern culture for millennia. In contrast, only recently in Western culture such an aspect has been considered a means of pain relief and control [26].

Attention modulation of pain experience correlates with changes in the activation of the Pain Neuromatrix. Concurrently, low attention corresponds to strong activation of the prefrontal cortex, and PAG, suggesting a significant interaction among brain systems involved in pain modulation. In contrast, hyper-vigilance amplifies pain intensity and is associated with interpreting harmless sensations as painfully unpleasant [29].

Cognition

Pain perception is based on a process of cognitive evaluation. The subject, consciously or unconsciously, assesses the meaning of sensory signals from the body to determine the extent to which they indicate the presence of actual or potential harm. Such assessment is purely subjective. The inherent variability in the cognitive assessment of pain may result from neurobiological dissociations between sensory and affective aspects of the pain experience [29].

Furthermore, is it well known that a bidirectional relationship then links cognition and pain, although the underlying physiological mechanisms are still not fully understood [32].

Changes in the intensity of pain result in altered activation of the cerebral cortex, while changes in the unpleasantness of pain result in altered activation of the anterior cingulate cortex. Perceived pain intensity is reduced when pain is perceived as controllable. The activation of the prefrontal cortex, also involved in emotion regulation, is negatively correlated with perceived pain intensity [29].

Emotion

The very nature of pain elicits a powerful emotional response, which is fuelled by modulating the perception of pain. Pain often provokes feelings of anger, sadness, and fear, depending on how it is cognitively evaluated. Notably, it has also been demonstrated the high incidence of affective disorders in chronic pain patients [30]. Negative emotions are associated with increased activation of the amygdala, anterior cingulate cortex, and anterior insular cortex. Fear of pain is also associated with hyper-vigilant states toward stimuli related to painful sensations. Negative emotions also affect the attention given to pain, which inevitably increases its unpleasantness. In addition, negative emotions and stress alter the functions of the prefrontal cortex, which then may reduce its effectiveness in regulating pain by using cognitive strategies such as reappraisal or seeing pain as controllable. Therefore, anger, sadness, and fear may result from pain, but they enter the feedback of biobehavioural processes that influence pain perception, exacerbating distress and suffering [29].

1.4.4 Modulation

Through the whole pain pathways, nociceptive or pain information is subjected to modulation events by up- or down-regulation [33]. Modulation can be distinguished into two major contributions: modulation during the ascending and descending pathways.

1.4.4.1 Modulation in the ascending pathway

The main contribution to modulation within the ascending pathway is based on the principles postulated by the “gate control theory”, presented earlier in this chapter. Although such theory extremely simplifies what happens during the transmission phase (based on a research conducted on 1965) [34], it provides a good point of view to appreciate how the nociceptive information is modulated before arriving in the brain.

By updating the original theory, it is now well accepted that A fibers non-nociceptive activity can lower the inflow of nociceptive information coming from C fibers by “shutting” the gate, thus partially inhibiting the nociceptive information arriving at the brain. This discovery had a significant clinical implication since several pain relief methodologies have been developed based on that, such as transcutaneous nerve stimulation or spinal cord stimulation [34], [35].

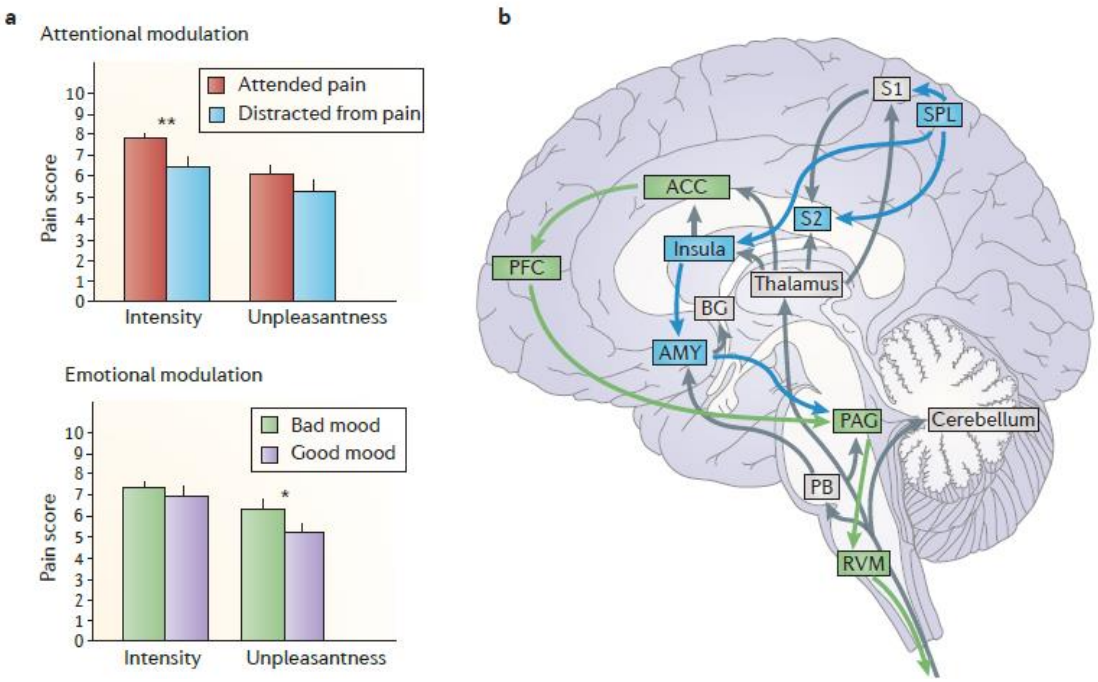


Figure 1.3 Afferent pain pathways and the mechanisms of activation underlying attention (in blue) and emotions (in green) [26]

1.4.4.2 Modulation in the descending pathway

Based on studies conducted by analyzing both the functional and morphological aspects of the central nervous system, it has been demonstrated that one of the main contributors to descending pain modulation arises by the activity of the brain stem, which receives input from higher level brain areas and can either down- or up-regulate the pain signal to be sent to the spinal cord. Precisely, the brain stem is linked with frontal brain areas, the amygdala, and the hypothalamus, which are associated with the cognitive and emotional dimensions of the human being [36].

The output from the PAG-RVM network gives another element involved in the modulation throughout the descending pathway. Specifically, RVM receives the information already processed by the PAG, and transmits the signal down the spinal cord, activating the endogenous opioid system that inhibits pain [37].

1.5 Pathological pain conditions

1.5.1 Chronic pain

Chronic pain moves from being a symptom to becoming a disease entity on its own, as also recognized by the International Classification of Diseases, 11th revision (ICD-11), when pain persists or recurs for more than 3 months [38]. Chronic pain conditions have

been rated as among the top leading causes of years lost to disability [39], and its prevalence has been estimated between 17% and 27% [40].

It has recently been demonstrated that chronic pain leads to significant functional and anatomical reorganizations of brain activities. For example, the mesocorticolimbic system, one of the principal dopaminergic pathways responsible for encoding motivation and rewards and not implicated in the acute pain pathways, has been shown to have a role in the development and persistence of chronic pain [40]

1.5.2 Neuropathic pain

Neuropathic pain arises because of lesions or diseases involving the somatosensory nervous system [41]. It basically represents a "*maladaptive response of the nervous system to damage*" [42]. While it is not that straightforward quantifying the incidence of neuropathic pain because of the lack of precise diagnostic criteria, from some analysis, it emerged that between 7 and 10% of the general population is affected by neuropathic pain [43].

Different mechanisms can be responsible for the generation of neuropathic pain symptoms, sometimes acting in combination:

- Peripheral mechanisms: the damage to sensory neurons can lead to different complications, such as spontaneous ectopic discharge of Aδ and C fibers or alterations of gene expression, which in turn cause a modification in the neurotransmitters released;
- Central mechanisms: the main part of the CNS involved in neuropathic pain has been proven to be the DH, where phenomena such as wind-up (i.e., the altered response of dorsal horn neurons to repeated stimuli from C fibers), central sensitization or disinhibition may happen [44].

A depiction of the main changes induced by neuropathic pain in the physiological pain pathways is given in Figure 1.4.

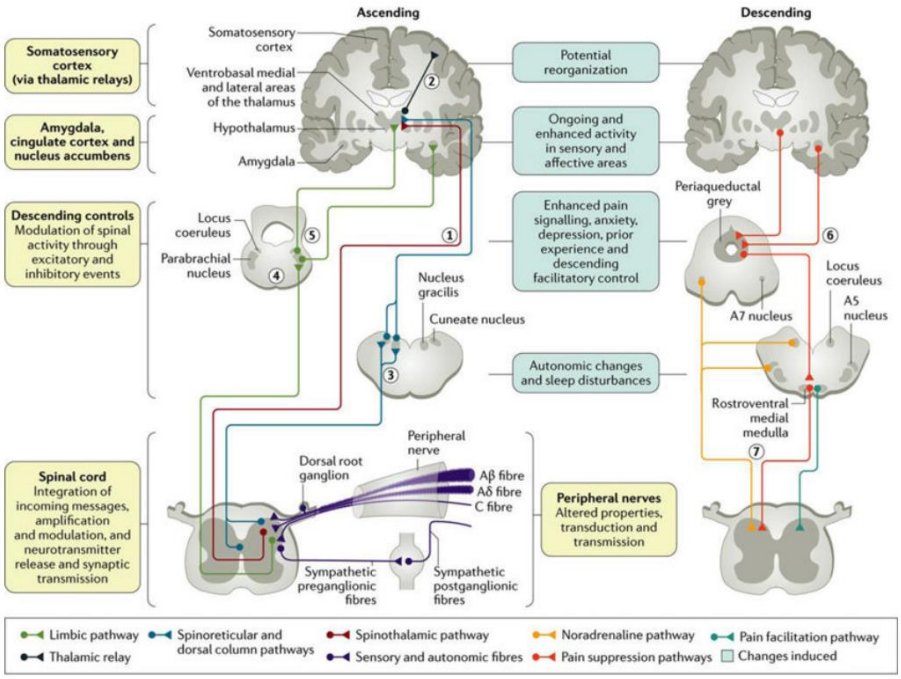


Figure 1.4 Modifications induced by neuropathic pain in pain pathways [44]

The main symptoms of neuropathic pain are essentially allodynia, which consists of feeling pain from an innocuous stimulus, and hyperalgesia, which consists of an excessive painful reaction to a painful stimulus [45].

1.6 Physiological consequences of pain

The presence of pain leads to several modifications throughout the body as a reaction to pain itself. The main elements impacted by pain are neuroplasticity and the autonomic nervous system activity.

1.6.1 Neuroplasticity

Neuroplasticity is the ability of the nervous system to reorganize its functions, structures, or connections based on repeated and/or intense stimuli the person is subjected to [46]. Several studies demonstrated significant changes in the CNS happening due to pain. Such permanent modifications are thought to be the main cause of the development of chronic pain, although the transition between acute and chronic pain is not yet fully known [47]. Neuroplastic changes can occur both in the spinal cord and in the brain:

- The DH can be subjected to neuroplastic changes due to, for example, damage occurring at the peripheral level that alters the somatosensory system. In this case, such modifications cause pain even in the presence of non-nociceptive stimuli

(i.e., allodynia) [48]. This manifestation can arise because of sensitization, wind-up, or altered reactions of neurons in the spinal cord [49].

- Different brain areas can be subjected to neuroplastic modifications after being repeatedly subjected to nociceptive stimuli. For example, it has been shown that there is a shift in neural circuits activation transitioning from acute to chronic pain, with a more pronounced activity of affective-emotional circuitry for chronic pain rather than a more sensory activation of the acute pain [50]

1.6.2 Autonomic Nervous System

The processes that are activated and contribute to the perception of pain are, in turn, intimately connected to the Autonomic Nervous System (ANS), which in turn consists of the Sympathetic Nervous System (SNS) and the Parasympathetic Nervous System (PNS). The body's response to pain is defined by changes in the ANS, although to date, it is still not completely clear how pain-ANS interactions may be reflected by functional connections in the brain [51]. On the other hand, alterations in the ANS may, in turn, influence the pain experience [52]. One possible mechanism underlying the pain-ANS interaction is baroreflex, the negative feedback that allows blood pressure to remain stable. Indeed, this mechanism has been associated with a reduction in pain perception both during spontaneous episodes of high blood pressure (i.e., when baroreceptors are activated) and during mechanical stimulation of baroreceptors [52].

The ANS is structurally and functionally positioned to interface between the internal and external environment, coordinating body functions to endure homeostasis and adaptive responses to stress. Central control of SNS and PNS activities, called Central Autonomic Network (CAN) involves several interconnected areas distributed throughout the neuroaxis. CAN plays a key role in the ongoing control of visceral function, homeostasis, and adaptation to internal or external conditions [51].

ANS, in turn, affects several physiological processes besides the volition of the subject, such as some functions of the cardiovascular system, respiration [53], and sweating [54].

2. PAIN ASSESSMENT

2.1 State-of-the-art pain assessment tools

Pain assessment is of paramount importance in providing appropriate and effective pain treatment. Representing the first step in the pain management procedure, it is essential not only to alleviate the pain itself, but also to be included in the patient's more complete clinical picture to get a sense of the progress of the disease and whether a treatment is effective.

Besides unstructured and structured interviews used at the beginning of the pain management path [55], pain assessment tools used in normal clinical practice are essentially scales and questionnaires. In special cases, instrumented tests are also used. All three tools are explored in more detail in the following paragraphs.

2.1.1 Scales

Scales provide a one-dimensional measure of pain, gathering information only about the intensity [56]. The three main scales used are the Numerical Rating Scale (NRS), the Visual Analogue Scale (VAS), and the Verbal Rating Scale (VRS).

The NRS consists of a 11-, 21-, or 101-point scale, where one end represents the absence of pain and the other the worst pain imaginable [57]. The 11-point NRS is the most used one, and its scoring is often interpreted as:

- 0 = no pain
- 1-3 = mild pain

- 4-6 = moderate pain
- 7-10 = severe pain [56]

A graphical depiction is given in Figure 2.1. Information provided by the NRS can be easily documented, are well understandable with minimal cultures and language barriers [56], and can be collected via face-to-face communication, in telephone interviews, and saved in pain diaries [58]. NRS has been proven to be a reliable scale, although its psychometric properties can depend on the body area where the pain is experienced [59]. Changes of 20% in the NRS can be considered clinically significant [58].

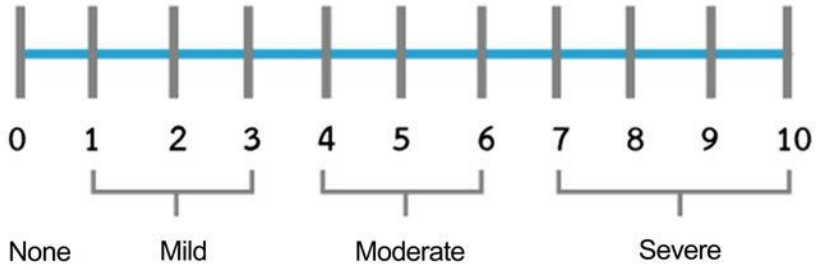


Figure 2.1 Numerical Rating Scale (NRS) and its intervals [60]

The VAS is represented by a straight line, with a length comprised between 5 cm and 20 cm (usually 10 cm), delimited by the two extremities representing "no pain" and "worst possible pain" [58]. The subject is asked to rate his or her pain by marking it on the line. The distance between "no pain" and the mark represents the subject's pain level. A representation is given in Figure 2.2.

Unlike NRS, VAS scoring is more time-consuming (it implies the measurement between the end and the mark), and it is more difficult to be understood, thus being more prone to misinterpretation [58], and it cannot be used in an emergency situation [56]. On the other hand, it has a high test-retest reliability, a strong correlation with other pain assessment tools, and a minimum clinically significant difference of 1.37 cm on a 10-cm line [61].

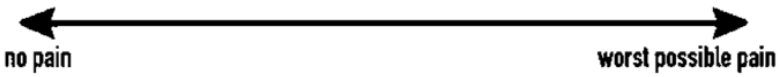


Figure 2.2 Visual Analogue Scale (VAS) [56]

The VRS is composed of a list of adjectives ordered in ascending order of magnitude describing different levels of pain. The subject is asked to choose the word that best describes his or her pain intensity. Usually, four descriptors are used: none, mild, moderate, and severe, but there are also VRS comprising six or fifteen descriptors [56], [58]. When a 4-descriptors VRS is used, the small number of options requires a large

change in the experienced pain before it can show up on the scale [57]. Furthermore, VRS is more likely to be misinterpreted by patients, and often pain intensity is interpolated from other pain dimensions [62].

2.1.2 Questionnaires

Questionnaires can help obtain a more comprehensive picture of the complex experience of pain. These tools are based on a multidimensional approach to tracking pain, going beyond the only pain intensity described by scales.

A preliminary distinction can be made by classifying questionnaires as multipurpose questionnaires that can be used in several pain conditions and in questionnaires specifically developed to assess neuropathic pain.

In the first group, the most used questionnaires are the McGill Pain Questionnaire and the Brief Pain Inventory.

The McGill Pain Questionnaire (MPQ) is so far the most widely used pain assessment tool [63]. First published in 1975 [64], the MPQ consists of three parts, each one measuring a different pain dimension: the first part is designed to mark the anatomical part where the pain is located; the second part consists of a VRS to rate the intensity of pain; the third part provides a list of 20 sets of 72 pain descriptors the subject can choose among to describe the sensory (10 sets), affective (5 sets), evaluative (1 set) and miscellaneous (4 sets) dimensions of pain [63], [65]. A few years after it was first published, the short-form MPQ was developed to have a more efficient tool in terms of compiling time. It comprises 15 descriptors for sensory and affective dimensions [66]. MPQ has been extensively used in clinical practice and has proven to be a reliable and valid tool to quantify the pain experience [65].

The Brief Pain Inventory (BPI) [67] was originally developed for cancer pain assessment, and it is currently used for many different pain conditions. The most used version consists of 9 items (the long version has 17 items), consisting in a single item related to the pain experienced on the day, a diagram where the subject marks the anatomical part where the pain is located, questions about the pain intensity, pain relief and pain interference with activities of daily living. BPI offers good reliability in pain intensity and interference and moderate correlation with other pain questionnaires [68].

Regarding questionnaires to assess neuropathic pain, they can be further divided into two categories: screening questionnaires, used to identify the presence of neuropathic pain,

and assessment questionnaires, developed to keep track of neuropathic pain symptoms [69].

Screening questionnaires are mainly based on verbal pain descriptors to discriminate between nociceptive and neuropathic pain. The most commonly used in clinical practice, which have the higher sensitivity and specificity in detecting neuropathic pain [69], are the following:

- Leeds assessment of neuropathic symptoms and signs (LANSS) Pain Scale, consisting in a patient-completed questionnaire and a short clinical assessment [70];
- painDETECT (PD-Q), comprising seven neuropathic pain descriptors, two items describing the timing of pain, and one item describing how pain radiates throughout the body [71];
- Doleur Neuropathic – 4 questions (DN4) questionnaire, which is composed of four questions with several yes/no items within each one of them [72].

Assessment questionnaires comprise questions about the sensorial description of neuropathic pain [69]. Questionnaires specifically developed to assess neuropathic pain are:

- Neuropathic Pain Scale, consisting of 10 questions to be rated on a 0-10 numerical rating scale and a question representing the time quality of pain [73]
- Neuropathic Pain Symptom Inventory (NPSI), comprising 5 scales, each one related to a specific characteristic of neuropathic pain (burning, pressing, paroxysmal, evoked, and paresthesia/dysesthesia); it is particularly useful for pain assessment in several patients who had brain trauma [74] and spinal cord injury [75].

2.1.3 Instrumented tests

In addition to scales and questionnaires, pain assessment in clinical practice can also be based on other methods that can be defined as “objective”. They make use of instrumentations aiming at defining the pain experience from a physiological point of view and providing more information to complete the clinical picture of the patient.

2.1.3.1 Quantitative Sensory Testing

Quantitative Sensory Testing (QTS) is a method used in clinical practice to standardly evaluate sensory function. It consists of applying several mechanical and thermal stimuli

to activate both large (Aβ) and small (Aδ and C) fibers. QST is used to assess whether there is an increase or decrease in nerve fiber activation. It is used especially in patients experiencing neuropathic pain [74].

2.1.3.2 Central metabolic activity

Although used only in selected cases, functional magnetic resonance imaging (fMRI) is one of the most widely used imaging techniques for pain assessment. The fMRI indirectly measures brain metabolic activity by detecting related changes in blood flow oxygenation [76].

fMRI can provide extremely useful information in acute pain or experimental pain when they are of short duration. Only recently has it proven effective in the study of chronic pain [77].

2.1.3.3 Evoked potential

Evoked potentials (EPs) test is usually implemented in clinical practice to evaluate the brain response to sensory stimuli [74].

EPs are recorded using electroencephalography (EEG), placing electrodes on the scalp. From the recorded signals, specific characteristics (i.e., amplitude, latency) are analyzed in order to obtain information about afferent sensory pathways based on which kind of stimuli has been applied [78].

The laser evoked potentials (LEPs) are mainly used to assess neuropathic pain. LEPs specifically activate Aδ and C fibers, inducing a first acute response mediated by Aδ fibers and a second diffuse response mediated by C fibers [79].

2.1.4 Limits of current tools

Scales and questionnaires suffer from several limitations:

- They require the patient to be able to communicate his/her pain, thus excluding clinical conditions such as coma or disorders of consciousness [80]
- In some cases, patients tend to underestimate their pain because it is assumed to be directly related to worsening disease [81]
- Healthcare professionals could neglect pain assessment to dedicate more time to diagnosis and treatment [82]
- Assessment is conducted occasionally and usually in clinical settings.

Regarding the instrumented tests, the negative aspects are mainly related to their elevated costs to be implemented and the complex setup they require. Also in this case, such tests can be carried out only in clinical settings.

2.2 Innovative methods

As highlighted in Section 1.6.2, the ANS is activated after the conscious perception of pain. The ANS, in turn, influences several physiological mechanisms whose activity can be monitored by recording Measurable Physiological Signals (MPS). The following are some of the MPS that resulted in being sensitive to pain.

2.2.1 Physiological signals

2.2.1.1 Photoplethysmography

Photoplethysmographic (PPG) signal represents the changes occurring in the blood volume at each heartbeat. It is a quasi-periodic signal, presenting the same elementary wave, called PPG pulse, in sync with the heart rate [83]. An example of PPG signal is given in Figure 2.3.

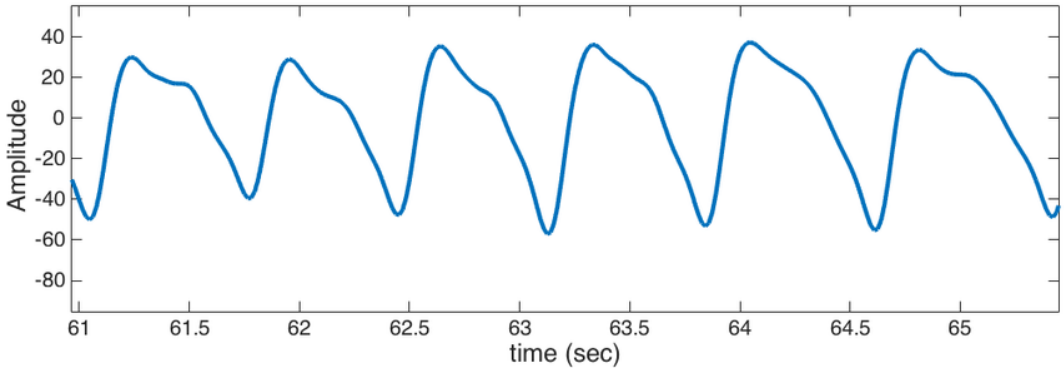


Figure 2.3 PPG signal

Its waveform is influenced by both SNS and PNS [84]. Information about the influence of ANS can be obtained either by estimating the Heart Rate (HR) and then by conducting a Heart Rate Variability (HRV) analysis, as a surrogate analysis usually conducted with the Electrocardiogram (ECG) [85] or by conducting a morphological analysis [86].

Several research studies demonstrated the possibility of extracting useful biomarkers for pain assessment by PPG, both by using HRV parameters [87], [88] and parameters derived from PPG morphology [89], [90].

2.2.1.2 Electrodermal Activity

Electrodermal Activity (EDA) measures the change in skin conductance occurring because of the activation of eccrine sweat glands. Sweat secretion alters the electrical properties of the skin, increasing electrical conductance until the sweat is absorbed or evaporates [91]. EDA signal consists of two components: the tonic, or slow-varying, component, referring to the smooth EDA changes that can also occur in the absence of any stimuli, and the phasic, or fast-varying component, which represents the rapid variations in EDA usually occurring as a response to a stimulus [92]. Figure 2.4 shows the EDA signal, together with its two components.

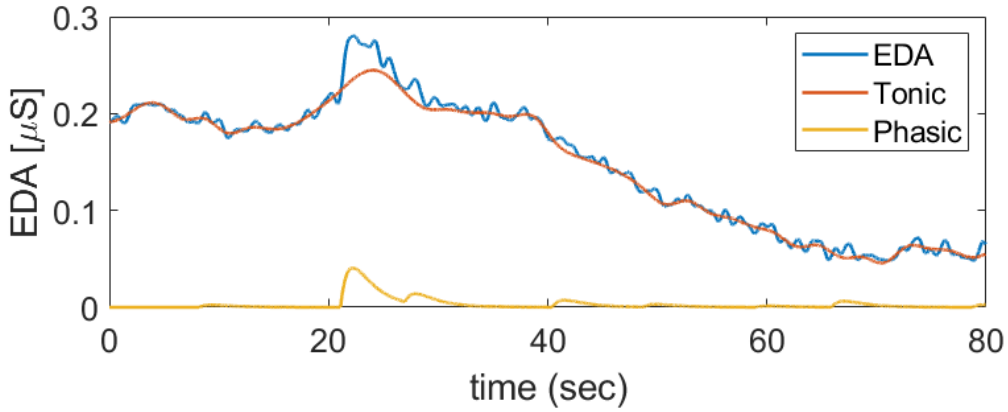


Figure 2.4 Example of EDA signal, together with its tonic and phasic components

Since sweat glands are exclusively innervated by the SNS [93], the increase in SNS activity associated with pain causes sweat to be discharged into pores on the skin surface. The association between EDA and pain has been studied for decades [94], with more recent studies proving its sensitivity to different pain conditions [95]–[98].

2.2.1.3 Electromyography

Electromyography (EMG) signals represent the electrical current generated by the contraction of skeletal muscles. The most used technique is the superficial EMG, recording muscle activity through electrodes placed on the skin, collecting the activity of several motor units [99]. Figure 2.5 shows an example of an EMG signal.

EMG signal was one of the first MPS to be evaluated for pain assessment. It is still used mainly to detect facial expressions related to pain, placing the electrodes on facial muscles well known to be activated in case of painful sensations, such as the corrugator muscles or the zygomaticus muscle [100], [101]. Surface EMG has also been widely used to

evaluate low back pain, recording the electrical activity of muscles around the spine [102], [103].

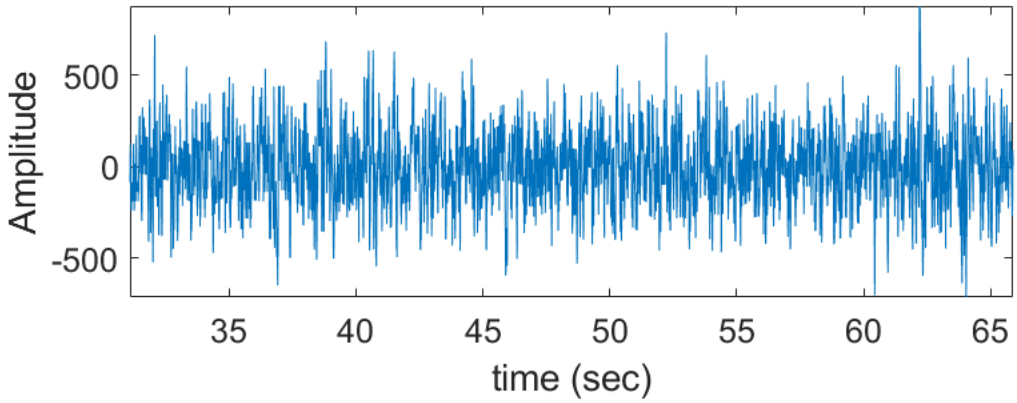


Figure 2.5 Example of EMG signal

2.2.1.4 Respiration

The respiratory signal is a quasi-periodic signal representing the fluctuations given by the inspiration and expiration phases. It is mainly exploited to estimate the breathing rate [104]. Figure 2.6 depicts a stream of a few seconds of a respiratory signal.

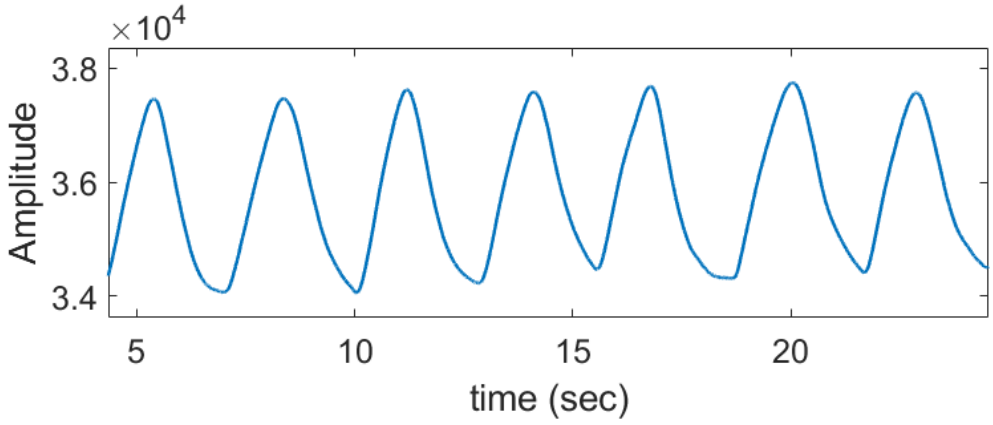


Figure 2.6 Example of a respiratory signal

The respiratory system is greatly affected by pain sensation, still through the ANS activity, resulting in increased respiration flow, frequency, and volume [105]. Respiration can be monitored either by using ad-hoc sensors placed on the chest or by extracting slow ECG or PPG modulations [106]. In both cases, it has already been assessed to be associated with pain intensity [107], [108].

2.2.1.5 Physical activity

Physical activity has been called a “vital sign” after its close link to health and wellness has been defined [109]. However, until a few years ago, assessment of physical activity

was aimed only at evaluation for clinical purposes, mainly conducted through questionnaires. Nowadays, with the consistent spread of accelerometer sensors, the objective evaluation of physical activity is within reach of all. An example of a tri-axial accelerometer signal is depicted in Figure 2.7.

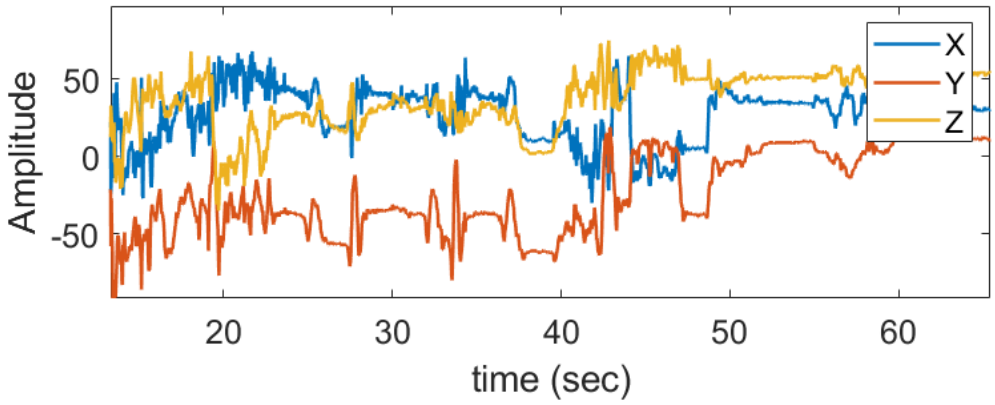


Figure 2.7 Triaxial accelerometer signal

Physical activity has been evaluated mainly as related to chronic pain since it is very likely that such a condition may impact how one can move [110]. The main applications of physical activity measures in relation to pain are related to musculoskeletal pain [111], [112], such as low back pain [113], [114].

2.2.2 Automatic pain assessment methods

Although the discoveries highlighted in the previous paragraphs concerning the sensitivity of certain physiological signals to the pain experience, the association between pain and physiological changes is not always so straightforward. Furthermore, the complexity of pain experience leads to a pattern of physiological modifications that can be difficult to fully grasp by using the standard statistical approach.

While standard statistical methods allow us to appreciate differences, for example, between two or more populations, or to build knowledge-driven models to estimate pain based on physiological signals, in the last years, Artificial Intelligence (AI) has been gaining momentum, allowing to build data-driven models, much more convenient than the statistical models when the underlying functioning principle is partially or completely unknown.

In particular, there has been a growing interest in the field of “emotion recognition”, the discipline encompassing all studies aiming at detecting emotions using technology, and automatic pain assessment through physiological signals falls within this category.

Such methods allow to develop three types of tools:

- Classifiers: given a number of classes, the AI methods allow the patient to be categorized in one of the classes if, for example, a particular event has occurred or represents the belonging to a particular pathological condition (e.g., a patient suffering from chronic pain)
- Regressors: a tool that establishes the functional relationship between the dependent variable (e.g., the intensity of the experienced pain) and one or more independent variables (e.g., parameters extracted from physiological signals)
- Clustering: it allows grouping homogeneous elements in the dataset provided based on the similarity between the elements.

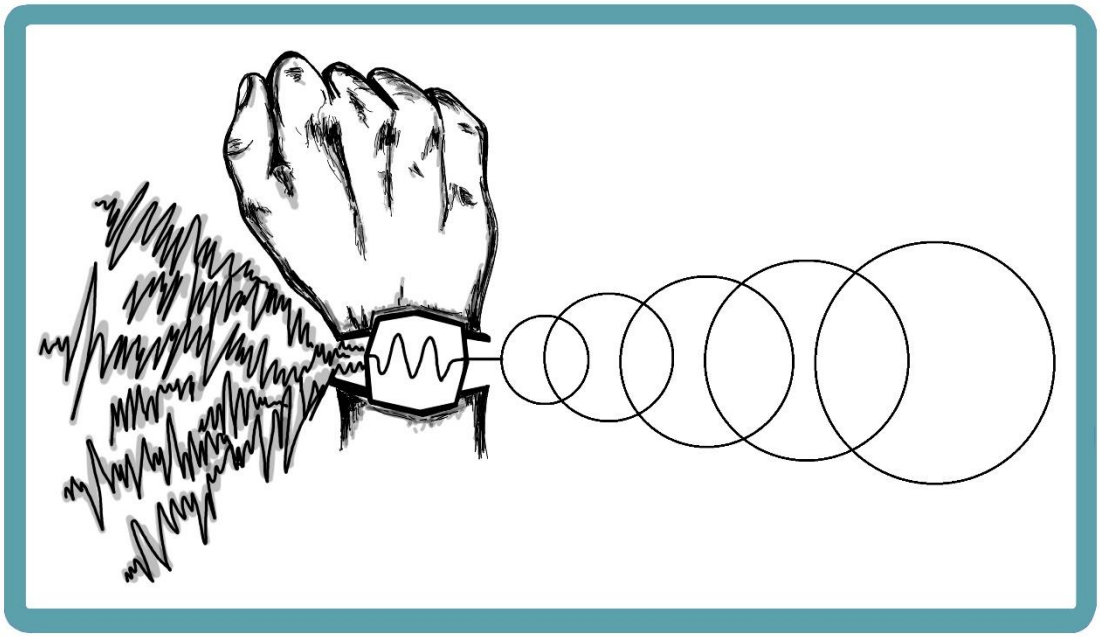
To date, there are several studies investigating the possibility of assessing pain by using physiological signals [115], [116], some of which by recording them through wearable devices [117], [118].

2.2.3 Advantages of automatic pain assessment methods

Compared to state-of-the-art pain assessment tools, methods to automatically assess pain by using physiological signals and AI algorithms can overcome several limitations:

- Such an approach can be potentially used also in those cases in which patients cannot verbally communicate their pain
- When even the monitored subject could verbally rate his or her pain, such methods would prevent the voluntary part from playing a role by underestimating or overestimating the experienced pain, giving a more objective evaluation
- Compared to the instrumented tests used in clinical practice, they would provide a more economical solution, both in terms of money and time, given the simple setup that can be used to record the physiological signals
- Some of the signals mentioned above have the extremely valuable aspect that they could be recorded using wearable sensors, thus allowing continuous and pervasive monitoring, having the possibility to record the physiological response exactly when and where the pain is experienced
- Measures obtained with such this approach could be used to observe and evaluate the efficacy of a treatment more reliably and systematically compared to the outcomes produced by scales and questionnaires

SECTION II: PREPROCESSING FOR PHYSIOLOGICAL SIGNALS RECORDINGS



3. Wrist Photoplethysmography Signal Quality Assessment for Reliable Heart Rate Estimate and Morphological Analysis

From the manuscript: Moscato S., Lo Giudice S., Massaro G., Chiari L.: "Wrist photoplethysmography signal quality assessment for reliable heart rate estimate and morphological analysis", Sensors, 2022

3.1 Abstract

Photoplethysmographic (PPG) signals are mainly employed for heart rate estimation but are also fascinating candidates in the search for cardiovascular biomarkers. However, their high susceptibility to motion artifacts can lower their morphological quality and, hence, affect the reliability of the extracted information. Low reliability is particularly relevant when signals are recorded in a real-world context, during daily life activities. We aim to develop two classifiers to identify PPG pulses suitable for heart rate estimation (Basic-quality classifier) and morphological analysis (High-quality classifier). We collected wrist PPG data over a 24h period from 31 participants. We defined four activity ranges based on accelerometer data and randomly selected an equal number of PPG pulses from each range to train and test the classifiers. Independent raters labelled the pulses into three quality levels. Nineteen features, including nine novel features, were extracted from

PPG pulses and accelerometer signals. We conducted ten-fold cross-validation on the training set (70%) to optimize hyperparameters of five machine learning algorithms and a neural network, and the remaining 30% was used to test the algorithms. Performances were evaluated using the full features and a reduced set, obtained downstream of feature selection methods. Best performances for both Basic- and High-quality classifiers were achieved using a Support Vector Machine (Acc: 0.96 and 0.97, respectively). Both classifiers outperformed comparable state-of-the-art classifiers. Implementing automatic signal quality assessment methods is essential to improve the reliability of PPG parameters and broaden their applicability in a real-world context.

3.2 Introduction

Wearable devices (WDs) are among the most widespread technologies introduced in recent years [119], potentially revolutionizing healthcare. With the aging population and the higher incidence of chronic diseases[120], [121], there is a growing need to provide healthcare services capable of reaching people who require frequent medical check-ups, especially those with low mobility and who live in remote areas. With their compact dimensions, high portability, and low manufacturing cost, WDs can efficiently perform long-term recordings outside healthcare facilities, allowing the remote, continuous monitoring of a user's health and, in turn, the early detection of anomalies [122], [123].
Commonly embedded in commercial smartwatches and fitness trackers worn at the wrist, one of the most used WD technologies is photoplethysmography (PPG), an optical technique that detects blood volume changes using a light source and a matched photodetector. The former illuminates a portion of the body surface, penetrating the skin and blood vessels. The latter detects the changes (using reflected or transmitted light, based on the PPG sensor design [124]) modulated by the pulsatile blood flow, which mainly depends on the heartbeat, vessel stiffness, and respiratory rate [125].
The PPG signal presents a quasi-periodic stereotyped waveform, commonly called PPG pulse, which occurs with each heartbeat [83]. Each PPG pulse can be divided into two phases: the anacrotic phase, which relates to the systolic heart contraction, and the catacrotic phase, which depends both on the diastolic heart phase and on the pulse wave reflected from the peripheral artery [126]. Within each PPG pulse, in ideal conditions, four fiducial points can be identified, as highlighted in Figure 3.1:

- Systolic foot: the beginning of the systolic phase and the minimum of the pulse;

- Systolic peak: the most prominent maximum;
- Dicrotic notch: most visible in healthy young subjects, it is supposed to represent the closure of the aortic valve [127];
- Diastolic peak: the second prominent maximum of the pulse.

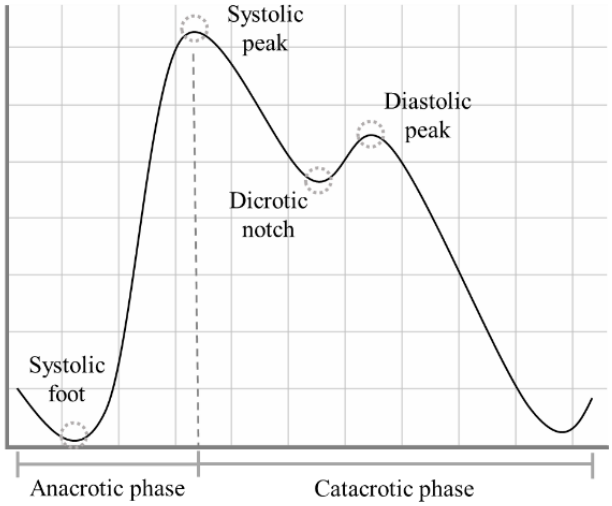


Figure 3.1 The shape of a typical PPG pulse

The PPG signal is strictly related to heart dynamics. Indeed, it is extensively used in commercial devices for heart rate (HR) estimation [121], [128] and subsequent HR variability (HRV) analysis [129], [130]. For example, HR can be estimated simply by detecting the systolic feet or peaks, calculating the time difference between two consecutive occurrences, and then calculating the ratio between 60 and the calculated time difference, expressing it in beats/min [131], [132].

Besides the HR estimation, it has long been recognized that the PPG signal carries valuable information in its morphology [133]. Recent research has corroborated this finding in emotion recognition [86], [115], [134] and cardiovascular measurements [135], [136].

In real-world applications, the preferred ground for PPG technology, obtaining reliable estimates both for HR and morphological features, is hampered by its high susceptibility to external noise and motion artifacts [137], [138]. Consequently, the information above cannot be used in clinical practice for diagnostic purposes. Before further processing, a signal quality analysis is essential to promote this signal's clinical use.

Based on the definitions provided by the recent literature [120], [139], the quality of a PPG pulse exploitable for further analysis can be expressed as:

- Basic-quality: systolic peaks are clearly identifiable;

- High-quality: the pulse waveform is clean and well-defined, with systolic and diastolic waves visible.

While HR and some morphological features related to detecting the systolic peak can be estimated from Basic-quality pulses, more sophisticated morphological features require the detection of both systolic and diastolic peaks [140]–[142], so only High-quality pulses are suitable.

Several researchers have already developed automatic methods for PPG signal quality assessment. Table I shows a selection of their works [139], [143]–[152].

Table 3.1 State of the art for the PPG signal quality algorithms

Ref	PPG sensor position	Settings	# subjects	Pulse-wise or segment-wise	Ground truth	Method	# quality levels
[143]	Finger and Wrist	Clinical	13 stroke patients + 500 patients retrospectively selected	30 s segments	Labels from 5 raters	Support Vector Machine with 42 features	2 + "not sure"
[145]	Wrist	Real-World	10 elderly subjects + 16 young subjects	10 s segments	Labels from 17 raters	Random forest with 9 features	5
[146]	Wrist	Real-World	50 healthy subjects	Pulse-wise	Labels from 1 rater	Signal similarity between adjacent pulses	3
[147]	Wrist	Real-World	17 epilepsy patients	7 s segments	Correspondence with RR from ECG	Support Vector Machine with PPG and accelerometer features	2
[148]	Finger	Clinical (public DB)	69 subjects from 3 public databases	Pulse-wise	Labels from 2 raters	Rules-based algorithm with 13 quality checks	2

[149]	Finger	Clinical (public DB)	44 patients from 2 public databases	Pulse-wise	Labels from 1 rater	Correlation with a template	2
[139]	Finger	Clinical	40 healthy subjects	60 s segments	Labels from 2 raters	Support Vector Machine with 1 feature	3
[150]	Finger	Clinical (public DB)	No info	10 s segments	Labels from 3 raters	Rules-based algorithm on HR estimate + correlation with a template	2
[151]	Finger	Clinical (public DB)	120 subjects	Pulse-wise	Labels from 1 rater	Non-linear scaling function based on adjacent pulses correlation	2
[152]	Finger	Clinical (public DB)	No info	6 s segments	No info	Deep learning algorithm with 4 features (based on the comparison with a template)	2
[144]	Finger	Lab	13 healthy subjects	60 s segments	Labels from 2 raters	Two-steps rules-based algorithm	2

Such studies significantly advanced the development of PPG signal quality algo-rithms, providing methods that can be used in real-time [148]–[150], trained on specific populations [143], [145]–[147], and validated by making use of publicly available datasets [148], [151], [152].

However, most previous studies only aim to detect PPG pulses for HR estimate, without rating their suitability for a more in-depth morphological analysis [144], [147]–[152]. Moreover, some base the quality estimation on a time window that includes several pulses [139], [143]–[145], [147], [150], [152] rather than pulse-wise, losing relevant information

that individual PPG pulses can convey as a result. Such a segment-wise analysis might also discard pulses suitable for analysis.

Although the publicly available datasets represent a considerable resource for training and testing automatic classifiers, they do not allow a proper quality characterization for real-world purposes. To the best of our knowledge, most of the currently available datasets are based on recordings of finger PPG signals in a clinical context, imposing several limitations. Since it is well-known that the morphology strongly depends on the measurement site [127], [153], the translation of a method based on signals recorded at the finger to signals recorded at the wrist (the preferred measurement site for real-world applications) is not feasible. Furthermore, the available datasets do not provide any ground truth information about the different quality of the signals (i.e., Basic and High), but only dichotomous labels (e.g., usable vs. non-usable). Finally, these datasets rely on hospital recordings, a context in which motion artifacts are far less frequent and less impactful than in the real world during daily life activities.

Recent works used PPG signals recorded by wrist-worn WDs in a real-world context and collected PPG pulses prone to lifelike motion artifacts [145]–[147] to overcome these limitations. Unfortunately, in these studies, no information is provided about the motion of the sensors, so it is unclear to what degree the related method is robust to daily-life motion artifacts.

This work aimed to develop two motion-aware classifiers:

- Basic-quality classifier: it detects all pulses with valid information content, exploitable for HR estimation and the extraction of basic morphological features;
- High-quality classifier: it detects all pulses with distinct systolic and diastolic waves, exploitable for the extraction of more in-depth morphological features.

We collected wrist PPG data for about 24 hours to design and test our classifiers in a real-world context. First, we defined different activity ranges to categorize the level of motor activity, which translates into motion artifacts in the PPG signals. Activity ranges were identified based on data from the accelerometer embedded in the same wrist-worn WD used to record the PPG signal. Then, for each range from each subject, we randomly selected PPG pulses to be classified. In this way, the classifiers could be trained using data subjected to different levels of motion artifacts, usually experienced in real-world contexts.

Such an approach could help in improving the reliability of the valuable biomarkers obtained by wrist PPG signals, minimizing the loss of information by conducting a pulse-wise analysis and selecting pulses suited for a specific analysis (i.e., HRV and fundamental morphological analysis or a more in-depth morphological analysis).

3.3 Materials and Methods

3.3.1 Wearable device

An Empatica E4 wristband was used to record the signals. Technical specifications are given at Appendix A.1.

3.3.2 Participants

A total of 31 recordings by as many participants were used. All the subjects were instructed to wear the Empatica E4 for 24 hours while carrying on with their normal daily activities. The participants were asked to provide their age and gender; other personal information was not collected.

3.3.3 PPG processing and pulse detection

The processing and pulse detection methods are reported in Appendix A.2.

3.3.4 Activity index and definition of activity ranges

In order to categorize pulses according to different amounts of movement, the activity index (A_{ind}) presented in [41] was calculated for each pulse. The complete processing pipeline is reported at Appendix A.3.

Once we estimated the $A_{ind}$for each recording, we defined four activity ranges (AR) based on the quartiles of all the A_{ind} values to label an equal number of pulses in each activity range.

3.3.5 Labelling procedure

Within each recording, we randomly selected a subset of 100 PPG pulses from each activity range, thus obtaining 400 pulses for each recording (12400 labelled pulses in total). Three independent raters (S.M., S.L., G.M.) then assigned a quality level to each pulse, selecting from one of the three levels defined below [154]:

- **Bad** (**B**): systolic and diastolic peaks cannot be easily distinguished from noise →The pulse is not suitable for further analysis.
- **Fair** (**F**): the systolic peak is clearly detectable; the diastolic peak is not → It is possible to estimate the HR and some basic morphological features
- **Excellent** (**E**): systolic and diastolic peaks are both clearly detectable → It is possi-ble to estimate the HR, basic morphological features, and perform in-depth morphological analysis

An example of the three quality levels is illustrated in Figure 3.2. A Matlab graphic user interface was developed to help the raters annotate the quality of the selected pulses, shown in Figure 3.3. The Matlab *findpeaks* function was applied to highlight the local maxima of the selected pulse and help detect the systolic and diastolic peaks.

Inter-rater agreement was assessed by calculating the overall Fleiss Kappa Score [155]. A majority voting approach was applied to determine the level if only two raters agreed. If there was no agreement among raters (i.e., each rater chose a different quality level), the pulse was automatically labelled as B.

3.3.6 Signal quality indices

We estimated nineteen signal quality indices (SQIs), listed in Table 3.2, corresponding to the selected and labelled pulses recorded in a real-world context. Specifically, we estimated:

- 2 SQIs from accelerometer data
- 17 SQIs from PPG pulses

Labelled PPG pulses were divided into training and test sets, with a proportion of 70% for the training set (22 subjects; 8800 pulses) and 30% for the test set (9 subjects; 3600 pulses).

SQIs from the training and test set pulses were then separately subjected to a Box-Cox transformation [156] and z-scored.

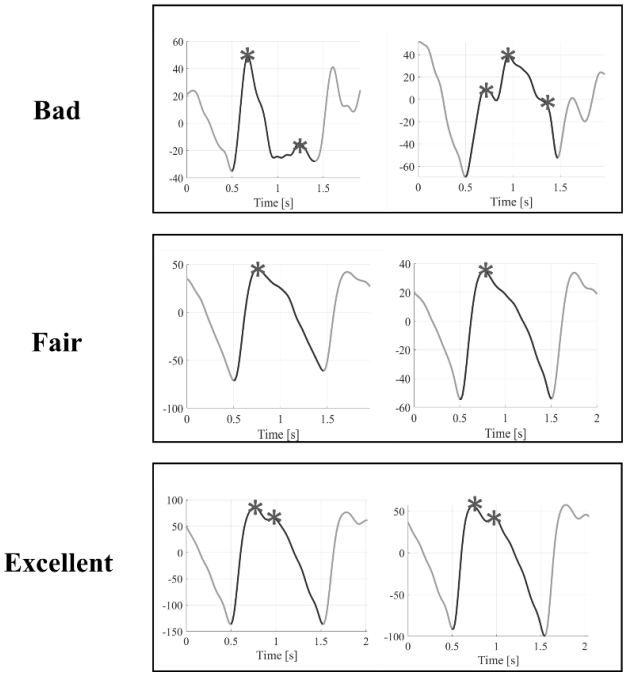


Figure 3.2 Examples of Bad, Fair, and Excellent quality pulses. Asterisks represent the local maxima for each pulse found by the Matlab findpeaks function

Figure 3.3 Matlab graphic user interface for PPG pulses annotation

3.3.7 SQIs selection

In order to limit the use of redundant SQIs, we applied a Neighborhood Component Analysis (NCA) separately for the two classifiers. NCA is a non-parametric method for selecting features to maximize a classifier's accuracy [157]. As output, NCA provides a weight for each feature: the higher the weight, the more influential the feature is for

solving the classification problem. We first tuned the NCA regularization parameter λ using ten-fold cross-validation on the training set to find the value that minimizes the classification loss. We then labelled those features with a weight greater than 20% of the maximum weight. In order to reach higher robustness of the selected features set, we ran the NCA ten times and then selected those features that were labelled at least 80% of the time.

Table 3.2 Signal quality indices (SQIs) for quality classification

SQI	Description	Source	Ref
Peak2peakACC	Peak to peak acceleration vector magnitude	ACC	This paper
MeanACC	Mean acceleration vector magnitude	ACC	This paper
SigSim	Correlation between consecutive PPG pulses	PPG	[146]
Entropy	Entropy	PPG	[139]
Kurtosis	Heavy tail and peakedness or a light tail and flatness distribution relative to the normal distribution	PPG	[139]
SNR	Signal-to-noise ratio	PPG	[139]
RelPower	Ratio of the power spectral density in the 1-2.25 Hz band compared to the overall power spectral density	PPG	[139]
Skewness	Measure of the symmetry of a probability distribution	PPG	[139]
ZR	Zero-crossing rate	PPG	[139]
Amplitude	Systolic peak amplitude	PPG	[144]
Width	Pulse width	PPG	[144]
TroughDepth	Systolic feet amplitude difference between consecutive systolic feet	PPG	[144]
MedianPulse	Median value of the z-scored PPG pulse	PPG	This paper
MedianPulse_noZ	Median value of the original PPG pulse	PPG	This paper
MeanPulse_noZ	Mean value of the original PPG pulse	PPG	This paper
StdPulse_noZ	Standard deviation of the original PPG pulse	PPG	This paper
SNR_Moody	Signal-to-noise ratio by Moody's algorithm	PPG	This paper
Npeaks	Number of detected local maxima	PPG	This paper
ZDR	First derivative zero-crossing rate	PPG	This paper

3.3.8 Basic and High quality classifiers

We designed the following classifiers:

- Basic-quality (BQ) classifier: it detects those pulses that can be used to esti-mate heart rate and for basic morphological analysis (i.e., the union of F and E pulses);
- High-quality (HQ) classifier: it detects those pulses that can be used for an in-depth morphological analysis (i.e., E pulses).

To develop the HQ classifier, we investigated two alternative strategies:

1. discern the union of B and F pulses against E pulses through a single-stage approach;
2. discern between F and E pulses downstream of a BQ classifier through a multi-stage approach.

A scheme illustrating the two strategies and the related classifiers is shown in Figure 3.4. In summary:

- the BQ classifier is trained to detect the F&E classes against the B class;
- the Type 1 HQ classifier (HQ1) is independent of BQ and is trained to de-tect the E class against the B&F class (Figure 3.4, panel A);
- the Type 2 HQ classifier (HQ2) is trained to detect the E class against the F class, having as an input the pulses selected by the BQ classifier (Figure 3.4, panel B).

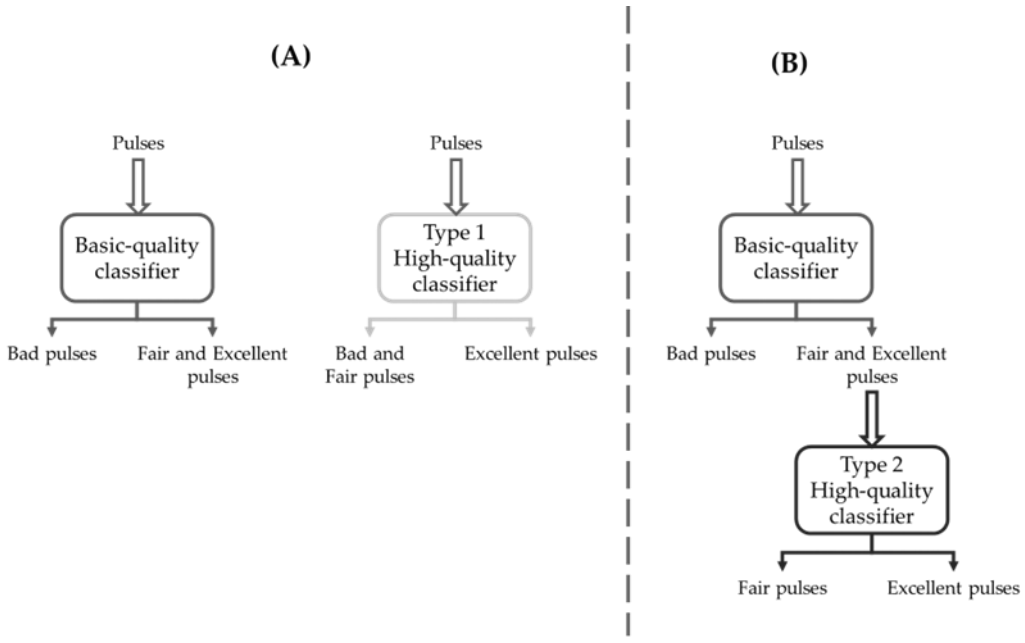


Figure 3.4 Schematic representation of the classification strategies. (A) Two independent classifiers: the Basic-quality classifier aiming at detecting Fair and Excellent pulses against Bad pulses, and the Type 1 High-quality classifier, aiming to detect Excellent

We first split the dataset into training (70%) and test (30%) sets both for BQ and HQ classifiers. We then conducted a ten-fold cross-validation on the training set with five machine learning (ML) algorithms (Tree, Naïve Bayes, Support Vector Machine, K-nearest neighborhood, Ensemble) and a neural network (NN) for hyperparameters optimization by using Bayesian optimization with 30 iterations. Finally, we trained and tested the classifiers with the full features set, and the SQIs selected features only.

We computed the following performance metrics on unseen data coming from the test set relative to the detection of eligible pulses (F&E pulses for the BQ classifier, E for HQ classifiers): area under the ROC curve (AUC), accuracy, sensitivity, specificity, precision, Matthew's correlation coefficient (MCC), F1 score, and Cohen's kappa (κ).

All the methods were implemented in Matlab 2021b [158]. The whole signal processing and classification pipeline is illustrated in Figure 3.5.

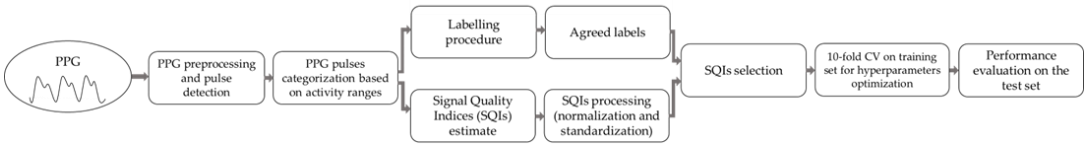


Figure 3.5 Signal processing and classification pipeline

3.3.9 State-of-the-art classifiers

We selected and adapted two classifiers from the literature to establish a benchmark for the performance of our classifiers.

i) Jang et al. [146] proposed two classifiers based on the signal similarity between adjacent PPG pulses, a parameter also used in our work (SigSim). Their study identified three quality levels (i.e., good, moderate, and low) based on detecting the PPG pulse second derivative's fiducial points [83]. Then, two dichotomous classifiers, conservative and non-conservative, were developed. The former compares the good-quality level pulses against the merge of moderate- and low-quality level pulses, while the latter compares the good- and moderate-quality level pulses against low-quality level pulses. Each classifier is based on a fixed threshold, determined using the equal training sensitivity and specificity criterion [159], meaning that the optimal threshold is obtained by minimizing the difference between sensitivity and specificity. Jang et al.'s non-conservative classifier is analogous to our BQ classifier, while their conservative classifier is analogous to both our HQ1 and HQ2 classifiers.

ii) The classifier proposed by Elgendi [139] is built on a Support Vector Machine that classifies 60-second PPG segments as belonging to one of three quality levels (i.e., excellent, acceptable, or unfit for diagnosis) based on the skewness property of the segment. We adapted this method to perform a pulse-wise analysis. Furthermore, since no information regarding the hyperparameters was reported, we applied the same approach described in Section 3.3.8 to find the best hyperparameters combination.

3.4 Results

3.4.1 Experimental data

We obtained real-world recordings of physiological signals from 31 subjects (15 males, 16 females), with a mean age of 37 years (±14) and an average recording length of 26:50 hours (±05:51). All subjects were Caucasian, except for one African subject.

3.2. Activity ranges

From the A_{ind} values estimated from the accelerometer signal, we obtained the following AR built on the quartile values of the A_{ind} distribution:

- AR_0: [0 – 0.0407]
- AR_1: (0.0407 – 0.4125]
- AR_2: (0.4125 – 1.3254]
- AR_3: (1.3254 to 6.7474]

According to the classification proposed by Lin et al. [160], the activity ranges 0–3 correspond to rest/sleep, rest/sleep/sedentary, light, and light/moderate activity, respectively. This means that the distribution of A_{ind} is skewed towards lower activity levels in our population.

3.4.2 Labelling results

A total of 12400 pulses were labelled by three independent raters, who agreed on 86% of the labels. Only 57 pulses (0.004%) were labelled differently by each rater and hence relegated to the B category. Overall, the inter-rater agreement was high, with a Fleiss Kappa Score of 0.84, representing perfect agreement according to Landis and Koch [161]. Using a majority voting approach, we set the final labels to train and test the classifiers: 5962 B pulses (48.08%), 4612 F pulses (37.19%), and 1826 E pulses (14.73%). The overall distribution of the three quality levels among the four activity ranges is shown in

Figure 3.6. As expected, as the A_{ind} (the amount of movement) increases, the percentage of B pulses gets higher, and the percentage of F and E pulses gets lower.

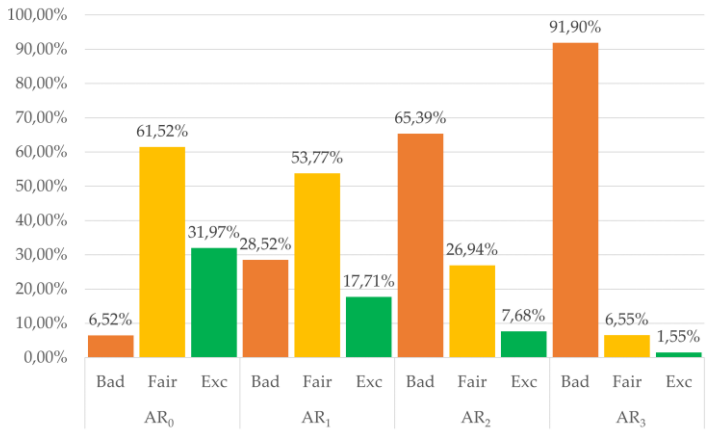


Figure 3.6 Distribution of the three quality classes among different activity ranges (AR). B = Bad, F = Fair, E = Excellent

3.4.3 SQIs selection

Said N the pulse length, the computational complexity to calculate the 19 features is approximately 37*N FLOPs. The computational complexity for each feature is reported in Appendix B, Table B.1.

We conducted SQIs selection separately for the BQ, HQ1, and HQ2 classifiers. In Table 3.3, the best λ values and their respective minimum classification loss values are reported for the three classifiers.

Table 3.3 Final best λ values for neighborhood component analysis and the related minimum classification loss

	BQ	**HQ1**	**HQ2**
Min classification loss	0.0498	0.0395	0.0575
Best λ	0.0017	0.0011	0.016

BQ = Basic Quality; HQ1 = High Quality 1; HQ2 = High Quality 2

The selection phase identified eight SQIs for the BQ classifier (*Peak2PeakACC*, *SigSim*, *TroughDepth*, *MedianPulse*, *StdPulse_noZ*, *SNR_Moody*, *Npeaks*, and *ZDR*), nine SQIs for the HQ1 classifier (*Peak2PeakACC*, *SigSim*, *Kurtosis*, *RelPower*, *Skewness*, *MedianPulse*, *StdPulse_noZ*, *Npeaks*, and *ZDR*), and nine SQIs for the HQ2 classifier (*Entropy*, *Kurtosis*, *RelPower*, *Skewness*, *MedianPulse*, *StdPulse_noZ*, *SNR Moody*, *Npeaks*, and *ZDR*). Results from each iteration of the NCA are reported in the Appendix B, Table B.2, Table B.3, and Table B.4 for the BQ, HQ1, and HQ2 classifiers, respectively.

3.4.4 Basic quality classifiers

A total of 5962 pulses belong to the B class (4260 used in the training set and 1702 in the test set), while 6438 pulses belong to the F&E class (4540 used in the training set and 1898 in the test set).

Table 3.4 presents the performances of the BQ classifiers on the test set. The best method using the full features set is the SVM with a Quadratic kernel, reaching an accuracy of 0.9606 and a well-balanced sensitivity (0.9603) and specificity (0.9547). On the other hand, the GentleBoost Ensemble reached the best performance among the methods trained and tested with the selected SQIs, with slightly lower values for accuracy (0.9536) and sensitivity (0.9384) but specificity (0.9706) higher than the best method using the full features set. Final hyperparameters are reported in Appendix B, Table B.5.

Concerning the state-of-the-art classifiers, the threshold based on the equal train-ing sensitivity and specificity criterion (identified in the work of Jang et al. [146]) is 0.922. Concerning the classifier proposed by Elgendi [139], the SVM with the Gaussian kernel function provided the best performance in terms of sensitivity (0.8398) and specificity (0.5764) with an accuracy of 0.7153. Our classifier outperformed both state-of-the-art classifiers for the selected performance measures. Results obtained with state-of-the-art classifiers are shown in the lower panel of Table 3.4.

3.4.5 High-quality classifiers

For the Type 1 High-quality classifiers, a total of 10574 pulses belong to the B&F class (7754 used in the training set and 1702 in the test set), while 1826 pulses belong to the E class (1046 used in the training set and 780 in the test set).

Table 3.5 presents the performances of the HQ1 classifiers on the test set. The best method for balancing sensitivity and specificity is the SVM, using all the features (Sens = 0.9244, Spec = 0.9784) or the subset of selected SQIs (Sens = 0.9192, Spec = 0.9702). In both cases, the SVM has a Quadratic kernel. Final hyperparameters are reported in Appendix B, Table B.6.

For the Type 2 High-quality classifiers, 4612 pulses belong to the F class (3494 used in the training set and 1118 used in the test set), while the distribution of pulses belonging to the E class is the same used to train and test the HQ1 classifiers

Table 3.6 presents the performances of the HQ2 classifiers on the test set. The kNN method using the subset of features selected by the NCA provided the best results

regarding sensitivity-specificity balance (Sens = 0.9321, Spec = 0.9195). Final hyperparameters are reported in Appendix B, Table B.7.

By comparing the best HQ1 and HQ2 classifiers, HQ1 achieved better performances in terms of accuracy and specificity (Acc = 0.9667, Spec = 0.9784) with respect to HQ2 (Acc = 0.9247, Spec = 0.9195), but slightly lower sensitivity (HQ1 Sens = 0.9244 vs. HQ2 Sens = 0.9321).

Concerning the state-of-the-art classifiers, the threshold identified for the HQ1 classifier with Jang's method [146] was 0.991. The linear SVM obtained the best performance in reproducing the classifier proposed by Elgendi [139]. However, both state-of-the-art classifiers performed worse than our classifier: the accuracy was 0.7090 for Jang's and 0.8406 for Elgendi's. Notably, the former reached moderate sensitivity (0.6301) and specificity (0.7245), while the latter showed a sensitivity closer to zero (0.0167).

The threshold for the HQ2 classifier with Jang's method [146] was 0.993. In reproducing the Elgendi's classifier [139], the quadratic SVM obtained the best performance. Also in this case, both state-of-the-art classifiers performed worse than our best HQ2 classifier, similar to what we observed for the HQ1 classifier.

3.5 Discussion

In this work, we developed automatic classifiers to detect PPG pulses suitable for further processing based on their peculiar morphological characteristics. First, using accelerometer data, we estimated the activity level of the subjects. We then detected four activity ranges based on the quartile values of aggregated A_{ind} from all the recordings. From each recording, we randomly selected 100 pulses for each activity range. Of the 19 SQIs estimated from each labelled pulse, eight and nine SQIs were selected to train and test the algorithms to develop the Basic- and the two High-quality classifiers, respectively. The best algorithms were then chosen, and the classifiers' performances were compared against two state-of-the-art classifiers.

Categorizing pulses by activity level allowed us to train the algorithms with pulses containing distinct amounts of motion artifacts. In this way, the ability of classifiers to detect PPG pulses suitable for heart rate estimate or for morphological analysis under various movement intensities could be achieved. However, it appears evident from Figure 5 that only a tiny portion of pulses in the highest activity range reached F or E quality levels, even if the highest activity range in our dataset corresponded to light/moderate

activity in the staging proposed by Lin et al. [160]. Several methods have been proposed to suppress the effect of motion artifacts on the PPG signals, either via software [162], [163] or hardware [164], [165] approaches. Our results suggest that future studies should combine algorithms for motion artifact suppression with a layer dedicated to signal quality analysis. This approach would be more conservative, allowing to obtain reliable parameters from a larger proportion of recorded pulses, even during intense physical activity.

The three independent raters reached a perfect agreement in the labelling procedure, probably thanks to the strict definitions given for each quality level. The high level of the inter-rater agreement also ensures the reliability of the resulting classifiers.

For each PPG pulse, we estimated 19 SQIs, calculated from two sources (i.e., PPG and ACC signals). Nine SQIs were novel and proposed for the first time in this study. The SQIs feature selection phase revealed that eight and nine SQIs were sufficient to solve the classification problem optimally for the BQ and for both types of HQ classifiers, respectively. It is worth noting that most of the selected SQIs are novel features. In particular, two of the newly introduced statistical parameters (*MedianPulse*, *StdPulse_noZ*) and two parameters related to the PPG pulse morphology (*Npeaks*, *ZDR*) were selected for all classifiers here presented, adding important information which helped better solve the classification problem.

Although the extraction of multiple features inevitably increases the computational complexity compared with the extraction of a single feature, the cost of the features presented in this work remains low and grows linearly with N. Moreover, it is interesting to note that the NCA selected features with increasing computational complexity for the BQ (5*N FLOPs), HQ1 (19*N FLOPs), and HQ2 (25*N FLOPs) classifiers, in line with the increasing complexity of the classification problem.

It is also worth noting that the *Peak2PeakACC* feature from the accelerometer data has been selected only for BQ and HQ1 classifiers, and not for the HQ2 classifier. This can be ascribed to the fact that B pulses (involved in both BQ and HQ1 classifiers) are generated because of motion artifacts, while the F and E pulses are largely independent of the movement.

All the implemented algorithms performed well to achieve BQ and HQ1 classifiers. Except for the Neural Network fed with the full features set, all the methods showed an accuracy higher than 0.90. However, the two classifiers differed in sensitivity and

specificity: BQ classifiers showed a balanced sensitivity and specificity, while the HQ classifiers had specificity higher than sensitivity (on average, 0.9728 compared to 0.9729). This difference can be ascribed to the imbalance in the number of pulses in the two classes (only 1826 pulses belonging to the E class compared to 10574 belonging to B&F classes), meaning that the algorithms are better trained in detecting pulses belonging to B&F class than to the E class.

Still regarding performance, some algorithms used to develop the HQ2 classifiers performed relatively poorly, except for the Ensemble and Tree algorithms. Again, the imbalance between F and E pulses (4612 F pulses against 1826 E pulses) may have played a role. However, as also pointed out by Elgendi [139], it was reasonable to expect that a classifier aiming at detecting E pulses against pulses belonging to a single quality level achieved worse performance than a classifier trained to detect E pulses against different quality pulses. In addition, it is necessary to consider the inevitable error propagation that a system of two cascaded classifiers entails.

Table 3.4 Performances for Basic-quality classifiers

Method	AUC		Acc		Sens		Spec		Prec		MCC		F1		κ	
	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs
Tree	0.9389	0.9413	0.9386	0.9406	0.9331	0.9283	0.9448	0.9542	0.9496	0.9576	0.8771	0.8814	0.9413	0.9428	0.877	0.881
NB	0.9242	0.9227	0.9219	0.92	0.883	0.8725	0.9653	0.973	0.966	0.973	0.8477	0.8455	0.9227	0.92	0.8442	0.8405
SVM	0.9606	0.9519	0.9603	0.9514	0.9547	0.9431	0.9665	0.9606	0.9695	0.9639	0.9205	0.9028	0.962	0.9534	0.9204	0.9026
KNN	0.9497	0.9455	0.9489	0.9453	0.9341	0.942	0.9653	0.9489	0.9678	0.9536	0.8983	0.8904	0.9507	0.9478	0.8977	0.8903
Ensemble	0.9546	0.9545	0.9539	0.9536	0.942	0.9384	0.9671	0.9706	0.9696	0.9727	0.9081	0.9078	0.9556	0.9552	0.9077	0.9071
NN	0.9513	0.9511	0.9508	0.9508	0.942	0.9457	0.9606	0.9565	0.9639	0.9604	0.9018	0.9016	0.9528	0.953	0.9015	0.9015
Jang et al. 2018	0.9265		0.9253		0.9025		0.9506		0.9532		0.8519		0.9272		0.8506	
Elgendi 2016	0.7081		0.7153		0.8398		0.5764		0.6886		0.4337		0.7567		0.4215	

Table 3.5 Performances for Type 1 High-quality classifiers (HQ1)

Method	AUC		Acc		Sens		Spec		Prec		MCC		F1		κ	
	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs
Tree	0,9144	0.9217	0.9494	0.9464	0.8526	0.8782	0.9762	0.9652	0.9085	0.8748	0.8484	0.8423	0.8796	0.8765	0.8477	0.8423
NB	0.8838	0.8838	0.9247	0.9283	0.8115	0.8051	0.956	0.9624	0.8362	0.8556	0.776	0.7848	0.8237	0.8296	0.7758	0.7843
SVM	0.9517	0.9447	0.9667	0.9592	0.9244	0.9192	0.9784	0.9702	0.922	0.8951	0.9019	0.881	0.9232	0.907	0.9019	0.8809
KNN	0.8996	0.9234	0.9386	0.9497	0.8308	0.8769	0.9684	0.9699	0.8792	0.8895	0.816	0.8512	0.8543	0.8832	0.8155	0.8511
Ensemble	0.9243	0.9107	0.9614	0.9539	0.859	0.8346	0.9897	0.9869	0.9585	0.9462	0.8839	0.8608	0.906	0.8869	0.8818	0.8581
NN	0.7556	0.9078	0.8881	0.9383	0.5218	0.8538	0.9894	0.9617	0.9314	5	0.6448	0.8178	0.6689	0.8571	0.6078	0.8178
Jang et al. 2018	0.7135		0.7292		0.6859		0.7411		0.4230		0.3685		0.5232		0.3486	
Elgendi 2016	0.5		0.7831		0		0.9906		0		0.0088		NaN		0.00005	

Table 3.6 Performances for Type 2 High-quality classifiers (HQ2)

Method	AUC		Acc		Sens		Spec		Prec		MCC		F1		κ	
	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs	All SQIs	Sel. SQIs
Tree	0.9223	0.8933	0.9278	0.9052	0.891	0.8269	0.9535	0.9597	0.9304	0.9348	0.8505	0.8046	0.9103	0.8776	0.8499	0.8006
NB	0.5	0.5	0.5885	0.5885	0	0	0.9991	0.9991	0	0	0.0192	0.0192	NaN	NaN	0.0011	0.0011
SVM	0.7313	0.9393	0.7713	0.9331	0.5064	0.9744	0.9562	0.9043	0.8896	0.8766	0.5376	0.8679	0.6454	0.9229	0.4948	0.8641
KNN	0.7889	0.9258	0.8145	0.9247	0.6449	0.9321	0.9329	0.9195	0.8702	0.8898	0.6177	0.8462	0.7408	0.9105	0.6013	0.8455
Ensemble	0.9358	0.9191	0.943	0.9273	0.8949	0.8731	0.9767	0.9651	0.9641	0.9458	0.8829	0.8499	0.9282	0.908	0.8812	0.8481
NN	0.5331	0.5258	0.6122	0.6096	0.0885	0.0551	0.9776	0.9964	0.734	0.9149	0.1499	0.1632	0.1579	0.104	0.0762	0.0601
Jang et al. 2018	0.5055		0.5042		0.4397		0.5492		0.405		0.0109		0.4216		0.0108	
Elgendi 2016	0.5		0.5885		0		0.9991		0		0.0192		NaN		0.9204	

There may be some B pulses wrongly classified within the F&E pulses by the first stage BQ classifier, so performances might be even worse than the ones reported in this study since the HQ2 classifier has been trained and tested only with real F and E pulses.

Our best classifiers outperformed the two state-of-the-art classifiers. Notably, the identified thresholds set for the Jang et al. [146] classifiers were higher than the values reported in the original work: 0.922 versus 0.673 for the BQ classifier, 0.991 (0.993) versus 0.796 for the HQ1 (HQ2) classifier. These discrepancies could be due to the higher quality levels of the F and E pulses identified in this work. However, the Jang et al. [30] BQ classifier attained good performance, with an accuracy of 0.9253, considering that a single SQI was used. On the other hand, the classifier proposed by Elgendi [24] demonstrated moderate performance for the BQ classifier (Sens = 0.8398, Spec = 0.5764) and poor performance for both HQ classifiers (Sens = 0.0167, Spec = 0.8406 for type 1; Sens = 0, Spec = 0.9991 for type 2).

The proposed classifiers can help extend the use of PPG signals recorded by wearable devices in the real world. On the one hand, the BQ classifier showed promising results, both in terms of sensitivity and specificity. Baek et al. [138] highlighted the detrimental effect on HRV analysis of missing inter-beat intervals. For this reason, a highly sensitive classifier is essential for detecting all pulses that can be used for HR estimation without losing discriminatory power by eliminating too many pulses because of their low quality. On the other hand, SVM selected as the best HQ classifier has high specificity with (relatively) low sensitivity. However, compared to other methods, it shows the best performance in terms of MCC, F1, and Cohen's κ. The importance of a HQ classifier is obvious, given the number of significant applications that have been proposed in the last few years. Features extracted from PPG morphology could be used, for example, for stress detection purposes [141], [166], [167] or blood pressure estimation [168]–[170], thus allowing continuous monitoring with a simple wrist-band. A large part of the population at risk of developing, e.g., burnout syndromes or cardiovascular disease would benefit from this achievement.

As a side result of this work, we built an annotated dataset that can be further exploited for future studies. As an ongoing activity, we are working on the preparation of the dataset to be publicly available.

This study has some limitations, most of which are related to the sample population used to train and test the algorithms. First, more robust classifiers could be obtained by

increasing the sample size: more subjects and labelled pulses would indeed be beneficial, preferably including subjects with arrhythmias or other cardiac pathologies. As this study was conceived, the classifiers we developed cannot discern arrhythmias from noise, thus potentially discarding arrhythmic beats that could also be useful for diagnostic purposes. Moreover, the algorithms' training phase could be re-fined considering subjects' age. As pointed out in [125], the dicrotic notch is more pronounced in healthy young than in older adults; and PPG morphology changes with age [140]. Therefore, a future study could collect and balance pulses belonging to different age groups both in the training and testing set. In addition, a further advancement of the method here proposed can be achieved by using recordings from different devices to train the signal quality algorithm. In fact, the method here proposed could be device-dependent, thus limiting its use on recordings conducted with different devices.

The classifiers developed in this study have not been tested in real-time. This is a crucial aspect to be assessed to understand whether the signal quality assessment can be smoothly embedded in the processing pipeline of wearable devices to provide reliable information with an acceptable delay [121]. Providing reliable health information in real-time would indeed facilitate the delivery of personalized treatments to the patient if and when needed [171].

3.6 Conclusions

This work aimed to develop two pulse-wise classifiers to detect reliable wrist PPG pulses that can be used in a real-world context for heart rate estimation and morphological analysis. We trained and tested several algorithms with a combination of features derived from different sources, including several novel features, and by selecting PPG pulses subjected to different levels of motion artifacts. The best performances were obtained by using subsets of features for both Basic- and High-quality classifiers. For both classifiers, the SVM with a Quadratic kernel achieved the best performance. Our results could help in improving the reliability and generalizability of the valuable biomarkers obtained by wrist PPG signals. Furthermore, the pulse-wise approach minimizes the loss of information by selecting all pulses suitable for either heart rate variability or morphological analysis. Future work can optimize the classifiers by increasing the sample size (both in terms of subjects and various cardiac health conditions) used to train the algorithms and explore the feasibility of embedding these methods in wearable devices for real-time application.

4. Quality Assessment and Morphological Analysis of Photoplethysmography in Daily Life

From the manuscript: Moscato S., Palmerini L., Palumbo P., Chiari L., "Quality assessment and morphological analysis of photoplethysmography in daily life", Frontiers in Digital Health, 2022, 4, 912353

4.1 Abstract

The photoplethysmographic (PPG) signal has been applied in various research fields, with promising results for its future clinical application. However, there are several sources of variability that, if not adequately controlled, can hamper its application in pervasive monitoring contexts. This study assessed and characterized the impact of several sources of variability, such as physical activity, age, sex, and health state on PPG signal quality and PPG waveform parameters (Rise Time, Pulse Amplitude, Pulse Time, Reflection Index, Delta T, and DiastolicAmplitude). We analyzed 31 24h recordings by as many participants (19 healthy subjects and 12 oncological patients) with a wristband wearable device, selecting a set of PPG pulses labeled with three different quality levels. We

implemented a Multinomial Logistic Regression (MLR) model to evaluate the impact of the aforementioned factors on PPG signal quality. We then extracted six parameters only on higher-quality PPG pulses and evaluated the influence of physical activity, age, sex, and health state on these parameters with Generalized Linear Mixed Effects Models (GLMM). We found that physical activity has a detrimental effect on PPG signal quality quality (94% of pulses with good quality when the subject is at rest vs 9% during intense activity), and that health state affects the percentage of available PPG pulses of the best quality (at rest, 44% for healthy subjects vs 13% for oncological patients). Most of the extracted parameters are influenced by physical activity and health state, while age significantly impacts two parameters related to arterial stiffness. These results can help expand the awareness that accurate, reliable information extracted from PPG signals can be reached by tackling and modeling different sources of inaccuracy.

4.2 Introduction

The digital healthcare revolution promises to switch from a hospital-centered model to a personal-centered model [172], offering the possibility to remotely and continuously monitor patients' health state, thus reducing the use of bulky instruments and complicated procedures [173]. One of the key elements of this revolution is represented by wearable devices, which are small electronic systems that can be worn during daily life [174]. However, such devices are not used as diagnostic tools yet for several reasons, including ethical aspects, limitations in the infrastructure, and concerns related to data protection [175]. Nonetheless, wearable sensors have been used in several applications for research purposes, ranging from rehabilitation [175] and sport [176] to cardiovascular monitoring [177] and emotion recognition [178].

In this context, PPG sensors are one of the most widespread technologies within wearable devices [179]. PPG signals consists of a repetition of a stereotyped waveform, as already pointed out in Chapter 3.1. An example of PPG signal with the highlighted fiducial points is depicted in Figure 4.1.

For various reasons, the fiducial points are not always traceable in the PPG pulses. Based on the fiducial points that can be detected, the quality of each PPG pulse can be expressed as [154]:

- Basic quality: systolic peaks are identifiable, so reliable heart rate, heart rate variability parameters, and some basic morphological features can be derived;

- High quality: systolic and diastolic peaks are visible, so a more in-depth morphological analysis can be conducted.

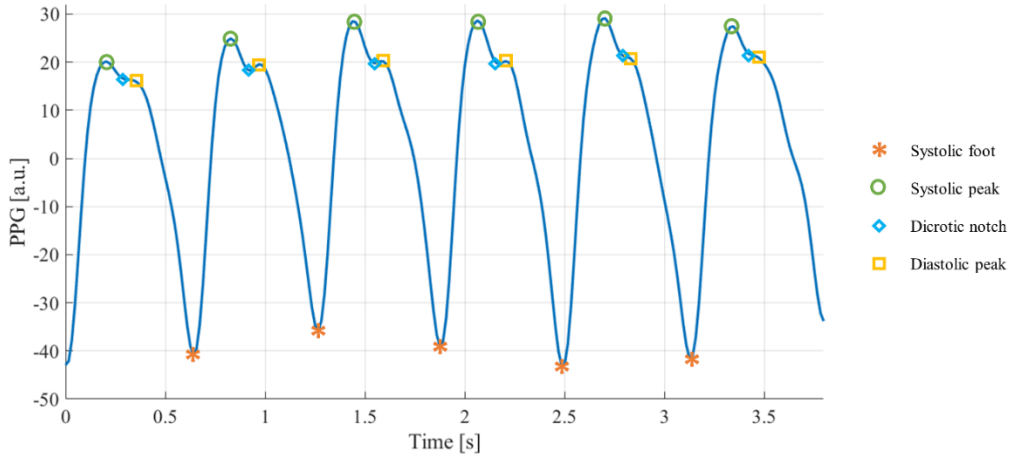


Figure 4.1 PPG fiducial points

Currently, PPG sensors are mainly used for heart rate estimation in a real-world context: the heart rate can be estimated by simply calculating the time distance between two consecutive systolic feet or peaks [83]. Still, the PPG waveform morphology also carries relevant information that can be exploited, e.g., for arterial stiffness [168], [180] and blood pressure [181] indirect estimation, or early detection of adverse cardiac events [182] or mental disorders signs [183].

Although the PPG signal has proven its potential as a helpful tool in different health domains, its clinical application is still hampered by its poor robustness to several sources of inaccuracy [184], which can be detrimental to the PPG signal quality or misleading for the interpretation of the extracted parameters [131]. This limitation is particularly emphasized in the real world, where the monitored subjects conduct their daily-life activities and are no longer in a controlled environment like in laboratory experiments.

The recent article of Fine et al. [184] offers a detailed review of the main factors that influence the PPG signal and its extracted features. If not adequately controlled, these factors may preclude the development of reliable PPG-based applications. Specifically, Fine et al. grouped the sources of inaccuracy in three categories: external perturbations, variations within and across individuals, and physiology. As an external perturbation, physical movement is the primary source of inaccuracy in the PPG signal; on the one hand, it is well recognized that physical movement leads to signal quality deterioration

[131]; on the other hand, it also influences the cardiovascular system, and in turn, the PPG morphology, inducing temporary changes, as the cardiovascular system must adapt to the physical stress [167]. Another external source of inaccuracy is given by the contact pressure between the PPG sensor and the skin, significantly influencing the quality and the morphology of the recorded signal [185], [186]. Individual subject variations can also play a role in modifying PPG signal quality and morphology. For example, it is well known that the dicrotic notch is less visible as age increases [125], making systolic and diastolic waves less pronounced. This factor can lower the probability of finding High quality pulses in older subjects, limiting the possibility of conducting an in-depth morphological analysis. Also, it is well known that sex can affect the cardiovascular system, thus, in turn, the PPG morphology [187]. Finally, the health state can also have an impact on PPG, even in those cases in which the pathology is not closely related to the cardiovascular system. For example, some recent studies demonstrated the link between cancer and cardiovascular alterations, which can origin both from the pathology itself or from cancer treatment [188]. Several studies have already investigated the association between cancer and heart rate variability, pointing out significant parameters' alterations in the oncological population [189], [190], also by using PPG [191]. In addition, some studies also revealed the impact of cancer [192] and related therapies [193] on PPG signal waveform. From this evidence, it is clear that a PPG-based system that is agnostic to the health state of the subject may lead to misinterpretation of the extracted parameters, failing its primary goal of providing continuous accurate monitoring [184].

Whatever the final application, all these factors, if not adequately controlled, can have a dual negative effect: from one point of view, they can have a different impact on PPG signal quality, hindering the extraction of meaningful PPG features (e.g., a small amount of High pulses prevents a reliable, in-depth morphological analysis), and, from another point of view, they can act as confounding factors, invalidating the interpretation and the reliability of the parameters extracted from the PPG morphology. Therefore, to obtain a "true health monitoring" [184] PPG-based application, a proper characterization of these factors is crucial.

This work aims to characterize the impact of these factors, namely physical activity, health state, age, and sex, both on PPG signal quality and PPG waveform parameters. We used a convenience sample of 31 participants, 19 healthy subjects and 12 oncological patients, monitored in a real-world scenario. For each subject, we selected an equal

number of PPG pulses for four different physical activity ranges (estimated by the accelerometer data) and labelled them with a quality level. Firstly, we evaluated the quality levels distribution based on the factors above. We then extracted six morphological parameters and appraised their behavior in relation to physical activity, health state, age, and sex.

4.3 Materials and Methods

4.3.1 Dataset

Thirty-one subjects (19 healthy subjects and 12 oncological patients) were monitored for 24 hours with the Empatica E4 wristband, whose technical specifications are provided in Appendix A.1.

Subjects were instructed to conduct their daily routine activities and remove the E4 wristband while showering. They were also asked to provide their age and sex.

The study was conducted according to the declaration of Helsinki, and each subject signed informed consent before participating in the study. The two datasets (healthy subjects and oncological patients) belong to two different studies: 1) healthy subjects' recordings were obtained from an internal data collection campaign involving researchers and students at the Personal Health Systems Lab of the University of Bologna; 2) oncological patients' recordings come from an interventional study approved by the Local Ethical Committee (Area Vasta Emilia Centro, Bologna, Italy; approval n° 542-2019-OSS-AUSLBO) (31).

4.3.2 Signal processing

PPG signals were processed by following the procedure reported in Appendix A.2, obtaining the segmentation in PPG pulses.

ACC signals' components were preprocessed by following the procedure reported in Appendix A.3, obtaining the A_{ind}.

Each PPG pulse was then associated with an A_{ind} value as a final signal processing step.

4.3.3 PPG pulse classification

We randomly selected a subset of 100 PPG pulses for each AR within each subject's recording, thus obtaining 400 pulses for each subject (12400 PPG pulses in total). We chose 400 pulses per subject as a good trade-off between the need to have a representative sample of all pulses and the clinical effort needed to evaluate and label them. It is also in

line with previous studies [140], [153]. Three independent raters (a cardiologist and two biomedical engineers, all three experts in cardiovascular signals) assigned to each pulse one of three quality levels: Bad (B), Fair (F), or Excellent (E), as reported in section 3.3.5 An example of the three different quality levels is presented in Figure 4.2.

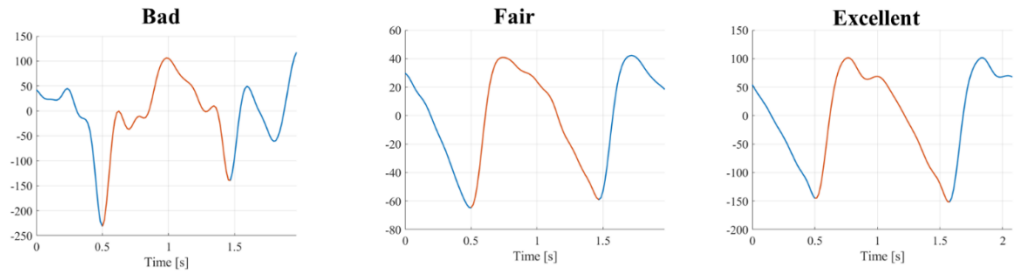


Figure 4.2 PPG pulses with three different quality levels (from left to right): bad, fair, and excellent

We adopted a majority voting approach to determine the level if only two raters agreed. If there was no agreement among raters, the pulse was automatically labelled as B.

Based on these quality levels, Basic PPG pulses were obtained as the union between F and E pulses, while the High quality PPG pulses coincide with the only E pulses [150].

4.3.4 PPG waveform parameters estimation

We estimated six PPG parameters only on those PPG pulses suitable for analysis (i.e., Basic and High quality pulses), thus discarding the B quality pulses. For both Basic and High quality pulses, the systolic peak is the highest value found using the Matlab *findpeak* function within each PPG pulse. For High quality pulses, the same Matlab *findpeak* function is applied, and the diastolic peak is found as the second-highest value.

From Basic quality pulses, we estimated:

- Rise Time (RT) [s]: time between the systolic foot and the subsequent systolic peak [194];
- Pulse Time (PT) [s]: time between two consecutive systolic feet [83], [140]
- Pulse Amplitude (PA) [a.u.]: height of the systolic peak, with the previous systolic foot as the reference [153].

From High quality pulses, we estimated:

- Reflection index (RI) [%]: ratio between diastolic and systolic amplitude [83], [140];

- Delta T (ΔT) [s]: time lag between systolic peak and the subsequent diastolic peak [195];
- Diastolic Amplitude (DA) [a.u.]: height of the diastolic peak, with the previous systolic foot as the reference [83].

A graphical representation of the PPG above parameters is proposed in Figure 4.3.

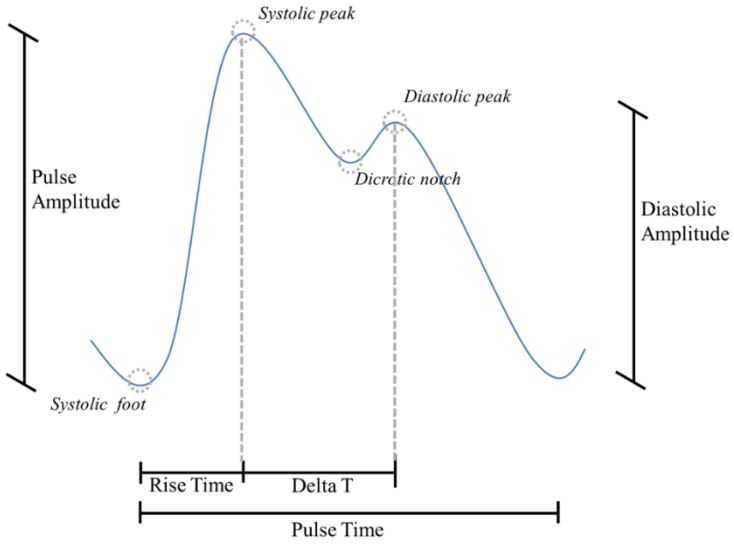


Figure 4.3 PPG morphology parameters

4.3.5 Statistical analysis

To qualitatively assess the impact of physical activity and health state on PPG signal quality, we evaluated the distribution of quality levels among the four AR and throughout the twenty-four hours separately for healthy and oncological subjects. To statistically assess the impact of physical activity, health state, age, and sex on PPG signal quality, we implemented a multinomial logistic regression (MLR) model. MLR is used to predict the relative probability of being in one category compared to being in a reference category, obtained with a linear combination of predictor variables that can be continuous or categorical. The logit function is usually employed as a link function for MLR models. Setting the K-th category as a reference, the MLR can be written as [196]:

$$\pi_j = Pr(y = j|\boldsymbol{x}) = \frac{exp(\boldsymbol{\beta_j}^T \boldsymbol{x})}{1 + \sum_{k=1}^{K-1} exp(\boldsymbol{\beta_k}^T \boldsymbol{x})} \quad (4.1)$$

where π_j is the j-th category membership probability against the reference category K, $\boldsymbol{\beta_j}$ is the regression coefficients vector, and $\boldsymbol{x}$ is the regressors vector. We set A_{ind} (continuous variable), health state (dichotomous variable, 0 = healthy subject, 1 =

oncological subject), age (continuous variable), and sex (dichotomous variable, 0 = male, 1 = female) as regressors vector.

To evaluate the influence of physical activity, health state, age, and sex on PPG waveform parameters, we fitted each PPG parameter with a Generalized Linear Mixed-Effects Model (GLMM). GLMMs extend the generalized linear models, allowing to model both fixed and random effects. A simple Linear Mixed-Effects model can be written as [197]:

$$E(y|\boldsymbol{X},\boldsymbol{Z}) = \boldsymbol{X\beta} + \boldsymbol{Zu} \quad (4.2)$$

where X is the matrix of the fixed effects, β is the vector of fixed effects regression coefficients, Z is the matrix of the random effects, u is the vector of random effects coefficients, and E(y|X,Z) is the expected outcome variable conditional on X and Z. In a "Generalized" Linear Mixed-Effects Model, the outcome variable can have a non-normal distribution so that a GLMM can be expressed as:

$$g(E(y|\boldsymbol{X},\boldsymbol{Z})) = \boldsymbol{X\beta} + \boldsymbol{Zu} \quad (4.3)$$

where g(·) is the link function for the outcome variable. The link function maps the relationship between the conditional, expected outcome and the linear combination of the predictors. There is an associated canonical link function for each distribution of the outcome variable.

GLMMs are particularly useful when data samples are non-independent, such as, e.g., in a hierarchical structure (i.e., different instances coming from a single subject) [198]. We fitted one GLMMs for each of the six PPG parameters, using the Basic pulses to determine RT, PT, and PA, and the High quality pulses to determine RI, ΔT, and DA. We set the PPG parameter as the outcome variable, the four factors as the fixed effects, while the "subject" variable was set as the random effect to consider the inter-subject variability. We tested three different distributions for the GLMMs (and the respective link functions): normal (identity), gamma (negative inverse), and inverse Gaussian (inverse squared), the last two suitable to model non-negative outcome variables. Table 4.1 presents the three distributions and the respective link functions. We then chose the best model based on the Akaike Information Criterion (AIC) [199] and evaluated the results, both for fixed and random effects [200]. We performed a marginal F-test to determine the significance of single fixed-effects coefficients. To test the significance of the random effects, we evaluated the 95% standard deviation's confidence interval as the estimated covariance parameter for the random effects (i.e., "subject"): if the interval does not contain 0, the

random effects are significant at the 5% significance level. The analyses were carried out on Matlab 2021b [158].

Table 4.1 Distribution of the outcome variable and respective link function

Distribution	Link function
Normal	$g(\mu) = \mu$
Gamma	$g(\mu) = -\mu^{-1}$
Inverse Gaussian	$g(\mu) = \mu^{-2}$

4.4 Results

4.4.1 Descriptive statistics

We analysed PPG recordings from 31 subjects, 19 healthy subjects and 12 oncological patients (1 bone and soft tissue, 4 gastrointestinal, 2 genital tract, 1 endocrine, 1 haematological, 2 breast, 1 urinary). The demographics of the sample are reported in Table 4.2. The average recording length was 26:50 (± 05:51) hours.

Table 4.2 Dataset descriptive statistics

	All	Healthy	Oncological
Sample Size	31	19	12
Age	37 ± 13.8	29.2 ± 7.1	49.5 ± 12.8
Gender	16 M, 15 F	10 M, 9 F	2 M, 10 F

Quartile values of the A_{ind} distributions were Q1 = 0.04, Q2 = 0.41, Q3 = 1.32, with a maximum value of 6.75. According to the classification made by Lin et al. (37), the four ARs correspond respectively to rest/sleep, rest/sleep/sedentary, light activity, and moderate activity.

For the 12400 randomly selected pulses, the three independent raters agreed on 86% of the labels. By applying a majority voting approach, we obtained the following labels distributions: 5962 (48.1%) B pulses, 4612 (37.2%) F pulses, and 1826 (14.7%) E pulses. Table 4.3 reports the distribution of quality levels for each subject..

4.4.2 Impact on PPG pulses quality

We evaluated the distribution of the three quality levels among the four ARs. As can be seen in Figure 4.5, panel a, the percentage of B pulses rises as the physical activity increases (ranging from 7% in AR_0 to 92% in AR_3), while the percentage of F and E

pulses decreases (from 62% in AR_0 to 7% in AR_3 for F pulses; from 32% in AR_0 to 2% in AR_3 for E pulses).

Table 4.3 Distribution of quality levels among healthy subjects and oncological

	Subject	B	F	E
Healthy Subjects	1	251	81	68
	2	157	203	40
	3	133	255	12
	4	227	45	128
	5	214	57	129
	6	239	85	76
	7	136	164	100
	8	144	192	64
	9	207	155	38
	10	208	84	108
	11	276	60	64
	12	170	128	102
	13	139	197	64
	14	231	113	56
	15	103	185	112
	16	124	239	37
	17	217	106	77
	18	316	31	53
	19	126	59	215
Oncological patients	1	203	197	0
	2	147	242	11
	3	160	239	1
	4	229	168	3
	5	222	171	7
	6	206	147	47
	7	205	123	72
	8	194	183	23
	9	127	248	25
	10	217	173	10
	11	189	162	49
	12	245	120	35

B = Bad; F = Fair; E = Excellent

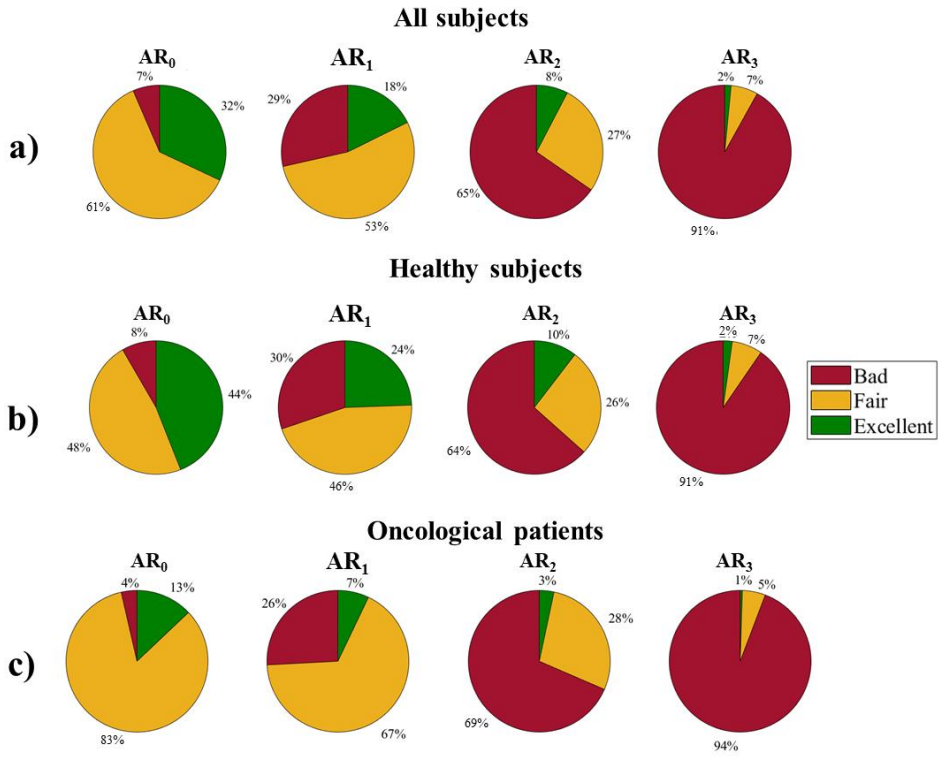


Figure 4.4 Distribution of the three quality levels among different activity ranges. a) All subjects b) Healthy subjects c) Oncological patients.

By analyzing separately healthy and oncological subjects, the different distribution of the three quality levels appears evident (see Figure 4.5, panels b and c): oncological patients present a lower percentage of E pulses in all the ARs and a higher percentage of F pulses in the lowest ARs (84% and 67% for oncological patients against 48% and 45% for healthy subjects in AR_0 and AR_1, respectively).

A graphical representation of the quality levels throughout the twenty-four hours is provided in Figure 4.6, together with the A_{ind} values. The figure shows the percentage of the different quality levels during each hour. A higher percentage of F and E pulses can be found during the night when the A_{ind} values are lower both evaluating the whole dataset (panel a) and dividing it into healthy (panel b) and oncological subjects (panel c). During the night, oncological patients present a lower number of B pulses (around 10%) compared to healthy subjects (around 20%).

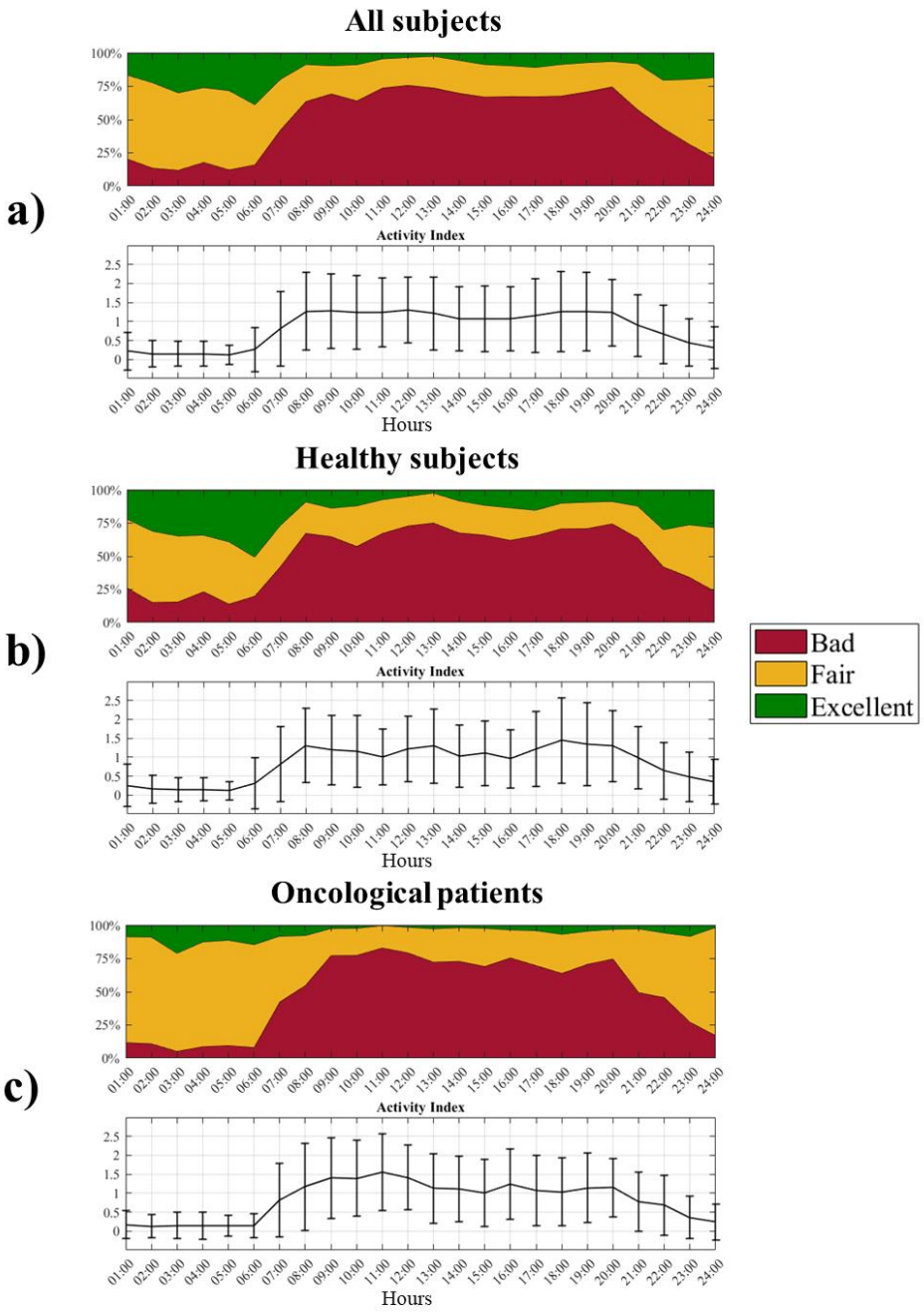


Figure 4. 5 Distribution of the three quality levels and related activity index profile over the 24 hours. a) All subjects. b) Healthy subjects. c) Oncological patients.

From the MLR model, we obtained the results reported in Table 4.4, setting the B quality level as the reference category. We present here the β coefficients for each regressor (i.e., A_{ind}, health state, age, and sex), whose interpretation is the following: positive β

coefficients represent a direct association between the regressor and the probability of belonging to that category compared to the reference one; higher values mean a stronger

Table 4.4 Multinomial Logistic Regression coefficients

	Fair*		Excellent*	
	Estimate	**p-value**	**Estimate**	**p-value**
Intercept	-0.4253	0.0001	0.7872	0
AI	-2.3154	0	-3.0329	0
HealthState	0.1555	0.042	-0.8738	0
Age	0.0327	0	-0.0088	0.0233
Gender	0.8385	0	0.4781	0

* against Bad quality level (set as reference category)

relationship between the regressor and the probability of belonging to that category compared to the reference one. A_{ind} has a significant impact on the relative probabilities (with respect to the B quality level) for both F and E quality levels: as A_{ind} increases, the relative probability of belonging to F and E quality levels decreases. Health state significantly influences the relative probability of having F and E quality pulses: oncological patients have an increased relative probability of having F pulses, while there is a lower relative probability for the same population of having E pulses. Finally, age significantly influences the relative probabilities of F and E quality levels: the coefficient has a positive value (0.03) for F quality level and a negative value (-0.0088) for E quality level. This means a higher relative probability of having F pulses and a lower relative probability of having E pulses as the age increases. Regarding sex, female subjects have an increased relative probability of having F (0.84) and E (0.48) pulses compared to males.

4.4.3 PPG waveform parameters

After grouping pulses into Basic (F+E) and High quality (E) pulses, we obtained the following proportions:

- 6438 Basic quality pulses, 3944 from healthy subjects (61.3%), and 2494 from oncological subjects (38.7%)
- 1826 High quality pulses, 1540 from healthy subjects (84.3%), and 286 from oncological subjects (15.7%)

We fitted six different GLMMs, one for each PPG parameter, using the Basic pulses to determine RT, PT, and PA, and the High quality pulses to determine RI, ΔT, and DA.

Table 4.5 shows the AIC values for all the six models using normal, gamma, and inverse Gaussian distributions. Four out of six models were best fitted with a normal distribution (RT, PT, RI, and ΔT), while two (PA and DA) were best fitted with an inverse gaussian distribution. The interpretation of β coefficients for the two distributions (and the related link functions) is the following: for the normal distribution (and identity link function), positive coefficients indicate that the outcome variable increases if the predictor increases; for the inverse Gaussian distribution (and inverse squared link function), positive coefficients indicate that the outcome variable increases if the predictor decreases.

Table 4.5 Akaike Information Criterion (AIC) for different models

AIC	Normal	Gamma	Inverse Gaussian
RiseTime	-18186	15611	41620
PulseTime	-10034	-7991.6	2326.1
PulseAmpl	-1337.1	-30577	-36263
RI	-2863.2	-729.06	2926.7
deltaT	-6182.6	4067.1	11774
DiaAmpl	1490.7	-4818.8	-5389.5

In Table 4.6 results from the GLMMs are shown. All PPG parameters, except RT, are significantly influenced by physical activity (A_{ind}). Specifically, all the parameters coefficients tend to have lower values as the A_{ind} increases. The health state significantly influences PT, PA, RI, and DA: these parameters assume lower coefficient values for oncological patients than healthy subjects. Age significantly influences ΔT and DA: ΔT is shorter as age increases, while DA increases with age progression. Sex does not have any significant effect on the analyzed parameters. Since the 95% random-effects confidence intervals for all the PPG parameters do not contain the 0 value, inter-subject variability is significant for all the tested PPG parameters. This means that part of the dataset variability is introduced on a subject-basis (i.e., the values of each parameter are not constant across different subjects), so the values of the PPG parameters are best modelled with a subject-specific intercept. Figure 4.7 shows the graphical representation of both Basic and High quality PPG pulses, as the mean of the analyzed pulses, for different ARs and health states.

Table 4.6 Generalized Linear Mixed Effects Models

	Fixed Effects					Random Effects σ(subject)		
		Estimate	Lower	Upper	F-Test	Estimate	Lower	Upper
RT§	A_{ind}	-0.00	-0.0055	0.001	0.21	0.03	0.03	0.04
	HealthState (Oncological)	-0.01	-0.03	0.005	0.14			
	Age	0.0005	-0.0008	0.002	0.47			
	Sex (Female)	-0.004	-0.02	0.01	0.61			
	Intercept [s]	0.24	0.19	0.30	0			
PT§	A_{ind}	-0.12	-0.13	-0.12	0	0.12	0.09	0.15
	HealthState (Oncological)	-0.12	-0.18	-0.05	0.0002			
	Age	-0.0005	-0.005	0.004	0.85			
	Sex (Female)	-0.03	-0.08	0.02	0.25			
	Intercept [s]	0.94	0.76	1.11	0			
PA*	A_{ind}	0.005	0.004	0.006	0	0.005	0.004	0.006
	HealthState (Oncological)	0.003	0.0005	0.006	0.02			
	Age	0.0001	-0.0001	0.0003	0.39			
	Sex (Female)	0.0004	-0.002	0.002	0.71			
	Intercept [a.u.]	0.10	0.09	0.11	0			
RI§	A_{ind}	-0.02	-0.04	-0.01	0.0003	0.09	0.07	0.01
	HealthState (Oncological)	-0.06	-0.11	-0.02	0.008			
	Age	0.003	-0.0007	0.006	0.12			
	Sex (Female)	0.02	-0.01	0.06	0.25			
	Intercept [ ]	0.67	0.53	0.80	0			
ΔT§	A_{ind}	-0.01	-0.02	-0.01	0	0.03	0.03	0.04
	HealthState (Oncological)	0.005	-0.01	0.02	0.58			
	Age	-0.001	-0.003	-0.0002	0.03			
	Sex (Female)	-0.001	-0.02	0.01	0.84			
	Intercept [s]	0.30	0.24	0.35	0			
DA*	A_{ind}	0.01	0.01	0.02	0.0001	0.04	0.03	0.06
	HealthState (Oncological)	0.04	0.02	0.05	0.0004			
	Age	-0.002	-0.004	-0.0004	0.016			
	Sex (Female)	-0.01	-0.03	0.004	0.11			
	Intercept [a.u.]	0.26	0.19	0.33	0			

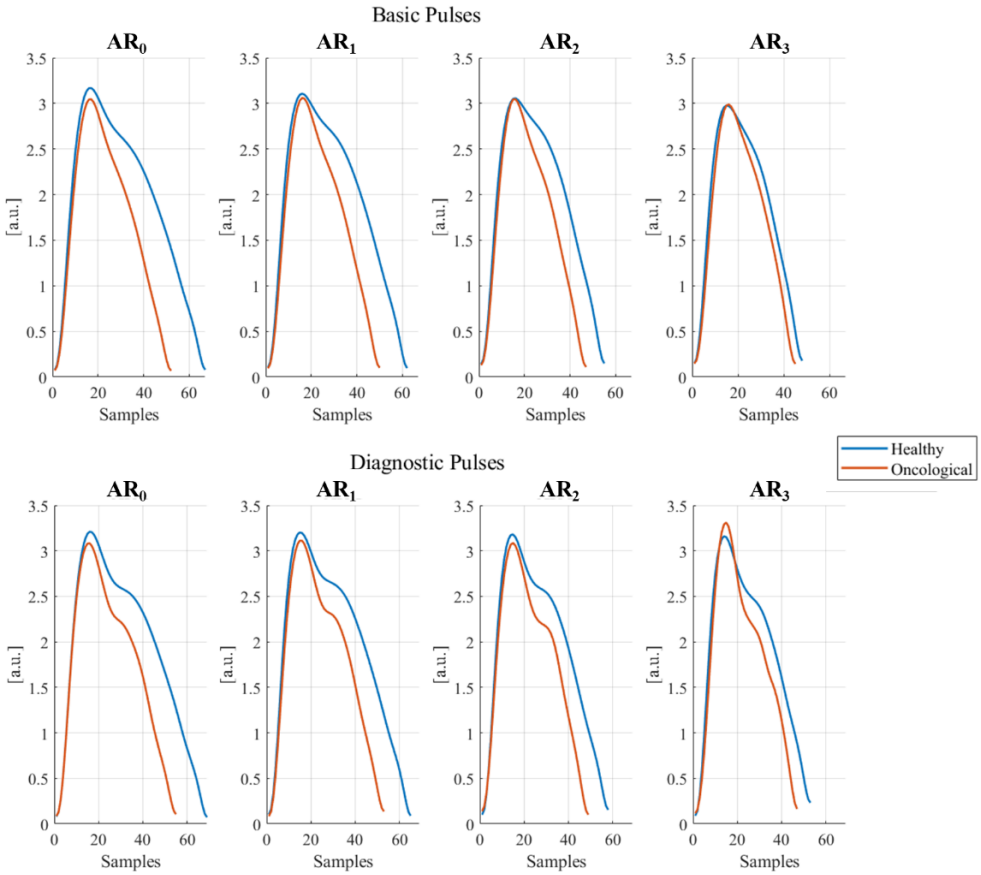


Figure 4.6 Basic and High quality pulses in different activity ranges (ARi, i=0,…,3), in healthy and oncological subjects. The represented pulses were obtained as the mean of all collected pulses for each AR and dividing them for healthy and oncological subjects.

4.5 Discussion

This study assessed the impact of several sources of inaccuracy on PPG signal quality and PPG waveform parameters by using 31 24h real-world recordings, 19 from healthy subjects and 12 from oncological patients. We randomly selected 400 pulses for each recording, 100 for each physical activity quartile and labelled them into three quality levels. We compared the quality levels distribution among different physical activity ranges throughout the 24 hours. We then used a Multinomial Logistic Regression model to quantitatively evaluate the impact of physical activity, health state, age, and sex on PPG signal quality. We finally estimated six PPG parameters only on higher-quality pulses (i.e., Basic and High quality) and fitted each of them into a Generalized Linear Mixed Effects model to evaluate their sensitivity to physical activity, health state, age, and sex.

Physical activity is well recognized as the main cause hindering the clinical application of PPG signals in daily life [201]. This study could demonstrate its detrimental effect by comparing the quality levels distribution among different physical activity ranges. As expected, as the physical activity got more intense, the percentage of higher quality pulses (i.e., F and E) got lower. Similar results were also obtained from the MLR model's fitting, confirming a lower relative probability of having F and E pulses against B pulses as the physical activity increased. Reliable information can thus be gathered in case of low physical activity, for example, when the subject is at rest or in sedentary conditions, corresponding to AR0 and AR1, in agreement with previous literature [202]. As also Pradhan et al. [145] highlighted, the best data quality could be obtained during the night when the subjects were likely to be asleep. However, a prodromic signal quality analysis appears necessary to obtain reliable data from PPG signal processing.

Another interesting aspect is the different quality distribution obtained by analyzing pulses of healthy and oncological subjects separately. The latter group showed a lower percentage of E pulses than the former in the lowest ARs, and concurrently a higher percentage of F pulses. In addition, cancer subjects were shown to have fewer negative pulses than healthy subjects. This could mean that the pathological condition (cancer in this case) does not affect the quality of PPG signal to such an extent that it completely loses its information content (i.e., B pulses) but rather negatively affects the morphology of PPG pulses by losing definition (i.e., F pulses). The MLR model confirms that for oncological patients and as the age increases, i) the relative probability of having F pulses (compared to B pulses) increases, and ii) the relative probability of having E pluses (compared to B pulses) decreases (positive coefficients). The lower probability of finding E pulses can be ascribed either to a deterioration of the cardiovascular health state because of the pathology itself and the related therapy [203] or to the higher age of the oncological patients compared to the healthy subjects [195], as can be seen in Table 4.1.

Sex had a significant impact on PPG pulses quality. The relative probability of having F and E pulses against B pulses was considerably higher for female subjects than males. Previous studies reported significant sex-based differences in the pulse transit time, that is, the time between the R peak recorded through the electrocardiogram and the consecutive PPG cycle [187], and it is well-known that the cardiovascular system differs between women and men, both in physiological and in pathological conditions [204]. A previous study has already found that commercial smartwatches are less accurate, for

heart rate measurement only, for men than for women [205]. This difference could be due to the different skin thickness, higher for males than females [206]: PPG sensor light has to pass through a larger thickness in male subjects, which could then lead to a deterioration in the PPG signal quality.

From the analysis of PPG waveform parameters, we found that all parameters, except RT, were significantly influenced by physical activity, lowering their values. This result is twofold: rise time RT can be used as a parameter independent of physical activity; conversely, other parameters must be interpreted in light of current physical activity levels. Furthermore, pathological states should also be considered when interpreting PT, PA, RI, and DA, negatively affecting their values. Finally, age had a significant impact only on ΔT and DA: as the age increases, the former assumes lower values while the latter increases. These findings agree with previous literature. Specifically, other authors have found that aging causes a reduction in the time between systolic and diastolic peaks (ΔT) [83], [207] and an increase in the diastolic amplitude (DA) [83], [140], mainly due to an increased arterial stiffness [208]. Since the diastolic peak depends on the reflection of the pressure wave on artery walls, a loss of elasticity (i.e., increased arterial stiffness) brings to a quicker and more intense wave reflection [209].

The PPG signal quality analysis results recommend using features extracted from a basic morphological analysis (i.e., using Basic quality pulses) rather than from an in-depth morphological analysis (i.e., using High quality pulses) in the real-world. This is remarkably advisable if the PPG-based application should be used by subjects at risk of cardiovascular system impairment or deterioration. Unfortunately, several experimental PPG-based applications use features that can be extracted only on High quality pulses [141], [170], [210], [211], thereby risking malfunctioning in this population, especially in real-world conditions, where the availability of High quality pulses is further lowered because of the presence of motion artifacts.

In addition, our results related to the PPG waveform parameters confirmed, as already pointed out by Fine et al. [184], that future PPG-based applications should accurately consider several personal and health-related factors, as these can act as sources of inaccuracy and limit the interpretability and generalizability of the results. PPG sensors undoubtedly have excellent qualities, as they can be easily embedded in wearable devices, are inexpensive, and can collect a variety of vital information. A proper characterization

of the various sources of inaccuracy influencing the PPG signal may expand its use in the clinical field, obtaining a powerful tool allowing pervasive and continuous recordings.

Besides the several sources of inaccuracy the PPG can be subjected to, it is worth remembering that different physiological activities can also influence PPG waveform parameters: in particular, for the time parameters (RT, PT, ΔT), changes in cardiac activity have a significant impact on the timing of the events reflected in the PPG waveform [212]; respiration can induce variations affecting both the pulsatile and non-pulsatile components of the PPG signal [213]; lastly, PPG is strongly affected by the autonomic nervous system, leading to significant changes especially in the time domain [214].

This study presents some limitations, primarily related to the sample size. As previously pointed out, age significantly differed between healthy and oncological subjects (mean age 29.2 vs 49.5 years), thus partly overlapping the effects due to age and health state. The range of physical activities gives another limitation: our dataset lacks vigorous activities, based on the classification of Lin et al. [160], although, based on our results, we can speculate that a tiny proportion of pulses in that category could be used for further processing (i.e., labelled as F or E).

We used a convenience sample, investigating the impact of cancer as a pathological state. We intended to raise the attention, by providing quantitative results, on how a pathology that apparently should have no impact on the PPG signal can lead to misinterpretations if not adequately considered. This study can help expand the knowledge about the impact of cancer on PPG, with the double objective of i) controlling the "health state" variable for a general purpose application, and ii) using PPG with a diagnostic and/or prognostic value for the oncological population. The transferability of these results to other pathologies should be investigated further. However, this work can pave the way to future studies aiming at evaluating the influence of different pathologies on PPG.

From a technical point of view, the low PPG sampling frequency (64 Hz) may limit the accuracy of those time-domain parameters with an order of magnitude comparable to the sampling period (0.0156 s), such as RT and ΔT (mean values equal to 0.26 s and 0.25 s respectively). In the present study, we obtained 34 different values for RT and 27 for ΔT, providing sufficient resolution for a valid analysis. Higher resolution may be used, with each parameter requiring higher temporal sensitivity at the cost of higher data and battery usage, which is a subject of a separate study.

Future studies can be conducted on larger datasets, with a more heterogeneous sampling by age and physical activity and deepening the effects of other personal and health factors, such as weight and height, or other pathological states.

4.6 Conclusions

This study aimed to evaluate the impact of different sources of inaccuracy both on PPG signal quality and on parameters extracted from the PPG morphology. We used a convenience sample of healthy subjects and oncological patients to assess the impact of physical activity, age, sex, and health state. As expected, we found that a higher percentage of good quality PPG pulses can be found during the night and when the subject is in sedentary conditions. Age, pathological state, and male sex are three factors that lower the probability of finding High quality pulses. Regarding the impact of these factors on PPG morphology parameters, physical activity and health state must be considered when interpreting parameter values, while age acts more on those PPG parameters closely related to arterial stiffness. Therefore, it is advisable to conduct further studies on this topic on larger datasets, investigating the effects of different pathological conditions on the PPG signal. Such an approach can help expand the use of PPG-based application, offering a greater robustness and, thus, a more reliable tool for continuous and pervasive monitoring.

5. Impact of Confounding Factors on Physiological Signals through a Cosinor Analysis

Adapted from the conference proceeding: Moscato S., Palumbo P., Sichi V., Giannelli A., Ostan R., Varani S., Chiari L. Impact of pain on physiological signals in home cancer patients, a cosinor analysis. IEEE EMBS in Biomedical and Health Informatics, 2021

5.1 Introduction

Monitoring physiological signals in a real-world context and using them to develop a method to assess pain automatically poses several challenges intrinsic to the uncontrolled environment the system is supposed to work in.

For instance, physiological signals can change throughout the day due to normal circadian fluctuations [215]. Therefore, analysing physiological signals at different times of the day and treating the extracted features regardless of the time of the day can be misleading.

Moreover, the signals that could be exploited for pain estimation can be influenced by other factors, confounding the results of the final application. Such confounding factors, thoroughly studied especially in the last decade, can originate from:

- Psychological state: it is well-known that physiological signals can be influenced by depression and anxiety [216], indeed the same signals exploited for pain assessment have been used for automatic depression [217] and anxiety [218] assessment as well.
- Pharmacological treatment: the use of opioids, and its withdrawal, significantly alter the normal physiological behaviour, being visible through the analysis of, for example, heart rate, electrodermal activity, and physical activity [219], [220].
- Health-related information: health conditions can significantly alter the behavior of physiological signals, also for those signals not strictly related to the pathology itself. An example of that is given in the previous Chapter of this thesis, where several cardiovascular parameters were shown to be impacted by the health condition of the subject (cancer in that case).
- Personal factors: gender may greatly impact the range values of several physiological parameters [221], [222].

If these confounding factors are not properly controlled, together with the information about the circadian rhythm, they can hinder the development of a reliable pain assessment method.

A depiction of the rationale of this approach, represented with conventional representation of Structural Equation Models, is given at Figure 5.1.

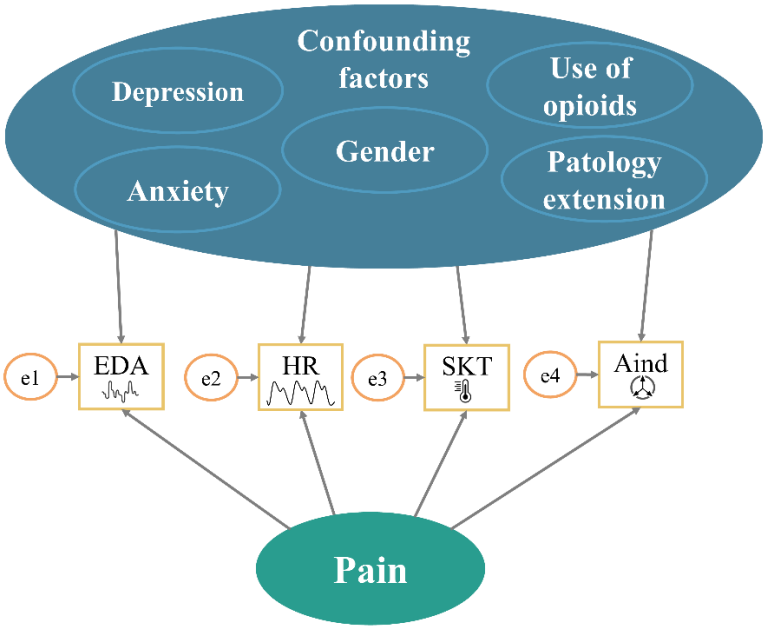


Figure 5.1 Impact of confounding factors on physiological signals

The aim of this study is to analyse the impact of pain and confounding factors on physiological signals on a 24-hour basis.

5.2 Materials and Methods

5.2.1 Study design

This study has been conducted on cancer patients involved in the Look of Life project [223]. An overview of the study, with related socio-demographic data, clinical data, and questionnaires, is given at Appendix C. For each patient the following information has been retrieved and further exploited:

- Pain: Absence vs Presence
- Anxiety: Normal vs Borderline/Abnormal
- Depression: Normal vs Borderline/Abnormal
- Use of opiods: Yes vs No
- Pathology extension: Localized vs Metastatic
- Gender: Male vs Female

5.2.2 Physiological signals

Patients were monitored with the Empatica E4 wristband, whose technical specifications are reported in Appendix A.1.

For the purpose of this study, we used the Heart Rate (HR) automatically estimated by Empatica algorithm, the Electrodermal Activity (EDA), the accelerometer (ACC) data, and the skin temperature (SKT).

As preprocessing phases, the following were the procedure used for the aforementioned signals:

- EDA has been subjected to the processing pipeline described in Appendix A.3
- ACC has been subjected to the preprocessing pipeline described in Appendix A.4 and then used to estimate the the Activity Index (A_{ind}) [160]
- SKT has not been subjected to any preprocessing procedure.

Non-wear segments were visually detected and excluded from further analysis.

The mean of each signal was determined on a 1-hour time window. If a 1-hour time window consisted for more than 50% of non-wear segments, it was discarded from further analysis.

5.2.3 Statistical analysis

A cosinor analysis was conducted to assess the impact of pain and other confounding factors on the physiological signals, considering the time of the day. The model can be written as [224]:

$$x(t) = M + M_p\delta + \left(A + A_p\delta\right) cos\left(\omega t - \left(\varphi + \varphi_p\delta\right)\right) + e(t) \qquad (5.1)$$

where

- M = MESOR
- A = amplitude
- ω = 2π/T, T = 24 hours
- φ = acrophase
- e(t) = error term
- δ is a dichotomous variable indicating pain presence (NRS > 0) or absence (NRS = 0)
- subscript p indicates parameters for the pain covariate on MESOR, amplitude, and acrophase.

We assessed the possible statistical differences based on the aforementioned binary factors by applying a Wald test.

5.3 Results

We enrolled 27 patients (11 M – 16 F, age 50 ± 12). A total of 1361 hours were available after non-wear periods removal. Table 5.1 shows the number of hours divided for each analysed binary factor.

Table 5. 1 Number of available hours divided for each binary factor

	No pain	**Pain**	**No Anx**	**Anx**	**No Dep**	**Dep**	**Female**	**Male**	**No Opioids**	**Opioids**	**Loc**	**Met**
# Hours	599	762	902	459	1179	182	803	558	630	731	608	753

Anx: Anxiety. Dep: Depression. Loc: Localized. Met: Metastatic.

Results from the cosinor analysis for all binary factors are shown in Table 5.2, while in Figure 5.2 the cosine models for the four physiological signals, divided for absence and presence of pain, are shown.

The circadian rhythm of all signals but the EDA resulted to be well modelled through a cosine wave.

Patients with pain had higher EDA and HR values and delayed HR acrophase.

Anxiety had a significant impact on HR, SKT, and A_{ind} values, also influencing the amplitude of SKT and A_{ind} circadian waves.

Depressed patients showed higher EDA both in the Mesor and amplitude, lower HR for Mesor but greater oscillations and lower values of SKT in terms of both Mesor and amplitude.

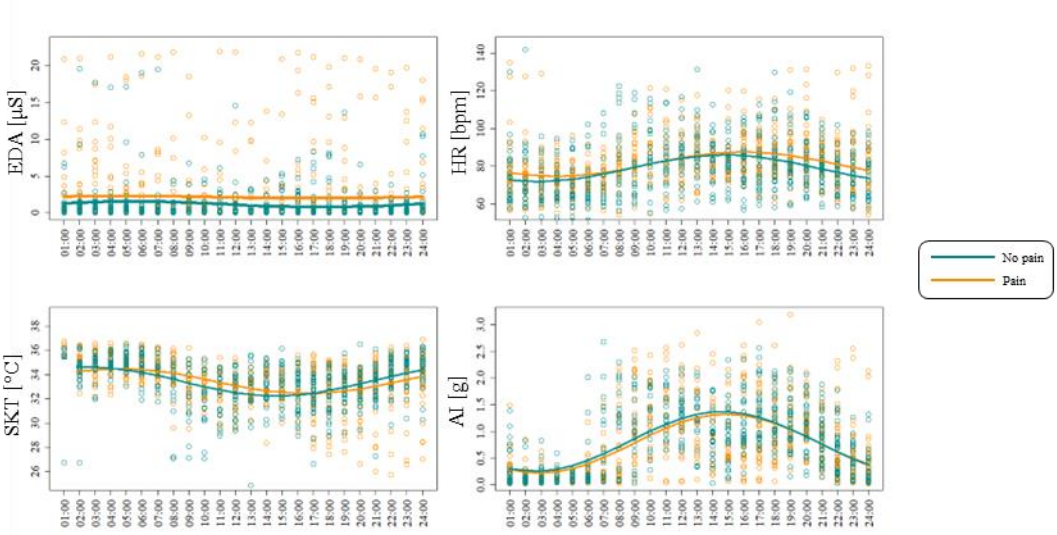


Figure 5. 2 Cosinor analysis depicting presence and absence of pain in four physiological signals

Gender only impacted SKT and AI, showing higher SKT Mesor for females and lower amplitudes, and higher Mesor and amplitude in A_{ind} for females.

Opioid intake significantly changed the behavior of all four physiological signals.

Finally, also the pathology extension played a role, increasing the Mesor values for EDA and decreasing the Mesor values for HR in metastatic cancer cases.

5.4 Discussion

We conducted a study to assess the possible impact of several binary factors on the circadian rhythmicity of four physiological signals, namely HR, EDA, A_{ind}, and SKT recorded in real-world contexts on oncological patients.

Circadian rhythmicity of all signals but EDA can be well represented through a cosine wave. Previous evidences showed a strong circadian rhythmicity for the EDA signal [225]. More complex models are needed to better represent EDA's circadian rhythm.

Physiological signals were different for patients according to the presence of pain. This is in line with previous literature, which has shown bidirectional interactions between pain and circadian rhythm [226]. Furthermore, significant changes in Mesor's EDA and

HR well represent the higher sympathetic activation expected from subjected experiencing pain [227].

Table 5.2 Cosinor analysis to assess the impact of pain and confounding factors on physiological signals' circadian rhythm

	EDA			HR			SKT			A_{ind}		
	M [µS]	**A [µS]**	**Φ [hh:mm]**	**M [bpm]**	**A [bpm]**	**Φ [hh:mm]**	**M [°C]**	**A [°C]**	**Φ [hh:mm]**	**M [g]**	**A [g]**	**Φ [hh:mm]**
No Pain	1.19	0.38	05:28	78.83	7.06	14:45	33.44	1.19	02:35	0.81	0.56	14:37
Pain	2.09 *	0.31	04:49	81.00 *	6.46	15:58 *	33.47	1.02	04:36 *	0.77	0.55	14:54
No Anx	1.64	0.21	04:28	80.69	6.64	15:05	33.35	1.20	03:17	0.75	0.53	14:40
Anx	1.90	0.38	05:27	78.84 *	6.60	15:58	33.65 *	0.82 +	04:35 +	0.14 *	0.61 +	15:03
No Dep	1.14	0.64	04:35	80.07	6.10	15:23	33.35	1.19	03:28	0.79	0.54	14:38
Dep	5.57 *	2.17 *	05:13	79.87 *	9.84 +	15:35	34.07 *	0.36 *	5:25 *	0.83	0.62	15:35 +
No Opioids	1.02	0.66	15:37	78.20	6.48	15:42	33.18	1.33	04:07	0.74	0.51	15:14
Opioids	1.36 *	0.20	17:50 +	81.62 *	7.01	15:19	33.72 *	0.80 *	3:01 +	0.84 *	0.60 *	14:28 *
Loc	2.41	0.24	02:42	77.87	7.20	15:05	33.57	0.94	02:56	0.79	0.51	15:28
Met	1.18 *	0.40	05:07	81.77 *	6.47	15:37	33.36	1.20	4:02 +	0.80	0.60 +	14:19 *
Female	1.57	0.58	15:37	80.50	6.60	15:21	33.63	0.94	02:45	0.92	0.67	14:35
Male	1.96	0.31	00:02	79.40	6.66	15:33	33.23 *	1.28 +	4:35 *	0.61 *	0.41 *	15:23 *

+ $p < 0.05$, * $p < 0.01$

Regarding anxiety and depression, also in this case the bidirectional interactions between mental health and circadian rhythm have been studied for decades [228]. In this context, we found results inconsistend with previous literature. Specifically, anxiety is supposed to accelerate HR [229], although a study revealed that the real HR proved to be unrelated to anxiety, while the perceived HR resulted to be significantly linked to anxiety [230]. Regarding physical activity, the lower activity level in patients with anxiety is in line with previous literature [231]. Results on depression also disagree with previous literature, stating a hypoactivity of EDA in depressed patients [232]. Further studies are needed to consolidate the behavior of physiological signals regarding mental health conditions.

In recent years, the study of physiological reseponses to opioids has grown consistently due to the increasing abuse of the substance itself [233], [234]. In this study, patients using opioids showed higher Mesor values for all the analysed physiological signals. In particular, increase in EDA and SKT are consistent with previous literature, while HR should be decreased with opioids intake [233].

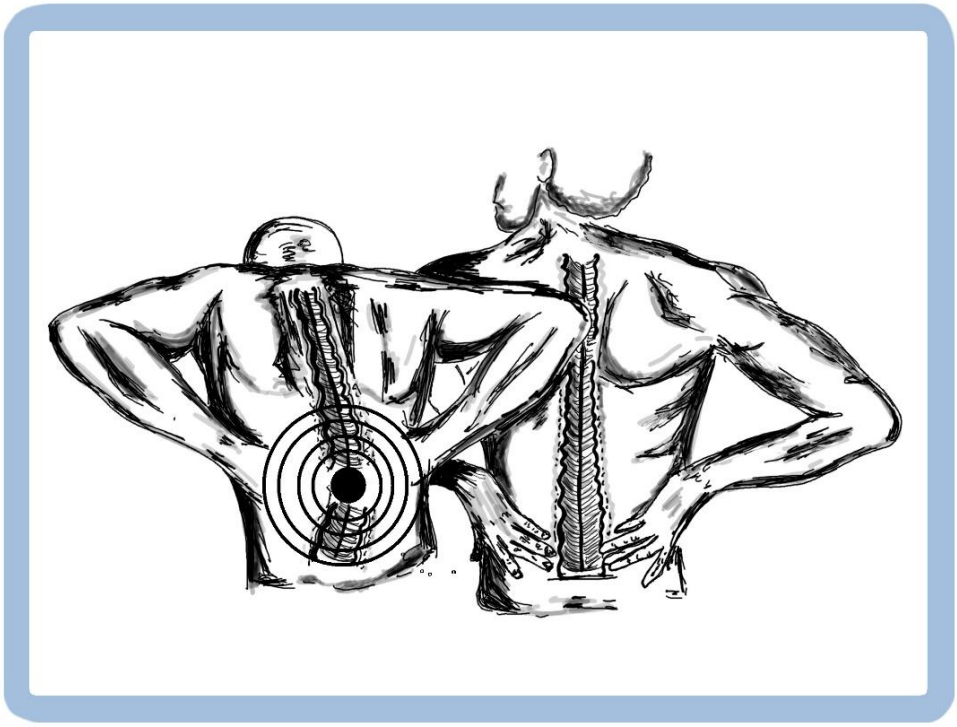
Also cancer stage had an impact on the analysed physiological signals. The study by Wu et al. [235] investigated the autonomic activity linked to cancer stage through HR and heart rate variability (HRV) analysis on a breast cancer population. They found no significant changes in HR, but they highlighted an increasing trend going from bening to advanced stage cancer, consistent with our results.

Sex showed to significantly impact only SKT and A_{ind}. Previous studies highlighted consistent differences in autonomic activity on a gender basis [236], but these did not emerge in this study.

Some limitations of this study need to be highlighted. The small dataset, together with a simple cosine model can hinder in finding all the significant differences there can be in the analysed binary factors.

Despite the limitations, results obtained from this study can help in expanding the knowledge about the factors that can have an impact on physiological signals. Moreover, having obtained these results on signals recorded in an ecological environment has a translational value and represents a step forward towards the development of real-world tools for automatic pain assessment.

SECTION III: PAIN ASSESSMENT ON EXTERNALLY-INDUCED NOXIOUS STIMULATION IN HEALTHY SUBJECTS AND LOW BACK PAIN PATIENTS



6. AUTONOMIC SIGNALS COMPARISON BETWEEN HEALTHY SUBJECTS AND CHRONIC LOW BACK PAIN PATIENTS AT REST AND DURING NOCICEPTIVE STIMULATION

Adapated from the conference proceeding accepted for GNB Congress: Moscato S., Zhu W., Guo Y., Kamarthi S., Colebaugh C. A., Schreiber K. L., Edwards R. R., Urman R. D., Xiao Y., Chiari L., and Lin Y. Comparison of autonomic signals between healthy subjects and chronic low back pain patients at rest and during noxious stimulation.

6.1 Abstract

Chronic pain is a major cause of disability worldwide. While acute pain may serve as a protective function, chronic pain and the associated neural processing changes negatively impact function and quality of life. This neural plasticity may include changes to the autonomic nervous system (ANS) potentially detectable as changes in various physiological signals. We aim to evaluate differences in the physiological signals reflecting ANS changes, by comparing healthy subjects and patients with chronic low

back pain during standardized pain stimuli. We extracted several features from photoplethysmography (PPG), electrodermal activity (EDA), and respiration, at rest and during a repeated pinprick test. We found sgnificant group differences in some PPG parameters at rest and in response to the two different noxious stimulations. In addition, chronic pain patients had consistently higher basal sympathetic activity, and a blunted autonomic response when subjected to nociceptive stimuli.

6.2 Introduction

Pain is an “unpleasant sensory and emotional experience associated with, or resembling that is associated with, actual or potential tissue damage” [8].

Different types of pain can be experienced. One of the main distinctions is between acute and chronic pain.

Acute pain arises from the activation of nociceptors when a thermal, mechanical, or chemical nociceptive stimulus affects the body.Noxious stimuli activate neural transduction at the peripheral, spinal, and brain levels, collectively falling under the “pain neuromatrix” concept. Acute pain has a protective function on the body, acting as an alarm bell for a potential threat [237].

On the other hand, chronic pain is defined as persistent or recurrent pain lasting for more than three months [238]. It is now well known that chronic pain leads to a significant neural plasticity [239], with substantial functional and anatomical changes to reach a new equilibrium different from that of healthy subjects [40]. Chronic pain does not serve a protective function, being maintained despite the absence of the inciting stimulus, and is associated with physiologic and psychosocial changes [240].

Manifestations of physiologic changes may include an Autonomic System (ANS) activity. Pain and ANS are anatomically and functionally linked, as pain influences the activity of ANS and vice versa [52]. In the case of chronic pain, previous studies have observed decreased ANS reactivity, potentially leading to a reduced ability of the body to respond promptly to internal and external stimuli [241].

ANS activity can be evaluated by monitoring some physiological signals, commonly called “autonomic” signals precisely because the ANS regulates them. Examples of these signals are Photoplethysmography (PPG), representing the blood volume changes occurring at each heartbeat, Electrodermal Activity (EDA), referring to variations of the electrical properties of the skin due to sweat secretion, and Respiration (Resp).

Autonomic signals have been extensively used in "emotion recognition" studies, representing a new way to analyse emotions, and increasingly bridging the gap in human-computer interaction [242]. Gaining greater facility with measurement of fluctuations and variability of autonomic signals under the condition of pain transduction in twofold: from a clinical point of view, this may allow insight into the impact of nociceptive activation on physiology, both in the acute and chronic phase; from a more technical point of view, it would allow tailoring algorithms better to detect a variety of painful stimuli in different individuals more sensitively, and to distinguish the extent of neural remodeling inherent in an individual with cchronic disease.

This study aims to explore differences in autonomic signals by comparing healthy control (HC) subjects and chronic Low Back Pain (cLBP) patients i) under rest conditions and ii) when stimulated with a noxious stimulus. This is a further work based on exploration work of the Intelligent Human Machine Systems laboratory, Northeastern University [243], [244].

6.3 Materials and Methods

6.3.1 Participants

The study involves the enrolment of both HCsubjects and cLBP patients. The inclusion criteria in common for the two samples are the following:

- 18 years or older
- No evidence of neurological and/or cognitive impairment
- No history of myocardial infarction or other serious cardiovascular condition in the prior 12 months
- Ability to speak English to complete the questionnair measures

In addition, HC subjects had to be free of any history or diagnosis of chronic pain. cLBP patients were required to have a history of cLBP for at least 3 months, with an average pain intensity higher than 3/10.

6.3.2 Experimental procedure

The study protocol, called the Novel Computation Methods for **C**ontinuous **O**bjective **M**ultimodal **P**ain **A**ssessment **S**ensing **S**ystem (COMPASS), was approved by the Institutional Review Board of Brigham and Women's Hospital, Boston, MA, USA

(protocol code 2019P002781, 18/11/2019). The whole experimental protocol revolves around administering of the Quantitative Sensory Testing (QST). This procedure tests both large and small sensory nerve fibers by providing nociceptive stimuli in a clinical lab setting [245]. A more in-depth discussion is given in Chapter 2.

A graphical representation of the study protocol is presented in Figure 6.1.

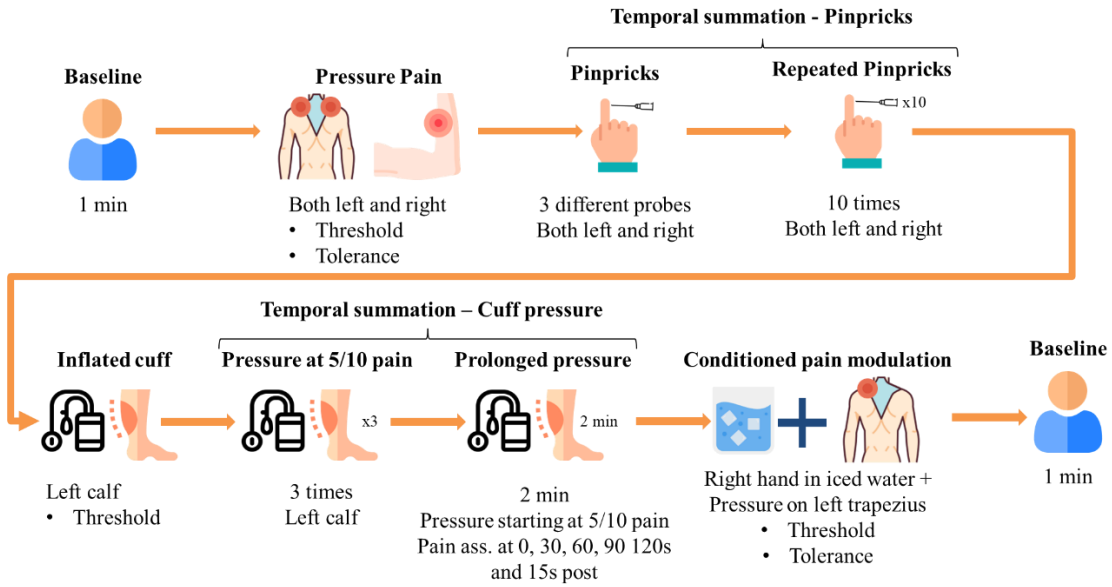


Figure 6.1 COMPASS protocol

The main stages of the study protocol are the following:

- Baseline – start of the experiment: the participant is asked to be at rest for one minute
- Pressure pain: the participants are stimulated with a pressure stimulus through a pressure algometer in three anatomical locations: right and left trapezius and right and left forearm. Mechanical pressure is increased at a steady rate of 30 kPA/s. The participant is asked to manifest when the pain threshold and pain tolerance are reached. Pressure stimulus is removed once the pain tolerance has been reached
- Pinpricks – 3 different probes: the participant is stimulated with standardized weighted pinprick applicators with three different forces (128 mN, 256 mN, and 512 mN). The three probes are applied consecutively one after the other in ascending order, and for each of them the participant is asked to rate his or her pain on a 0-10 NRS. The stimuli are applied on the index finger of both hands, between the first and second interphalangeal joint

- Repeated pinpricks: the probe that induced some pain in the previous step is chosen for repeated testing. A train of 10 stimuli has been applied with a rate of 1 stimulation/second. The participant is asked to rate his or her pain on a 0-10 NRS after the first, the fifth, and the tenth stimulus. The stimuli are applied in the same anatomical location as for the previous test
- Inflated cuff – calibration: a blood pressure cuff is placed around the left gastrocnemius muscle and inflated with a rapid cuff inflator until the participant manifests that the pain threshold has been reached
- Inflated cuff – three repetitions: the blood pressure cuff around the left gastrocnemius muscle is inflated three times in a row until the participant manifests that a 5/10 on a 0-10 NRS is reached
- Inflated cuff – prolonged pressure: the blood pressure cuff is inflated and left for 2 minutes
- Conditioned pain modulation: the participant is asked to immerse the right hand in a circulating water bath maintained at a temperature of 4-8 °C and simultaneously stimulated with the pressure algometer on the left trapezius. The participant is asked to manifest when pain threshold and pain tolerance are reached
- Baseline – end of the experiment: the participant is asked to be at rest for one minute

During the entire duration of the experimental procedure, the following biomedical devices are used to record physiological signals:

- FlexComp, an FDA-cleared system that allows to record with a fixed sampling frequency of 256 Hz:
    - Electrodermal Activity (EDA), whose electrodes are placed on the left hand's ring and index finger
    - Photoplethysmography (PPG), with a sensor placed on the left hand's middle finger
    - Electromyography (EMG), with two electrodes placed on the left forearm
    - Skin Temperature (ST), whose sensor is placed on the back of the hand
    - Respiratory Rate (RR), with a band around the trunk above the stomach
- Enobio, an FDA-cleared portable electrophysiology system for the recording of electroencephalogram (EEG) with a sampling frequency of 1000 Hz

- Tobii: an FDA-cleared vision system intended for eye movement recording

6.3.3 Physiological signal processing

In this study, only the physiological signals recorded with the FlexComp system during the following three conditions have been evaluated:

- Baseline – start of the experiment
- Repeated pinpricks on the right hand
- Inflated cuff – prolonged pressure

All the signals recorded with the system mentioned above have been subjected to the following procedure:

- Signal preprocessing
- Signal segmentation
- Feature extraction

6.3.3.1 EDA

EDA signal has been subjected to a 5th-order Butterworth low-pass filter with a cut-off frequency of 1 Hz. EDA signal has also been normalized with a z-score procedure. The normalized signals were then subjected to the cvxEDA algorithm [246] to decompose the whole signal into the tonic and phasic components.

6.3.3.2 PPG

PPG signal has been subjected to the preprocessing pipeline depicted in Appendix A.2. Since PPG signals in this dataset are clean enough (with respect to more noisy data for those signals recorded in real-world context), a simple minimum finding approach has been implemented to detect the systolic feet and divide the signals into PPG pulses based on those systolic feet, with the restriction of finding minimum values if distant for more than 0.5 s.

6.3.3.3 Respiration

The respiration signal has been subjected to a 4th-order high-pass Butterworth filter, with a cut-off frequency of 0.1 Hz. This filter has been applied to remove the slow fluctuations, which are useless for this study. The Advanced Counting method [247] has been applied to the filtered signal to detect the respiration cycles and estimate the respiration rate.

6.3.4 Autonomic parameters extraction

For the baseline recording, features have been extracted by the whole recording (i.e., 1 minute).

For the repeated pinpricks test, the recording has been segmented into four phases:

- Pre-test: 5 seconds before the beginning of the test
- 1^{st}-5^{th} rep: Between 1st and 5th repetition (5 seconds)
- 5^{th}-10^{th} rep: Between 5th and 10th repetition (5 seconds)
- Post-test: 5 seconds after the end of the test

For the inflated cuff – prolonged pressure, the recording has been segmented into six phases:

- Pre test: 30 seconds before the beginning of the test
- 30 sec: 0-30 seconds
- 60 sec: 30-60 seconds
- 90 sec: 60-90 seconds
- 120 sec: 90-120 seconds
- Post-test: 30 seconds after the end of the test

Twelve parameters have been extracted from the analysis of PPG pulses:

- Heart Rate Variability (HRV) analysis - we estimated the Interbeat Intervals (IBIs) as time differences between two consecutive systolic feet. The obtained IBIs time series has been filtered using the approach described in [248]. Extracted HRV parameters are the mean value of IBIs (*meanIBI*), standard deviation of normal heartbeats (*SDNN*), root mean squared of successive differences (*RMSSD*), Poincaré plot standard deviation perpendicular (*SD1*), and along (*SD2*) the line of identity.
- Morphological analysis - by analysing the morphology of each PPG pulse, we estimated PPG pulse amplitude (*PulseAmpl*), area under the curve between systolic foot and successive systolic peak (*A1*), area under the curve between the systolic peak and the successive systolic foot (*A2*), area under the PPG pulse (*A*), time between systolic foot and the successive systolic peak (*T1*), time between systolic peak and the successive systolic foot (*T2*) [86].

For the EDA, we estimated a total of 8 parameters:

- Whole EDA signal: mean (*meanEDA*) and standard deviation (*stdEDA*), and the symbolic information entropy (*SIE*) [249]
- Tonic component: mean (*meantonic*) and standard deviation (*stdtonic*)
- Phasic component: relevant peaks have been retrieved as those peaks with a slope (by analysing the first derivative) higher than 0.01. From the relevant peaks, the mean (*meanampEDR*) and standard deviation (*stdampEDR*), together with the frequency (*freqEDR*) expressed as the number of peaks per minute have been retrieved.

For the Resp signal, we retrieved the mean respiration rate (*meanRR*). Since we used an automatic method to detect the respiration cycle, we discarded those estimates exceeding 30 breaths/min.

6.3.1 Statistical analysis

We used a Mann-Whitney U test to compare parameters between HC subjects and cLBP patients during the rest condition. We used a Wilcoxon signed-rank test for paired data to assess any statistical differences in different phases during the two tests,.

6.4 Results

6.4.1 Dataset

Twenty-four subjects have been enrolled in this study, 15 HC subjects (age 27.20 ± 11.58, 4 M, 11 F) and 9 cLBP patients (age 43.67 ± 14.97, 4 M, 5 F). Recordings failed for one cLBP patient, who was therefore discarded. Some participants have been subjected to the same protocol once, but for the purpose of this study only the first trial for each participant has been analysed.

6.4.2 Baseline recording

Results of the comparison between HC subjects and cLBP patients are presented in Table 6.1.

By analysing the baseline recordings, *meanIBI*, *A2*, *A*, and *T2* were significantly different between HC subjects and cLBP patients. All these parameters are significantly lower for cLBP patients than for HC subjects.

Table 6. 1 Baseline recordings' results. HC = Healthy Controls, P = patients

	HC mean (std)	cLBP mean (std)	p-value
meanIBI [ms]	865.51 (93.80)	720.20 (133.36)	0.048
SDNN [ms]	58.68 (12.99)	26.14 (27.99)	0.067
RMSSD [ms]	66.76 (29.69)	38.16 (34.16)	0.057
SD1 [ms]	47.19 (21.00)	26.96 (24.13)	0.057
SD2 [ms]	65.30 (12.20)	41.55 (31.94)	0.078
PulseAmpl [a.u.]	3.27 (0.15)	3.28 (0.26)	0.944
A1 [a.u.*sample]	69.01 (20.25)	61.93 (20.25)	0.159
A2 [a.u.*sample]	269.55 (42.91)	211.62 (62.26)	0.014
A [a.u.*sample]	338.56 (53.13	2743.55 (64.91)	0.034
T1 [ms]	156.78 (30.17)	146.72 (44.24)	0.290
T2 [ms]	710.48 (104.42)	575.70 (110.74)	0.024
meanEDA [n.u.]	-1.32 (0.55)	-0.91 (0.97)	0.438
stdEDA [n.u.]	0.11 (0.11)	0.16 (0.25)	0.944
SIE [ ]	0.78 (0.29)	0.74 (0.25)	0.833
meantonic [n.u.]	-1.33 (0.55)	-0.91 (0.97)	0.438
stdtonic [n.u.]	0.10 (0.10)	0.16 (0.25)	0.888
meanampEDR [n.u.]	0.04 (0.06)	0.01 (0.02)	0.488
stdampEDR [n.u.]	0.04 (0.08)	0.01 (0.02)	0.384
freqEDR [peaks/min]	3.14 (3.04)	1.80 (1.42)	0.438
meanRR [breaths/min]	16.17 (2.58)	17.40 (5.03)	0.672

6.4.3 Repeated pinpricks test

We separately compared the four different phases of the repeated pinpricks test for HC subjects and cLBP patients.

By analysing the reaction of both HC subjects and cLBP patients, we found significant changes in:

- *meanIBI*, for HC subejcts: significantly higher values were detected in the 5^{th}-10^{th} rep phase, compared to pre- and post-test, and a significant decrease from 1^{st}-5^{th} rep to post-test
- *SDNN*, RMSSD, and SD1 for cLBP patients: significantly lower value from the pre-test to post-test
- *A* for cLBP patients, significant decrease from the pre-test to 1^{st}-5^{th} rep and 5^{th}-10^{th} rep

- *T2* for HC subjects: significant increase by comparing 1st-5th rep with 5th-10th rep and post-test
- *meanEDA* and meantonic for HC subjects: significantly lower values in the pre-test compared to all the other phases
- *stdEDA* and stdtonic for HC subjects: significant decrease from 1st-5th rep to post-test
- *SIE* for cLBP patients: significant decrease from1 st-5th rep to post-test
- *stdampEDR* for HC subjects: significant decrease from 5th-10th rep to the "after phase"
- *freqEDR* for cLBP patients: significant decrease from pre-tets to 1st-5th rep and 5th-10th rep, and a significant increase by comparing 1st-5th rep and 5th-10th rep with post-test
- *meanRR* for HC subjects: significant decrease by comparing pre-test with post-test. For cLBP patients: significant increase from 1st-5th rep to post-test

Significant changes are reported in Figure 6.2, while numerical values are reported in Appendix D, Table D.1 and Table D.2 for HC subjects and cLBP patients, respectively.

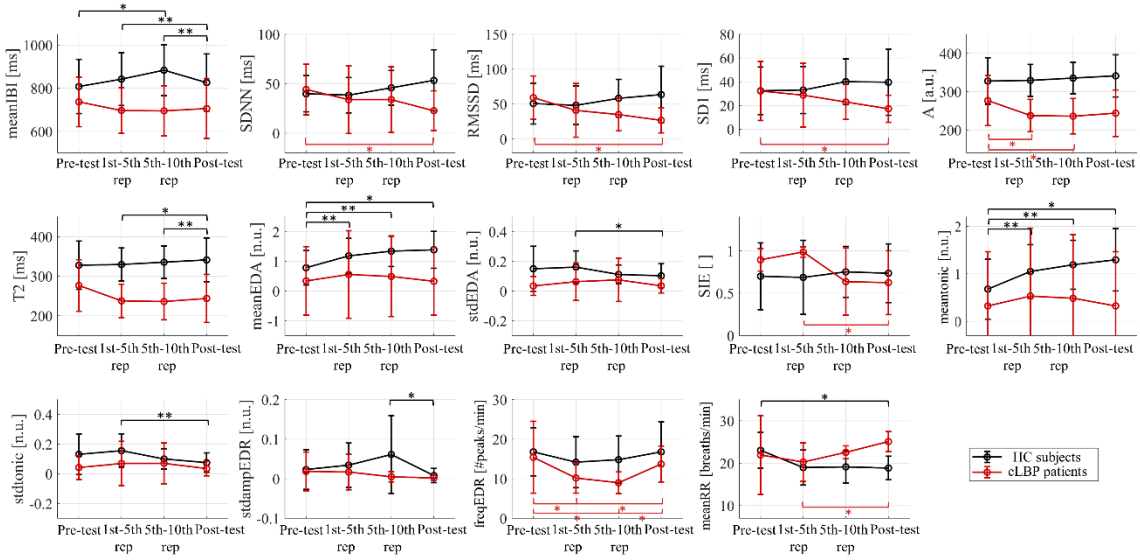


Figure 6.2 Significant changes during Repeated pinpricks test

6.4.4 Inflated Cuff – prolonged pressure test

We separately compared the four different phases of the inflated cuff-prolonged pressure test for HC subjects and cLBP patients.

By analysing the reaction of both HC subjects and cLBP patients, we found significant changes in:

- *meanIBI* for HC subjects: significant decrease by comparing 120 sec with post-test
- *SDNN* for cLBP patients: constant decrease by comparing pre-test, 60 sec, and 90 sec to 120 sec
- *SD2* for HC subjects: significant decrease from 30 sec to 120 sec and from 90 sec to 120 sec. For cLBP patients: significant decrease from 30 sec to 120 sec and from 60 sec to 120 sec
- *PulseAmpl* for HC subjects: significant decrease by comparing pre-test and 30 sec with post-test and pre-test with 90 sec, and a significant increase by comparing 60 sec, 90 sec, and 120 sec to the "after" phase. For cLPB patients: significant decrease from pre-test to 30 sec, and from pre-test, 30 sec and 60 sec to post-test
- *A1* for HC subjects: significant decrease from pre-test to 30 sec, significant increase from pre-test and 60 sec, and a significant increase by comparing 30 sec to 60 sec, 90 sec, 120 sec, and post-test
- *A2* for HC subjects: significant decrease from 30 sec to 60 sec, 90 sec, 120 sec, post-test. For cLBP patients: significant increase from 30 sec to 90 sec
- *A* for HC subjects: significant decrease by comparing pre-test and 60 sec with post-test. For cLBP patients: significant decrease from 60 sec to 120 sec
- *T1* for HC subjects: significant increase from pre-test to post-test, and a significant decrease from 90 sec and 120 sec to post-test
- *T2* for HC subjects: significant increase from pre-test to 30 sec, a significant decrease from 30 sec to 60 sec and post-test, and a significant increase from 30 sec to 90 sec and 120 sec
- *meanEDA* for HC subjects: significant decrease from 120 sec to post-test
- *stdEDA* for HC subjects: significant increase from pre-test to 30 sec, a significant decrease from 30 sec to all the following phases, a significant increase from 60 sec to 90 sec and a significant decrease from 60 sec to 120 sec and post-test, a significant decrease from 90 to all the following phases and from 120 sec to post-test.For cLBP patients: significant decrease from 30 sec to 60 sec, from 90 sec to all the following phases, and from 120 sec to post-test

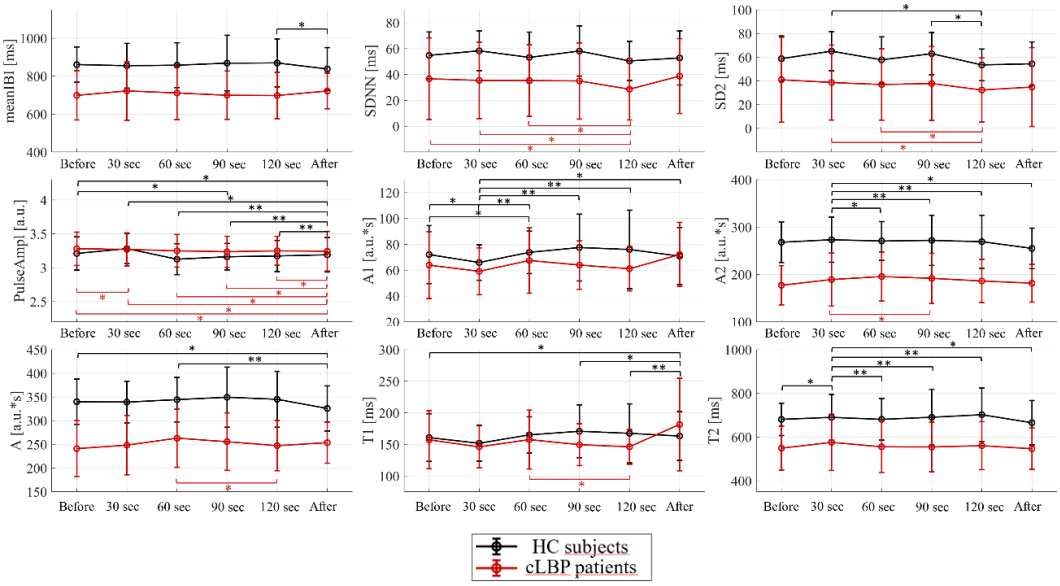


Figure 6.3 Significant changes in PPG parameters from the Inflated cuff - pressure pain test

- *SIE* for HC subjects: we found a significant increase from pre-test to 30 sec, a significant decrease from 30 sec to all the following phases, from 60 sec to 120 sec and post-test, and a significant increase from 90 sec to post-test. For cLPB patients: decrease from 30 sec to 60 sec, a significant increase from 30 sec to post-test, and a significant increase from 60 sec and 90 sec to post-test
- *meantonic* for HC subjects: significant increase from pre-test to 30 sec and 60 sec, a significant decrease from 30 sec to all the following phases, a significant increase from 60 sec to 90 sec and a significant decrease from 60 sec to 120 sec and post-test, a significant decrease from 90 to all the following phases and from 120 sec to psot-test. For cLBP patients: decrease from 60 sec and 90 sec and from 120 sec and post-test
- *stdtonic* for HC subjects: there was a significant increase from pre-test to 30 sec and 60 sec, a significant decrease from 30 sec to all the following phases, and a significant decrease from 60 sec to 120 sec and post-test. For cLBP patients: significant decrease from 30 sec to 60 sec and from 120 sec to post-test
- *meanampEDR* for HC subjects: significant increase from pre-test to 30 sec and 60 sec, a significant decrease from 30 sec to all the following phases and from 60 sec to 120 sec. For cLPB patients: significant decrease from the “before” phase to 30 sec and 60 sec and post-test, and a significant increase from 60 sec to post-test
- *stdampEDR* for HC subjects: significant increase from the “before” phase to 30 sec, and a constant decrease from 30 sec to the following phases

- *freqEDR*: only for HC subjects there was a significant constant decrease from pre-test to all the following phases, from 30 sec to 60 sec, 90 sec, and 120 sec, and a significant increase from 120 sec to post-test.

Significant changes are depicted in Figure 6.3 and 6.4 respectively for features extracted from PPG and EDA. Numerical values are reported in Appendix D, Table D.3 and D.4 for HC subjects and cLPB patients respectively.

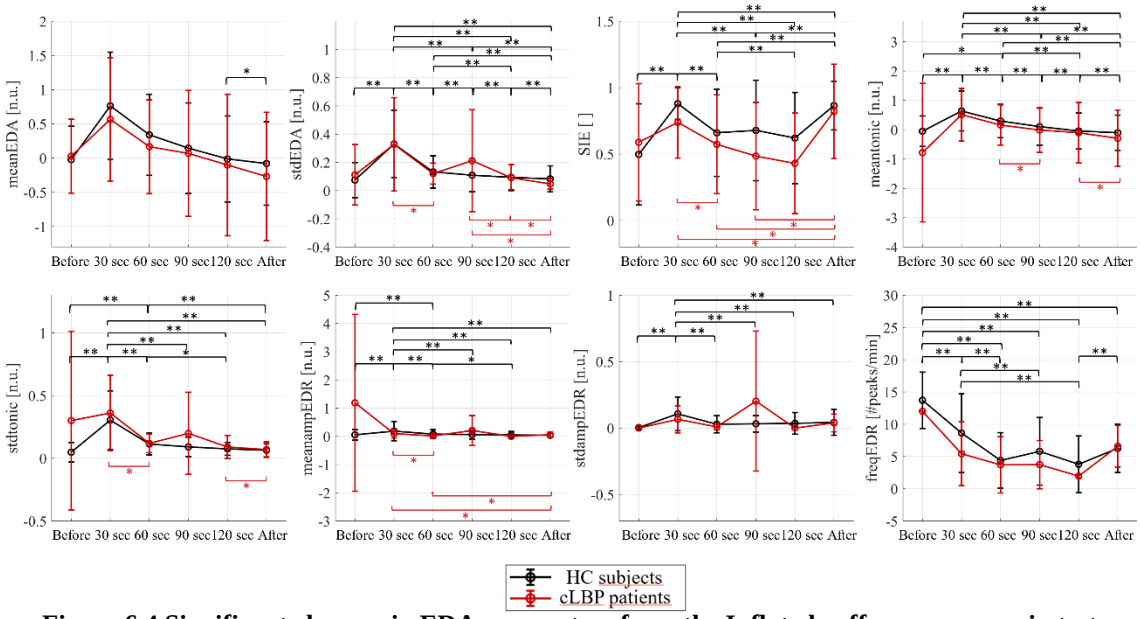


Figure 6.4 Significant changes in EDA parameters from the Inflated cuff - pressure pain test

6.5 Discussion

We conducted a study to assess the different autonomic activity by comparing healthy control subjects and chronic low back patients in rest condition and when subjected to noxious stimulation.

Under the rest condition, only specific parameters extracted from the PPG were statistically different between HC subjects and cLBP patients. In particular, *meanIBI* was significantly lower in cLBP patients. This is in line with previous study findings on chronic pain patients [250]. Greater basal sympathetic nervous system outflow could potentially lead to a higher basal heart rate (corresponding to a lower *meanIBI*) [251].

Some PPG morphological parameters, namely *A2*, *A*, and *T2*, were significantly lower in cLBP patients than HC subjects. Both *A2* and A are related to *T2*, which is the time difference between the systolic peak and the successive systolic foot. This can also be the basis for lowering the *meanIBI*: chronic pain induces a change in the second part of the

cardiac cycle, during the diastolic phase. In addition, morphological parameters have been studied in relation to stress, finding similar results as the ones reported in this study [252]. Regarding the analysis of physiological parameters during the repeated pinpricks test, HC subjects and cLBP patients showed a different autonomic reaction. Overall, cLBP patients appeared to exhibit a blunted degree of change, in agreement with some previous studies [253].

The cuff pressure test induces more significant changes rather than the repeated pinprick test. This could be due either to a more consistent autonomic reactivity to nociceptive stimuli or to the longer time window for which the nociceptive stimulus has been imposed. In particular, there were more significant changes in EDA parameters in this experimental condition, at least for HC subjects, and the trend was very similar in this case for the two different populations.

It is interesting to note the behaviour of autonomic parameters extracted from the EDA: while HC subjects showed a dynamic in several parameters, cLBP patients showed a significant change only for the *SIE* parameter, related to the complexity of the signal, and the frequency of EDA peaks (*freqEDR*). Specifically for *freqEDR*, since EDA is influenced only by the sympathetic branch of the autonomic nervous system, it is supposed that a nociceptive stimulus should increase the number of EDA peaks [95]. Conversely, for cLBP patients the frequency of EDA peaks diminishes when stimulated with a nociceptive stimulus. This is proof that a functional neural rearrangement occurred in cLBP patients.

The study presents some limitations that can hamper the generalizability of the results. Firstly, our dataset consisted of an unbalanced number of subjects for the two populations (15 HC subjects vs 8 cLBP patients). Some changes may not have been detected because of the small number of patients with cLBP. More subjects for both populations should be involved in this study.

Still regarding the sample population, HC subjects and cLBP patients presented a different age distribution, with the former group being significantly younger than the latter. Age-related differences in PPG parameters are well recognized in previous literature [254], [255]. Therefore, a strategy could be to control such parameter in future works, or to enroll HC subjects and cLBP patients from the same age group.

Another limitation are the short phases related to the nociceptive stimuli test. A strategy to gain more robust results could be repeating the same nociceptive stimulus or making them last longer.
In future studies, we plan to compare reactions in the two populations to different noxious stimuli and to develop automatic methods to assess pain.

6.6 Conclusion

We carried out a study to explore differences in the autonomic activity measured by a set of physiologic measures, between HC subjects and cLBP patients at rest and during noxious stimulations Our findings suggest a higher basal sympathetic activation at rest for cLBP patients, but a less dynamic response when subjected to a noxious stimulus.

7. AUTOMATIC CLASSIFICATION OF HEALTHY CONTROLS AND CHRONIC LOW BACK PAIN PATIENTS

7.1 Introduction

As the previous chapter shows, healthy control (HC) subjects and chronic Low Back Pain (cLBP) patients consistently differences in several physiological parameters. In order to obtain robust and reliable methods for objective pain assessment, a first step could be to correctly discern between these two populations, following the rationale depicted in Figure 7.1.

This chapter presents the development of a binary classification to distinguish between HC subjects and cLBP patients as a prodromic phase for pain assessment based on physiological signals. Different strategies for window segmentation and the number of selected features are shown using five different machine learning (ML) approaches.

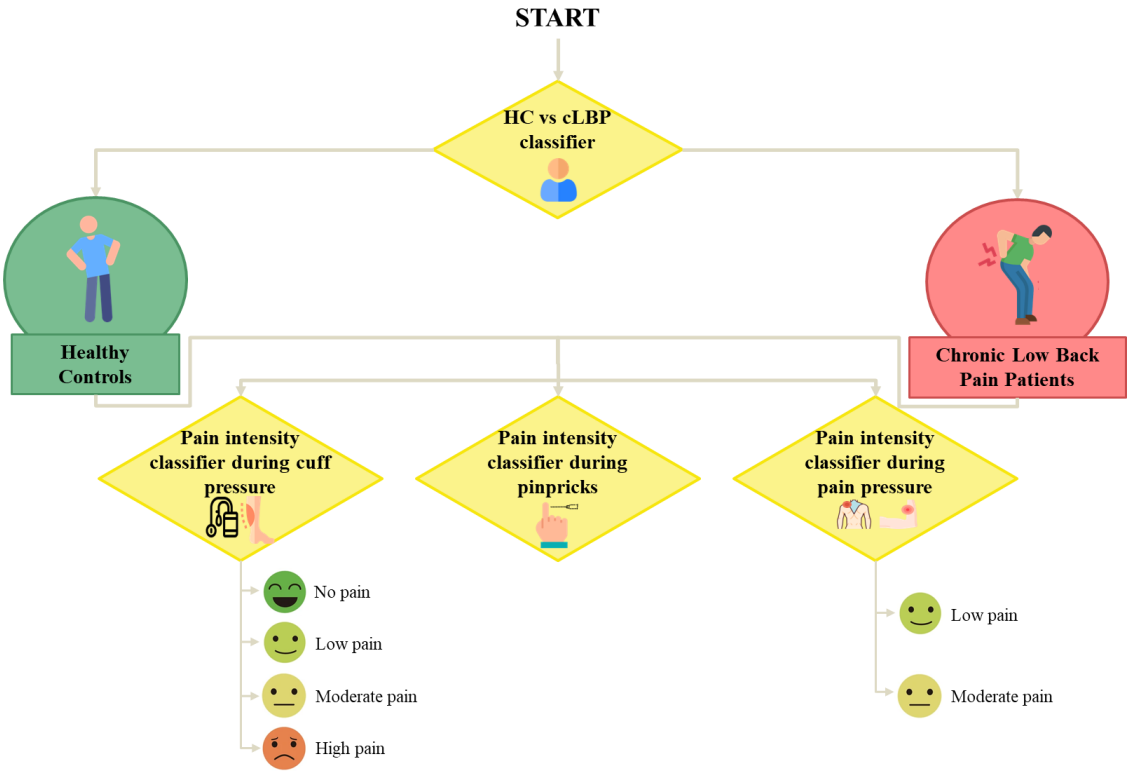


Figure 7.1 Rationale to develop a robust and reliable automatic pain assessment method involving healthy control subjects and chronic low back pain patients

7.2 Materials and Methods

7.2.1 Dataset

Specifications about the population involved in this study are given at 6.3.1. Some participants have been subjected to the experimental protocol twice. In this case, we retained all the trials.

For the HC subjects / cLBP patients classification purpose, baseline sessions, recorded at the beginning and at the end of the experimental protocol (see Figure 6.1), have been used. We checked through linear mixed effects models that there are no statistically significant differences in the features extracted in these two-time recordings of the experimental protocol.

7.2.2 Physiological signals processing

Since there is no standardized time window to extract the features from, we segmented the two baseline sessions into time windows with different length, namely 3, 5, and 10 seconds. For each window, 61 features have been extracted.

7.2.2.1 EDA

The preprocessing pipeline applied to the EDA signal is the same as reported in 6.3.2.

A total of 15 features have been extracted from the EDA:

- Whole EDA signal: *Fmax* [256], Symbolic Information Entropy (*SIE*) [249]
- Tonic component: mean, and standard deviation
- Statistical features: mean (*meanEDA*), standard deviation (*stdEDA*), median (*medianEDA*), interquartile range (*iqrEDA*), maximum value (*maxEDA*), minimum value (*minEDA*), range (*rangeEDA*), kurtosis (*ktsEDA*), skewness (*skewEDA*)

7.2.2.2 PPG

The preprocessing pipeline applied to the EDA signal is the same as reported in 6.3.2.

A total of 18 features have been extracted from the PPG:

- Heart Rate Variability (HRV) analysis: after having estimated the inter-beat intervals (IBIs) at the time difference between consecutive systolic feet, and checking for outliers applying the suggestion of Banhalmi et al. 2018 [248], the following HRV parameters have been extracted: *meanHR*, *SDNN*, *RMSSD*
- Basic morphological analysis: *PulseAmpl*, *A1*, *A2*, *A*, *T1*, *T2*
- Statistical features: mean (*meanPPG*), and standard deviation (*stdPPG*), median (*medianPPG*), interquartile range (*iqrPPG*), maximum value (*maxPPG*), minimum value (*minPPG*), range (*rangePPG*), kurtosis (*ktsPPG*), skewness (*skewPPG*)

7.2.2.3 EMG

EMG signal has been subjected to a 5^{th}-order Butterworth high-pass filter, with 0.5 Hz as the cut-off frequency. Nineteen features have been extracted from the EMG [257]:

- EMG-specific features: Integrated EMG (*IntEMG*), Mean Absolute Value (*MAV*), Simple Square Integrated (*SS_Int_EMG*), Root Mean Square (*RMS*), Variance (*Var*), Waveform length (*Waveform_length*), Different Mean Absolute Value (*Diff_MAV*), Second-order moment (*SecondOrd_mom*), Difference Variance Version (*Diff_Var*), Difference Absolute Standard Deviation (*Diff_Abs_SD*)
- Statistical features: mean (*meanEMG*), standard deviation (*stdEMG*), median (*medianEMG*), interquartile range (*iqrEMG*), maximum value (*maxEMG*),

minimum value (*minEMG*), range (*rangeEMG*), kurtosis (*ktsEMG*), skewness (*skewEMG*)

7.2.2.4 Respiration

The respiration signal was subjected to a 4th-order high-pass filter, with a cut-off frequency of 0.1 Hz. Nine features have been exctracted:

- Statistical features: mean (*meanResp*), standard deviation (*stdResp*), median (*medianResp*), interquartile range (*iqrResp*), maximum value (*maxResp*), minimum value (*minResp*), range (*rangeResp*), kurtosis (*ktsResp*), skewness (*skewResp*)

7.2.3 Feature selection

We applied the Maximum Relevance – Minimum Redundancy (MRMR) feature selection algorithm [258] to select a subset of features to optimize the classification process. We tested the performance of different classifiers by using the first 10, 20, and 30 features selected from the MRMR method.

7.2.4 Machine learning algorithms

The dataset has been divided into training (70%) and test (30%) sets. The division has been carried out on a subject basis via a pseudorandomization, to have a balanced number of HC subjects and cLPB patients in both training and test sets.

We trained and tested five different ML algorithms, whose hyperparameters are the following:

- Support Vector Machine (SVM) (Kernel, C, Gamma)
- Decision Tree (DT) (Criterion, Max depth)
- K- nearest neighborhood (k-NN) (Algorithm, Leaf size, # neighbors)
- Stochastic Gradient Descent (SGD) (Alpha, L1 ratio, Loss, Penalty)
- AdaBoost (Max depth, Min ample leaf, # estimators, Learning rate)

The training set has been used for hyperparameters optimization of the aforementioned ML algorithms using the GridSearch method with a 5-fold cross-validation, using the accuracy as a performance metric to be maximised.

Once the hyperparameters have been optimized for all time-window lengths and all the feature sets (10, 20, and 30 features), the ML algorithms have been tested on the test set.

As performance metrics, we used accuracy, sensitivity, and specificity. In Figure 7.2 the pipeline of the approach is shown.

The preprocessing, feature extraction and application of the feature selection algorithm have been carried out on Matlab 2022b [259], while ML algorithms training and testing have been carried out by using the Python scikit-learn package [260].

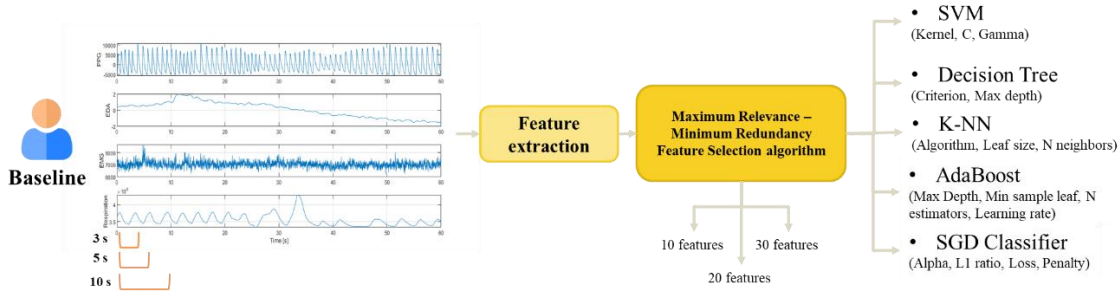


Figure 7.2 Approach to develop the binary classifier to distinguish between HC subjects and cLBP patients

7.3 Results

7.3.1 Sample data

A total of 38 recordings have been used, 25 coming from HC subjects and 13 from cLBP patients. In order to have the same proportion of recordings from HC subjects and cLBP patients in the training and test sets, we used recordings from 11 HC subjects and 6 cLBP patients as the training set, and 5 HC and 2 cLBP patients as the test set.

By dividing the baseline sessions into 3, 5, and 10 seconds window length, the number of instances obtained is reported in Table 7.1.

Table 7.1 Number of instances by dividing baseline sessions into 3, 5, and 10 seconds

	Total dataset			**Training set (70%)**			**Test set (30%)**		
	Tot	**HC**	**cLBP**	**Tot**	**HC**	**LBP**	**Tot**	**HC**	**cLBP**
3 sec	1230	819	411	898	577	321	332	242	90
5 sec	732	486	246	535	343	192	197	143	54
10 sec	359	238	121	264	169	95	95	69	26

7.3.2 Feature selection

By applying the MRMR feature selection method to the three datasets based on different window lengths and selecting the first 10, 20, and 30 features, we obtained the selected features reported in Table 7.2.

7.3.3 Classification

The optimized hyperparameters, obtained as output from the 5-fold cross-validation using the GridSearch method, are depicted in Table 7.3.

Figure 7.3 presents the values of validation accuracy, together with test accuracy, sensitivity, and specificity for the three different sets of features extracted by windows of A) 3 seconds, B) 5 seconds, and C) 10 seconds.

Table 7.2 Selected features extracted from different time window lengths

		EDA	**PPG**	**EMG**	**Resp**
3 sec	10 features	minEDA, Fmax	A2, PulseAmpl, A1, RMSSD,T2	skewEMG	maxResp, skewResp
	20 features	minEDA, Fmax, slopeEDA, maxEDA, meantonic, meanEDA, rangeEDA, medianEDA	A2, PulseAmpl, A1, RMSSD,T2, meanHR, A, maxPPG	skewEMG	maxResp, skewResp, iqrResp
	30 features	minEDA, Fmax, slopeEDA, maxEDA, meantonic, meanEDA, rangeEDA, medianEDA, stdtonic, iqrEDA, stdEDA	A2, PulseAmpl, A1, RMSSD,T2, meanHR, A, maxPPG, iqrPPG, minPPG	skewEMG, iqrEMG, ktsEMG, minEMG	maxResp, skewResp, iqrResp, rangeResp, medianResp
5 sec	10 features	meanEDA, slopetonic	meanHR, SDNN, minPPG, A1	ktsEMG	maxResp, ktsResp
	20 features	meanEDA, slopetonic, meantonic, maxEDA, minEDA, medianEDA	meanHR, SDNN, minPPG, A1, RMSSD, T2, maxPPG, A	ktsEMG	maxResp, ktsResp, skewResp, rangeResp
	30 features	meanEDA, slopetonic, meantonic, maxEDA, minEDA, medianEDA,	meanHR, SDNN, minPPG, A1, RMSSD, T2, maxPPG, A, rangePPG	ktsEMG, skewEMG, minEMG, Diff MAV, meanEMG	maxResp, ktsResp, skewResp, rangeResp, medianResp, stdResp

		rangeEDA, iqrEDA, skewEDA			
10 sec	10 features	meanEDA	SDNN, stdPPG, A1, A2, meanHR	iqrEMG, skewEMG	maxResp, skewResp
	20 features	meanEDA, rangeEDA, minEDA, maxEDA, meantonic	SDNN, stdPPG, A1, A2, meanHR, RMSSD, minPPG, A, T2	iqrEMG, skewEMG	maxResp, skewResp, medianResp, ktsResp
	30 features	meanEDA, rangeEDA, minEDA, maxEDA, meantonic, medianEDA, iqrEDA, skewEDA, stdtonic	SDNN, stdPPG, A1, A2, meanHR, RMSSD, minPPG, A, T2, iqrPPG, maxPPG, meanPPG	iqrEMG, skewEMG, ktsEMG	maxResp, skewResp, medianResp, ktsResp, rangeResp, stdResp

The best performance in terms of test accuracy has been achieved by the k-NN algorithm fed with 10 features extracted by segmenting the recordings into 10 seconds time windows (Accuracy = 77.89%, Sensitivity = 92.31%, Specificity = 72.46%).

Table 7.3 Optimized hyperparameters for the 5 ML algorithms, for different features set and extracted on 3, 5, and 10 seconds time windows

		Kernel			C			Gamma					
		3 sec	5 sec	10 sec	3 sec	5 sec	10 sec	3 sec	5 sec	10 sec			
SVM	10f	sigmoid	poly	rbf	1000	10	100	0.01	0.1	0.1			
	20f	rbf	poly	rbf	10	1	10	0.1	0.1	0.1			
	30f	rbf	rbf	rbf	10	10	1	0.1	0.1	0.1			
		Criterion			Max Depth								
		3 sec	5 sec	10 sec	3 sec	5 sec	10 sec						
DT	10f	Gini	Entropy	Entropy	150	150	6						
	20f	Entropy	Entropy	Entropy	8	150	8						
	30f	Entropy	Entropy	Entropy	5	12	11						
		Algorithm			Leaf size			N neighbors					
		3 sec	5 sec	10 sec	3 sec	5 sec	10 sec	3 sec	5 sec	10 sec			
KNN	10f	Ball tree	Ball tree	Ball tree	1	1	1	4	2	2			
	20f	Ball tree	Ball tree	Ball tree	1	1	1	2	2	2			
	30f	Ball tree	Ball tree	Ball tree	1	1	1	2	8	2			
		Alpha			L1 ratio			Loss			Penalty		
		3 sec	5 sec	10 sec	3 sec	5 sec	10 sec	3 sec	5 sec	10 sec	3 sec	5 sec	10 sec
SGD	10f	10^{-3}	10^{-4}	10^{-1}	0.15	0.07	0.1	Perceptron	log	Squared error	L2	L2	Elasticnet
	20f	10^{-3}	10^{-3}	10^{-4}	0.15	0.14	0.06	Perceptron	Hinge	Perceptron	L1	L1	Elasticnet
	30f	10^{-2}	10^{-2}	10^{-2}	0.05	0.15	0.06	Perceptron	Perceptron	Modified Huber	Elasticnet	Elasticnet	L1
		Max Depth			Min samples leaf			N estimators			Learning rate		
AdaBoost		3 sec	5 sec	10 sec	3 sec	5 sec	10 sec	3 sec	5 sec	10 sec	3 sec	5 sec	10 sec

10f	10	8	4	5	5	10	100	10	50	0.5	1	0.5
20f	4	4	4	10	5	10	1000	1000	100	0.5	1	0.5
30f	6	4	8	5	5	10	1000	250	50	0.1	1	1

SVM = Support Vector Machine. DT = Decision Tree. KNN = k-Nearest Neighbors. SGD = Stochastic Gradient Descent.

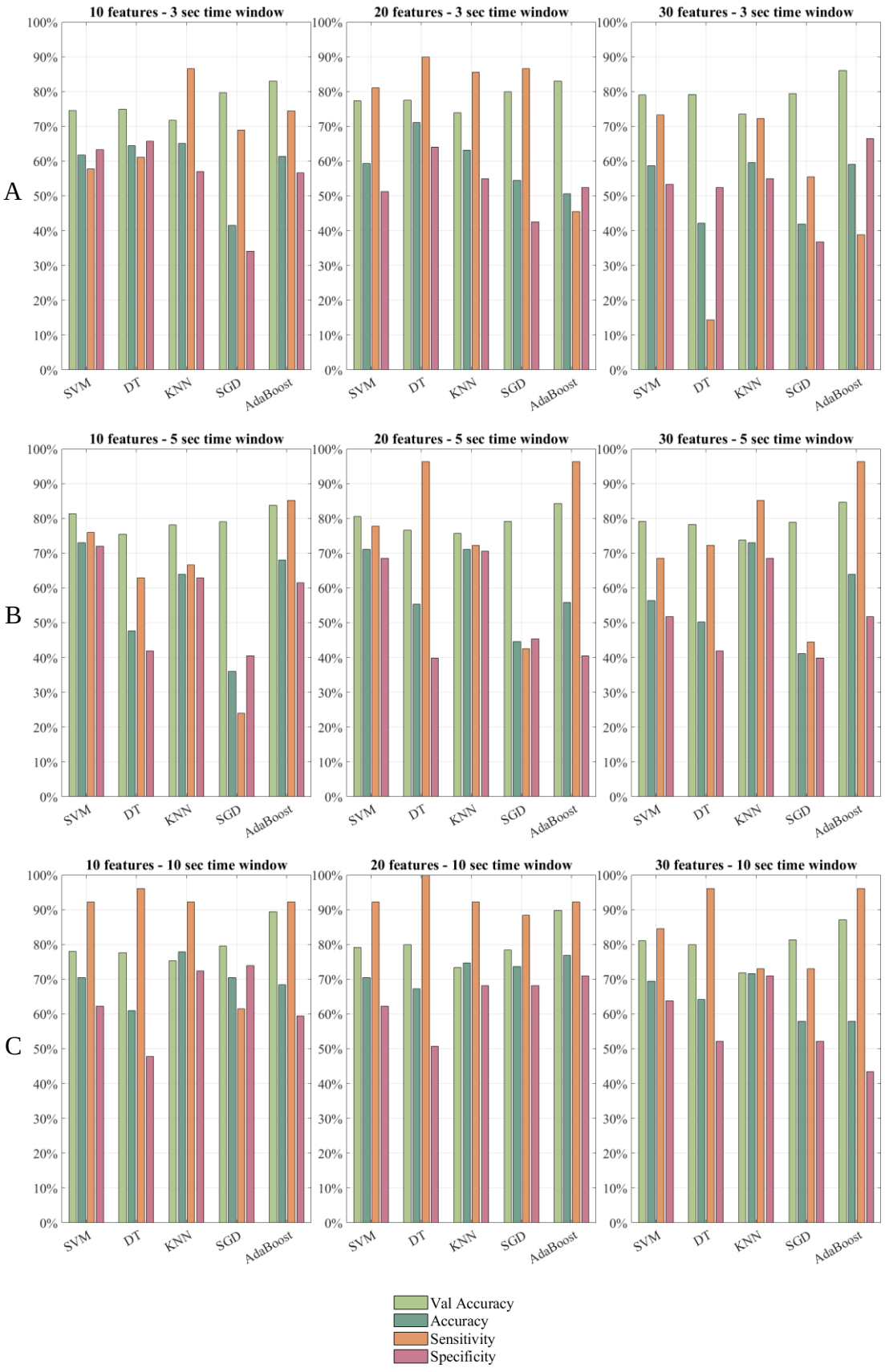


Figure 7.3 Classification performance for different features set extracted on 3, 5, and 10 seconds time windows

7.4 Discussion

This work used several ML algorithms as classifiers to discriminate between healthy control subjects and chronic low back pain patients using physiological signals recorded under resting conditions. Different configurations have been assessed in terms of the time window length to segment the physiological signals from which to extract parameter and in terms of the number of features to train and test the ML algorithm. We obtained the best performance to discriminate between the two populations by using a k-NN fed with 10 features extracted by segmenting the physiological signals into 10-second time windows.

We tried different time window lengths for two main reasons: on the one hand, in the literature, there is no standard for the time window in which to extract parameters; on the other hand, as the time window length gets shorter, a higher number of instances can be obtained, so the ML algorithms can likely be better trained. Based on the results we obtained, the best performance was achieved by using the 10-second time windows, that is also the length which provides the fewest instances. This can be interpreted as features extracted on a longer time window being more meaningful for this kind of classification than features extracted on shorter time windows.

Evaluating the performance of all classifiers, a common trend can be appreciated: higher sensitivity rather than specificity. Having more HC instances available than cLBP, we expected the opposite behavior. By performing a principal component analysis on the training set consisting of 10 features extracted over 10-second time windows (i.e., the best-performing dataset) and plotting the first two components, as shown in Figure 7.4, it can be seen that HC subjects have a wider distribution than cLBP patients. Therefore, the higher sensitivity than specificity can be attributed to the more homogeneous behavior of cLBP patients than HC subjects.

This work presents the same limitations reported in the previous chapter. The small sample size and the different age distributions in the two populations can hinder the generalizability of the results. Furthermore, the unbalance between HC subjects and cLBP patients in the dataset represents another weakness of this study.

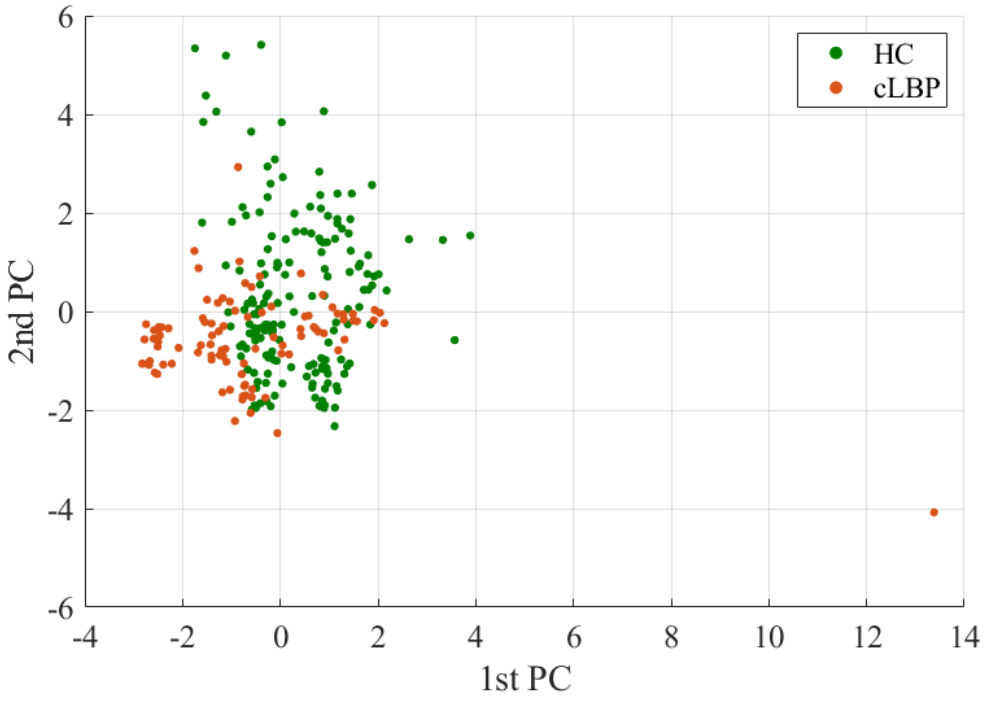


Figure 7.4 First and Second Principal Component Scores, divided for Healthy Controls (HC) and chronic Low Back Pain (cLBP) patients

SECTION IV: PAIN ASSESSMENT ON CANCER PATIENTS

8. PHYSIOLOGICAL RESPONSE TO PAIN IN CANCER PATIENTS

From the manuscript: Moscato S., Cortelli P., Chiari L., "Physiological responses to pain in cancer patients: a systematic review", Computer Methods and Programs in Biomedicine, 2022

8.1 Abstract

Pain is one of the most debilitating symptoms in persons with cancer. Still, its assessment is often neglected both by patients and healthcare professionals. There is increasing interest in conducting pain assessment and monitoring via physiological signals that promise to overcome the limitations of state-of-the-art pain assessment tools. This systematic review aims to evaluate existing experimental studies to identify the most promising methods and results for objectively quantifying cancer patients' pain experience.

Four electronic databases (Pubmed, Compendex, Scopus, Web of Science) were systematically searched for articles published up to October 2020.

Fourteen studies (528 participants) were included in the review. The selected studies analyzed seven physiological signals. Blood pressure and ECG were the most used

signals. Sixteen physiological parameters showed significant changes in association with pain. The studies were fairly consistent in stating that heart rate, the low-frequency to high-frequency component ratio (LF/HF) , and systolic blood pressure positively correlate with the pain.

Current evidence supports the hypothesis that physiological signals can help objectively quantify, at least in part, cancer patients' pain experience. While there is much more to be done to obtain a reliable pain assessment method, this review takes an essential first step by highlighting issues that should be taken into account in future research: use of a wearable device for pervasive recording in a real-world context, implementation of a big-data approach possibly supported by AI, including multiple stratification factors (e.g., cancer site and stage, source of pain, demographic and psychosocial data), and better-defined recording procedures. Improved methods and algorithms could then become valuable add-ons in taking charge of cancer patients.

8.2 Introduction

Cancer pain is an umbrella term that comprises many heterogeneous pain conditions with different physiological characteristics [261]. Pain can be due to the presence of the tumor itself, oncological treatments (e.g., chemotherapy, surgery, or immunotherapy) [81], or tissue damage [262].

The International Association for the Study of Pain (IASP) defined pain as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage" [263]. IASP also disclosed general guidelines for pain classification [264] that can be used for cancer pain evaluation. They are based on four significant features:

- pathophysiological mechanism: cancer pain can arise both as
- nociceptive pain that can be either
- somatic, the most frequent type of pain in the cancer population [82], or
- visceral, usually manifested after abdominal or thoracic surgery [265];
- neuropathic pain, with a prevalence of 20% in the cancer population [261];
- mixed pain, a combination of the two;
- duration: cancer patients usually suffer from chronic pain, which persists or recurs for more than three months [238], or breakthrough pain [266];
- etiology;

- anatomic location.

Regardless of its cause, pain is one of the most debilitating symptoms experienced by persons with cancer. On average, one-half of all cancer patients suffer from pain [267], and this percentage tends to become higher with the progression of the disease [268]. Furthermore, pain is detrimental to the psychological well-being of the subject. The reduced quality of life, in turn, reduces the adherence to therapy, which inevitably leads to adverse outcomes [269]. In addition to the personal and social impact, cancer pain also represents an economic burden [266]: the healthcare costs for oncological pain relief are almost five times higher than in the healthy population [270]. It should be added that, in this estimation, indirect costs related to patients' and caregivers' lower productivity are not taken into account [266] [271].

According to the American Pain Society and the European Task Force on Cancer Pain, an appropriate pain treatment starts with an appropriate pain assessment [272] [273]. Nowadays, in routine clinical practice, pain is assessed using scales and questionnaires to overview the pain experience. Scales are unidimensional ratings of pain intensity, while questionnaires give a more comprehensive evaluation as they keep track of different aspects of the pain experience, like the anatomical location and the time the pain is experienced [274]; some questionnaires are developed specifically for a particular type of cancer pain, as for the cancer-related neuropathic pain [262].

Although scales and questionnaires are state-of-the-art pain assessment tools (PAT), they suffer from several limitations. Since pain sensation is inherently subjective [275], the patient must be cooperative and non-cognitively impaired to communicate it [276]. Moreover, the memory of pain tends to be inaccurate and is often influenced by several context factors [277]. Specifically for cancer pain, patients tend to underrate their pain because it is supposed to be directly related to the worsening of the pathology [81]. Paradoxically, this could deteriorate the subject's health since pain acts as an alarm bell that alerts the body to take action to protect itself [278]. It can also happen that healthcare professionals leave out the pain assessment from their routine because they are more concerned about cancer diagnosis and treatment [82], so the evaluation is only carried out occasionally, usually in clinical settings.

Pain is a phenomenon of the utmost complexity that involves different physiological mechanisms at both the central and peripheral levels [279]. The conscious perception of pain is a result of higher brain center processes, collectively called the "Pain

Neuromatrix" [25], which modulate the pain sensation based on the subject's attention, affective dimension, and cognitive appraisal [51]. These processes, in turn, disrupt the ordinary functioning of some physiological mechanisms.

The physiological systems mainly affected by the pain experience are the Central Nervous System (CNS) and the Autonomic Nervous System (ANS). The CNS can be monitored using non-invasive electrophysiological or brain-imaging techniques to detect activated brain areas and the connection patterns established following the pain experience [280]. As a result, the brain's processes in response to a stimulus can be reconstructed [281]. On the other hand, the effects of ANS activation in response to pain can be monitored indirectly by measuring several physiological functions [282], collectively called autonomic signals. The ANS, composed of the sympathetic and parasympathetic systems, represents the bridge between the central nervous system and the internal organs. The ANS actions follow the "fight or flight" principle [283]: they occur involuntarily to preserve the integrity of the subject. Autonomic signals are often exploited in the research field of "Emotion recognition" [84], and objective pain assessment is one of its branches. One of the main advantages of exploiting autonomic signals is collecting them through wearable devices. With their relatively low-cost technology and progressive miniaturization [123], wearable sensors have become a valuable source of information about the health status both for healthy and diseased subjects, allowing continuous and unobtrusive monitoring even in real-world conditions [122]. In the last few decades, several research groups have focused on the link between pain and measurable physiological signals reflecting the disrupted mechanisms. Monitoring these signals could indeed provide additional tools for cancer pain assessment. Unlike scales and questionnaires, they would not require the subject's cooperation. Because physiological mechanisms are not affected by the subject's psychology, they also represent the pain status more objectively. Moreover, by using wearable devices, pain assessment could be carried out also outside the clinical context, when and where the pain is actually experienced. As an added benefit, healthcare professionals could dedicate the time of the visit to diagnosing and treating cancer instead of assessing pain.

On the other hand, such an approach imposes several challenges from a technical point of view. Physiological signals, especially those recorded by wearable devices, can be subjected to external noise and motion artifacts [121]. Thus, to overcome this issue, they must be subjected to a proper preprocessing step, in which signals are cleaned, and the

effects of possible artifacts are mitigated. Another critical step is represented by the feature extraction [284]: once the signals are ready to be processed, it is crucial to extract those features that can capture the phenomenon of interest (pain in this case). Added to this is the feature selection step, in which only the features that best describe the phenomenon are selected and then used [285]. Lastly, complex algorithms needed to understand the underlying relationship between features and pain since it is likely not to be simply linear [286]. In this case, it is possible to use either classical statistical methods or artificial intelligence (AI) algorithms. The former can be applied to appreciate differences in different painful conditions, and, consequently, they can be used to develop models that link a given painful condition to a precise physiological response. The latter use machine learning and deep learning algorithms [287] that automatically learn physiological responses patterns linked to a given painful condition.

Pain monitoring and assessment through physiological signals are still in an exploratory phase. To date, several studies have investigated aspects of the relationship between pain and physiological systems, nicely recapped by three recent reviews. The latest review by Chen et al. [118] summarizes the most common pain and stress assessment methods, followed by a synopsis of the main physiological signals that could be recorded through wearable devices. Next, the paper by Naranjo-Hernandez et al. [116] offers an overview of sensors that can potentially be used for chronic pain assessment, offering fascinating insights on the signals to be exploited and the relevant algorithms for their processing. Finally, the survey by Werner et al. [115] is more technical, including details about AI algorithms developed for automatic pain recognition through physiological signals. However, these reviews did not address the association between pain and changes in physiological signals specific to the cancer population and the feasibility of conducting these assessments in real-world settings.

For these reasons, we aim to conduct a systematic review of studies investigating the effect of pain on cancer patients' physiological signals. Our specific objectives are:

- To assess which physiological signals have been investigated in relation to cancer pain;
- To understand which statistical methods have been used to assess the association between cancer pain and physiological signals;
- To compare (whenever possible) the results of different studies;

- To evaluate the diffusion of instrumental pain assessment also in real-world settings

All the studies on this topic will be collected, and the link between physiological signals and cancer pain experience will be critically appraised.

8.3 Materials and Methods

8.3.1 Search strategy

We adopted the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines for the review protocol. PRISMA is "an evidence-based minimum set of items for reporting in systematic reviews" [288], [289].

The primary research question of the current review is: "What are the physiological responses to pain in cancer patients?".

The secondary questions are:

- Which physiological signals are currently monitored in this patient group?
- Which methods are used to investigate the association between subjective pain experience and physiological responses?
- Do different studies provide consistent evidence about the role of physiological signals in relation to pain?
- In real-world settings, what is the diffusion of studies investigating the physiological effect on cancer pain?

The following eligibility criteria have been defined using the SPIDER search tool [290]:

- Sample: Cancer patients
- Phenomenon of Interest: Pain experience
- Design: Pain assessment
- Evaluation: Recording of physiological signals
- Research type: Quantitative Research

A systematic literature search of PubMed, Compendex, Web Of Science, and Scopus databases was conducted to October 2020. We limited searches to 1990 onwards and included only studies published in English and Italian.

Based on the eligibility criteria, the search string was: ((Pain* OR Nocicept*) [Title] AND (Automat* OR Predict* OR Measur* OR Evaluat* OR Recognition OR Estimat* OR Classif* OR Assess* OR Examin* OR Detect* OR Effect) [Title + Abstract] AND

((Physiologic* OR Peripheral OR Autonom*) [Title + Abstract] AND (Signal OR Signals OR Parameter* OR Variable* OR Measure* OR Result OR Results OR Nervous System)) [Title + Abstract] AND (Cancer OR Oncolog*) [Title + Abstract].

8.3.2 Study selection

We used the following inclusion criteria:

- Cancer patients (including comparisons with a control group)
- Measure of physiological signals
- Pain assessment through scales/questionnaires, or information regarding the intensity of a nociceptive stimulus, or expected change due to an intervention (painful or therapeutic)

The following are the exclusion criteria:

- Animal experiments
- Scales or questionnaires only
- Case reports
- Assessment using biomolecules.

Duplicate publications were removed using Mendeley software [291]. In the first phase, two review authors (SM and LC) independently screened retrieved titles and abstracts and excluded irrelevant studies using Rayyan [292]. Disagreements were resolved by consensus. In the second phase, a reviewer (SM) searched the reference lists of the selected studies and other systematic reviews on similar topics to find additional papers.

8.3.3 Data extraction and quality assessment

The following information was extracted from each study:

- Year of publication
- Study objective
- Settings (clinical or real-world)
- Number of subjects and demographic data (age, gender)
- Cancer diagnosis
- Pain information and type of external pain stimulus (if any)
- Study type
- Pain ratings through scales or questionnaires, or intensity of the nociceptive stimulus, or pre-post intervention assessment

- Recorded physiological signals
- Recording procedure (e.g., duration, sampling frequency)
- Statistical methods (e.g., correlation, intergroup differences, before-after intervention) used to study the association between pain and physiological response

Based on the study design, included studies were divided into two categories:

- Concurrent validity studies, comparing state-of-the-art PAT and physiological signals;
- Sensitivity to change studies, evaluating physiological signals before and after an intervention (painful or antalgic).

Articles using the same physiological measures were clustered to investigate consistency across studies within these two categories.

To assess the quality of the studies, we selected two different tools for the two categories:

- Concurrent validity studies: Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool [293]. The physiological outcome resulting from the pain sensation is the new diagnostic tool to be assessed in terms of accuracy and reliability. At the same time, state-of-the-art PAT (i.e., scales, questionnaires) or the intensity of the nociceptive stimuli represent the ground truth: the term of comparison of the instrumental measurements. QUADAS-2 consists of thirteen questions related to four key domains:
    - patient selection: describe methods of patient selection and the included patients;
    - index test: describe the index test and how it was conducted and interpreted;
    - reference standard: describe the reference standard and how it was conducted and interpreted;
    - flow and timing: describe any patients who did not receive the index tests or reference standard; describe the interval and any interventions between index tests and the reference standard.
- Sensitivity to change studies: NIH Quality Assessment Tool (NIH-QAT) for before-after (pre-post) studies with no control group [294] . This tool consists of twelve signaling questions. We removed the Q12 (related to group interventions) since it is out of this systematic review scope.

For both tools, the risk of bias is assessed based on the answer to each signaling question (yes/no/info not available). We assigned an overall dichotomous risk of bias indicator based on the majority of answers, yes (low risk of bias) or no/not available (high risk of bias).

8.4 Results

8.4.1 Study selection

Searching the databases produced 1,181 records. Once duplicates were removed, 1,155 records were screened based on title and abstract, and 1,102 were excluded because they were not within the scope of this review. We assessed the full text of 53 studies and retained 13 of them that met the inclusion criteria [295]–[307]. One additional article (Buvanendran et al., 2010; [308]) was identified during the final manual search among the references of [304]. In total, 14 journal articles were included (see Figure 9.1).

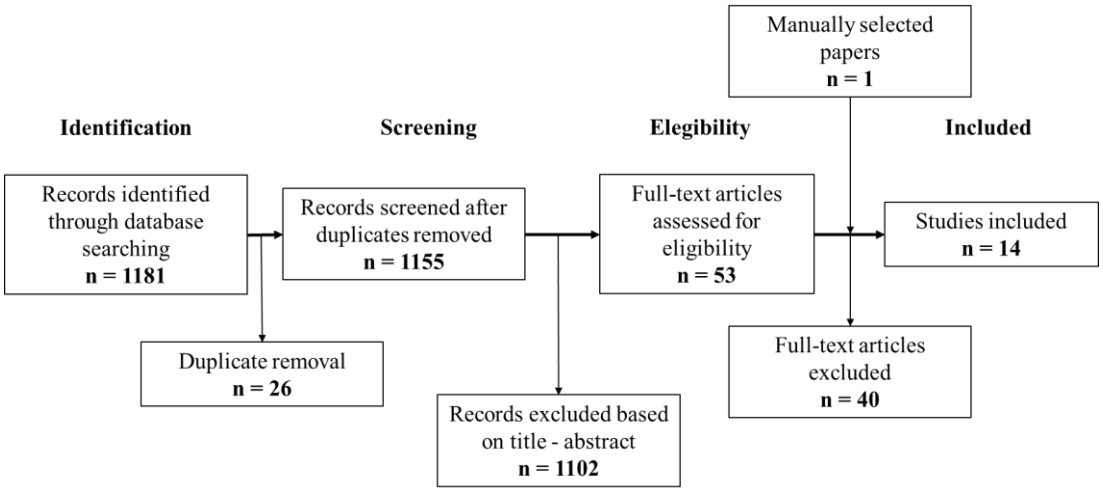
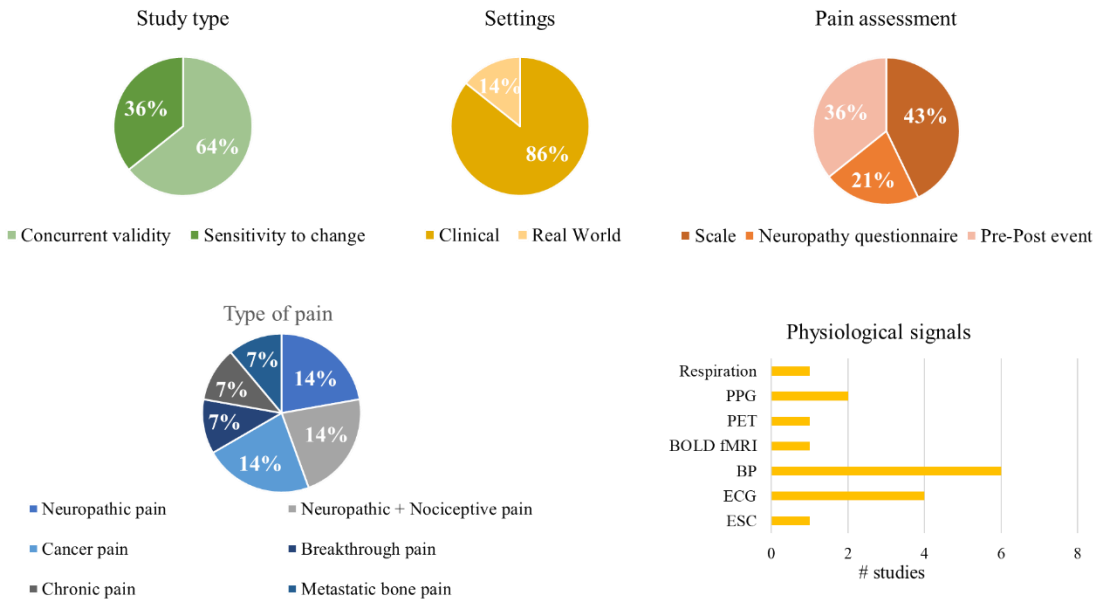


Figure 8.1 PRISMA flow diagram

8.4.2 Study characteristics

Descriptive characteristics of the fourteen studies are presented in Table 9.1. The included papers were published between 1993 and 2018. Three studies were based in the USA [299], [307], [308], two in Italy [297], [305], and one each in France [295], Turkey [296], Japan [300], South Korea [301], Austria [302], Poland [303], UK [304], Israel [306], and Lebanon [298]. Five hundred twenty-eight subjects were enrolled in the selected studies (172 males and 356 females, 516 oncological patients and 12 healthy volunteers), with an average of 38 subjects per study (range: 9-100) and a mean age of 53 years (range: 4-75

years). Three studies are focused exclusively on breast cancer (hence the high number of women), while eleven include different cancer sites (Figure 9.2).



With regard to the pain type, four studies enrolled patients suffering from neuropathic pain [295], [296], [304], [306], mostly due to chemotherapy treatment (3/16); five studies involved patients subjected to a nociceptive stimulus (surgery [300], [303], dental stimulation [297], and invasive procedures [298], [301]). Two studies looked at cancer pain in general [299], [305], and one each at breakthrough pain episodes [302], chronic pain [308], and metastatic bone pain [307].

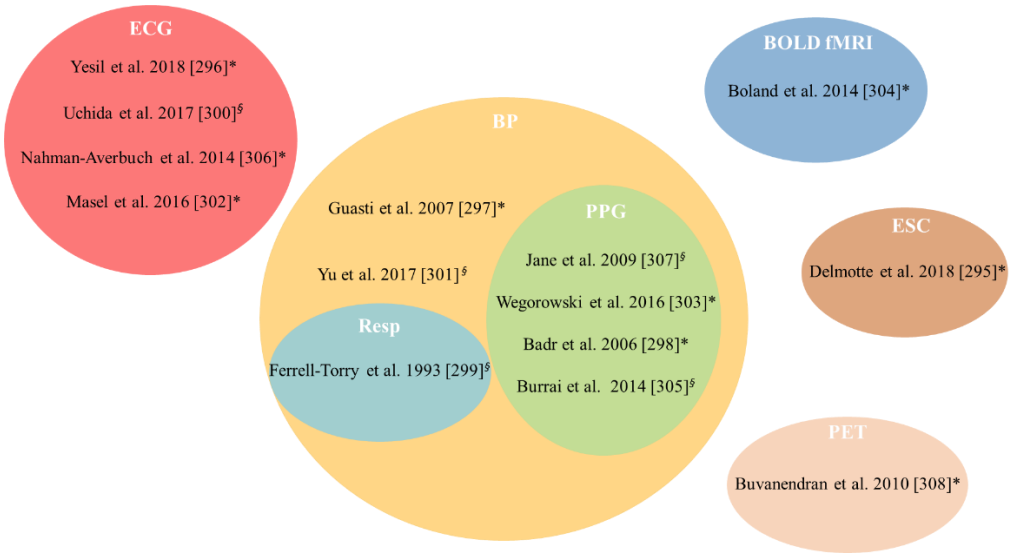
Concerning the experimental settings, twelve studies were carried out in clinical settings and only two in real-world conditions. Furthermore, only five studies reported information about the duration of the recordings, while none of the studies included information about the sampling frequency of the physiological signals.

Physiological signals under investigation in the selected studies and the relative extracted physiological parameters are the following:

- blood pressure (BP) in seven studies [297]–[299], [301], [303], [305], [307]
    - systolic blood pressure (sysBP): maximum blood pressure during ventricles contraction
    - diastolic blood pressure (diaBP): minimum blood pressure before the next contraction
- electrocardiogram (ECG) in four studies [296], [300], [302], [306]

- parameters from Heart Rate Variability (HRV) analysis: they are related to the time variations between consecutive heartbeats

- photoplethysmography (PPG) in two studies [298], [307]
  - parameters from HRV analysis
  - oxygen saturation (SpO2): the amount of oxygen carried by the hemoglobin in the blood



ECG: Electrocardiogram, BP: Blood Pressure, BOLD fMRI: Blood-Oxygenated-Level-Dependent functional Magnetic Resonance Imaging, ESC: Electrochemical Skin Conductance, PET: Positron Emission Tomography

Figure 8.2 Graphical representation of the salient features of selected studies: * concurrent validity studies; § sensitivity to change studies

- respiration (Resp) in one study [299]
  - respiration rate: the pace at which breathing occurs
- electrochemical skin conductance (ESC) in one study [295]
- positron emission tomography (PET) imaging in one study [308]
- blood-oxygen-level-dependent functional magnetic resonance imaging (BOLD-fMRI) in one study [304].

A brief description of the physiological parameters computed in the selected studies is presented in Appendix B, Table B.1.

8.4.3 Concurrent validity studies

Only two [296], [302] concurrent validity studies were carried out in a real-world context. In four studies [295], [296], [304], [308], patients were divided into two or more groups based on their pain experience. The most-used tools (five studies [298], [302], [303], [306], [308]) were scales such as the Numerical Rating Scale (NRS) (3/9), Visual Analogue Scale (VAS) (1/9), and FACES scale (1/9). For neuropathic pain evaluation, ad-hoc questionnaires were used (3/9) [295], [296], [304], as the Leeds Assessment of Neuropathic Symptoms and Signs and the Neuropathic Pain Symptom Inventory. A graphical depiction about the main features of the concurrent validity studies is given in Appendix B, Figure B.1.

In combination with scales and questionnaires, six physiological signals were exploited, four of which are autonomic signals: ECG (3/9), BP (3/9), PPG (2/9), ESC (1/9), BOLD-fMRI (1/9), and PET (1/9).

In eight out of nine studies, the features extracted from physiological signals were compared to pain assessment tools through correlation analysis, using Pearson's correlation coefficient [298], [302], [303], [306], Spearman's correlation coefficient [295], [297], both based on the distribution of the parameter [296] or a linear regression analysis [304], while five studies assessed the differences between two or more groups, using Wilcoxon test [295], [306], the Mann-Whitney U test [304], [308], or t-test and chi-square test for numerical and categorical variables respectively [296]. Only in one article did researchers investigate whether it is possible to use the extrapolated physiological parameters to classify patients by whether they had pain or not by analyzing the Receiver Operating Characteristic (ROC) curve [295].

A total of 19 different physiological parameters were assessed in association with pain; 14 were significantly related to pain in at least one study. The most-used physiological parameters are those extrapolated by HRV analysis and BP.

We further divided the concurrent validity studies by physiological signal type: monodimensional signal (i.e., time-series) and neuroimaging techniques. The main findings of the selected studies about physiological parameters are presented in Table 8.2.

8.4.3.1 Monodimensional signal

Four monodimensional signals were used, from which 17 physiological parameters were derived. Twelve of the parameters were statistically significantly associated with state-of-the-art PAT, both in terms of correlation and intergroup differences.

Four of the physiological parameters were used in more than one study: heart rate, LF/HF ratio, total power, systolic blood pressure. Specifically, heart rate showed to be higher in patients with neuropathic pain [296] and correlated with a state-of-the-art PAT [298], but in the other two studies, it was not significantly associated with pain scales [302], [303]. The LF/HF ratio, obtained as an output from HRV Analysis, was exploited in three different studies, resulting significantly higher in neuropathic patients and positively

Table 8.1 Descriptive characteristics of the included studies

Ref.	Study objective	Sett.	No subjects	Age mean (std) Gender	Cancer diagnosis (no.)	Information on pain	Study type	Pain assessment	Physiol. signals(s)	Rec duration	Stat. analyses
Delmotte et al. 2018 (France) [295]	To investigate how Electrochemical Skin Conductance (ESC) could be helpful in Oxaliplatin-Induced Peripheral Neuropathy (OIPN) diagnosis	Clinical	36	64 (11) – 18 F	Colon (12), Stomach (6), Liver (1), Pancreas (9), Rectum (7), Peritoneum (1). All with Oxaliplatin-Induced Peripheral Neuropathy	Neuropathic pain	Concurrent validity	Neuropathic Pain Symptom Inventory (NPIS)	ESC	-	Correlation, Inter-group differences, Classification
Yesil et al. 2018 (Turkey) [296]	To investigate whether neuropathic pain is associated with changes in cardiac sympathovagal activity in patients with breast cancer (BC)	Real World	70	48.2 (7.04) - 70 F	Breast cancer with chemotherapy-induced neuropathy	Neuropathic pain	Concurrent validity	Leeds Assessment of Neuropathic Symptoms and Signs (LANSS)	ECG	24 hours	Correlation, Inter-group differences
Uchida et al. 2017	To examine the effects of low-dose remifentanil on	Clinical	20	59 (7) exp. group, 60 (11)	Patients undergoing breast cancer surgery with pain (13),	Nociceptive pain (breast	Sensitivity to change	Before and after the surgery,	ECG	-	Pre-Post intervention, Inter-

(Japan) [300]	post-operative pain relief and heart rate variability after surgery			control group - 20 F	and without pain (7 age-matched)	cancer surgery)		group comparison			group differences
Yu et al. 2017 (South Korea) [301]	To test the effect of inhalation of lavender oil or linalyl acetate on pain relief	Clinical	66	60.9 - 29 F	Colorectal cancer (66)	Nociceptive pain (colorectal cancer surgery)	Sensitivity to change	Before and after antalgic therapy	BP	1 minute before and after therapy	Pre-Post intervention
Masel et al. 2016 (Austria) [302]	To study the changes in the ANS by measuring HRV during opioid therapy for cancer breakthrough pain (CBTP) in palliative-care patients with cancer and compare these changes with patient-reported pain levels on a NRS	Real World	10	62 (5.2) - 4 F	Advanced cancer (10)	Breakthrough pain	Concurrent validity	NRS	ECG	6 variable-length time windows before and after therapy	Correlation
Wegorowski et al. 2016 (Poland) [303]	To assess the possibilities of modifying the intensity of post-operative pain evaluated with VAS after	Clinical	100	58.58 (12.01) - 100 F	Breast cancer (100)	Nociceptive pain (breast cancer surgery)	Concurrent validity	VAS	BP	-	Correlation

	surgical 2treatment for breast neoplasm offered by pre-emptive analgesia										
Boland et al. 2014 (UK) [304]	To compare areas associated with central pain processing in patients with multiple myeloma who had chemotherapy-induced peripheral neuropathy with those from healthy volunteers	Clinical	24	58 (IQR:35-67) - 10 F	Cancer patients with Multiple myeloma chemotherapy-induced peripheral neuropathy (MM-CIPN) (12) – healthy controls (12)	Neuropathic pain + Nociceptive pain (heat pain stimulation)	Concurrent validity	Total Neuropathy Score reduced version (TNS-reduced) - MM-CIPN vs healthy volunteers	BOLD fMRI	-	Correlation, Inter-group differences
Burrai et al. 2014 (Italy) [305]	To test the differences in physiological parameters, pain-level, and mood-level between an experimental group subjected to live sax music and a control group who experienced only standard care	Clinical	52	64.3 (12.9) exp. group - 25 F, 64.6 (12.8) control group - 18 F	Metastatic cancer (45), non-metastatic cancer (7)	Cancer pain	Sensitivity to change	Before and after antalgic therapy	BP	-	Pre-Post intervention, Inter-group differences

Nahman-Averbuch et al. 2014 (Israel) [306]	To evaluate the relationships between autonomic parasympathetic function and the perception of (i) spontaneous pain, (ii) experimental non-painful sensations, (iii) painful experimental sensations in chemotherapy-induced neuropathy patients	Clinical	27	56.5 (7.9) - 20 F	Breast (11) , Lungs (2), Breast and Lungs (1), Ovary (2), Myeloma (2), Stomach (1), Oesophagus (1), Colon (3), Leukaemia (2), Hodgkin's Lymphoma (1), Sarcoma (1). All with peripheral neuropathy	Neuropathic pain + Nociceptive pain (heat pain stimulation)	Concurrent validity	NRS	ECG	5 minutes at rest, 1 minute for deep breathing test, 15 seconds for Valsalva manoeuvre	Correlation, Inter-group differences
Buvanendran et al. 2010 (USA) [308]	To determine the difference in brain activity in cancer patients with moderate to severe chronic pain versus no pain	Clinical	20	50.15 (19.79) - 17 F	Lymphoma (11), Breast (2), Lung (5), Pancreas (1), Esophageal (1)	Chronic pain	Concurrent validity	NRS	PET	17 minutes scan	Intergroup differences
Jane et al. 2009 (USA) [307]	To examine the effects of massage therapy	Clinical	30	51.7 (11.6) - 19 F	Lung cancer (11), breast cancer (11), head and neck (2), gastrointestinal (4), genitourinary (2)	Metastatic bone pain	Sensitivity to change	Before and after antalgic therapy	PPG, BP	-	Pre-Post therapy

Guasti et al. 2007 (Italy) [297]	To test the pain sensitivity in athyreotic patients followed for differentiated thyroid carcinomas during profound, short-term hypothyroidism induced for clinical reasons and during LT4-replacement treatment, focusing on the potential interferences of blood pressure-mediated changes in pain perception that may occur in the two clinical conditions.	Clinical	19	49 (15) - 14 F	Thyroid carcinoma (19)	Nociceptive pain (electrical stimulation)	Concurrent validity	Dental pain sensitivity	BP	-	Correlation
Badr et al. 2006 (Lebanon) [298]	To study the relationship between different indicators of pain, including self-reports, behavioral observations, and physiological	Clinical	45	4-10 (range) - 17 F	Leukemia (23), Solid Tumors (22)	Nociceptive pain (access of a subcutaneous central venous port)	Concurrent validity	FACES rating scale, DOLLS rating scale	PPG, BP	-	Correlation

	measures, in children with cancer undergoing invasive procedures										
Ferrell-Torry et al. 1993 (USA) [299]	To assess the effect of massage therapy on anxiety, relaxation, and the perception of pain in hospitalized cancer patients	Clinical	9	56.6 (range:23 -77) - 0 F	Esophageal (2), rectum (1), prostate (1), stomach (1), lung (1), lymphocytic leukemia (1), mixed nodular lymphoma (1), poorly differentiated cancer with an unknown primary site (1)	Cancer pain	Sensitivity to change	Before and after antalgic therapy	BP, Respiration	-	Pre-Post therapy

related to neuropathic scale only in one study [296]. In comparison, the results gathered by the other two studies were not statistically relevant [302], [306]. Two studies assessed the behavior of total power, another parameter obtained by the HRV Analysis, which showed to be significantly lower in patients with neuropathic pain [296] and not significantly correlated with pain scale [302]. Finally, the systolic blood pressure was assessed in three studies, showing a significant positive correlation with two different pain scales [298], [303].

8.4.3.2 Neuroimaging techniques

Two studies employed brain imaging techniques (BOLD fMRI and PET imaging) to assess pain. The information regarding the association between measured activity in specific brain areas and pain (i.e., positive correlation with pain, or higher in patients' group with pain) is presented in Table 8.2.

Table 8.2 Main findings of the selected studies

Concurrent validity studies								
Monodimensional signals								
Ref.	Pain assessment	Physiological signal(s)	Physiological parameter(s)	Main results	Correlation with pain	Intergroup differences		Significance level
						Pain	No pain	
Delmotte et al. 2018 [295]	Neuropathic Pain Symptom Inventory (NPIS)	ESC	Hands ESC	Significantly lower in the presence of painful neuropathy		55.4 (19.7) µS	77.6 (7.9) µS	p = 0.0003
				Significantly correlated with NPSI score	R = -0.69			p < 0.0001
			Feet ESC	Significantly lower in the presence of painful neuropathy		55 (15) µS	78.1 (6.6) µS	p < 0.0001
				Significantly correlated with NPSI score	R = -0.79			p < 0.0001
Yesil et al. 2018 [296]	Leeds Assessment of Neuropathic Symptoms and Signs (LANSS)	ECG	SDNN	Significantly higher in NP patients		116.44 (26.44) ms	141.21 (26.02) ms	p = 0.001
				Negatively related to LANSS score	R = -0.391			p < 0.01
			SDAAN	Significantly lower in NP patients		109.78 (26.04) ms	132.12 (26.89) ms	p = 0.003
				Negatively related to LANSS score	R = -0.36			p < 0.01
			SDNNindex	Significantly lower in NP patients		41.5 (9.51)	49.33 (10.41)	p = 0.007
				Negatively related to LANSS score	R = -0.278			p < 0.05
			Total power	Significantly lower in NP patients		1764 (795.61) ms^2	2455.25 (991.02) ms^2	p = 0.009
				Not related to LANSS score	Data not shown			
			LF/HF	Significantly higher in NP patients		4.68 (1.82)	3.1 (1.34)	p < 0.001
				Positively related to LANSS score	0.256			p < 0.05

			HR	Significantly higher in NP patients		86.83 (9.28) bpm	80.63 (7.61) bpm	$p < 0.05$
Masel et al. 2016 [302]	NRS	ECG	Log LF/HF	Non-significant reduction after treatment	Data not shown			
				Decreased in all patients who had a reduction in pain of > 2 points, but remained unchanged in patients who had reductions of up to 2 points	Data not shown			
			HF	Remained unchanged No significant correlation with NRS	Data not shown	Data not shown	Data not shown	
			LF/HF	Remained unchanged. No significant correlation with NRS	Data not shown	Data not shown	Data not shown	
			Total power	Remained unchanged. No significant correlation with NRS	Data not shown	Data not shown	Data not shown	
			pNN50	Remained unchanged. No significant correlation with NRS	Data not shown	Data not shown	Data not shown	
			HR	Remained unchanged. No significant correlation with NRS	Data not shown	Data not shown	Data not shown	
Wegorowski et al. 2016 [303]	VAS	PPG, BP	HR	No significant correlation with pain intensity	$R = 0.143$			$p = 0.157$
			SysBP	Positive correlation with pain intensity	$R = 0.386$			$p < 0.001$
			DiaBP	Positive correlation with pain intensity	$R = 0.446$			$p < 0.001$
Nahman-Averbuch et al. 2014 [306]	NRS	ECG	rMSSD	No difference between painful and non-painful-PNP groups		17.9 (range 3.3-24.8) ms	7.7 (range 4.5-26.6) ms	$p = 0.237$
				In non-painful-PNP group, lower rMSSD correlated with lower heat pain threshold. Not significant in painful-PNP group, significant correlation considering all the patients (with and without pain)	$R = 0.433$			$p = 0.05$

			LF/HF	No difference between painful and non-painful-PNP groups	3.2 (range 0.9-9.1)	4.4 (range 0.4-10.8)		p = 0.878
			Deep breathing ratio	No difference between painful and non-painful-PNP groups		1.12 (range 1.02-1.41)	1.14 (range 1.05-1.43)	p = 0.951
			Valsalva ratio	No difference between painful and non-painful-PNP groups		1.39 (range1.04-1.97)	1.42 (range 1.17-1.74)	p = 0.972
				Lower Valsalva ratio correlated with a lower level of pain change value	R = -0.495			p = 0.023
				Negative correlation with average pain ratings to the "test stand-alone" stimulus	R = -0.559			p = 0.008
Guasti et al. 2007 [297]	Dental pain sensitivity	BP	SysBP	No association between blood pressure changes and pain sensitivity variations	Data not shown			
Badr et al. 2006 [298]	FACES rating scale, DOLLS rating scale	PPG, BP	HR	Significant correlations with 3 time points for FACES	R = 0.82, 0.71, 0.85			p < 0.01
				Significant correlations with 3 time points for DOLLS	R = 0.78, 0.97, 0.76			p < 0.001
			SysBP	Significant correlations with 3 time points for FACES	R = 0.59, 0.78, 0.91			p < 0.001
				Significant correlations with 3 time points for DOLLS	R = 0.75, 0.81, 0.79			p < 0.001
			SpO2	Not correlated to either the FACES or DOLLS scores	Data not shown			
Neuroimaging techniques								
Ref.				Main results	Correlation with pain	Intergroup differences		

	Pain assessment	Physiological signal(s)	Physiological parameter(s)			Pain	No pain	Significance level
Boland et al. 2014 [304]	Total Neuropathy Score reduced version (TNS-reduced) - Patients vs healthy subjects	BOLD fMRI	BOLD response	Heat-pain stimulation evoked a BOLD response in healthy volunteers and patients		=		p < 0.001
				Patients demonstrated significantly less activation in R superior frontal gyrus		-	+	p = 0.03
				Patients demonstrated significantly greater activation in L precuneus		+	-	p = 0.01
				Significant correlation of BOLD response with TNS-reduced version in the left operculo-insular cortex	+			p = 0.03
Buvanendran et al. 2010 [308]	NRS	PET	Brain activation	Patients with moderate-to-severe pain had increased regional glucose metabolism bilaterally in the prefrontal cortex		+	-	z-score > 3
				Unilateral activation was found in the right parietal precuneus cortex.		+	-	z-score > 3
				No areas of the brain showed decreased activity due to moderate-to-severe pain.	=			-

Sensitivity to change studies

Ref.	Intervention	Physiological signal(s)	Physiological parameter(s)	Main results	Longitudinal differences		Intergroup differences		Significance level
					Before intervention	After intervention	Experimental Group	Control Group	
Uchida et al. 2017 [300]	Antalgic therapy	ECG	HF	Significant increase before and after antalgic therapy for the experimental group	35.6 (14.3) ms^2	49.4 (3.0) ms^2			p = 0.01

				No statistical differences between groups			36.1 (9.0) ms^2	35.6 (14.3) ms^2	p = 0.933
			LF/HF	No statistical differences before and after antalgic therapy for the experimental group	2.3 (1.4)	1.6 (1.4)			p = 0.104
				No statistical differences between groups			1.9 (0.8)	2.3 (1.4)	p = 0.476
Yu et al. 2017 [301]	Antalgic therapy	BP	SysBP	Reduction after the antalgic therapy, not significant	Data not shown	Data not shown			
			DiaBP	Reduction after the antalgic therapy, not significant	Data not shown	Data not shown			
			HR	Reduction after the antalgic therapy, not significant	Data not shown	Data not shown			
Burrai et al. 2014 [305]	Relaxation therapy	BP, PPG	SysBP	No statistical differences before and after therapy for the experimental group	100 (80-160) mmHg	110 (80-130) mmHg			p = 0.644
				No statistical differences before and after therapy for the control group	100 (70-130) mmHg	100 (80-130) mmHg			p = 0.139
				No statistical differences between groups			110 (80-130) mmHg	100 (80-130) mmHg	p = 0.253
			DiaBP	No statistical differences before and after therapy for the experimental group	70 (50-110) mmHg	70 (60-90) mmHg			p = 0.868
				No statistical differences before and after therapy for the control group	70 (50-80) mmHg	70 (60-80) mmHg			p = 0.120
				No statistical differences between groups			70 (60-90) mmHg	70 (60-80) mmHg	p = 0.223
			HR	No statistical differences before and after therapy for the experimental group	74 (56-84) bpm	74.5 (50-104) bpm			p = 0.672

				Significant differences before and after therapy for the control group	74.5 (50-104) bpm	74 (55-98) bpm			p = 0.018
				No statistical differences between groups			74.5 (50-104) bpm	74 (55-98) bpm	p = 0.486
			SpO2	No statistical differences before and after therapy for the experimental group	98 (94-100) %	99 (94-100) %			p = 0.192
				Significant differences before and after therapy for the control group	97 (94-100) %	97 (91-100) %			p = 0.319
				Significant differences between groups			99 (94-100) %	97 (91-100) %	p = 0.003
Jane et al. 2009 [307]	Relaxation therapy	PPG, BP	HR	No statistical differences before and after therapy	83.7 (17.2) bpm	82.9 (15.1) bpm			p = 0.35
			MAP	No statistical differences before and after therapy	88.8 (14.2) mmHg	90.1 (14.5) mmHg			p = 0.26
Ferrell-Torry et al. 1993 [299]	Relaxation therapy	BP, Respiration	HR	No significant differences before and after therapy	80.4 (16.5) bpm	77.2 (17.3) bpm			p > 0.05
				Significant decrease before and 10-minutes after therapy	80.4 (16.5) bpm	75.9 (16.3) bpm			p < 0.05
			RR	Significant decrease before and after therapy	22.6 (2.2) breaths/min	19.7 (2.5) breaths/min			p < 0.05
				Significant decrease before and 10-minutes after therapy	22.6 (2.2) breaths/min	19.8 (2.3) breaths/min			p < 0.05
			SysBP	Significant decrease before and after therapy	120.9 (14.7) mmHg	114.7 (16.8) mmHg			p < 0.05
				No significant decrease before and 10-minutes after therapy	120.9 (14.7) mmHg	115.1 (15.1) mmHg			p > 0.05

			DiaBP	Significant decrease before and after therapy	74.9 (8.6) mmHg	69.1 (7.0) mmHg			$p < 0.05$
				No significant decrease before and 10-minutes after therapy	74.9 (8.6) mmHg	73.1 (7.2) mmHg			$p > 0.05$
			MAP	Significant decrease before and after therapy	90.2 (7.0) mmHg	84.3 (6.7) mmHg			$p < 0.05$
				No significant decrease before and 10-minutes after therapy	90.2 (7.0) mmHg	87.1 (5.1) mmHg			$p > 0.05$

NRS = numerical rating scale, VAS = visual analogue scale, ESC = electrochemical skin conductance, ECG = electrocardiography, HR = heart rate, SDNN = standard deviation of NN intervals, SDAAN = standard deviation of the mean of NN intervals, rMSSD = root mean square of successive differences, pNN50 = percentage of successive NN intervals that differ from one another by > 50 milliseconds, LF = low frequency, HF = high frequency, LF/HF = low frequency/high frequency ratio, SpO2 = oxygen saturation, BP = blood pressure, SysBP = systolic blood pressure, DiaBP = diastolic blood pressure, MAP = mean arterial pressure, RR = respiration rate, BOLD fMRI = blood-oxygen level dependent functional magnetic resonance imaging, PET = positron emission tomography

8.4.3.3 Study quality

The risk of bias is reported in Table 9.3. Seven out of nine studies in this category present an overall low risk of bias [296], [297], [302]–[304], [306], [308]. Both studies with an overall high risk of bias [295], [298] lack information concerning the patients' selection procedures, the index test [295], and the reference standard [298]. Due to lack of information, the risk of bias about index test and flow and timing remains unclear for most studies.

8.4.4 Sensitivity to change studies

All five studies in the sensitivity-to-change group were carried out in a clinical setting. Four of them assessed the behavior of the physiological parameters before and after a therapy [299], [301], [305], [307]. One of them evaluated differences before and after a painful intervention (i.e., surgery) [309]. BP was used in four out of five works, ranking first among physiological signals, followed by PPG (2/5), ECG (1/5), and Resp (1/5). In two studies, participants were divided into an experimental and a control group. Thus the intergroup differences could be assessed, besides the differences before and after the intervention. Pre-post differences were assessed by using t-test [300], [307], analysis of variance (ANOVA) [299], [301], or Wilcoxon test [305], while inter-group differences were assessed by using t-test [300] or Mann-Whitney U test [305]. The main findings of the selected sensitivity-to-change studies are reported in Table 2. A graphical depiction about the main features of the sensitivity to change studies is given in Appendix B, Figure B.2.

All the sensitivity-to-change studies exploited monodimensional signals. A total of 8 different physiological parameters were assessed: six of them were found to differ significantly before and after the intervention in at least one study. The most used parameters were heart rate and blood pressure-related properties (systolic, diastolic, or mean arterial pressure), both exploited in four studies.

Four physiological parameters were used in more than one study: heart rate, systolic blood pressure, diastolic blood pressure, and mean arterial pressure. Specifically, heart rate was used in four studies, showing a common trend to decrease after an antalgic therapy in two of them [299], [305] and no statistically significant change in the other two studies [301], [307]. Parameters from blood pressure were extensively used in different studies. Results on systolic blood pressure show a reduction of different intensities after an antalgic

Table 8.3 Risk of bias assessment

	QUADAS-2 – Concurrent validity studies													
	Patients selection				**Index test**			**Ref. Standard**			**Flow and Timing**			**Risk of bias**
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q11**	**Q12**	**Q13**	
Delmotte et al. 2018 [295]	NA	N	NA	Y	N	NA	Y	Y	Y	Y	NA	Y	NA	High
Yesil et al. 2018 [296]	NA	N	Y	Y	N	NA	Y	Y	Y	Y	NA	Y	Y	Low
Masel et al. 2016 [302]	Y	Y	Y	Y	NA	Y	Y	Y	Y	Y	Y	Y	NA	Low
Wegorowski et al. 2016 [303]	Y	Y	Y	Y	NA	Y	Y	Y	Y	Y	Y	Y	NA	Low
Nahman-Averbuch et al. 2014 [306]	NA	N	Y	Y	Y	NA	Y	Y	Y	Y	NA	Y	NA	Low
Boland et al. 2014 [304]	N	N	Y	Y	N	NA	Y	Y	Y	Y	Y	Y	Y	Low
Buvanendran et al. 2010 [308]	Y	Y	Y	Y	N	NA	Y	Y	Y	Y	NA	Y	Y	Low
Guasti et al. 2007 [297]	Y	Y	Y	Y	N	NA	Y	Y	Y	Y	Y	Y	Y	Low
Badr et al. 2006 [288]	NA	N	NA	Y	Y	NA	Y	NA	NA	NA	NA	Y	N	High

	NIH QAT – Sensitivity to change studies											
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q11**	**Risk of bias**
Uchida et al. 2017 [300]	Y	N	Y	N	Y	Y	Y	NA	Y	Y	N	Low
Yu et al. 2017 [301]	Y	N	Y	Y	Y	Y	Y	NA	Y	Y	N	Low
Burrai et al. 2014 [305]	Y	Y	Y	N	Y	Y	Y	NA	Y	Y	Y	Low
Jane et al. 2009 [307]	Y	Y	NA	N	NA	N	Y	NA	Y	Y	Y	High
Ferrell-Torry et al. 1993 [299]	Y	N	N	Y	N	Y	Y	NA	Y	Y	Y	Low

therapy in two studies [299], [305], and no significant results in another one [301]. Diastolic blood pressure was found to decrease after antalgic therapy in only one [52] out of three studies [54], [58] significantly. Mean arterial pressure was used in two studies, decreasing significantly after an antalgic therapy in one study [299], while it did not significantly change in the other study [307].

8.4.4.1 Study quality

The risk of bias is reported in Table 3. Four out of five studies [299]–[301], [305] present a low risk of bias. The overall high risk of bias for [307] is mainly due to patients' selection criteria (Q3, Q4, Q5). For all the studies, no information regarding the blindness of the examiner (Q8) was reported. All studies clearly defined the objective (Q1) and the outcome measures (Q7), had a loss of follow-up less than 20% (Q9), and the statistical methods were applied correctly to assess differences before and after the intervention, providing the related p-values (Q10).

8.5 Discussion

The primary purpose of this systematic review is to clarify the effects of pain on cancer patients' physiological signals. To do so, we investigated which signals are currently used, their concurrent validity with routine pain assessment tools, their sensitivity to change in pain levels, and the diffusion of instrumental pain assessment in real-world settings.

A majority of selected studies assessed the concurrent validity of physiological parameters against scales and/or questionnaires, while other studies evaluated the sensitivity-to-change of the physiological parameters to an intervention. The two categories of studies present consistent primary objectives: concurrent validity studies mainly aimed to assess the behavior of the physiological signals in relation to the patients' painful state, either via comparisons with state-of-the-art PAT or clustering subjects based on the different levels of pain. Sensitivity to change studies all focused on finding evidence through physiological signals of the efficacy of antalgic therapies.

8.5.1 Monodimensional signals

Most of the included studies use monodimensional signals to quantify some aspects of ANS activation. In concurrent validity studies, only four physiological parameters are present in more than one selected study. Despite the small sample size, both in terms of

the number of studies and participants, and the lack of methodological information, it was still possible to find consistent results among these studies. When limiting to statistically significant results, heart rate, LF/HF ratio, and systolic blood pressure values consistently correlated with pain positively: the higher the self-reported pain, the higher the parameter's value. The same association was also found in sensitivity to change studies, which showed a decrease in heart rate and systolic and diastolic blood pressure after an antalgic therapy or an experimental relaxation procedure. Higher values of these parameters are attributable to a more active sympathetic than parasympathetic branch of the ANS *[85]*, which is in line with what is expected during a pain experience. However, one should keep in mind that more complex interactions may play a role [286]: pain sensation triggers a complex net of neurological paths, which are not always activated linearly (if the relationship were linear, the higher the pain perception, the greater the activation of the monitored physiological function would be). It follows that a correlation analysis can only partially explain the relationship between pain and physiological signals. The assessment of non-linear relationships and more complex models involving different physiological signals might help address this problem. Indeed, an increasing number of studies in the emotion recognition field use artificial intelligence algorithms [310]–[312] because they are capable of extrapolating complex relationships between several inputs (physiological parameters) and a single output (pain perception).

Within the monodimensional signals set, a noticeable application is represented by the Electrochemical Skin Conductance, exploited only in one study [295], aiming to classify cancer patients with painful neuropathy. The Electrodermal Activity is a quantitative measure of the sympathetic nervous system [91], and it is widely used in the emotion recognition field, particularly in automatic pain recognition algorithms [313]–[315]. Future studies could involve using such signal, which proved to be well suited and can also be recorded by means of wearable devices.

8.5.2 Heart Rate and Heart Rate Variability

Heart rate is the most-used physiological parameter linked to pain assessment. The reference gold standard is the ECG signal, although recent research has focused more on heart rate estimated from the PPG signal. Its better convenience and pervasiveness justify this choice. Heart rate and the derived HRV parameters [85] extrapolated by the PPG are currently used in stress detection algorithms [166], [316]. However, there are some

concerns regarding the reliability of the PPG signals, which can be easily corrupted by external noise and motion artifacts leading to inaccurate HRV estimates [138].
HRV analysis is well-suited for real-world applications: Holter ECG is a long-established technique in routine clinical practice, while PPG can be easily embedded into a wearable device (e.g., a smartwatch or ring). Indeed, the only two studies run in a free-living context are based on HRV analysis using a 24-hours Holter ECG device. Notably, the work of Masel et al. [302] showed the possibility of detecting pain episodes without active cooperation from patients. This result clearly highlights the possible disruptive role of automatic pain assessment in real-world settings: a tool to detect pain timely, even in unconscious patients, and, in turn, provide antalgic therapy at pain onset.

8.5.3 Neuroimaging techniques

Relevant information can also be deduced by neuroimaging techniques, giving the possibility to explore brain areas involved in pain perception. Such methods are valuable for research purposes since pain-activated CNS processes are still not fully understood. On the other hand, they cannot represent an alternative cancer pain assessment solution because of their bulky instrumentations and expensive procedures.

9.5.4 Study quality

The study quality assessment revealed a low risk of bias for eleven out of fourteen studies. We considered it appropriate to select two study quality tools, QADAS-2 and NIH-QAT, to assess the different sources of risk of bias for the two different study designs. QADAS-2 proved to be well suited for highlighting the primary sources of bias for the concurrent validity studies: it is worth noting the significant lack of information, especially for index test (i.e., physiological parameter) and flow and timing sections, for which the risk of bias remains unclear. For sensitivity to change studies, we chose NIH-QAT, conceived as a tool for before-after (pre-post) studies with no control group. Although some studies in this category were presented as randomized control trials, we were solely interested in physiological signals changes before and after an intervention. As a result, sensitivity to change studies proved to be less prone to bias than concurrent validity studies, except that there is no information available for any study on the blindness of the examiners (Q8).

8.5.5 Limitations

The studies included in this review displayed a marked heterogeneity in the type of cancer and pain source to be evaluated. Such heterogeneity is a direct consequence of the broad range of painful conditions grouped under the umbrella term “cancer pain” that includes all the painful conditions related to cancer, regardless of the primary cancer sites and the painful stimulus. More so, concurrent validity studies highlight the wide range of the currently available state-of-the-art PAT (e.g., NRS, VAS, neuropathy-specific questionnaires). Altogether, fragmentation and lack of accepted guidelines in pain assessment represent another hint that there is the need to promote and standardize this very delicate aspect of cancer patient management, as also highlighted in [317].

Available literature also shows a considerable lack of methodological information regarding the experimental procedures and measurement setups that partly prevent the replicability of studies. Even if physiological parameters are clearly time-dependent and often non-stationary, only five studies out of fourteen report the recording duration, and no one specifies the sampling frequency of the collected physiological signals. This underreporting represents a significant limitation and prevents accurate comparisons or meta-analyses for those studies based on time-dependent variables (e.g., parameters from HRV analysis) since the same parameter can assume a different meaning in different time frames [59].

A limitation in using physiological signals to assess pain is that physiological mechanisms can be also affected by personal factors, like gender [318], age [319], and health status (which is particularly true in cancer). In some cases, cancer pathology itself can lead to a change in physiological mechanisms, which can be misinterpreted and related to the pain experience [320]. This limitation should be overcome, in the future, by analyzing larger patient cohorts.

8.5.6 Future directions

Cancer pain is a remarkably complex and multidimensional phenomenon that impacts the patients' psychological, social, and spiritual well-being [321]. As highlighted in the European Society for Medical Oncology position paper [322], a patient-centered approach is needed for cancer treatment, and this approach should also be translated to pain assessment. To reach this goal, implementing a biopsychosocial model [320] for pain assessment could overcome the limitations imposed by the current tools, providing a

complete picture of the pain state that considers all the different aspects that converge in the pain experience.

The challenge of assessing cancer patients' pain in a free-living context is relevant but still largely unaddressed. Several smartphone apps have been developed to date that can provide pain management for cancer patients. Most are based on self-rated pain assessments [323], [324] , but a brand-new app— whose feasibility and acceptability are being assessed in a pilot study—exploits the smartphone hardware to record some physiological parameters like heart rate and activity level [325].

In this scenario, wearable and mobile technologies can represent a game-changer, offering a valuable source of novel information. Pervasive monitoring systems allow the collection of long-term multimodal physiological recordings in a real-world context, giving the possibility to extrapolate information that could not be otherwise obtained in clinical settings. Such an approach is perfectly suitable for those clinical trials that involve interventions in free-living scenarios, in which participants can freely conduct their daily activities while their physiological functions are being monitored. Moreover, such an approach allows recording the natural physiological response to pain, unlike those studies that analyze the acute reaction when an external nociceptive stimulus is applied.

In order to efficiently elaborate the information gathered by wearable sensors, a proper approach should be applied, and AI algorithms could offer a viable solution. AI algorithms can indeed help identify complex patterns in long-term multimodal physiological recordings [287]. These are the prominent aspects that should be considered before these methods can be translated into a tool to tailor and personalize the antalgic interventions:

- Big data approach: AI algorithms reach good performances if they are trained on large datasets. Thus, when setting a study protocol, it is necessary to collect a reasonable number of instances that will be used to train the algorithms to avoid the curse of dimensionality [326].
- The individual variability of the physiological response: we have a limited mechanistic understanding of interindividual differences in pain and analgesia response. This issue can be considered, for example, by conducting a leave one subject out cross-validation to train the AI algorithms in order to correctly manage the inter-subject variability [327].

- Confounding factors: physiological signals can be affected by several other factors in addition to pain. These factors can be accounted for by enriching the AI algorithms with information related, for example, to personal (i.e., age, gender, weight, height) or health data (i.e., pathology, depression, and anxiety levels). Based on the specific type of pain being investigated, researchers should then collect other information besides pain, including patient-reported outcomes, that can help in better understanding the physiological response linked to the experienced pain [328].

We are currently working on the design of innovative clinical trials, carried out in residential facilities by monitoring patients in their free-living, and thus collecting the pain response when and where it is experienced.

8.6 Conclusions

This systematic review collected and pooled the knowledge regarding the behavior of physiological parameters in response to cancer patients' pain. Although the included studies were characterized by considerable heterogeneity, it was still possible to identify promising results relevant to develop new pain assessment tools for cancer patients based on physiological signals and, possibly, wearable sensors, paving the way to real-world scenarios.

9. AUTOMATIC PAIN ASSESSMENT ON CANCER PATIENTS

From the conference proceeding: Moscato S., Orlandi S., Giannelli A., Ostan R., Chiari L., “Automatic pain assessment on cancer patients”, IEEE Engineering in Medicine and Biology Conference, 2022

9.1 Abstract

Pain assessment represents the first fundamental stage for proper pain management, but currently, methods applied in clinical practice often lack in providing a satisfying characterization of the pain experience. Automatic methods based on the analysis of physiological signals (e.g., photoplethysmography, electrodermal activity) promise to overcome these limitations, also providing the possibility to record these signals through wearable devices, thus capturing the physiological response in everyday life. After applying pre-processing, feature extraction and feature selection methods, we tested several machine learning algorithms to develop an automatic classifier fed with physiological signals recorded in real-world contexts and pain ratings from 21 cancer patients. The best algorithm achieved up to 72% accuracy. Although performance can be improved by enlarging the dataset, preliminary results proved the feasibility of assessing pain by using physiological signals recorded in real-world contexts.

9.2 Introduction

Pain is an extremely complex phenomenon that involves both the sensory and emotional systems in association with actual or potential tissue damage [263]. Pain is a subjective experience and it is affected by cognition, attention, and affective dimension [51]. Therefore, current pain assessment methods used in clinical practice (i.e., scales and questionnaires [279]) often fail to provide a complete picture of the pain experienced by the patient, hence being detrimental for pain treatment.

Recently, there has been growing attention in the research of "emotion recognition", the discipline that gathers all the studies aiming at detecting emotions from humans via automatic technological approaches [329]. The rationale behind this field of application is that pain and emotion elicit a response in the autonomic nervous system, disrupting the ordinary functioning of some physiological mechanisms [282]. Therefore, monitoring the behavior of these physiological mechanisms, by recording and processing signals (e.g., photoplethysmography, electrodermal activity) when pain is experienced could represent a valuable source of information for improving pain assessment. In this scenario, wearable devices appear to be the natural tools to be used, since they allow monitoring physiological signals in a continuous and pervasive manner and recording the physiological reaction to pain when and where it is actually experienced [123].

Two studies aimed at developing automatic methods for pain assessment based on physiological signals recorded through wearable devices [330], [331]. The work by Johnson et al. [330] showed the feasibility of collecting physiological signals with wearable devices in patients with sickle cell diseases in a hospital setting and exploiting them to automatically predict pain intensity, based on pain scores given by patients. Badura and colleagues [331] demonstrated that the same approach can be applied also in physiotherapy settings, monitoring patients and asking them to rate their pain during a session of fascial therapy. Although these studies used wearable devices, physiological signals were recorded in controlled environments, in which the patient was not free to perform his/her daily life actions.

The aim of our study is to develop an automatic pain assessment method bansed on physiological signals recorded in a real-world context by means of wearable devices.

9.3 Materials and Methods

Look of Life is a project whose aim is to evaluate the effects of virtual reality (VR) as a non-pharmacological intervention to improve cancer-related symptomatology [223]. The experimental protocol is given at Appendix C. The study was approved by the Ethical Committee of Area Vasta Emilia Centro (Bologna, Italy; approval n° 542-2019-OSS-AUSLBO). Participants were provided with a VR headset to be freely used at home for four days. Before and after each VR session, each participant had to fill the Edmonton Symptoms Assessment Scales (ESAS) [332], a multi-item instrument that rates nine common symptoms (including pain) experienced by cancer patients on a 0-10 numerical rating scale (NRS). The ESAS was digitally inserted into the VR headset. Participants were also asked to wear the Empatica E4 wristband, whose technical specifications are given at Appendix A.1.

9.3.1 Pre-processing and data aggregation

We used the pain ratings obtained before each VR session. Based on the pain ratings, two classes were identified: no pain (NRS = 0) and pain (NRS ≥ 1). We matched the start of the VR session (i.e., timestamp saved in the VR headset) with the Empatica E4 recordings. For each physiological signal, we selected a 2-min time window before the beginning of each VR session. Afterwards, we detected the "available instances", namely those time windows of physiological signals with concurrent Empatica E4 recording and VR headset usage. If a patient performed two consecutive VR sessions, the 2-min time window of physiological signals recorded before the second VR session was not analyzed if included in the previous VR session.

9.3.2 Signal processing

PPG signals were processed following the procedure described in Appendix A.2. The quality of each pulse was assessed by using the automatic algorithm presented at Section 3. For this study, only basic pulses were further exploited. To conduct the HRV analysis, basic pulses were used by firstly estimating the inter-beat intervals (IBIs) as the difference between two consecutive systolic feet.

The accelerometer data was processed following the procedure described at Appendix A.3.

EDA signals were processed with the procedure described at Appendix A.4..

9.3.3 Feature extraction

A total of 39 features were extracted from each available instance. Specifically, we estimated:

- 12 features from the HRV analysis: mean heart rate (*meanHR*), standard deviation of IBI (*SDNN*), root mean squared difference of successive IBI (*RMSSD*), percentage of successive IBIs that differ from the previous one for more than 50 ms (*pNN50*), Poincaré plot standard deviation perpendicular and along the line of identity (*SD1*, *SD2*), and its ratio (*SD1/SD2*), low-frequency power (*LF*), high-frequency power (*HF*), total power (*TP*), and ratio between LF and HF (*LF/HF*), and the approximate entropy (*ApEn*).
- 5 features from the PPG morphological analysis: area between the systolic foot and the successive systolic peak (*A1*), area between the systolic peak and the successive systolic foot (*A2*), total area under the PPG pulse (*A*), time between systolic foot and the successive systolic peak (*T1*), time between systolic peak and the successive systolic foot (*T2*) [86].
- 17 features from the EDA: 5 features from the whole EDA signal, which are mean (*meanEDA*), standard deviation (*stdEDA*), slope (*slopeEDA*), maximum frequency (*Fmax*) [333], and symbolic information entropy (*SIE*) [249]; the cvxEDA algorithm [246] was applied to divide the EDA signal into tonic (also known as electrodermal level, EDL) and phasic (also known as electrodermal response, EDR) components. Three features were estimated from the EDL: mean (*meanEDL*), standard deviation (*stdEDL*), and slope (*slopeEDL*). Nine features were estimated from the EDR: mean (*meanampEDR*), standard deviation (*stdampEDR*), maximum (*maxampEDR*) and minimum (*minampEDR*) amplitude, frequency of EDR (*freqEDR*), dynamic range (*DR EDR*), integral (*INSC*), normalized average power (*APSC*), and normalized root mean square (*RMSC*) [334].
- 3 features from the TEMP: mean (*meanTEMP*), standard deviation (*stdTEMP*), and slope (*slopeTEMP*).
- 2 features from the AI: mean (*meanAI*) and standard deviation (*stdAI*).

9.3.4 Feature selection

Features selection methods were applied only on the "valid instances", representing those instances in which all the features had a finite value. This depends on the availability of PPG pulses suitable for HRV and for morphological analysis: if any PPG pulse was not detected, derived features cannot be estimated.

The feature set was normalized through the Box-Cox transformation [156] and standardized with the z-score procedure.

In order to evaluate the linear dependency among the features, we carried out a correlation analysis. The correlation between pairs of features was calculated as the Pearson correlation coefficient. We removed one of the two features with a correlation coefficient higher than 0.9.

We then applied four different feature selection methods using algorithms implemented in Waikato Environment for Knowledge Analysis (WEKA) [335]. We applied the correlation-redundancy feature set analysis (Cfs), correlation (Corr), gain ratio (GainRatio), and reliefF. The first algorithm used a greedy approach, while the latter three a ranked approach. A 10-fold cross-validation was applied. We selected a number of features equal to 10% of the total valid instances.

9.3.5 Classification

Classification was performed using WEKA. We used five different classifiers: Support Vector Machine (SVM), Random Forest (RF), Multilayer Perceptron (MP), Logistic (Log), and Adaptive Boosting Algorithm (AdaBoost). A 10-fold cross-validation repeated 10 times on a newly selected random set was applied. The following metrics were used to evaluate the classifier performance: accuracy (Acc), sensitivity (Sens), and specificity (Spec), and area under the ROC curve (AUC).

9.4 Results

9.4.1 Descriptive statistics

Thirty-five patients were enrolled in the study (14 M, 21 F, age 49.7 ± 12.8), collecting a total of 196 VR sessions. The available instances (i.e., sessions with concurrent recorded physiological signals and not within previous VR sessions) consisted of 92 sessions (47% of the total sessions) from 29 patients. Valid instances (i.e., sessions with concurrent recorded physiological signals in which all the features have a finite value) consisted in

44 sessions (22% of the total sessions) only from 21 patients (9 M, 12 F, age 49.4 ± 13.8), with 20 instances corresponding to the “no pain” class and 24 to the “pain” class.

9.4.2 Feature selection

From the correlation analysis, we removed 14 features, using the following measures for further analysis: *meanHR*, *SDNN*, *pNN50*, *SD1/SD2*, *LF*, HF, *LF/HF*, *ApEn*, *A1*, *T1*, *T2*, *meanTEMP*, *stdTEMP*, *slopeTEMP*, *meanAI*, *stdAI*, *Fmax*, *SIE*, *meanEDA*, *stdEDA*, *slopeEDA*, *slopeEDL*, *meanampEDR*, *stdampEDR*, and *minampEDR*. The correlation matrix is graphically represented in Figure 10.1.

We then applied four different feature selection methods to the features selected from the correlation analysis:

- Cfs: meanHR, T1, slopeTEMP, meanAI
- Corr: meanHR, LF/HF, stdTEMP, meanAI
- GainRatio: ApEn, A1, meanTEMP, minampEDR
- ReliefF: pNN50, SD1/SD2, meanAI, minampEDR

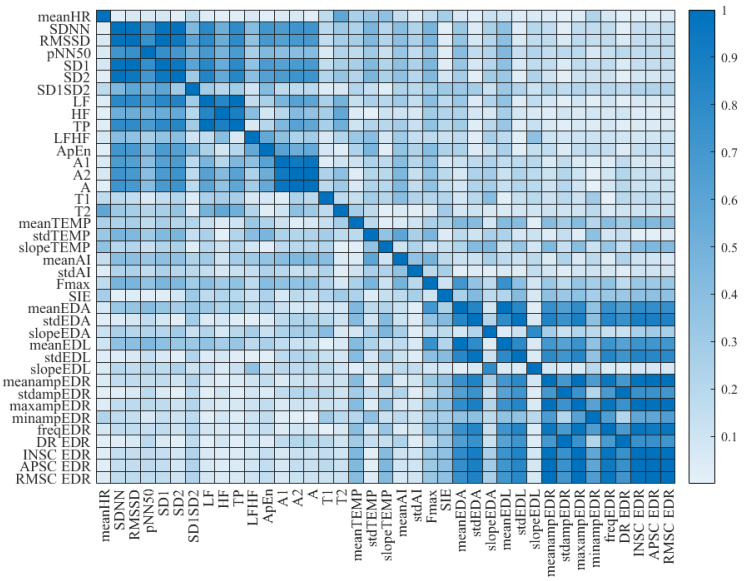


Figure 9.1 Correlation matrix

9.4.3 Classification

We applied four different classification algorithms fed with the features set selected by the four different feature selection methods. Results are reported in Table 10.1.

For three classification algorithms out of four, the features selected by the Corr method provided the best performances in terms of Acc and AUC. Despite the SVM (Corr)

achieved the highest accuracy rate (up to 73%), overall, the best performance in terms of Acc, Sens, and Spec was achieved with the MP algorithm using the GainRatio as feature selection method (Acc: 72%; Sens: 77%; Spec: 68%).

Table 9.1 Classification results

Classifier	FS method	Acc mean (std)	Sens Mean (std)	Spec mean (std)	AUC mean (std)
SVM	Cfs	0.64 (0.02)	0.93 (0.03)	0.29 (0.06)	0.61 (0.02)
	Corr	0.73 (0.01)	0.90 (0.03)	0.52 (0.03)	0.71 (0.01)
	GainRatio	0.67 (0.04)	0.85 (0.02)	0.45 (0.07)	0.65 (0.04)
	ReliefF	0.66 (0.02)	0.85 (0.03)	0.43 (0.06)	0.64 (0.03)
RF	Cfs	0.59 (0.05)	0.66 (0.06)	0.51 (0.07)	0.61 (0.04)
	Corr	0.64 (0.03)	0.72 (0.06)	0.54 (0.06)	0.65 (0.04)
	GainRatio	0.57 (0.04)	0.58 (0.05)	0.56 (0.07)	0.62 (0.04)
	ReliefF	0.63 (0.03)	0.67 (0.05)	0.59 (0.03)	0.63 (0.04)
MP	Cfs	0.58 (0.06)	0.69 (0.07)	0.46 (0.09)	0.58 (0.06)
	Corr	0.61 (0.03)	0.68 (0.02)	0.52 (0.07)	0.59 (0.03)
	GainRatio	0.72 (0.02)	0.77 (0.03)	0.68 (0.06)	0.72 (0.05)
	ReliefF	0.58 (0.05)	0.80 (0.06)	0.43 (0.05)	0.56 (0.06)
Log	Cfs	0.62 (0.02)	0.73 (0.04)	0.49 (0.02)	0.61 (0.01)
	Corr	0.65 (0.02)	0.73 (0.05)	0.56 (0.03)	0.66 (0.02)
	GainRatio	0.63 (0.03)	0.76 (0.04)	0.48 (0.07)	0.58 (0.03)
	ReliefF	0.67 (0.03)	0.78 (0.04)	0.55 (0.03)	0.60 (0.02)
AdaBoost	Cfs	0.66 (0.05)	0.80 (0.08)	0.50 (0.04)	0.67 (0.05)
	Corr	0.52 (0.06)	0.63 (0.09)	0.40 (0.42)	0.51 (0.08)
	GainRatio	0.52 (0.07)	0.60 (0.08)	0.42 (0.11)	0.53 (0.06)
	ReliefF	0.59 (0.05)	0.72 (0.08)	0.43 (0.05)	0.56 (0.05)

9.5 Discussion

This study shows the feasibility of developing methods for automatic pain assessment. To the best of our knowledge, this is the first study that used recordings conducted in a real-world context to develop an automatic pain assessment classifier.

We used a multimodal approach by exploiting features from different physiological signals. The importance of including several sources of information is enforced by the results obtained from the feature selection step, in which all the algorithms selected features from different physiological signals. Given the complexity of the phenomenon, the multimodal approach allows looking at pain from different perspectives.

Comparing the results obtained from the study by Badura et al. [331], our best classifier reached lower values in terms of all performance indicators. Moreover, we obtained opposite results with respect to the classification algorithms: while Badura et al. showed better performance with the AdaBoost algorithm compared to the SVM, in our case AdaBoost algorithm provided worse performance than the SVM one. One important distinction between our work and Badura et al. [331] 's one relies on the fact that our signals, recorded in a real-world context, are affected to a greater extent by external noise and motion artifacts. Moreover, in our study the classification problem is based on no pain vs pain, while the work by Badura et al. [331] relates to the binary classification between three different levels of pain (i.e., no pain, moderate pain, and severe pain): results comparing two extreme levels of pain attained excellent performance, while our classifier is based on differences between two adjacent levels (NRS=0 vs NRS≥1).

As already mentioned, this study presents some limitations: the small datasets, the possibility to use a small set of features (given in turn by the small dataset), the use of 10-fold CV, instead of a more appropriate "leave-one-subject-out" validation, in order to take into account the inter-subject variability, poor quality of recorded physiological signals, which in turn limited the use of all the available instances because of the lack of good quality PPG signal segments.

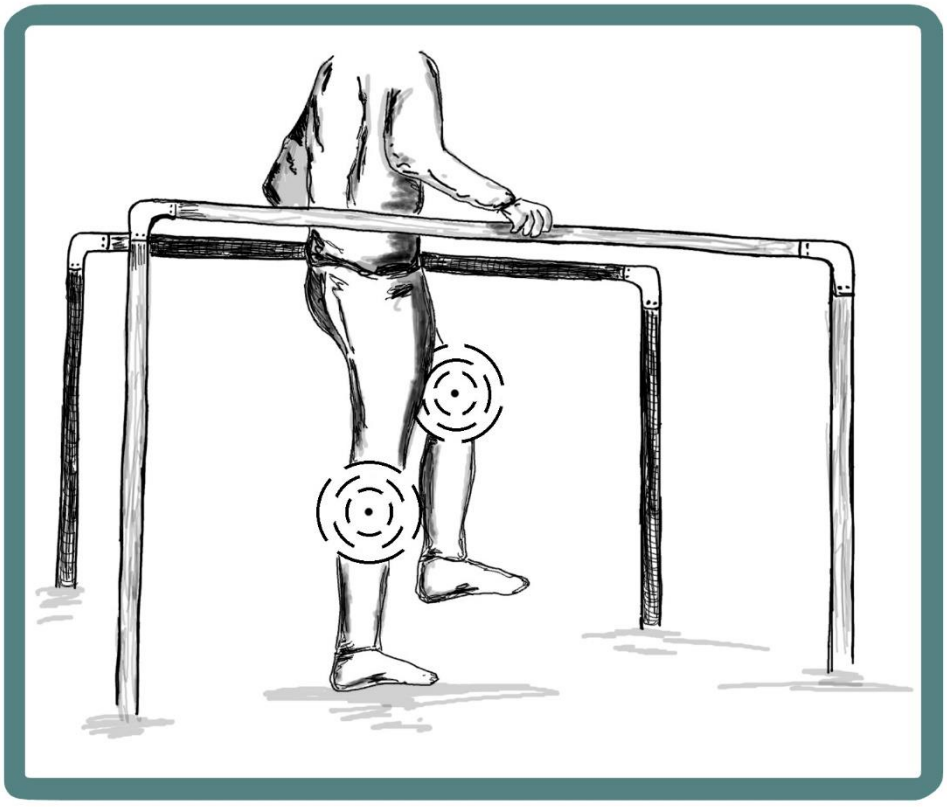
Future automatic methods for pain assessment can be improved by using larger datasets and more valid instances for each participant. An interesting strategy is to test the performance of the classifiers fed with features extracted from longer time windows: this strategy can also be used to retrieve more valid instances since the probability of finding, for example, good quality PPG signal segments increases. Another approach can be to change the way to represent pain classes. For instance, it can be interesting to apply different thresholds to define no pain vs. pain, to test classification performance on a multiclass problem, and to also use regression algorithms.

9.6 Conclusion

Automatic pain assessment in real-world scenarios can represent a decisive means to improve and optimize pain management. Moreover, we showed that it is possible to achieve good classification performance using machine learning algorithms trained with data collected in real-world contexts. These preliminary results show the feasibility,

although on a small dataset, of such an approach, that could be included in the clinical decision making to improve the pain management path.

SECTION V: PAIN ASSESSMENT ON PATIENTS UNDERGOING NEUROREHABILITATION



10. STUDY PROTOCOL FOR AN EXPLORATORY INTERVENTIONAL STUDY INVESTIGATING PAIN IN NEUROREHABILITATION THROUGH WEARABLE SENSORS (PAINLESS)

From the paper being sent to BMJ Open: Moscato S., Orlandi S., Di Gregorio F., Lullini G., Pozzi S., Sabattini L., Chiari L., La Porta F., Study protocol for an exploratory interventional study investigating PAIN in neurorehabilitation through wearable SensorS (PAINLESS)

10.1 Abstract

Millions of people survive injuries to the central or peripheral nervous system for which neurorehabilitation is required. Unfortunately, in addition to the physical and cognitive impairments associated with neurological deficits, many neurorehabilitation patients experience pain, often not widely recognized and inadequately treated. This is particularly true for Multiple Sclerosis (MS) patients, for whom pain is one of the most common

symptoms. In clinical practice, pain assessment is usually conducted based on a subjective estimate of the patient's pain experience, mainly using self-administered questionnaires or scales. However, these tools can lead to evaluations that are not always accurate due to the influence of numerous factors, including emotional or cognitive aspects.

To date, no objective and simple-to-use clinical methods allow objective quantification of the subjective pain experience and a diagnostic differentiation between the two main types of pain (nociceptive vs. neuropathic pain). Wearable technologies are increasingly being applied in various clinical settings for monitoring patients' health parameters, allowing the real-time collection and processing of health data. As such, we aim to develop a novel automatic tool fueled by artificial intelligence (AI) to assess the presence of pain and its characteristics during neurorehabilitation treatments by evaluating the feasibility of using physiological signals collected by wearable sensors.

We aim to recruit 15 participants suffering from MS who will undergo physiotherapy treatments. During the study, participants will wear a wearable sensor (i.e., a wristband) for three consecutive days and be monitored before and after their physiotherapy sessions. Measurements of traditionally used pain assessment questionnaires and scales (i.e., painDETECT, DN4 questionnaire, EuroQol 5-dimension 3-level) and physiological signals (photoplethysmography, electrodermal activity, skin temperature, accelerometer data) will be collected. The parameters of interest from the physiological signals will be identified, and automatic classification methods will be developed using AI algorithms.

10.2 Introduction

According to the definition of the "International Association for the Study of Pain" (IASP), pain is "an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage" [1]. When pain arises from actual tissue damage, it is called nociceptive, and it has a clear protective function as it alerts the nervous system of potential threats to which it has to react adequately [4]. However, another type of pain (i.e., neuropathic pain) occurs without actual tissue damage as it is secondary to central or peripheral nervous system lesions. In this respect, neuropathic pain, which usually manifests as electric shocks, unpleasant perception of intense cold, and feelings of pressure or constriction, can occur at almost any site; it is generally chronic and, as such, can be extremely disabling [2].

Pain is one of the most common complaints of Persons with Multiple Sclerosis (PwMS) [3], an autoimmune disease characterized by inflammation, selective demyelination, and gliosis of central nervous system white matter. In particular, PwMS patients describe their pain as often widespread, chronic, and debilitating, and, as such, it may be associated with psychological distress and decreased daily functioning [4]. Since MS affects approximately 2.1 million people worldwide [5], and the prevalence of pain in this condition is between 30% and 85% [6], it can be estimated that from 630,000 to 1,800,000 PwMS around the world are likely to suffer from disabling pain. Furthermore, nociceptive and neuropathic pain may coexist in PwMS, thus posing a diagnostic and therapeutic challenge as nociceptive pain, mainly due to spasticity or other musculoskeletal impairments, may limit the effectiveness of physical therapies [2]. To make things even more complicated, the subjective experience of pain in PwMS often requires a biopsychosocial approach for assessment and treatment, where the goal is to treat the manifestations of pain at the sensory level as well as its related psychological and social aspects [7]. Hence, for appropriate and successful pain treatment in PwMS, the availability of a tool that could assess pain in its intensity and nature as objectively as possible would be highly beneficial.

In clinical practice, pain assessment is often based on subjective estimates obtained by interviewing patients, mainly using self-administered questionnaires [8]. Several self-report scales are available for the overall evaluation of pain intensity. The *Numerical Rating Scale* (NRS) is the most used, given its reported excellent reliability and validity. It consists of a 0-10 scale, where 0 is “absence of pain” and 10 is “the worst pain possible” [9]. Other scales are the *Pain Severity Subscale* of the *Multidimensional Pain Inventory* (MPI), consisting of three items on pain severity and the suffering related to pain, and the *Neuropathic Pain Scale Inventory*, which includes questions about the intensity and the quality of pain [8]. In addition, other questionnaires were specifically devised to assess symptom severity arising from neuropathic pain. Examples are the *Neuropathic Pain Symptoms Inventory* (NPSI), used for pain assessment in several populations of neurotrauma patients [8], the *painDETECT* (PD-Q), developed to measure pain’s neuropathic components [10], and *Neuropathic Pain–4 questions* (Douleur Neuropathique, DN4) [11]. There are also more general questionnaires aimed at assessing the health-related quality of life in which one of the subdimension is dedicated to assessing pain, such as the EuroQoL 5-dimension 3-level (EQ-5D-3L) [12]. Finally, in

addition to scales and questionnaires, pain can be assessed through “objective” instrumented methods. Some of these methods are the Quantitative Sensory Testing (QST), a battery of tests aiming at identifying pain threshold and changes in sensory function [8], the analysis of electromyographic (EMG) signals to record facial emotional expressions, voice analysis [13], functional magnetic resonance imaging (fMRI) and functional near-infrared spectroscopy (fNIRS) to monitor the main metabolic activity [13,14], or the analysis of evoked potentials recorded by the electroencephalography (EEG) [8].

Despite the availability of different tools for assessing pain, several limitations should be highlighted. First, scales and questionnaires, although undoubtedly helpful for capturing the subjective dimension of the experience of pain, can lead to inaccurate assessments due to the influence of numerous factors, not least those related to emotional or cognitive aspects. Furthermore, they can be administered reliably only to patients who are cooperative enough and not suffering from severe mental and/or communication impairments [15]. Furthermore, beyond the lack of objectivity, existing pain measurement methods may be inaccurate in discriminating between nociceptive and neuropathic pain [16]. Instrumented methods currently available could partially overcome this limitation [17,18]. Still, they can hardly be used on large populations because of the expensive costs in terms of money, time, and complex setup. Given the limitations and barriers of the existing methods, there is a need to develop new and efficient strategies for objective pain assessment. These new tools can be considered complementary to state-of-the-art pain assessment methods or new methodologies to be applied in cases where scales and questionnaires fail, such as in non-communicative patients.

Some insights potentially helpful in developing novel tools to measure pain objectively may be gleaned from the current knowledge of the neurophysiological mechanisms of pain. Indeed, pain perception involves the activation of neural mechanisms, including the Autonomic Nervous System (ANS) [19]. The ANS represents the interface between the human body’s internal and external environment, acting to maintain homeostasis and respond to stress stimuli [20]. In turn, its activity influences the normal functions of several physiological mechanisms, such as skin conductance [21], heart rate, and the cardiovascular system in general [22,23]. Thus, monitoring these physiological mechanisms may provide a novel method for objective pain assessment since it would eliminate the influence of subjectivity and the impossibility of verbally communicating

it. In this context, a new opportunity may be given by combining two currently widespread technologies already available in clinical and research fields: wearable sensors and artificial intelligence (AI) algorithms. The former allows us to continuously and passively record physiological signals in pervasive contexts, while the latter would enable the development of data-driven models to detect particular conditions automatically.

Several studies examined the relationship between pain and physiological signals [13,24]. Specifically, Johnson et al. [25] showed the feasibility of developing novel methods to assess pain by collecting physiological signals with wearable devices on 27 patients with sickle cell disease in a hospital setting using machine learning classifiers and regressors. In another work, Badura et al. [26] applied the same approach in a physiotherapy setting, monitoring 35 patients who rated their pain during a session of fascial therapy. In addition, our group developed an automatic dichotomous classifier for pain assessment in oncological patients in a previous study [27]. Together with pain evaluations, real-world recordings from 31 patients were used to feed the classifier for detecting “pain” and “no pain” conditions. Best classification performances were obtained using four features extracted from photoplethysmography and electrodermal activity with the AdaBoost algorithm, reaching an accuracy equal to 72% [27]. However, despite these encouraging initial studies, the literature on the diagnostic accuracy of pain measurements involving wearable sensors is still scarce [28,29]. Furthermore, none of the previous studies explicitly focused on PwMS.

Thus, based on this preliminary evidence, the present feasibility study aims to investigate the use of physiological signals recorded by wearable sensors to achieve the following specific objectives: 1) to evaluate the feasibility of developing a differential diagnosis method to assess the absence or presence of pain; 2) to evaluate the feasibility of developing a regression model to assess pain intensity; 3) to evaluate the feasibility of developing a differential diagnosis method to discern the type of pain (nociceptive vs. neuropathic pain).

10.3 Materials and Methods

10.3.1 Design

The project “PAIN in neurorehabilitation through wearabLE SensorS (PAINLESS)” is a feasibility, single cohort, interventional study.

10.3.2 Participants

We aim to recruit 15 participants aged between 18 and 75, undergoing neurorehabilitation motor treatments in the Neurorehabilitation Unit of IRCSS Istituto delle Scienze Neurologiche di Bologna (ISNB). Inclusion and exclusion criteria are detailed in Table 10.1. Before enrollment in the study, the principal investigator (PI) will check the eligibility criteria. In particular, after verifying the eligibility criteria, the PI (or a delegate) will provide the potentially eligible person with all the information and details relative to the study in simple language during an interview that will preferably take place in the presence of a caregiver. The participant is then asked to give his or her informed consent to participate in the study.

10.3.3 Intervention

For all enrolled participants, the intervention is represented by objective monitoring of physiological parameters, continuously recorded for 48 hours with the wearable medical device Empatica E4 [30], and concurrent subjective monitoring via specific questionnaires digitally administered via Microsoft Forms™. In particular, the intervention will be articulated across four main stages:

- t_0-t_{1a}: baseline monitoring (24h)
- t_{1a}-t_{1b}: device recharging and data downloading (1h max)
- t_{1b}-t_2: monitoring during a physiotherapy treatment session (1h)
- t_2-t_3: post- physiotherapy treatment monitoring (23 hours)

At t_0, t_{1b}, t_2, and t_3, participants will fill in subjective pain questionnaires (described in detail in the next section) to carry out a stratification and to keep monitoring it throughout the intervention in one of the following three categories: 1) absence of pain; 2) nociceptive pain; 3) neuropathic pain. A graphical depiction of the protocol is shown in Figure 10.1.

Table 10.1 Inclusion and exclusion criteria

Inclusion criteria
• Age between 18 and 75 years • Diagnosis of certainty of Multiple Sclerosis for at least three months • Prescription of a physiotherapy-based motor rehabilitation program • Signature of the informed consent to participate in the study
Exclusion criteria
• Heart rhythm modifying disease and/or factors such as arrhythmogenic heart disease (e.g., atrial fibrillation), presence of pacemakers and/or use of drugs capable of affecting heart rhythms, such as beta blockers (C07) or other antiarrhythmic drugs (C01) • Cognitive impairments that preclude the possibility of providing valid informed consent, such as a disorder of consciousness or confusional state, the latter defined by temporal and/or spatial disorientation detected during ordinary conversation. In case of doubt, a simple confusional state assessment test (4AT) will be administered before enrollment • Language comprehension skills lower than 75% in an ordinary conversation due to aphasic disorder of severe deafness despite the use of a hearing aid. In case of doubt, a simple language comprehension test (token test) will be administered before enrollment • Linguistic expression less than 75%. In case of doubt, a simple verbal fluency test (verbal fluency by phonemic category) will be administered before enrollment • Severe psychiatric comorbidity that may interfere with adherence to the study protocol (e.g., severe personality disorders, severe psychomotor agitation) • History or current use of narcotic drugs (including marijuana) • Modification in the two weeks prior to enrollment or foreseeable modification during enrollment of any chronic pain management program, both pharmacological (cortisone for systemic use, H02; antirheumatics, M01; analgesics, N02; antiepileptics, N03; antidepressants tricyclics, N06AA; atypical antidepressants such as duloxetine or venlafaxine, N06AX) and non-pharmacological (e.g., acupuncture or other manual therapies, physical therapies, such as tecar therapy)

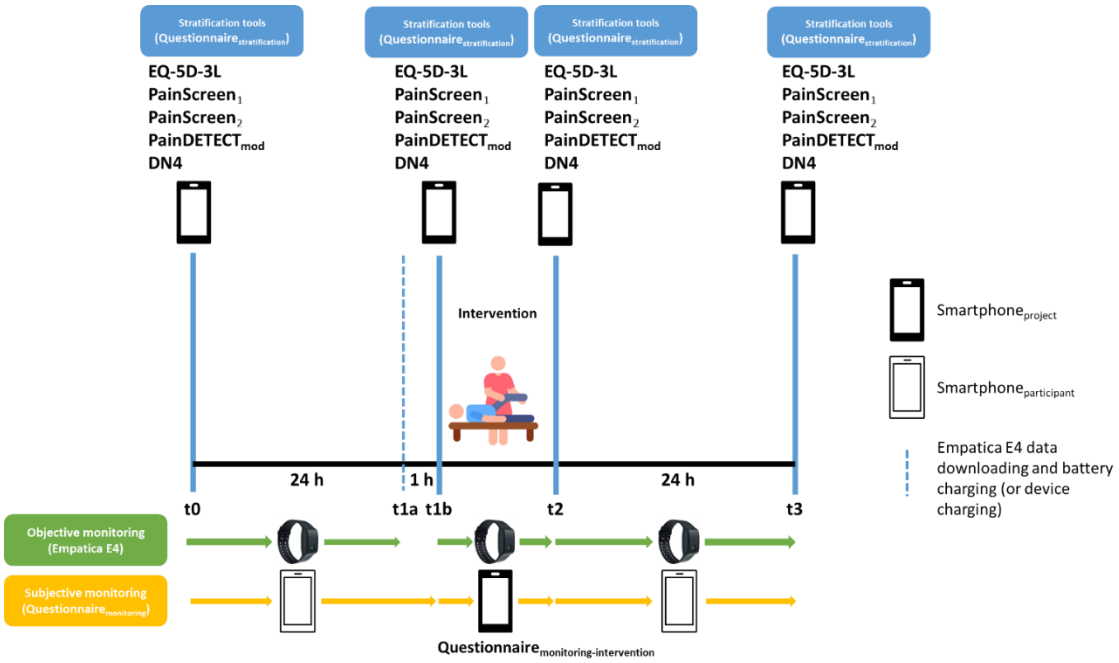


Figure 10.1 Study protocol

10.3.3.1 Reference measurements

The reference measurements, which will be taken for each participant, will be included in the following Case Report Form (CRF):

1. a Recruitment CRF, which will contain the demographic information, the Expanded Disability Scale [31] information about the disease and drugs;
2. a Sleep-wake questionnaire CRF, which the PI will administer to set reminders for each participant to fill in the monitoring questionnaire CRF.
3. a Stratification questionnaire CRF will allow the classification of patients into the three previously mentioned categories (absence of pain, nociceptive pain, or neuropathic pain) following the procedure described in Figure 2. In particular, this CRF will include the following tools: a) two screening questions (Pain Screen1 and Pain Screen2) to respectively assess the presence of current pain or in the past four weeks; b) the painDETECT questionnaire [10]; c) the Doleur Neuropathique 4 Questions (DN4) [11]; d) the Euro Quality of Life 5-dimension 3-level (EQ-5D-3L) [12] to evaluate the health-related quality of life.
4. a Monitoring questionnaire CRF, which each participant will fill in through the smartphoneparticipant during the 48h-monitoring, including information about any experienced pain.

5. a Monitoring-treatment questionnaire CRF will be administered by the PI (or his delegate) through the smartphoneproject to each participant during the motor neurorehabilitation treatment. It is a reduced version of the Monitoring questionnaire CRF.

10.3.3.2 Wearable devices and physiological signals

Each participant will be asked to wear the Empatica E4 wristband, a medical wearable device that records the following physiological signals:

- Photoplethysmography (PPG): it records variations in blood volume flow that occur with each heartbeat, affected by both the sympathetic and the parasympathetic nervous systems. PPG signal can be exploited to estimate the heart rate, thus allowing the heart rate variability (HRV) analysis and, moreover, interesting features can be extracted by conducting a more in-depth morphological analysis [83];
- Electrodermal Activity (EDA): it represents the activation of the eccrine sweat glands, innervated by the sympathetic nervous system, representing an arousal index. Features related to pain sensations can be extracted either from the whole signal or from the two principal components, the tonic (slow changes) and the phasic (fast changes) components [91];
- Skin temperature (SKT): it is an index of sympathetic activation, mainly depending on the amount of superficial blood flow;
- Accelerometer data (ACC): it records physical activity and movement.

10.3.3.3 Experimental pipeline

- t0: The CRF Stratification questionnaire is administered through a smartphone by the PI (or his delegate). The participant is then asked to wear the Empatica E4 wearable medical device and the smartphoneparticipant will be provided, which will be used to fulfill the CRF Monitoring questionnaire. Reminders will be set to fulfill the questionnaire based on the CRF Sleep-Wake questionnaire, administered in this phase.
- t0-t1a: The participant wears the Empatica E4 wearable device and fulfills the CRF Monitoring questionnaire. Reminders are set hourly during wakeness.
- t1a-t1b: The participant comes back to the clinic 24 hours after t0 and drops off the Empatica E4 and the smartphoneparticipant for data downloading and device

recharging. After about an hour, the participant is asked again to wear the Empatica E4. Then, the CRF Stratification questionnaire is administered, and the motor neurorehabilitation treatment starts.

- t1b-t2: The participant undergoes the motor neurorehabilitation treatment, and every 10 minutes the PI (or his delegate) administers the CRF Monitoring-treatment questionnaire, a reduced version of the CRF Monitoring questionnaire, through the smartphoneproject,.
- t2: The CRF Stratification questionnaire is administered, and the participant receives the smartphoneparticipant.
- t2-t3: The participant wears the Empatica E4 wearable device and fulfills the CRF Monitoring questionnaire. Reminders are set hourly during wakeness.
- t3: The participant comes back to the clinic 24 hours after t2 and drops off the Empatica E4 and the smartphone.

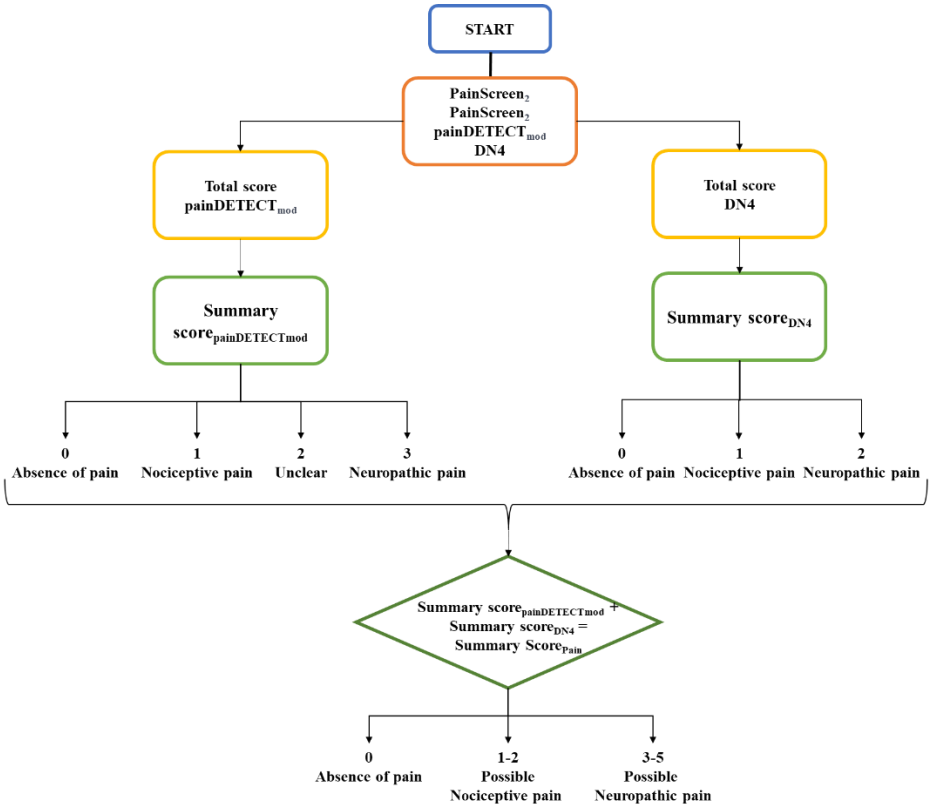


Figure 10.2 Stratification algorithm

10.3.2 Signal and data analysis pipeline

Physiological signals recorded through the Empatica E4 wristband will be analyzed in four successive phases: 1) Preprocessing (artifact mitigation, filtering); 2) Segmentation (time-windows detection of physiological signals linked to the assessments); 3) Signal processing and feature extraction; 4) Feature selection. Following this pipeline, we will implement AI algorithms to develop the classifiers and regressors methods indicated in Table 10.2. Classifiers and regressors will be trained and tested based on the outcomes from the Stratification questionnaire CRF, Monitoring questionnaire CRF, and Monitoring-treatment questionnaire CRF. Validation will be conducted by testing the Leave-One-Subject-Out cross-validation and 10-fold cross-validation. We will also consider adding covariates, either from the Monitoring questionnaire CRF or personal data (e.g., age, information about the pathology, and use of drugs). This will allow verifying, both on a quantitative and qualitative basis, whether there are differences in physiological parameters related to these specific covariates.

The performance of the classifiers will be assessed using the following indicators: accuracy, sensitivity, specificity, and area under the Receiving Operating Characteristic (ROC) curve (or precision and recall when a multi-class classification is applied). Instead, the regression models' performance will be assessed using the following indicators: root mean squared error, absolute error, relative error, and correlation.

Table 10. 2 Classifiers and regressors methods for pain assessment

Pain class	**Absence vs Presence of pain**
	Nociceptive vs Neuropathic pain
	Absence of pain vs Nociceptive pain vs Neuropathic pain
Pain intensity	Multi-class classifier, based on literature guidelines
	Regression model

10.3.3 Objectives and related endpoints

1. Feasibility of developing a differential diagnosis method based on physiological signals recorded using wearable sensors to assess the absence or presence of pain. The related primary endpoint will be evaluated based on the number of available instances to be processed for determining the absence/presence of pain, which means the number of concurrent physiological signals registrations and pain assessments. If this endpoint is met, a predictive test will be developed based on AI techniques and physiological parameters. The diagnostic performance of this test

will be evaluated against the gold standard (questionnaires) by evaluating standard performance indicators (i.e., sensitivity, specificity, predictive values). The endpoint will be considered achieved if at least 80% of the instances are available. The diagnostic accuracy will be calculated using the CRF Stratification and CRF Monitoring questionnaires as a reference. The threshold for the diagnostic accuracy to define the endpoint achieved is set at 75%.

2. Feasibility of developing a regression model based on physiological signals recorded using wearable sensors to assess pain intensity (secondary endpoint). The related secondary endpoint will be evaluated based on the number of available instances to be processed to assess pain intensity, i.e., the number of concurrent physiological signals registrations and pain assessments. If this endpoint is met, a regression model will be developed based on AI techniques and physiological parameters. The diagnostic performance of this test will be evaluated against the gold standard (questionnaires) by evaluating standard performance indicators (i.e., accuracy, mean squared error). The endpoint will be achieved if at least 80% of the instances are available. The coefficient of determination of the regression model will be calculated using the CRF Stratification questionnaire and CRF Monitoring questionnaire as a reference. The threshold for the coefficient of determination to define the endpoint achieved is set at 0.5.
3. Feasibility of developing a differential diagnosis method based on physiological signals recorded using wearable sensors to discern between nociceptive and neuropathic pain (secondary endpoint). The related secondary endpoint will be assessed based on the number of available instances to be processed to distinguish between nociceptive and neuropathic pain, i.e., the number of concurrent physiological signals registrations and pain assessments. If this endpoint is met, a predictive test will be developed based on AI techniques and physiological parameters. The diagnostic performance of this test will be evaluated against the gold standard (questionnaires) by evaluating standard performance indicators (i.e., sensitivity, specificity, predictive values). The endpoint will be considered achieved if at least 80% of the instances are available. The diagnostic accuracy will be calculated using the CRF Stratification and CRF Monitoring questionnaires as a reference. The threshold to define the endpoint achieved is set at 75%.

10.3.5 Sample size

Given the study's exploratory nature, the effect size is unknown; thus, it is not possible to calculate the sample size accurately.

10.3.4 Ethics and dissemination

The study will be conducted according to the ethical principles established in the Declaration of Helsinki and has been subjected to approval by the local Ethical Committee (285-2022-SPER-AUSLBO). Any changes to the protocol will be proposed to the local Ethical Committee as a request for amendment. Although it is not foreseen that there will be a direct short-term benefit to participants, the research protocol presents minimal risks for the participants and no burden, as required by Article 28 of the Declaration of Helsinki.

Personal data will be retained in agreement with the GDPR guidance for ten years. Specifically, the PI and co-PIs will be responsible for archiving and preserving the essential study documents before, during, and after the completion of the study, according to the timeframe required by the current regulations and good clinical practice.

Researchers involved in the study will disseminate the results in a timely and complete manner, participating in conferences and writing scientific articles for submission to international journals. In addition, the findings from the study will form part of a doctoral dissertation for one of the authors (SM). The researchers will scrupulously, objectively, and impartially provide as much evidence and information as possible on aspects such as the state-of-the-art literature before the study, the original purpose, and methods defined before conducting the research, any changes in objectives and methods since the study were commenced, the significant results achieved, including negative or null results and, finally, the possible interpretations, applicability, and limitations of the findings.

10.4 Discussion

In regular clinical practice, pain assessment is usually carried out by administering subjective scales and questionnaires. Although their usefulness for the subjective quantification of pain, these tools can lead to inaccurate assessments due to the influence of many factors, such as emotional and cognitive factors. In addition, they cannot be administered to those patients unable to communicate verbally. Therefore, identifying optimal physiological parameters recorded through wearable devices and using artificial

intelligence algorithms would allow the development of automatic methods capable of determining the absence or presence of pain in MS patients, its intensity, and distinguishing pain as nociceptive or neuropathic.

Such continuous and objective pain monitoring in everyday life activities and during treatments would overcome the limitations imposed by the tools currently used in clinical practice. In particular, continuous and objective monitoring would bring about several advantages. First, this pain assessment disregards the patients' ability or willingness to communicate their pain verbally. Second, this approach is supposed to provide a completely automatic method that would not require spending time ad hoc to administer scales and questionnaires, as it could be used in hospital or daily life contexts while patients are involved in other activities. Lastly, having a more reliable method to discriminate between nociceptive and neuropathic pain would allow a better personalization of the antalgic therapy.

The long-term goal is to integrate such an innovative method into regular clinical practice as a tool for clinical decision-making for the antalgic therapy to be chosen. Implementing this method would allow PwMS to be monitored both during neurorehabilitation treatment and in a pervasive context. This would allow for a timelier assessment of the patient's pain, ultimately aiming to ameliorate their quality of life. Prospectively, if properly calibrated, such a method could allow quantification and monitoring of pain in patients unable to express it verbally, such as patients with severe brain injury, in a minimally conscious state, or with aphasia.

An innovative aspect of this study relies on the possibility of overcoming the "etiological" boundaries of pain at the measurement level. This would be extremely useful, considering that, in many pathologies, different types of pain may coexist. For example, in brain injury, there may be a mix of nociceptive and neuropathic pain, both of central and peripheral origin. This study could bring initial insights into how pain can be measured by recording a minimum set of physiological parameters based on physiological indicators invariant to the pathology. In other words, we will be able to assess whether the parameters to be measured are independent of the underlying pathology, precisely as is the case for different physiological parameters such as body temperature or heart rate. For the latter, differences of quantitative nature (e.g., fever) give rise to specific diagnostic profiles only in combination with other data (e.g., body temperature changes and other diagnostic indicators), being the measurement of the temperature parameter independent

of the pathology that modifies it. Similarly, from the combination of physiological parameters of pain, diagnostic combinations ("profiles") could be identified for specific pathologies.

The proposed study is also relevant for health systems because it aims to improve the pain assessment phase, which is necessary to choose the most appropriate antalgic therapy for the patient. In addition, such a system would allow the prescription of more personalized pain treatment plans, make efficient use of resources, and minimize the waste resulting from the incorrect choice of ineffective strategies to improve the patient's pain status. In addition, the proposed protocol is also relevant in terms of research, as the availability of an objective system of pain quantification, together with the already available subjective assessment tools, would make the quantification of treatment effects in the context of RCTs and other studies undoubtedly more accurate and less prone to interpretive bias.

The methodology presented here may suffer from several limitations. First, being designed as an exploratory feasibility study, the limited sample size may hinder the development of robust and reliable methods for objectively assessing pain and, consequently, achieving reliable results and good performance. Furthermore, additional specific personal, contextual, or health-related factors (e.g., age, sex, physical activity level, type of disability) can significantly impact the physiological parameters used to develop automatic pain assessment methods. Thus, our models may not be robust enough to properly assess pain should these factors not be adequately controlled.

10.5 Conclusion

In this paper, we presented a protocol to evaluate the feasibility of developing automatic methods for pain assessment in Persons with Multiple Sclerosis based on physiological signals and AI algorithms. In addition, we illustrated the intervention by highlighting the state-of-the-art and innovative tools to obtain reliable and robust methods for automatic pain assessment. Such an approach, if proven feasible, can lead to significant progress in the field of pain management by providing a better characterization of pain and, therefore, more timely and efficient interventions to control it.

11. PILOT STUDY ON PAINLESS PROTOCOL

11.1 Introduction

The PAINLESS protocol presented in the previous chapter has been approved by Comitato Etico Area Vasta Emilia Centro (285-2022-SPER-AUSLBO). Data collection officially started in January 2023 and it is still ongoing. Data from the first three participants were collected and analyzed to conduct a pilot study and assess protocol feasibility. Preliminary results will be reported in this chapter along with the information related to these three case-studies. Our preliminary results were used to implement a practical pipeline to optimize the data collection phase.

11.2 Materials and Methods

11.2.1 Participants

Participants are Multiple Sclerosis (MS) patients enrolled in a rehabilitation program at Ospedale Bellaria (Bologna). In the following, the main characteristics of the three participants are presented.

Participant A

Participant A is a 60 years old woman, diagnosed with MS 10 years ago, with an Expanded Disaibility Status Scale (EDSS) score of 4. The type of MS is relapsing-

remitting. At pathology onset, participant A presented neuropathic pain episodes, recurrent headache, pathologic fatigue and muscle exhaustion, along with cognitive fatigue on long-term visuo-spatial memory and concentration deficit. At the last visit, the pathologic fatigue persisted, but she performed well in coordination tests, although presenting gait with widened base, stiffness and uncertainty in changes of direction and the impossibility of performing tandem march. She also presents some associated pathology: extrinsic asthma, autoimmune thyroiditis, and disc herniation in the lumbar region.

Participant B

Participant B is a 53 years old woman, diagnosed with MS 5 years ago, with an EDSS score of 6.5. The type of MS is secondary progressive. At pathology onset, participant B presented undifferentiated connectivitis because of diffuse polyarticular pain associated with fatigue. During the last visit, participant B walked with a 4-wheel walker, presented trigger points at the right scapulohumeral, sacro iliac and knees, and gait was performed with widened base.

Participant C

Participant C is a 60 years old woman, diagnosed with MS 31 years ago, with an EDSS score between 1 and 3.5. The type of MS is relapsing-remitting. At pathology onset, and 8 years ago presented right lower limb stiffness and occasionally right upper limb stiffness. At the last visit, participant C presented widespread pain, especially joint pain, and complained of fatigue and insomnia. After the COVID-19 infection, participant C began to have an intense headache and musculoskeletal pain, along with right sciatica with worsening gait.

11.2.2 Experimental procedure

We applied the pipeline described at 10.3.2 – *Experimental pipeline* for each participant. On admission, after having completed the CRF 1 – Recruitment questionnaire, CRF 2 – Sleep-wake questionnaire, and CRF 3 – Stratification questionnaire (administered with the $\text{smartphone}_{\text{project}}$), participant were provided with the Empatica E4 wristband, to be worn on the non-dominant arm, and a smartphone ($\text{smartphone}_{\text{participant}}$), through which fulfill the CRF 4 – Monitoring questionnaire during the following 24 hours. Reminders

to fulfill the questionnaire were set based on the answers obtained from CRF 2 – Sleep-wake questionnaire.

After 24 hours, each participant returned to the hospital one hour before the scheduled neurorehabilitation session. Physiological signals recordings were downloaded from the E4 wristband. Both the E4 wristband and smartphone were put on charge.

Just before the beginning of the neurorehabilitation session, each participant was administered the CRF 3 – Stratification questionnaire and was asked to wear again the E4 wristband. During the neurorehabilitation session, the participant had to answer three questions composing the CRF 5 – Monitoring-treatment questionnaire (through $\text{smartphone}_{\text{project}}$) about the experienced pain every 10 minutes. At the end of the neurorehabilitation session, the participant fulfilled again the CRF 3 – Stratification questionnaire and was given the smartphone before leaving the hospital. After 24 hours, the participant returned to the hospital, fulfilling the CRF 3 -. Stratification questionnaire for the last time and returns the devices.

11.2.3 Signal and data analysis

For the purpose of this pilot study, we proceeded with the analyses whose pipeline is shown in Figure 11.1.

From the E4 wristband, the following procedures were applied:

- Photoplethysmograhy (PPG): after being subjeted to the preprocessing pipeline presented in Appendix A.2, the basic and high-quality classifiers shown in Chapter 3 was applied to obtain basic and high-quality pulses, to be further exploited to obtain heart rate variability (HRV) features, basic and higher morphological features. At least 10 good quality PPG pulses are needed to extract HRV features
- Electrodermal Activity (EDA): after being subjected to the preprocessing pipeline presented in Appendix A.4, the quality check of Taylor et al. [355] was applied. Then features from the full EDA and the tonic and phasic componentswere extracted by applying the algorithm by Greco et al. [246].
- Skin Temperature (SKT): this signal was not subjected to any preprocessing phase, it was used to extract statistical features directly

- Accelerometer (ACC): it was subjected to the procedure described in Appendix A.3 to obtain the Activity Index (AI), which was then further exploit to extract statistical features

A list of the features is given in Appendix F, Table F.1.

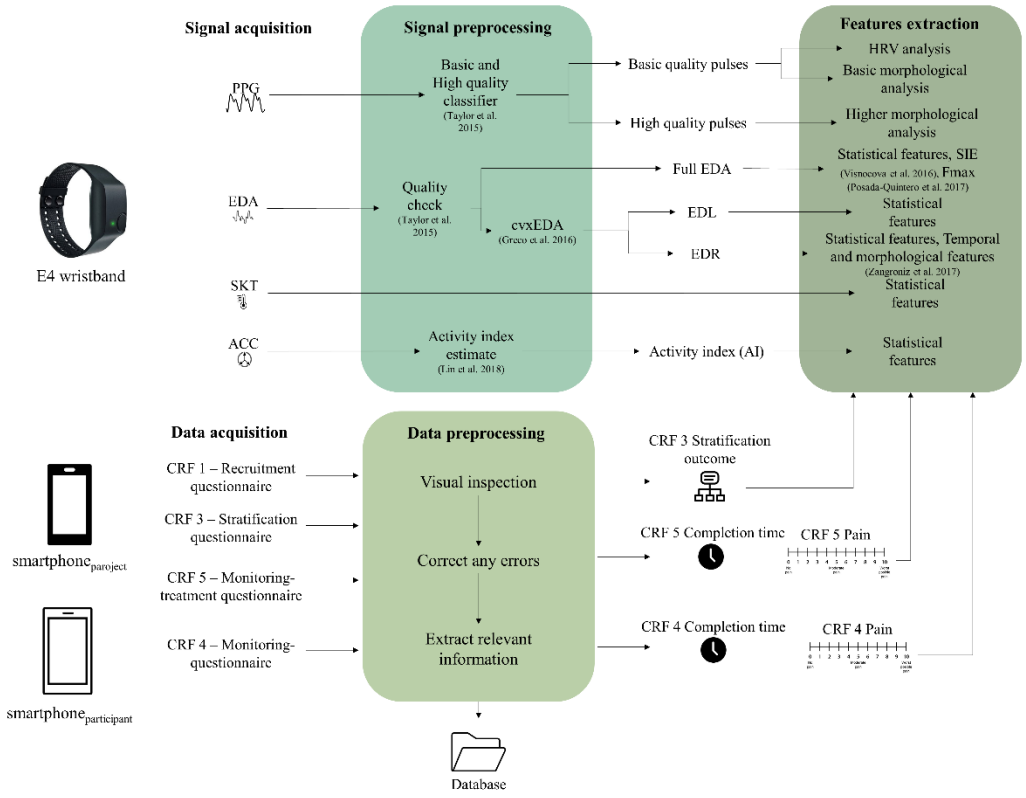


Figure 11.1 Signal and data analysis pipeline

All the questionnaires (CRF 1, CRF 3 and CRF 5 fulfilled through $\text{smartphone}_{\text{project}}$ and CRF 4 fullfiled through $\text{smartphone}_{\text{participant}}$) were downloaded and subjected to a three-step data preprocessing: visual inspection, errors correction, and information extraction. Information extracted in this pilot study was:

- CRF 3 Stratification outcome
- CRF 5 completion time and pain intensity
- CRF 4 completion time and pain intensity

CRF 3 Stratification outcomes were used to qualitatively compare different groups (*i.e.*, absence of pain, nociceptive pain, neuropathic pain) in a representative time series for each physiological signals (*meanHR* for the PPG, EDA, SKT, and AI themselves for the remaining signals) throughout the two monitoring days.

Information from both CRF 4 and CRF 5 werere used to extract features in time windows synchronized with completion time. Time window length used in this pilot study is 5 minutes, with the end corresponding to CRF4 and CRF 5 completion time. Other time window lengths will be assessed when more patients are enrolled. Features were then qualitatively analysed in relation to pain intensity. As an example, in this work the following features were used along with pain intensity ratings given through CRF4 and CRF5:

- PPG: *meanHR*, standard deviation of normal heartbeats (*SDNN*), and pulse amplitude (*PulseAmpl*)
- EDA: *meanEDA*, mean amplitude of phasic component's peaks (*meanampEDR*), and mean of EDA's tonic component (*meanEDL*),
- SKT: mean skin temperature (*meanTEMP*)
- ACC: mean AI (*meanAI*)

11.3 Results

11.3.1 Participants

Information about participants regarding the monitoring time with E4 wristband, CRF3 Stratification outcome, number of instances collected through CRF 4 – Monitoring questionnaire and CRF 5 – Monitoring-treatment questionnaire are reported in Table 11.1. Figure 11.2 and 11.3 show the histogram of the pain intensity ratings given respectively through CRF 4 – Monitoring questionnaires and CRF 5 – Monitoring-treatment questionnaies.

Each participant was monitored for at least 22 hours, and fulfilled the CRF 4 questionnaire at least 19 times throughout the whole experimental procedure. Neurorehabilitation treatment lasted about 1 hour per participant, obtaining 4 CRF 5 instances for each participant from this period.

Participant A showed contrasting results from the CRF 3, while participants B and C had the same results (neuropathic and nociceptive respectively) for all four times the CRF 3 was administered.

Table 11. 1 Descriptive statistics on the information obtained by participants

	Monitoring Time (hh:mm)		CRF 3 Stratification outome	CRF 4 # instances		CRF 5 # instances
	Day 1	Day 2		Day 1	Day 2	
Participant A	22:49	24:07	2 neuropathic, 2 nociceptive	7	13	4
Participant B	23:17	23:57	4 neuropathic	6	14	4
Participant C	22:51	23:59	4 nociceptive	10	14	4

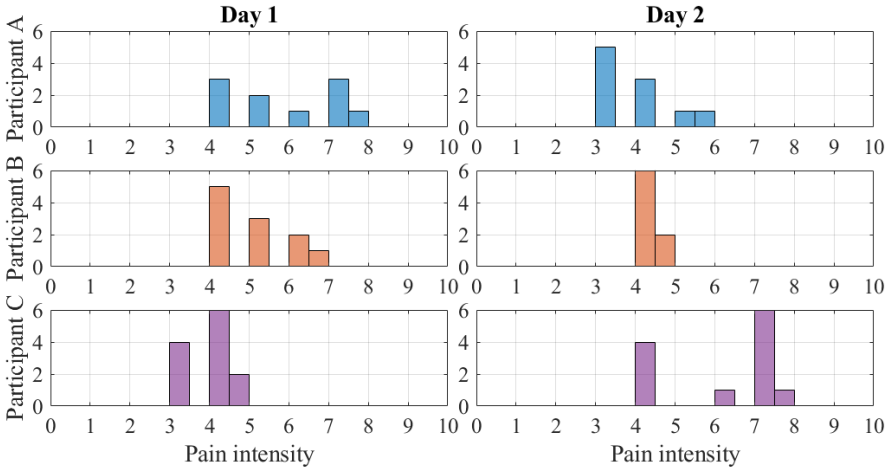


Figure 11.2 Histogram of pain intensity ratings from CRF 4 - Monitoring questionnaires

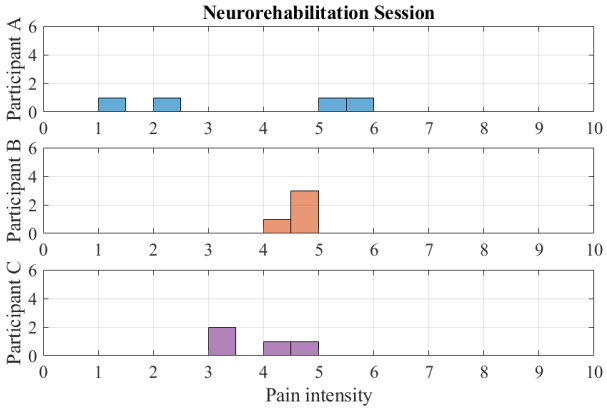


Figure 11.3 Histogram of pain intensity ratings from CRF 5 – Monitoring-treatment questionnaires

11.3.2 Signals and data analysis

11.3.2.1 CRF3 Stratification outcome

Since participant A showed inconsistent results as CRF 3 Stratification outcome, we did not consider these data in this qualitative analysis.

Figure 11.3 shows the behavior of meanHR, EDA, SKT, and AI throughout the whole monitoring period. All the physiological signals are in the same values range. Notably, a remarkable *meanEDA* peak for participant C (who experienced nociceptive pain)can be seen during Day 1. Another noteworthy detail is given by the SKT, with a negative peak around the 10th hour of Day 1: this is probably due to the removal of the E4 wristband from the wrist.

From this qualitative analysis, any consideration can be drawn based on the CRF3 Stratification outcome since only two participants are involved and the differences in the physiological signals cannot be ascribed solely to the type of experienced pain.

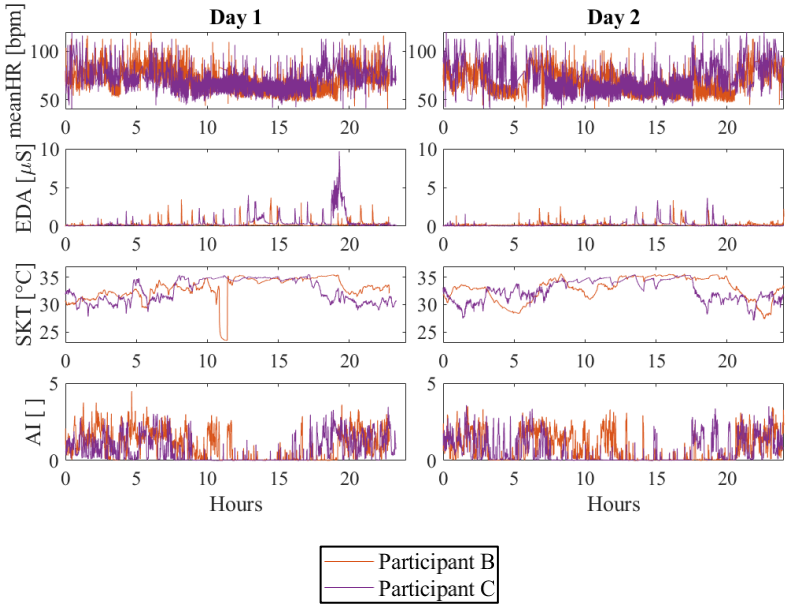


Figure 11.4 Physiological signals throughout the monitoring period

11.3.2.2 CRF 4 Pain intensity

Figure 11.5, 11.6, and 11.7 show examples of features extracted respectively from PPG, EDA, SKT and ACC along with the pain intensity ratings given when fulifilling the CRF 4 – Monitoring questionnaire. In particular for PPG features, since the signal is subjected to a signal quality check, it might be possible that there is any good quality PPG pulse for a certain time window, so features cannot be estimated.

11.3.2.3 CRF 5 Pain intensity

All three participants underwent a neurorehabilitation session. Physiological recordings of participant C are compromised because the investigator left in place the USB dock used to charge the device, preventing the proper recordings of PPG and SKT signals (the USB dock covers these sensors) and a misplace of the device on the wrist. For this reason,

signals from participant C during the neurorehabilitation sessions are not presented. Figure 11.8, 11.9, and 11.10 show examples of features extracted respectively from PPG, EDA, SKT and ACC during the neurorehabilitation session, along with the pain intensity ratings given when administered with the CRF 5 – Monitoring – treatment questionnaire. It is worth noting that Participant B does not show any *meanHR* or *SDNN* value because of the lack of at least 10 good-quality PPG pulse.

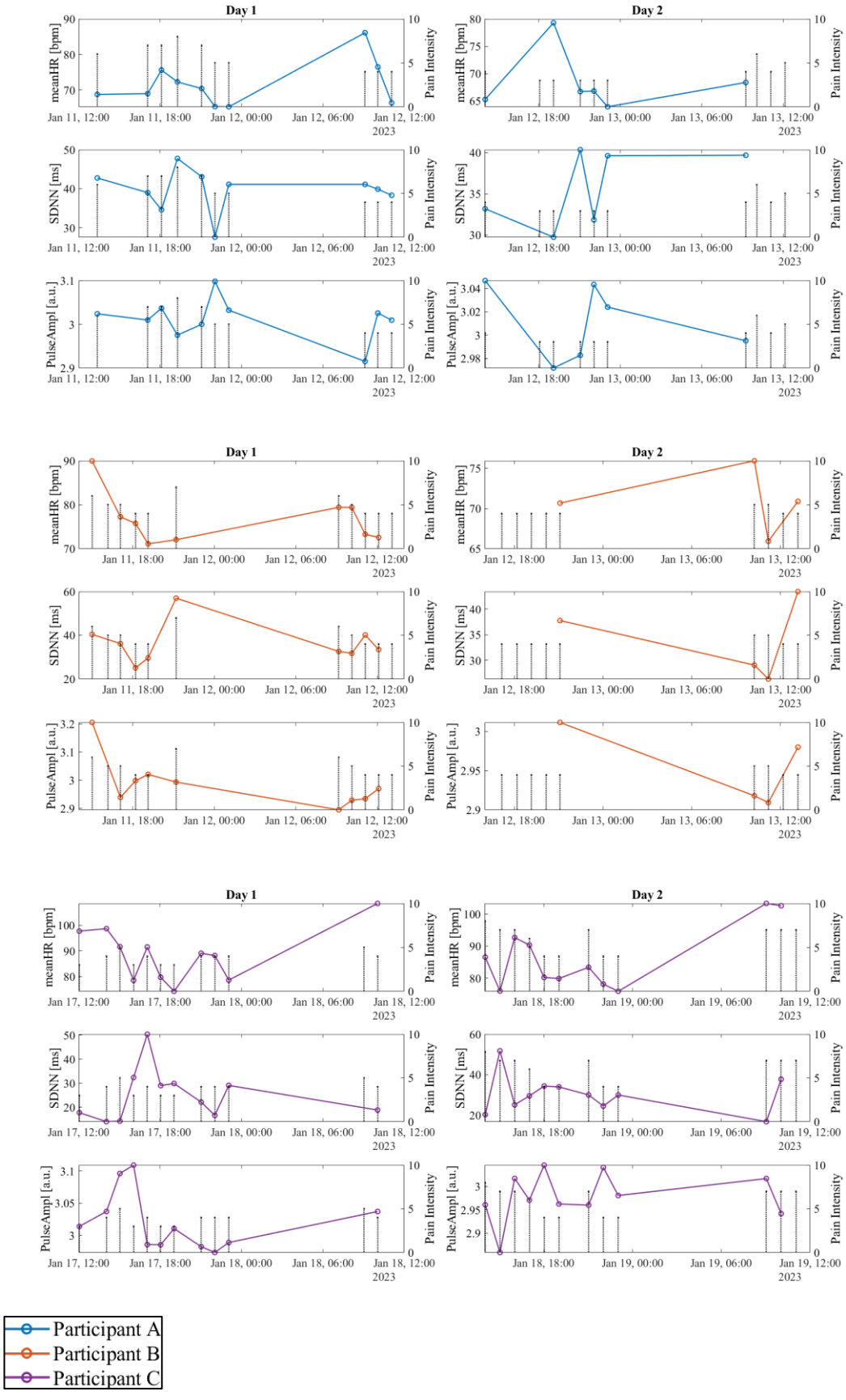


Figure 11.5 PPG features trend during the monitoring period

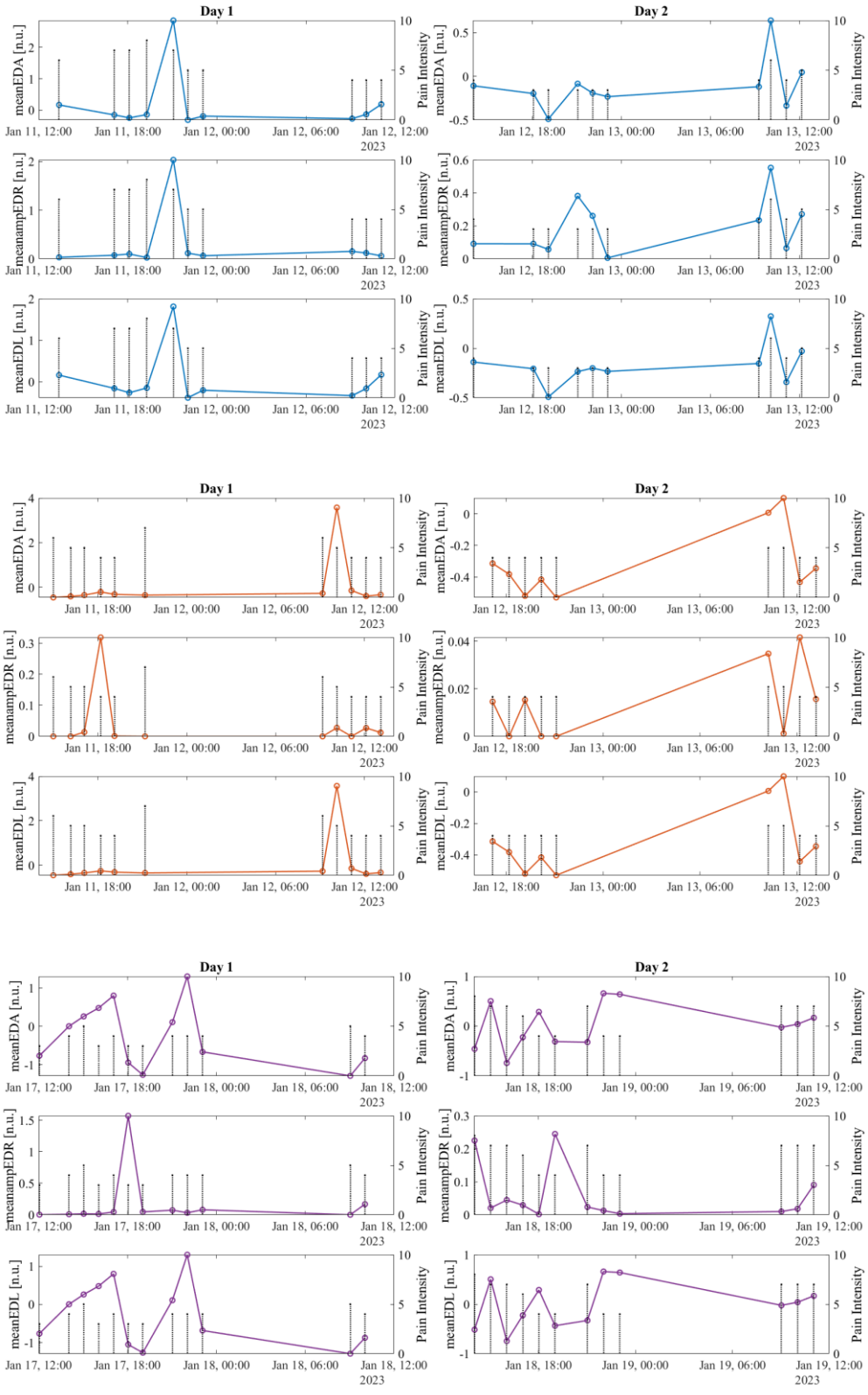


Figure 11.6 EDA features trend during the monitoring period

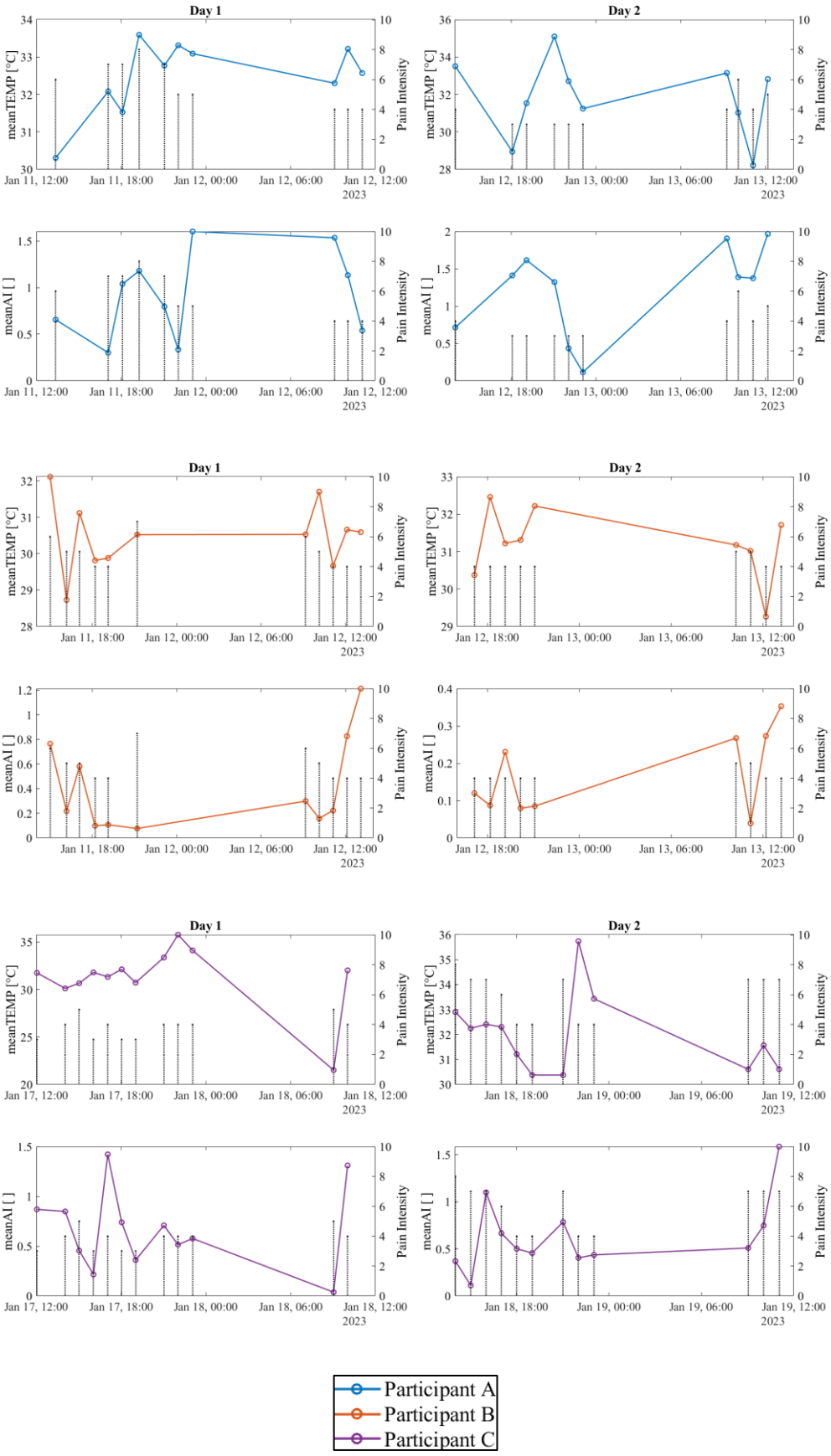


Figure 11.7 TEMP and ACC features trend during the monitoring period

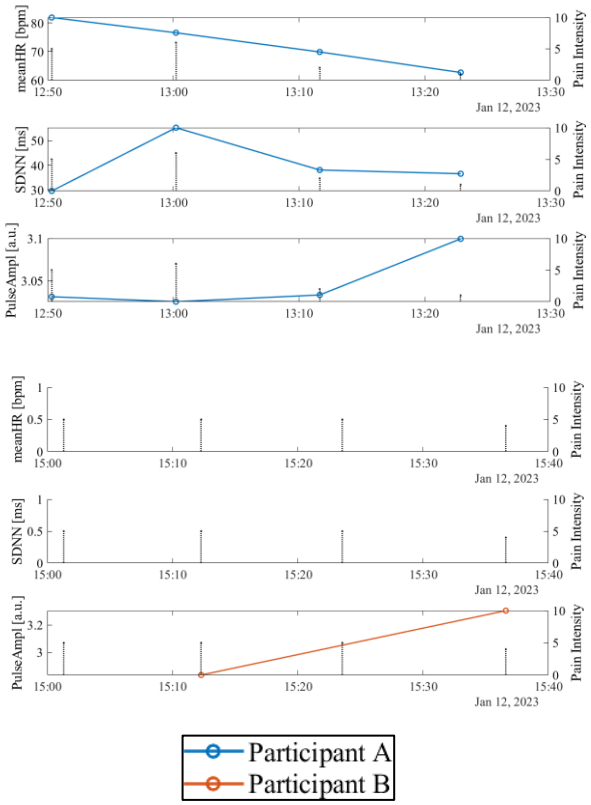


Figure 11.8 PPG features trend during neurorehabilitation session

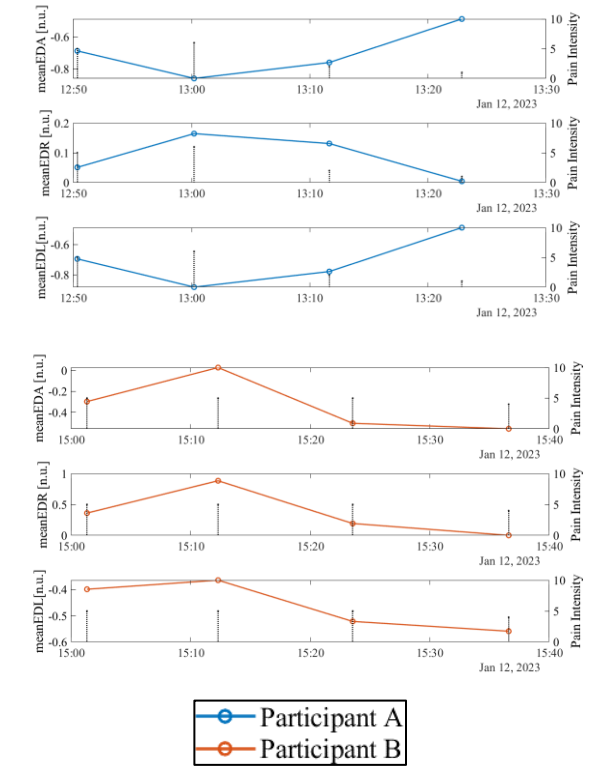


Figure 11.9 EDA features trend during neurorehabilitation session

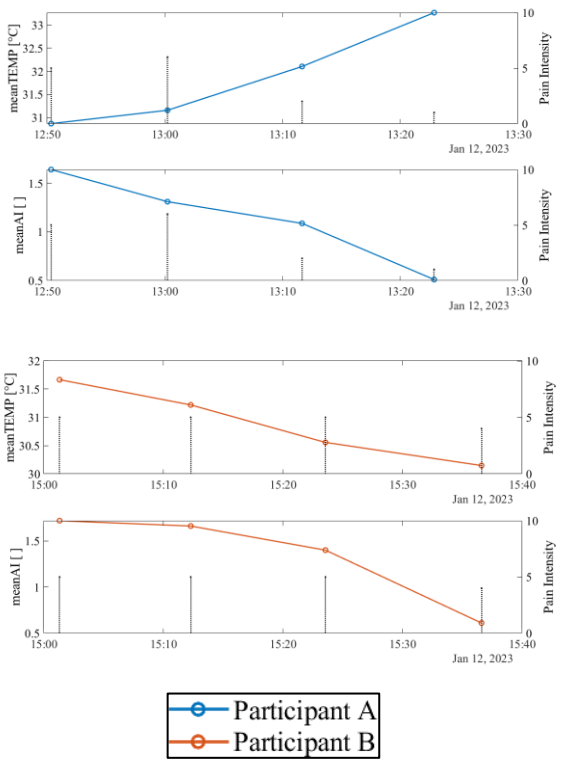


Figure 11.10 SKT and ACC features trend during neurorehabilitation session

11.3 Discussion

In this pilot study, we investigated the feasibility of the PAINLESS clnical study protocol in order to implement a practical pipeline to opimize the data collection phase.

We retrieved information from the questionnaires, administered to participants or self-administered by participants themselves, to highlight possible errors encountered during the practical application of the protocol. Questionnaires' outcomes were also used to link some of them to physiological parameters extracted from the E4 wristband recordings to assess their behavior qualitatively. In this case, we were also interested in exploring the availability of finite physiological parameter values coupled with pain intensity ratings, fundamental elements for developing classifiers and regressors to estimate pain intensity automatically.

From the CRF 3 – Stratification questionnaire, administered four times during the whole experimental procedure, two participants (B and C) were consistently classified as experiencing neuropathic and nociceptive pain respectively. In comparison, participant A was classified as experiencing neuropathic pain for the first two times, and neuropathic

pain for the last times the questionnaire was administsered. The CRF 3 questionnaire was conceived as the union between two neuropathic pain screening questionnaires, painDETECT and DN4, with the aim of reducing the possibility of inconsistent results. This aspect should be deepened to avoid excessive data removal, with the risk of not having enough data to train and test the classifiers to discern between nociceptive and neuropathic pain.

By analysing the physiological signals recorded throughout the two days of the experiment, we noticed that Participant B removed the E4 wristband thanks to a drop in the skin temperature. When processing signals, a prodromic phase must be devoted to a visual inspection to remove possible "non-wear" periods.

From the CRF 4 – Monitoring questionnaire we obtained between 19 and 24 instances, in line with the numerosity we expected. Participants were all able to self-administered the questionnaire through the smartphone. Regarding the coupling between pain intensity ratings from CRF 4 and physiological signals, it can be appreciated from Figures 11.5 that PPG parameters are not present for every pain intensity ratings. This is due to a lack of good-quality signals from which physiological parameters can be reliably extracted. A strategy to improve the possibility of finding good quality signal segments is to enlarge the time window from which the physiological parameters can be extracted. The trade-off between the availability of good-quality PPG pulses and the accuracy of pain estimation (it may happen that a wider time window is not representative of the pain rating) will be explored.

From the CRF 5 – Monitoring-treatment questionnaire, we collected 4 instances per participant, in line with the duration of the neurorehabilitation session (about 1 hour). We had a technical error by leaving the USB dock for participant C in place, thus loosing physiological recordings. Regarding Figure 11.8, we observed a lack of finite HRV parameters values. We will deal with this aspect by applying a similar strategy described for CRF 4 – Monitoring questionnaire, although in this case we are limited in the time window length to 10 minutes (since CRF 5 is administered every 10 minutes in this case). Overall, the protocol was demonstrated to be feasible. This pilot study also allowed to promptly identify critical issues that will be better controlled in the continuation of the data collection phase. As an output of this preliminary study, we drafted a practical pipeline presented in Appendix F, Table F.2.

CONCLUSION

The genereal aim of this PhD thesis was the development of automatic methods to asssess pain by using physiological signals and artificial intelligence algorithms. This aim was set to meet the clinical demand of implementing tools that provide a timely, personalized, and reliable characterization of the pain experience.

To pursue this aim, three practical objectives were set.

The first objective was to select and/or develop preprocessing methods to condition physiological signals and control confounding factors appropriately. I developed a signal quality check method to detect basic and high-quality photoplethysmography pulses based on characteristics extracted from the photoplethysmographic signal itself and from accelerometer data and machine learning algorithms. Such an approach has reached very high accuracy levels, above 90%, together with well-balanced sensitivity and specificity values. This means that the detection of good-quality pulses exploitable for further analysis is well-optimized. Still related to the preprocessing phase, I conducted two studies to characterize the impact of several confounding factors on some photoplethysmography parameters and different physiological signals recorded in a real-world context. For the first study, I found that several personal, contextual, and health factors play a role in modifying both the percentage of available basic and high-quality photoplethysmography pulses and the values of the extracted physiological parameters. For the second study, involving oncological patients only, it turned out that the circadian rhythms of heart rate, electrodermal activity, skin temperature, and physical activity

strongly depend on pain, but also on other factors that can hinder the possibility of reliably assessing pain, such as anxiety, depression, use of opiods, and sex.

These studies revealed the need to apply rigorous methods for signal preprocessing, the basis on which the method of pain assessment rests, and to employ proper methods to control possible confounding factors.

The second objective was to define an ideal pipeline for objective pain assessment. By combining the information gathered about the signals in Section I, their related processing algorithms, and artificial intelligence techniques, along with the findings reported in Section II, it was possible to implement a general strategy to implement pain assessment methods. A depiction is given in Figure III.

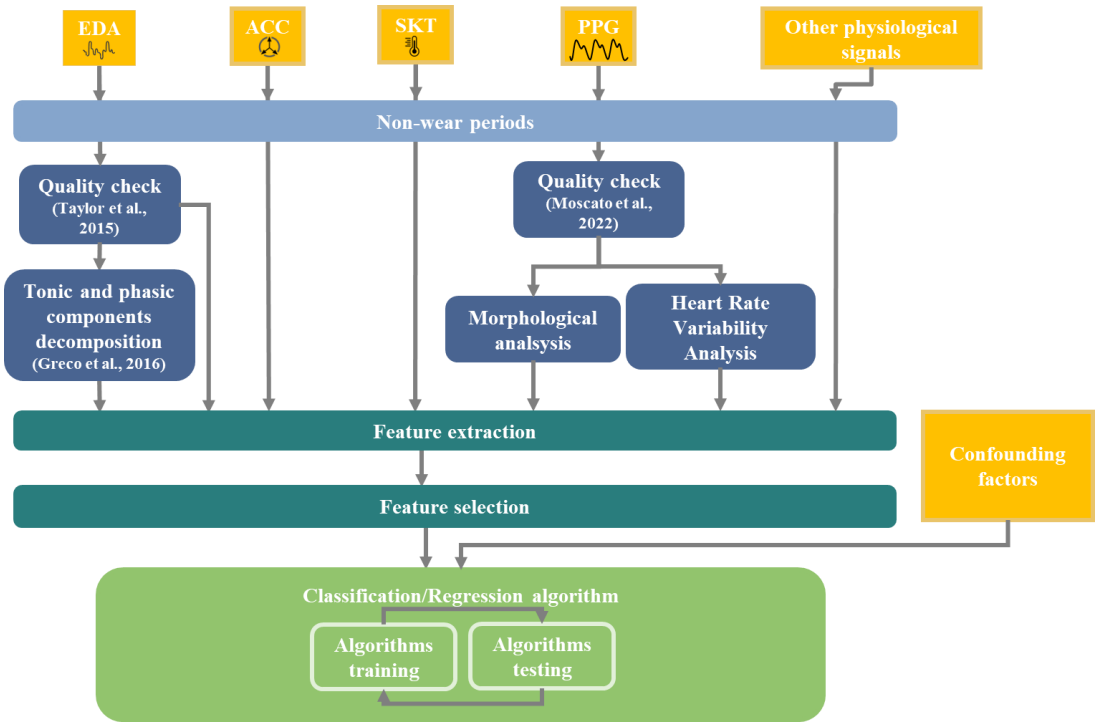


Figure III Ideal pipeline for automatic pain assessment

Such strategy is meant to be then modified on the specific case, taking into account the type of pain to be investigated and some practical considerations, such as the number of available instances. The higher the numerosity of the sample, the more features and confounding factors can be included to tailor the model even better.

The third objective was to apply the ideal pipeline in different health domains. In this thesis, three different pain conditions have been explored.

I investigated the physiological reaction to different noxious stimulations on healthy subjects and patients suffering from chronic low back pain. This study made it possible to appreciate significant differences in the basal physiological activity and in response to

different external noxious stimuli in the two populations. This study corroborated the finding that chronic pain causes substantial changes in the autonomic nervous system, presenting a higher basal sympathetic activity and a blunted autonomic response when subjected to nociceptive stimuli. Thus, based on these results, I trained and tested several machine learning algorithms to automatically classify healthy subjects and chronic low back pain patients. I tested the inclusion of a different number of features extracted from different time window lengths from recordings under rest conditions. This approach was feasible, although performance can be increased by increasing the number of subjects involved in the study.

Delving into cancer pain, I conducted a systematic review to gather studies on the physiological reaction to cancer pain. Despite the marked heterogeneity of the included studies, it was still possible to conclude that some parameters related to the cardiovascular system, such as the heart rate and systolic blood pressure, positively correlates with pain. However, the most important conclusion drawn from this study was the need to implement more rigorous and big-data-based methods to assess pain in the oncological population. Using these findings as a basis, I conducted a study involving cancer patients whose physiological signals were monitored in a real-world context with a wearable device. Such physiological recordings were used along with pain intensity ratings given by the patients to train and test machine learning algorithms to detect the absence and presence of pain. Such an approach was demonstrated to be feasible, but also in this case more instances would be needed to reach better performance.

Lastly, I presented the protocol for a clinical trial based on pain assessment in patients undergoing neurorehabilitation by using wearable devices. For this exploratory study, the research team chose to include only multiple sclerosis patients among the whole cohort of neurorehabilitation patients to reduce the variability possibly introduced by the pathology itself. A pilot study on three patients has been carried out to define a practical pipeline to optimize the data collection process. Aside from a few technical problems, these preliminary results were promising and in line with what we theoretically expected regarding the number of instances to be collected.

In conclusion, all the presented studies showed encouraging outcomes in applying such an approach to assess pain through physiological signals and artificial intelligence algorithms. Further studies on larger populations are needed to consolidate better the results presented here. Nevertheless, such methodology can enrich the routine clinical

pain assessment procedure, by making it possible to monitor pain when and where it is actually experienced by using wearable devices, and in general without the involvement of explicit communication. This, in turn, would enable better characterization of the pain experience, improve antalgic therapy personalization, and bring timely relief, with the ultimate goal of ameliorating the quality of life of patients suffering from pain.

APPENDICES

APPENDIX A

A.1 Empatica E4 wristband

The Empatica E4 wristband [356], presented in Figure A.1, is a CE medical-grade device that allows the continuous, simultaneous recording of several physiological signals. It has the following technical specifications:

- Photoplethysmography (PPG) sensor is equipped with four light sources (two green, two red) and two photodetectors, and with a sampling frequency of 64 Hz.
- Electrodermal activity (EDA) is recorded by two dry electrodes placed on the ventral part of the wrist, with a sampling frequency of 4 Hz.
- Skin temperature (TEMP) is recorded through a thermopile with a sampling frequency of 4 Hz.
- Accelerometer data is recorded through a tri-axial accelerometer with a sampling frequency of 32 Hz.

Figure A.1 Empatica E4 wristband

A.2 PPG processing

A fifth-order Butterworth band-pass filter with cut-off frequencies of 0.5 and 12 Hz is applied for each PPG recording [147]. The algorithm by Elgendi et al. [357], originally developed to detect second derivative PPG fiducial points, was adapted to detect the systolic peak and systolic foot of each pulse in order to segment the signal into single pulses. Each pulse was then normalized with the z-score procedure:

$$\text{pulse}_{\text{norm}} = \frac{\text{pulse} - \text{mean(pulse)}}{\text{std(pulse)}} \quad \text{(A.1)}$$

A.3 ACC processing

Each accelerometer (ACC) component (x, y, z) is resampled at $fs_{ACC-RES}$=64 Hz with linear interpolation (to match the PPG sampling frequency) and converted to g units. Next, a fourth-order band-pass filter was applied, with cut-off frequencies of 0.025 and 10 Hz [358], [359]. The ACC vector magnitude was then calculated for each sample j as:

$$A_j = \sqrt{{ACC_{x_j}}^2 + {ACC_{y_j}}^2 + {ACC_{z_j}}^2} \quad \text{(A.2)}$$

The A_{ind} was estimated using the algorithm of Lin et al. [41]:

- Standard deviation of A_j for 5-second epochs:

$$\sigma = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(A_j - \mu)^2} \quad \text{(A.3)}$$

  Where

  $\mu = \frac{1}{N}(A_1 + A_2 + \cdots + A_N)$

  $\text{N} = 5\ \text{s} * \text{fs}_{\text{ACC-RES}}$

- Minute-wise A_{ind}:

$$A_{ind} = \sum_{k=1}^{M}\sigma_k \quad \text{(A.4)}$$

  where M is set to 12 to obtain a minute-wise A_{ind} by summing 12 5-second epochs.

A.4 EDA processing

EDA is filtered with a 4th order Butterworth low-pass filter, with a cut-off frequency of 1 Hz and then subjected to a quality check analysis with the automatic algorithm by Taylor et al. [355]. Bad quality segments were replaced by a linear interpolation between the last and the following good quality segment.

APPENDIX B

Table B.1 Computational complexity for each feature. N = pulse's length

Feature	Computational complexity
Peak2peakACC	1
MeanACC	N
SigSim	N
Entropy	7*N
Kurtosis	4*N
SNR	8*N
RelPower	2*N
Skewness	6*N
ZR	3
Amplitude	1
Width	1
TroughDepth	1
MedianPulse	N
MedianPulsenoZ	N
MeanPulse_noZ	N
StdPulse_noZ	4*N
SNR_Moody	8
Npeaks	1
ZDR	N

Table B.2 Results from neighborhood component analysis for the Basic-quality classifier (BQ) applied ten times

Repetition	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
Peak2peakACC	x	x	x	x	x	x	x	x	x	x	100%
MeanACC											0%
SigSim	x	x	x	x	x	x	x	x	x	x	100%
Entropy											0%
Kurtosis						x	x				20%
SNR		x				x	x				30%
RelPower							x				10%
Skewness							x				10%
ZR							x				10%
Amplitude											0%
Width						x	x				20%
TroughDepth	x	x	x	x	x	x	x	x	x	x	100%
MedianPulse		x	x	x	x	x	x		x	x	80%
MedianPulsenoZ											0%
MeanPulse_noZ											0%
StdPulse_noZ	x		x	x	x	x	x	x	x	x	90%
SNR_Moody	x	x	x	x		x	x	x	x	x	90%
Npeaks	x	x	x	x	x	x	x	x	x	x	100%
ZDR	x	x	x	x	x	x	x	x	x	x	100%

Table B.3 Results from neighborhood component analysis for the Type 1 High-quality classifier (HQ1) applied ten times

Repetition	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
Peak2peakACC	x		x	x	x	x	x	x	x	x	90%
MeanACC											0%
SigSim	x	x	x	x	x	x	x	x	x	x	100%
Entropy											0%
Kurtosis	x		x	x	x	x	x	x	x	x	90%
SNR	x		x	x	x		x		x		60%
RelPower	x	x	x	x	x	x	x	x	x	x	100%
Skewness	x	x	x	x	x	x	x	x	x	x	100%
ZR											0%
Amplitude											0%
Width											0%
TroughDepth	x		x	x	x						40%

MedianPulse		x		x	x	x	x	x	x	x	x	90%
MedianPulsenoZ												0%
MeanPulse_noZ												0%
StdPulse_noZ		x	x	x	x	x	x	x	x	x	x	100%
SNR_Moody				x	x	x	x	x		x	x	70%
Npeaks		x	x	x	x	x	x	x	x	x	x	100%
ZDR		x	x	x	x	x	x	x	x	x	x	100%

Table B.4 Results from neighborhood component analysis for the Type 2 High-quality classifier (HQ2) applied ten times

Repetition	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
Peak2peakACC		x					x				20%
MeanACC		x	x			x	x		x		50%
SigSim		x				x	x		x		40%
Entropy	x	x	x	x	x	x	x	x	x	x	100%
Kurtosis	x	x	x	x	x	x	x	x	x		90%
SNR		x					x		x		30%
RelPower	x	x	x	x	x	x	x	x	x	x	100%
Skewness	x	x	x	x	x	x	x	x	x	x	100%
ZR		x					x				20%
Amplitude		x					x		x		30%
Width		x					x		x		30%
TroughDepth											0%
MedianPulse	x	x	x	x	x	x	x	x	x		90%
MedianPulsenoZ		x					x		x		30%
MeanPulse_noZ		x									10%
StdPulse_noZ	x	x	x	x	x	x	x	x	x	x	100%
SNR_Moody	x	x	x	x	x	x	x	x	x	x	100%
Npeaks	x	x	x	x	x	x	x	x	x	x	100%
ZDR	x	x	x	x	x	x	x	x	x	x	100%

Table B.5 Hyperparameters for the Basic-quality classifiers

Algorithms	Hyperparameters	All features	SQIs selection
Tree	Maximum number of splits	39	38
	Split criterion	Gini's diversity index	Gini's diversity index
Naïve Bayes (NB)	Distribution names	Kernel	Kernel
	Kernel type	Gaussian	Gaussian
Support Vector Machine (SVM)	Kernel function	Quadratic	Gaussian
	Kernel scale	1	27.494
	Box constraints	0.025078	119.112
	Standardize data	True	False
K-nearest neighborhood (KNN)	Number of neighbors	10	5
	Distance metrics	Correlation	Chebyshev
	Distance weight	Inverse	Squared iinverse
	Standardize data	True	False
Ensemble	Ensemble method	GentleBoost	Bag
	Maximum number of splits	21	1025
	Number of learners	400	285
	Learning rate	0.0093026	-
Neural Network	Number of fully connected layers	2	1
	Activation function	Sigmoid	Tanh
	Regularization strength	1.18E-06	5.45E-05
	Standardize data	No	No
	1st layer size	5	3
	2nd layer size	204	-
	3rd layer size	-	-
Elgendi 2016 (SVM)	Kernel function	Gaussian	
	Kernel scale	0.48469	
	Box constraints	0.015755	
	Standardize data	True	

Table B.6 Hyperparameters for the Type 1 High-quality classifiers

Algorithms	Hyperparameters	All features	SQIs selection
Tree	Maximum number of splits	39	38
	Split criterion	Gini's diversity index	Gini's diversity index
Naïve Bayes (NB)	Distribution names	Kernel	Kernel
	Kernel type	Gaussian	Gaussian
Support Vector Machine (SVM)	Kernel function	Quadratic	Gaussian
	Kernel scale	1	27.494
	Box constraints	0.025078	119.112
	Standardize data	True	False
K-nearest neighborhood (KNN)	Number of neighbors	10	5
	Distance metrics	Correlation	Chebyshev
	Distance weight	Inverse	Squared iinverse
	Standardize data	True	False
Ensemble	Ensemble method	GentleBoost	Bag
	Maximum number of splits	21	1025
	Number of learners	400	285
	Learning rate	0.0093026	-
Neural Network	Number of fully connected layers	2	1
	Activation function	Sigmoid	Tanh
	Regularization strength	1.18E-06	5.45E-05
	Standardize data	No	No
	1st layer size	5	3
	2nd layer size	204	-
	3rd layer size	-	-
Elgendi 2016 (SVM)	Kernel function	Gaussian	
	Kernel scale	0.48469	
	Box constraints	0.015755	
	Standardize data	True	

Table B.7 Hyperparameters for the Type 2 High-quality classifiers

Algorithms	Hyperparameters	All features	SQIs selection
Tree	Maximum number of splits	69	53
	Split criterion	Gini's diversity index	Gini's diversity index
Naïve Bayes (NB)	Distribution names	Gaussian	Gaussian
	Kernel type	Epanechnikov	Triangle
Support Vector Machine (SVM)	Kernel function	Linear	Quadratic
	Kernel scale	1	1
	Box constraints	0.8367	55.859
	Standardize data	False	False
K-nearest neighborhood (KNN)	Number of neighbors	3	39
	Distance metrics	City block	Chebyshev
	Distance weight	Squared inverse	Squared inverse
	Standardize data	False	False
Ensemble	Ensemble method	GentleBoost	Bag
	Maximum number of splits	494	223
	Number of learners	116	480
	Learning rate	0.44184	-
Neural Network	Number of fully connected layers	3	1
	Activation function	Tanh	Sigmoid
	Regularization strength	4.02E-06	2.81E-09
	Standardize data	False	False
	1st layer size	2	4
	2nd layer size	1	-
	3rd layer size	175	-
Elgendi 2016 (SVM)	Kernel function	Quadratic	
	Kernel scale	1	
	Box constraints	0.0017795	
	Standardize data	False	

APPENDIX C

C.1 Look of Life project

Look of Life is a project designed by Fondazione ANT and whose aim is to evaluate the effects of Virtual Reality (VR) as a non-pharmacological intervention to improve cancer-related symptomatology (e.g., pain, anxiety, depression).

C.2 Experimental protocol

Cancer patients assisted by Fondazione ANT who accepted to participate in this study are provided with a VR headset to be freely used at home for four days. Questionnaires and scales are administered to patients to evaluate the effects of VR:

- At the beginning (T0) and at the end (T1) of the observational period, the Hospital Anxiety and Depression Scale (HADS) and the Brief Pain Inventory (BPI) have been administered
- Before and after each VR session, the patient has to fill the Edmonton Symptoms Assessment Scale (ESAS), digitally inserted within the VR headset

Patients are also asked to wear the Empatica E4 wristband to collect physiological signals that can be used to evaluate the effects of the VR.

Figure B.1 gives a visual representation of the experimental protocol.

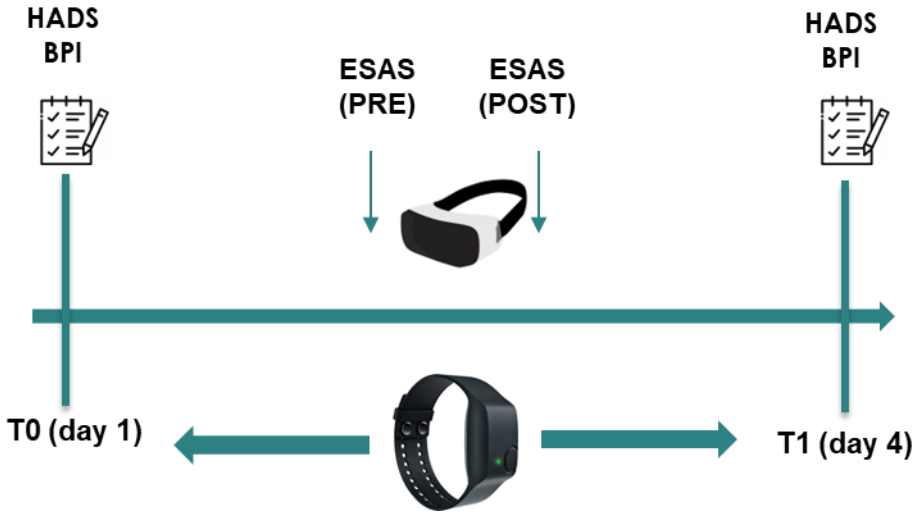


Figure C.1 Look of Life experimental protocol

C.3 Socio-demographic, clinical data, and questionnaires,

The gathered socio-demographic and clinical data are the following:

- Sex
- Civil status
- Cohabitation
- Occupation
- Education
- Confidence with technology
- Cancer diagnosis
- Karnofsky Performance Score
- Extension of pathology
- Therapies

In the following a detailed description of the questionnaires used in this study:

- Hospital Anxiety and Depression Scale (HADS): it is a simple and brief scale consisting of 14 items exploring both anxiety and depression symptoms. The main characteristic of the scale is that it excludes the somatic symptomatology from the evaluation. HADS allows to detect three different levels of anxiety and depression, separately: normal, borderline, abnormal.
- Brief Pain Inventory (BPI): it is an easy-to-use questionnaire assessing the intensity of pain, the interference of pain with the patient's life, pain relief, pain quality, and patient perception of the cause of pain.

- Edmonton Symptom Assessment Scale: it is a reliable multi-item instrument developed to rate the intensity of nine common symptoms experienced by cancer patients (pain, tiredness, nausea, depression, anxiety, drowsiness, appetite, well-being, and shortness of breath.

APPENDIX D

Table D.1 Results from Repeated pinpricks test - Healthy controls

	Pre-tets	**1st-5th rep**	**5th-10th rep**	**Post-test**	**p-value**					
	mean (std)	**mean (std)**	**mean (std)**	**mean (std)**	**Pre-test - 1st-5th rep**	**Pre-test - 5th-10th rep**	**Pre-test – Post-test**	**1st-5th rep - 5th-10th rep**	**1st-5th rep – Post-test**	**5th-10th rep – Post-test**
meanIBI [ms]	807.89 (125.40)	841.93 (122.36)	883.35 (117.97)	826.21 (132.24)	0.208	0.018	0.934	0.303	0.008	0.001
SDNN [ms]	39.85 (18.27)	38.40 (17.89)	45.83 (17.53)	53.34 (30.96)	0.890	0.599	0.135	0.151	0.135	0.639
RMSSD [ms]	50.68 (29.20)	48.30 (27.68)	58.30 (26.97)	63.61 (40.55)	0.934	0.277	0.389	0.169	0.083	0.720
SD1 [ms]	32.41 (20.05)	33.01 (19.82)	40.07 (19.19)	39.56 (27.85)	0.639	0.229	0.524	0.169	0.303	0.934
SD2 [ms]	30.39 (16.00)	33.88 (18.40)	40.09 (18.40)	40.61 (26.79)	0.720	0.135	0.151	0.252	0.489	0.847
PulseAmpl [a.u.]	3.10 (0.30)	3.12 (0.35)	3.21 (0.32)	3.20 (0.38)	0.762	0.489	0.208	0.524	0.252	0.847
A1 [a.u.*s]	71.75 (30.40)	67.70 (23.72)	64.56 (17.49)	81.98 (68.27)	0.229	0.679	1.000	0.978	0.720	0.720
A2 [a.u.*s]	256.08 (39.74)	261.83 (39.21)	270.90 (37.33)	259.31 (55.60)	0.421	0.252	0.303	0.121	0.599	0.389
A [a.u.*s]	327.82 (60.93)	329.52 (41.91)	225.45 (41.20)	341.29 (55.08)	0.561	0.489	0.489	0.639	0.599	0.599
T1 [ms]	158.31 (46.45)	151.87 (48.27)	145.04 (27.80)	168.57 (95.40)	0.498	0.296	0.715	0.498	1.000	0.679
T2 [ms]	667.04 (104.71)	686.57 (106.65)	706.39 (101.79)	646.90 (146.47)	0.720	0.135	0.815	0.421	0.035	0.003
meanEDA [n.u.]	0.79 (0.58)	1.19 (0.59)	1.34 (0.52)	1.39 (0.62)	0.001	0.008	0.010	0.303	0.489	0.599
stdEDA [n.u.]	0.15 (0.15)	0.16 (0.11)	0.11 (0.06)	0.10 (0.08)	0.762	0.679	0.277	0.135	0.015	0.389

SIE [ ]	0.70 (0.39)	0.69 (0.44)	0.75 (0.30)	0.74 (0.35)	0.804	0.934	0.847	0.761	0.847	0.978
meantonic [n.u.]	0.68 (0.63)	1.05 (0.57)	1.19 (0.52)	1.30 (0.66)	0.002	0.005	0.010	0.421	0.303	0.252
stdtonic [n.u.]	0.13 (0.14)	0.16 (0.11)	0.10 (0.07)	0.08 (0.07)	0.489	0.804	0.454	0.095	0.003	0.073
meanampEDR [n.u.]	0.12 (0.25)	0.18 (0.23)	0.17 (0.29)	0.08 (0.14)	0.080	0.330	0.735	1.000	0.110	0.268
stdampEDR [n.u.]	0.02 (0.05)	0.03 (0.06)	0.06 (0.10)	0.01 (0.02)	0.492	0.123	0.461	0.232	0.160	0.014
freqEDR [#peaks/min]	16.80 (6.09)	14.18 (6.45)	14.81 (5.98)	16.80 (7.59)	0.352	0.421	1.000	0.761	0.454	0.551
meanRR	23.04 (4.25)	18.98 (4.13)	19.14 (3.80)	18.86 (2.77)	0.054	0.193	0.008	0.787	1.000	1.000

Table D.2 Results from Repeated pinpricks test - cLBP patients

	Pre-tets	**1st-5th rep**	**5th-10th rep**	**Post-test**	**p-value**					
	mean (std)	**mean (std)**	**mean (std)**	**mean (std)**	**Pre-test - 1st-5th rep**	**Pre-test - 5th-10th rep**	**Pre-test – Post-test**	**1st-5th rep - 5th-10th rep**	**1st-5th rep – Post-test**	**5th-10th rep – Post-test**
meanIBI [ms]	736.68 (114.50)	696.85 (105.38)	695.07 (116.07)	705.71 (139.08)	0.469	0.297	0.688	0.578	0.688	0.813
SDNN [ms]	44.14 (25.50)	33.84 (34.19)	33.94 (33.18)	22.70 (20.03)	0.375	0.375	0.031	0.578	0.375	0.375
RMSSD [ms]	59.21 (31.07)	41.00 (38.36)	35.06 (23.40)	26.69 (18.04)	0.375	0.109	0.031	0.688	0.219	0.297
SD1 [ms]	32.18 (24.82)	28.69 (26.85)	22.91 (14.67)	17.22 (11.41)	1.000	0.219	0.047	0.688	0.219	0.297
SD2 [ms]	29.52 (27.68)	33.89 (38.79)	35.02 (40.05)	18.05 (22.20)	0.938	0.813	0.156	0.578	0.219	0.219
PulseAmpl [a.u.]	3.32 (0.20)	3.21 (0.33)	3.24 (0.29)	3.25 (0.26)	0.297	0.297	0.375	1.000	0.578	0.469
A1 [a.u.*s]	70.63 (25.96)	55.90 (9.53)	55.85 (5.61)	67.06 (23.94)	0.109	0.469	0.688	0.938	0.297	0.578
A2 [a.u.*s]	206.00 (63.57)	181.81 (41.17)	180.24 (43.95)	176.80 (61.10)	0.109	0.078	0.219	1.000	0.375	0.297
A [a.u.*s]	276.64 (65.19)	237.71 (42.30)	236.09 (46.28)	243.86 (60.79)	0.047	0.031	0.219	1.000	0.813	0.813
T1 [ms]	180.40 (73.95)	137.99 (20.50)	149.14 (28.34)	171.05 (61.56)	0.109	0.688	0.813	0.156	0.297	0.813
T2 [ms]	567.84 (122.49)	560.06 (99.24)	545.67 (98.77)	548.06 (106.24)	0.938	0.219	0.469	0.219	0.219	0.813
meanEDA [n.u.]	0.34 (1.16)	0.56 (1.48)	0.49 (1.35)	0.33 (1.14)	1.000	1.000	0.688	0.688	0.375	0.219

stdEDA [n.u.]	0.03 (0.06)	0.06 (0.13)	0.08 (0.15)	0.04 (0.05)	0.297	0.156	0.688	0.375	0.578	0.469
SIE [ ]	0.89 (0.13)	0.99 (0.06)	0.64 (0.40)	0.62 (0.38)	0.156	0.219	0.297	0.078	0.016	0.813
meantonic [n.u.]	0.32 (1.14)	0.53 (1.43)	0.49 (1.34)	0.33 (1.14)	1.000	0.938	0.813	0.688	0.375	0.219
stdtonic [n.u.]	0.04 (0.08)	0.07 (0.15)	0.07 (0.14)	0.04 (0.05)	0.578	0.156	0.688	0.938	0.688	0.469
meanampEDR [n.u.]	0.02 (0.06)	0.04 (0.08)	0.003 (0.01)	0.003 (0.0.01)	0.250	1.000	1.000	0.250	0.375	1.000
stdampEDR [n.u.]	0.02 (0.05)	0.02 (0.05)	0.001 (0.004)	0.005 (0.008)	1.000	1.000	1.000	1.000	1.000	1.000
freqEDR [#peaks/min]	15.43 (9.07)	10.20 (3.81)	8.99 (2.73)	13.71 (4.54)	0.016	0.031	1.000	0.203	0.016	0.031
meanRR	21.94 (9.32)	20.27 (4.56)	22.56 (1.53)	25.11 (2.37)	1.000	1.000	1.000	0.625	0.031	0.187

Table D.3 Results from the Inflated cuff - pressure pain test – Healthy Control subjects

	Pre-test	30 sec	60 sec	90 sec	120 sec	After	p-value														
	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	Pre-test – 30 sec	Pre-test – 60 sec	Pre-test – 90 sec	Pre-test – 120 sec	Pre-test – Post-test	30 sec – 60 sec	30 sec – 90 sec	30 sec – 120 sec	30 sec - Post-test	60 sec – 90 sec	60 sec – 120 sec	60 sec - Post-test	90 sec – 120 sec	90 sec - Post-test	120 sec – Post-test
meanIBI [ms]	860.94 (93.03)	855.23 (117.66)	857.85 (118.05)	869.01 (146.63)	869.62 (126.03)	837.87 (112.70)	0.626	0.903	0.952	0.542	0.058	0.326	0.358	0.217	0.463	0.326	0.173	0.241	0.808	0.091	0.025
SDNN [ms]	54.78 (18.08)	58.28 (15.36)	53.18 (19.57)	58.11 (19.42)	50.44 (15.07)	52.75 (20.87)	0.217	0.358	0.463	0.326	0.358	0.091	1.000	0.091	0.217	0.241	0.626	1.000	0.058	0.135	0.502
RMSSD [ms]	66.38 (27.74)	65.23 (25.98)	62.11 (32.98)	67.85 (38.61)	59.98 (33.76)	58.79 (31.83)	0.626	0.296	1.000	0.194	0.217	0.626	0.761	0.326	0.426	0.326	0.903	0.583	0.217	0.326	1.000
SD1 [ms]	46.88 (19.61)	46.06 (18.37)	43.85 (23.35)	47.94 (27.28)	42.39 (23.87)	40.89 (22.14)	0.626	0.296	1.000	0.194	0.194	0.626	0.761	0.326	0.358	0.326	0.903	0.502	0.241	0.326	0.952
SD2 [ms]	58.78 (19.19)	65.10 (16.57)	57.76 (19.47)	63.01 (17.86)	53.50 (13.40)	54.43 (18.48)	0.173	0.761	0.358	0.268	0.326	0.119	0.583	0.035	0.091	0.217	0.358	0.715	0.042	0.135	0.670
PulseAmpl [a.u.]	3.21 (0.25)	3.28 (0.22)	3.13 (0.23)	3.16 (0.20)	3.17 (0.23)	3.19 (0.25)	0.104	1.000	0.042	0.104	0.003	0.296	1.000	0.855	0.003	0.173	0.078	0.005	0.583	0.000	0.001
A1 [a.u.*s]	72.07 (22.38)	65.91 (13.71)	73.76 (16.44)	77.50 (25.84)	75.97 (30.37)	70.91 (22.11)	0.020	0.011	0.194	0.296	0.626	0.001	0.001	0.001	0.049	0.583	0.808	0.358	0.542	0.463	0.502
A2 [a.u.*s]	268.01 (42.71)	273.30 (47.53)	270.56 (40.79)	271.95 (53.03)	269.12 (55.84)	254.90 (42.86)	0.268	1.000	0.296	0.326	0.952	0.035	0.007	0.009	0.049	0.153	0.583	0.903	0.903	0.358	0.426
A [a.u.*s]	340.07 (47.85)	339.21 (44.08)	344.32 (46.98)	349.45 (63.51)	345.09 (58.97)	325.81 (47.71)	0.583	0.542	0.761	0.808	0.030	0.952	0.952	0.626	0.058	0.903	0.808	0.025	0.426	0.058	0.058
T1 [ms]	160.74 (37.74)	151.83 (28.71)	164.93 (28.69)	170.71 (41.79)	167.40 (46.81)	163.15 (38.55)	0.952	0.583	0.426	0.268	0.042	0.391	0.391	0.715	0.194	0.903	0.808	0.091	0.358	0.020	0.007
T2 [ms]	680.83 (73.95)	690.01 (104.40)	681.21 (95.32)	690.52 (127.78)	702.43 (122.13)	665.65 (101.29)	0.017	0.903	0.194	0.268	0.761	0.002	0.001	0.002	0.017	0.194	0.542	0.463	0.463	0.296	0.583
meanEDA [n.u.]	-0.02 (0.49)	0.76 (0.78)	0.34 (0.59)	0.14 (0.66)	-0.01 (0.63)	-0.08 (0.61)	0.808	1.000	1.000	0.463	0.296	0.735	1.000	0.715	0.173	0.808	0.241	0.391	0.358	0.241	0.042
stdEDA [n.u.]	0.07 (0.12)	0.33 (0.24)	0.13 (0.11)	0.11 (0.12)	0.09 (0.01)	0.08 (0.09)	0.000	0.058	0.670	0.808	0.583	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.002

SIE [ ]	0.50 (0.38)	0.88 (0.12)	0.66 (0.33)	0.68 (0.38)	0.62 (0.34)	0.86 (0.18)	0.000	0.009	0.104	0.194	0.268	0.000	0.000	0.000	0.000	0.135	0.013	0.001	0.626	0.296	0.241
meantonic [n.u.]	-0.05 (0.52)	0.64 (0.68)	0.30 (0.57)	0.11 (0.63)	-0.04 (0.61)	-0.10 (0.60)	0.000	0.042	0.626	0.808	0.626	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.007
stdtonic [n.u.]	0.05 (0.08)	0.30 (0.23)	0.11 (0.09)	0.09 (0.08)	0.08 (0.05)	0.06 (0.05)	0.000	0.001	0.078	0.104	0.104	0.000	0.000	0.000	0.000	0.296	0.017	0.001	0.502	0.173	0.217
meanampEDR [n.u.]	0.05 (0.18)	0.19 (0.34)	0.08 (0.16)	0.06 (0.16)	0.05 (0.12)	0.04 (0.07)	0.000	0.064	0.461	0.945	0.496	0.003	0.001	0.001	0.000	0.557	0.037	0.064	0.578	0.734	0.820
stdampEDR [n.u.]	0.004 (0.01)	0.11 (0.13)	0.03 (0.06)	0.03 (0.06)	0.04 (0.08)	0.04 (0.10)	0.001	0.250	0.063	0.188	0.219	0.001	0.014	0.010	0.014	0.688	0.844	0.156	0.688	0.938	0.438
freqEDR [#peaks/min]	13.71 (4.36)	8.63 (6.10)	4.39 (4.30)	5.80 (5.24)	3.80 (4.41)	6.26 (3.74)	0.007	0.000	0.001	0.000	0.001	0.001	0.001	0.002	0.119	0.715	0.670	0.068	0.358	0.761	0.009
meanRR	24.45 (4.95)	17.59 (3.65)	17.74 (2.68)	16.93 (5.16)	17.78 (2.82)	18.91 (2.66)	0.125	0.125	0.250	0.125	0.125	0.670	0.626	0.670	0.104	0.715	0.855	0.104	1.000	0.241	0.078

Table D.4 Results from the Inflated cuff - pressure pain test – cLBP patients

	Pre-test	30 sec	60 sec	90 sec	120 sec	Post-test	p-value														
	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	Pre-test – 30 sec	Pre-test – 60 sec	Pre-test – 90 sec	Pre-test – 120 sec	Pre-test – Post-test	30 sec – 60 sec	30 sec – 90 sec	30 sec – 120 sec	30 sec - Post-test	60 sec – 90 sec	60 sec – 120 sec	60 sec - Post-test	90 sec – 120 sec	90 sec - Post-test	120 sec – Post-test
meanIBI [ms]	698.78 (129.34)	722.21 (154.87)	711.28 (140.26)	699.52 (127.60)	698.19 (122.31)	721.97 (93.38)	0.11	0.69	1.00	0.578	0.813	0.219	0.109	0.156	0.813	0.219	0.297	0.813	0.813	1.000	0.688
SDNN [ms]	36.65 (31.59)	35.42 (29.62)	35.31 (27.74)	34.95 (29.46)	28.57 (23.73)	38.64 (28.94)	0.94	1.00	0.47	0.031	0.938	0.688	0.938	0.031	0.813	0.813	0.016	0.688	0.297	0.688	0.469
RMSSD [ms]	43.07 (27.86)	42.44 (36.78)	43.10 (35.19)	42.18 (36.86)	32.64 (28.24)	52.03 (36.68)	0.81	0.94	0.30	0.109	0.938	1.000	0.688	0.219	0.688	0.578	0.109	0.469	0.375	0.219	0.078
SD1 [ms]	30.45 (26.77)	29.97 (25.95)	30.46 (24.88)	29.82 (26.07)	23.06 (19.96)	36.56 (25.68)	0.81	0.94	0.30	0.109	0.938	1.000	0.813	0.219	0.688	0.578	0.109	0.469	0.375	0.219	0.078
SD2 [ms]	41.00 (35.89)	38.65 (31.63)	37.01 (30.10)	37.81 (31.14)	32.32 (27.04)	34.81 (33.24)	0.94	0.58	0.58	0.156	0.156	0.938	0.813	0.016	0.578	0.375	0.016	0.469	0.219	0.813	0.813
PulseAmpl [a.u.]	3.28 (0.25)	3.27 (0.24)	3.25 (0.24)	3.24 (0.23)	3.25 (0.21)	3.24 (0.29)	0.02	0.11	0.30	0.219	0.031	0.078	0.156	0.078	0.031	0.688	0.813	0.031	0.375	0.031	0.031
A1 [a.u.*s]	63.93 (25.84)	59.11 (18.17)	67.43 (25.23)	63.94 (18.67)	61.10 (17.13)	72.09 (24.79)	0.47	0.37	0.30	0.375	0.375	0.578	0.375	0.469	0.375	0.375	0.813	0.688	0.688	0.938	0.938
A2 [a.u.*s]	177.13 (41.33)	189.23 (56.07)	195.57 (51.63)	191.77 (53.09)	186.17 (45.26)	181.51 (39.83)	0.22	0.81	0.69	0.813	0.688	0.297	0.016	0.219	0.078	1.000	1.000	0.578	0.578	0.469	0.297
A [a.u.*s]	241.06 (58.72)	248.35 (62.62)	263.00 (61.24)	255.71 (60.73)	247.28 (52.85)	253.60 (43.20)	0.16	0.08	0.22	0.156	0.375	0.219	0.938	0.813	0.688	0.297	0.016	0.219	0.219	0.375	0.578
T1 [ms]	157.16 (45.89)	146.05 (33.26)	157.54 (46.51)	149.51 (33.19)	146.14 (27.58)	181.51 (73.22)	0.69	0.16	0.22	0.375	0.469	0.219	0.813	0.813	0.578	0.078	0.031	0.578	0.156	0.688	0.938
T2 [ms]	549.30 (99.78)	575.95 (128.86)	555.51 (116.53)	554.55 (113.27)	560.69 (109.75)	546.71 (94.34)	0.11	0.47	0.81	0.219	0.688	0.938	0.813	0.938	0.219	0.688	1.000	0.297	0.469	0.156	0.156
meanEDA [n.u.]	0.03 (0.54)	0.56 (0.90)	0.17 (0.69)	0.07 (0.92)	-0.10 (1.03)	-0.27 (0.94)	0.11	0.37	0.69	0.109	0.938	0.688	0.078	0.375	0.375	0.938	0.813	0.688	0.688	0.813	0.375

stdEDA [n.u.]	0.11 (0.21)	0.33 (0.33)	0.12 (0.07)	0.21 (0.36)	0.09 (0.0'9)	0.05 (0.04)	0.109	0.578	0.813	0.469	0.219	0.031	0.219	0.156	0.078	0.297	0.297	0.156	0.047	0.016	0.047
SIE [ ]	0.59 (0.44)	0.74 (0.27)	0.57 (0.37)	0.48 (0.40)	0.43 (0.38)	0.82 (0.36)	0.297	0.469	0.469	0.813	0.938	0.047	0.297	0.109	0.016	0.297	0.375	0.016	0.078	0.047	0.156
meantonic [n.u.]	-0.78 (2.36)	0.51 (80.90)	0.16 (0.68)	-0.01 (0.76)	-0.10 (1.03)	.0.29 (0.96)	0.688	0.938	0.688	0.156	0.688	0.469	0.297	0.109	0.688	0.031	0.375	0.219	1.000	0.297	0.047
stdtonic [n.u.]	0.30 (0.71)	0.36 (0.30)	0.12 (0.07)	0.20 (0.33)	0.09 (0.09)	0.07 (0.06)	0.078	0.219	0.578	0.938	0.938	0.031	0.156	0.156	0.109	0.297	0.297	0.156	0.297	0.078	0.047
meanampEDR [n.u.]	1.19 (3.13)	0.10 (0.015)	0.02 (0.04)	0.21 (0.53)	0.001 (0.002)	0.06 (0.10)	0.469	0.469	0.578	0.688	0.813	0.016	0.297	0.078	0.016	0.297	0.375	0.031	0.078	0.578	0.813
stdampEDR [n.u.]	0 (0)	0.07 (0.10)	0.01 (0.02)	0.20 (0.53)	0 (0)	0.04 (0.07)	0.813	0.875	0.750	0.500	0.875	0.438	0.813	0.250	0.625	0.500	0.250	0.625	0.500	0.875	0.250
freqEDR [#peaks/min]	12 (0)	5.43 (4.95)	3.71 (4.37)	3.64 (3.74)	1.96 (0.20)	6.59 83.23)	0.250	1.000	0.500	1.000	0.500	0.375	0.813	0.250	0.625	0.500	1.000	0.500	0.500	1.000	0.500
meanRR	23.44 (4.76)	18.90 (4.46)	18.39 (5.37)	17.73 (4.28)	18.93 (4.35)	20.59 (4.62)	0.750	1.000	0.250	0.750	0.750	0.469	0.297	0.469	0.844	0.813	0.375	0.313	0.297	0.063	0.063

APPENDIX E

Table E.1 Glossary of physiological signals and computed parameters

Physiological signal	Physiological parameter acronym	Explanation	Unit of measure
Electrocardiogram (ECG) / Photoplethysmography (PPG)	HR	Heart Rate	bpm
	SDNN	Standard deviation of NN intervals	ms
	SDAAN	Standard deviation of the mean of NN intervals	ms
	SDDNindex	Mean of the standard deviation of NN intervals	[ ]
	rMSSD	Root mean square of successive differences	ms
	pNN50	Percentage of successive NN intervals that differ from one another by > 50 milliseconds	%
	Total power	Total power	ms^2
	LF	Low Frequency	ms^2
	HF	High Frequency	ms^2
	LF/HF	Low frequency/High frequency ratio	[ ]
	Log LF/HF	Logarithm of Low Frequency/High Frequency ratio	[ ]
	Deep breathing ratio	Ratio of the longest RR interval over the shortest RR interval during deep breathing test	[ ]
	Valsalva ratio	Ratio of the longest RR interval after Valsalva manoeuvre over the shortest RR interval during the manoeuvre	[ ]

Photoplethysmography (PPG)	SpO2	Amount of oxygen-carrying hemoglobin in the blood relative to the amount of non-oxygen-carrying hemoglobin	%
Electrochemical Skin Conductance (ESC)	Hands ESC	Hands Electrochemical Skin Conductance	µS
	Feet ESC	Feet Electrochemical Skin Conductance	µS
Blood Pressure (BP)	SysBP	Systolic Blood Pressure	mm Hg
	DiaBP	Diastolic Blood Pressure	mm Hg
	MAP	Mean Arterial Pressure	mm Hg
Respiration	RR	Respiration Rate	breaths/min

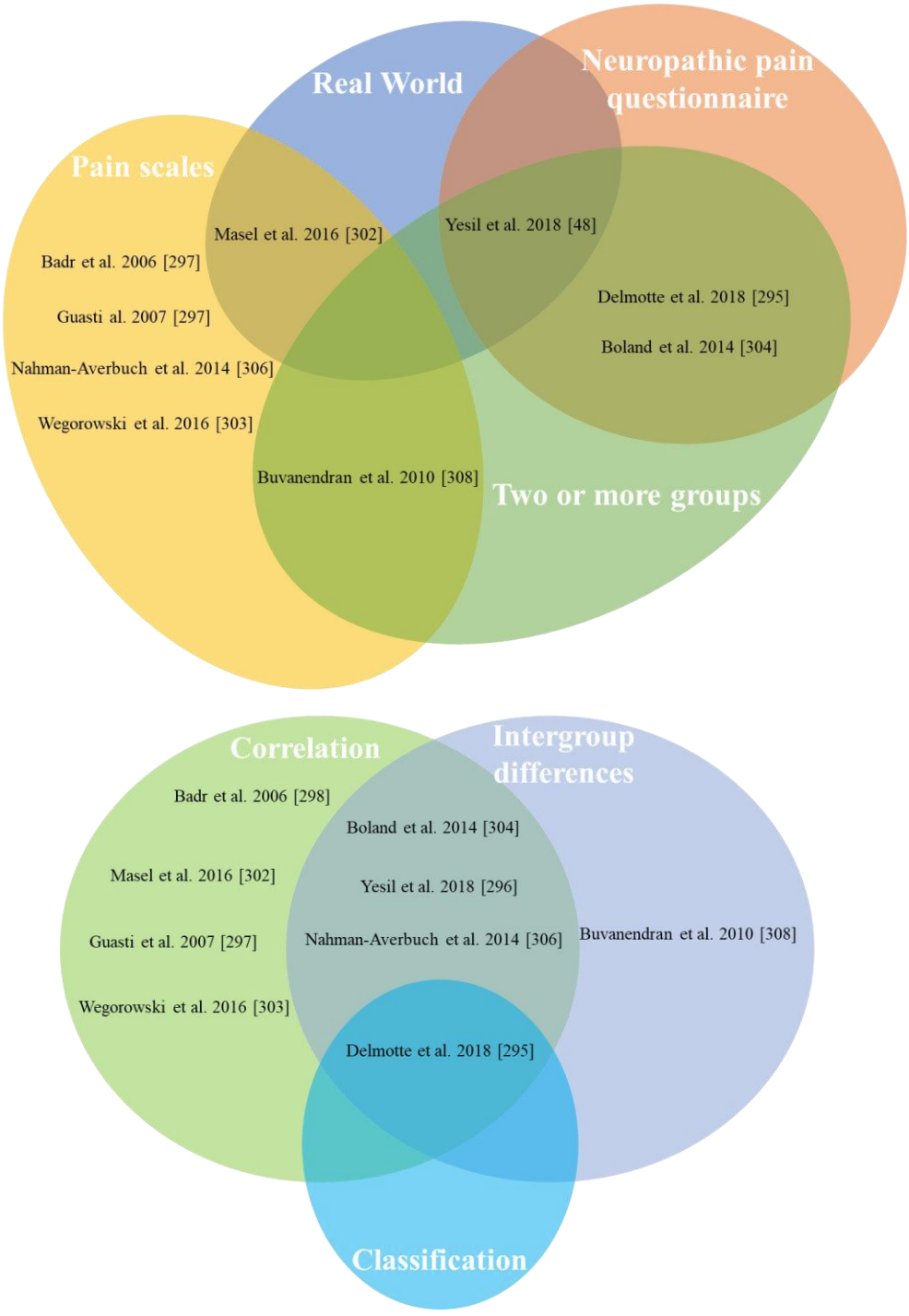


Figure D.1 Graphical representation of the salient features of concurrent validity studies

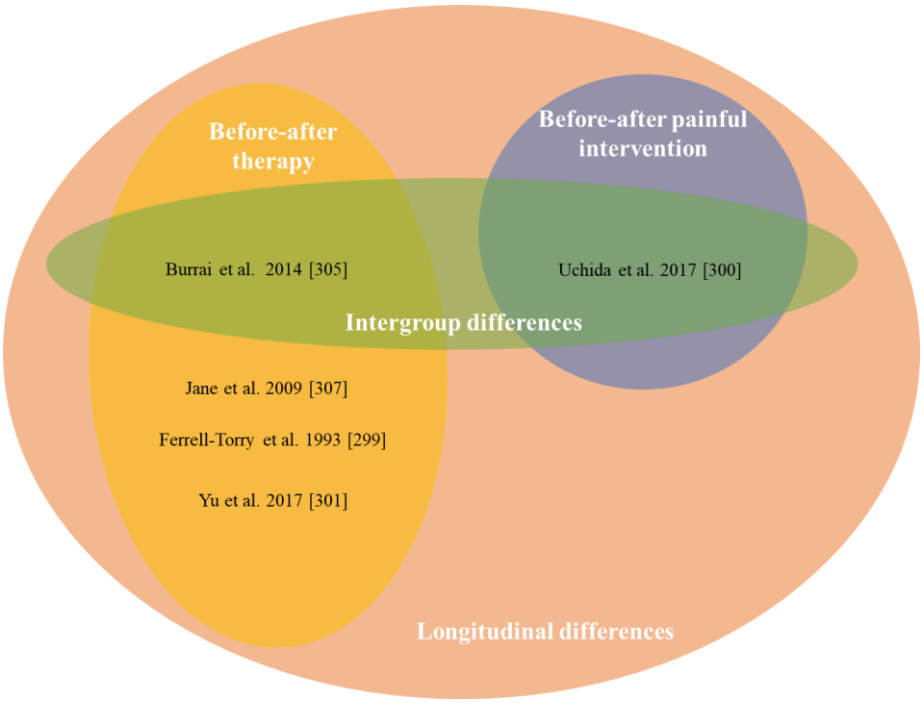


Figure D.2 Graphical representation of the sensitivity to change studies

APPENDIX F

Table F.1 Physiological signals recorded with Emaptica E4 wristband and related features

Signal	Analysis	Acronym	Description	u.o.m
PPG	Heart Rate Variaibility Analysis	HR	Heart Rate	[bpm]
		RMSSD	Root Mean Squared of Successive Differences	[ms]
		SDNN	Standard Deviation of Normal Heartbeats	[ms]
		pNN50	Percentage of successive heartbeats that differ more than 50 ms	%
		LF	Absolute power of the low frequency band [0.04-0.15] Hz	[ms^2]
		HF	Absolute power of the high frequency band [0.15-0.4] Hz	[ms^2]
		LF/HF	Ratio of LF to HF power	[ ]
		SD1	Poincaré plot standard deviation perpendicular the line of identity	[ms]
		SD2	Poincaré plot standard deviation along the line of identity	[ms]
		SD1/SD2	Ratio of SD1 to SD2	[ ]
		ApEn	Approximate Entropy	[ ]
	Basic Morphological analysis	Pulse Amplitude	Systolic peak's amplitude	[a.u.]
		A1	Area under the waveform between systolic foot and successive systolic peak	[a.u.]
		A2	Area under the waveform between systolic peak and successive systolic foot	[a.u.]
		A	Area under the waveforme	[a.u.]
		T1	Time between systolic peak and successive systolic foot	[ms]
		T2	Time between systolic foot and successive systolic peak	[ms]
		IPA	Inflection Point Area, rapporto tra A2 e A1	[ ]
		deltaT	Intervallo di tempo fra systolic peak e il successivo diastolic peak	[ms]

	Higher Morphological analysis	AI	Augmentation Index, rapporto tra altezza diastolic peak e systolic peak	%
EDA	Full EDA	meanEDA	Mean of EDA signal	[n.u.]
		stdEDA	Standard deviation of EDA signal	[n.u.]
		minEDA	Minimum of EDA signal	[n.u.]
		maxEDA	Maximum of EDA signal	[n.u.]
		Fmax	Maximum Frequency	[n.u.]
		SlopeEDA	Slope of EDA signal	[n.u./s]
		SIE	Symbolic Information Entropy	[ ]
	Electrodermal Level (EDL) - Tonic component	meanEDL	Mean EDA tonic component	[n.u.]
		stdEDL	Standard deviation of EDA tonic component	[n.u.]
		minEDL	Minimum of EDA tonic component	[n.u.]
		maxEDL	Maximum of EDA tonic component	[n.u.]
		slopeEDL	Slope of EDA tonic component	[n.u./s]
	Electrodermal Response (EDR) - Phasic component	meanampEDR	Mean amplitude of phasic component's peaks	[n.u.]
		stdampEDR	Standard deviation of amplitude of phasic component's peaks	[n.u.]
		minampEDR	Minimum of amplitude of phasic component's peaks	[n.u.]
		maxampEDR	maximum of amplitude of phasic component's peaks	[n.u.]
		freqEDR	Frequency of phasic component's peaks	[#/min]
SKT		meanTEMP	Mean Temperature	[°C]
		stdTEMP	Standard deviation of Tempetrature	[°C]
ACC	Vector Magnitude	meanVM	Mean of Vector Magnitude	[g]
		stdVM	Standard deviation of Vector Magnitude	[g]
	Activity Index	meanAI	Mean Activity Index	[ ]
		stdAI	Standard Deviation of Activity Index	[ ]

Table F.2 Practical pipeline for PAINLESS clinical study

Day 1: Enrollment and devices delivery
The participant arrives to the hospital
Place the E4 wristband on the non-dominant wrist (make sure the USB dock has been correctly removed
Administer the questionnaires with smartphoneproject with the following order:
CRF 1 - Recruitment
CRF 2 - Sleep-wake questionnaire

Set reminders on smartphoneparticipant based on CRF 2 questionnaire's outcomes

CRF 3 - Stratification questionnaire

Ask the participant to fulfill the CRF 4 - Monitoring questionnaire to see if he/she has any problems

Provide the participant with the smartphoneparticipant

Set the appointment for Day 2, approx. one hour before the already scheduled neurorehabilitation session

---

Day 2: Neurorehabilitation session

The participant arrives to the hospital

Pick up the E4 wristband

Download data

Put the device on charge

Pick up the smartphoneparticipant

Put the device on charge

Reset the reminders

Neurorehabilitation session (approx. 1 hour after participant arrival)

Place the E4 wristband on the non-dominant wrist (make sure the USB dock has been correctly removed

Administer the CRF 3 - Stratification questionnaire with smartphoneproject at the beginning of the neurorehabilitation session

Neurorehabilitation session starts

Administer the CRF 5 - Monitoring treatment questionnaire every 10 minutes

Note down exercises done during the neurorehabilitation session

Neurorehabilitation session ends

Administer the CRF 4 - Stratification questionnaire with smartphoneprokect at the end of the neurorehabilitation session

Provide the participant with smartphoneparticipant

Set the appointment for Day 3, approx. after 24 hours from the beginning of the neurorehabilitation session

---

Day 3: Exit from the study

The participant arrives to the hospital

Administer the CRF 3 - Stratification questionnaire with smartphoneproject

Pick up the E4 wristband

Download data

Put the device on charge

Pick up the smartphoneparticipant

Put the device on charge

Remove the reminders

Ask for opinions and advice from the participant

REFERENCES

[1] "IASP Terminology - Definition of Pain," 2019. https://www.iasp-pain.org/resources/terminology/#pain (accessed Sep. 27, 2021).

[2] T. Hadjistavropoulos *et al.*, "A biopsychosocial formulation of pain communication.," *Psychol. Bull.*, vol. 137, no. 6, pp. 910–939, Nov. 2011, doi: 10.1037/a0023876.

[3] J. D. Loeser, "Concepts of pain," in *Chronic Low Back Pain*, Stanton-Hi., New York: Raven Press, 1982, pp. 145–8.

[4] K. A. Sluka and S. Z. George, "A New Definition of Pain: Update and Implications for Physical Therapist Practice and Rehabilitation Science," *Phys. Ther.*, vol. 101, no. 4, Apr. 2021, doi: 10.1093/ptj/pzab019.

[5] M. Cohen, J. Quintner, and S. Van Rysewyk, "Reconsidering the International Association for the study of pain definition of pain," *Pain Reports*, vol. 3, no. 2, pp. 1–7, 2018, doi: 10.1097/PR9.0000000000000634.

[6] R.-D. Treede, "The International Association for the Study of Pain definition of pain: as valid in 2018 as in 1979, but in need of regularly updated footnotes," *PAIN Reports*, vol. 3, no. 2, p. e643, Mar. 2018, doi: 10.1097/PR9.0000000000000643.

[7] A. C. de C. Williams and K. D. Craig, "Updating the definition of pain," *Pain*, vol. 157, no. 11, pp. 2420–2423, Nov. 2016, doi: 10.1097/j.pain.0000000000000613.

[8] S. N. Raja *et al.*, "The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises," *Pain*, vol. 161, no. 9, pp. 1976–1982, Sep. 2020, doi: 10.1097/j.pain.0000000000001939.

[9] World Health Organization, WHO Guidelines on the pharmacological treatment of persisting pain in children with medical illnesses. Geneva (Switzerland), 2012.

[10] A. P. Trouvin and S. Perrot, “New concepts of pain,” *Best Pract. Res. Clin. Rheumatol.*, vol. 33, no. 3, p. 101415, 2019, doi: 10.1016/j.berh.2019.04.007.

[11] J. Nijs *et al.*, “Nociplastic Pain Criteria or Recognition of Central Sensitization? Pain Phenotyping in the Past, Present and Future,” *J. Clin. Med.*, vol. 10, no. 15, p. 3203, Jul. 2021, doi: 10.3390/jcm10153203.

[12] P. M. Orr, B. C. Shank, and A. C. Black, “The Role of Pain Classification Systems in Pain Management,” *Crit. Care Nurs. Clin. North Am.*, vol. 29, no. 4, pp. 407–418, Dec. 2017, doi: 10.1016/j.cnc.2017.08.002.

[13] A. Abd-Elsayed and T. R. Deer, “Different Types of Pain,” in *Pain*, Cham: Springer International Publishing, 2019, pp. 15–16. doi: 10.1007/978-3-319-99124-5_3.

[14] I. Burmistr, “Theories of pain, up to Descartes and after neuromatrix: what role do they have to develop future paradigms?,” *Pain Med.*, vol. 3, no. 1, pp. 6–12, Mar. 2018, doi: 10.31636/pmjua.v3i1.81.

[15] L. A. Trachsel, S. Munakomi, and M. Cascella, “Pain Theory,” in *StatPearls [Internet]*, Treasure Island (FL), Ed. StatPearls Publishing, 2022.

[16] M. Moayedi and K. D. Davis, “Theories of pain: from specificity to gate control,” *J. Neurophysiol.*, vol. 109, no. 1, pp. 5–12, Jan. 2013, doi: 10.1152/jn.00457.2012.

[17] S. Marchand, “Mechanisms Challenges of the Pain Phenomenon,” *Front. Pain Res.*, vol. 1, Feb. 2021, doi: 10.3389/fpain.2020.574370.

[18] M. Shaikh, A. J. Hakim, and N. Shenker, “The physiology of pain,” in *Hypermobility, Fibromyalgia and Chronic Pain*, Elsevier, 2010, pp. 35–52. doi: 10.1016/B978-0-7020-3005-5.00003-3.

[19] R. Melzack, “Phantom limbs and the concept of a neuromatrix,” *Trends Neurosci.*, vol. 13, no. 3, pp. 88–92, Mar. 1990, doi: 10.1016/0166-2236(90)90179-E.

[20] M. F. Yam, Y. C. Loh, C. S. Tan, S. K. Adam, N. A. Manan, and R. Basir, “General pathways of pain sensation and the major neurotransmitters involved in pain regulation,” *Int. J. Mol. Sci.*, vol. 19, no. 8, 2018, doi: 10.3390/ijms19082164.

[21] A. I. Basbaum and T. M. Jessell, “Pain,” in *Principles of Neural Science*, 5th editio., E. R. Kandel, J. H. Schwartz, T. M. Jessel, S. A. Siegelbaum, and A. J. Hudspeth, Eds. New York: McGraw-Hill, 2013, pp. 530–555.

[22] C. E. Steeds, “The anatomy and physiology of pain,” *Surg. (United Kingdom)*, vol. 34, no. 2, pp. 55–59, 2016, doi: 10.1016/j.mpsur.2015.11.005.

[23] S. A. Woller, K. A. Eddinger, M. Corr, and T. L. Yaksh, “An overview of pathways encoding nociception.,” *Clin. Exp. Rheumatol.*, vol. 35 Suppl 1, no. 5, pp. 40–46, 2017.

[24] S. B. O’Sullivan, T. J. Schmitz, and G. D. Fulk, *Physical Rehabilitation*. 2014.

[25] R. Melzack, “Pain and the neuromatrix in the brain.,” *J. Dent. Educ.*, vol. 65, no. 12, pp. 1378–1382, 2001.

[26] M. C. Bushnell, M. Čeko, and L. A. Low, “Cognitive and emotional control of pain and its disruption in chronic pain,” *Nat. Rev. Neurosci.*, vol. 14, no. 7, pp. 502–511, Jul. 2013, doi: 10.1038/nrn3516.

[27] K. Rajneesh and R. Bolash, “Pathways of Pain Perception and Modulation,” in *Fundamentals of Pain Medicine*, Cham: Springer International Publishing, 2018, pp. 7–11. doi: 10.1007/978-3-319-64922-1_2.

[28] E. Brodin, M. Ernberg, and L. Olgart, “Neurobiology: General considerations - from acute to chronic pain,” *Nor Tann. Tid.*, vol. 126, pp. 28–33, 2016.

[29] E. L. Garland, “Pain Processing in the Human Nervous System: A Selective Review of Nociceptive and Biobehavioral Pathways,” *Prim Care*, vol. 39, no. 3, pp. 561–571, 2013, doi: 10.1016/j.pop.2012.06.013.Pain.

[30] S. Yang and M. C. Chang, “Chronic Pain: Structural and Functional Changes in Brain Structures and Associated Negative Affective States,” *Int. J. Mol. Sci.*, vol. 20, no. 13, p. 3130, Jun. 2019, doi: 10.3390/ijms20133130.

[31] H. K. Beecher, “Pain in men wounded in the battle,” *Ann. Surg.*, vol. 123, no. 1, pp. 96–105, Jan. 1946, doi: 10.1097/00000658-194601000-00008.

[32] T. Khera and V. Rangasamy, “Cognition and Pain: A Review,” *Front. Psychol.*, vol. 12, May 2021, doi: 10.3389/fpsyg.2021.673962.

[33] M. Yam, Y. Loh, C. Tan, S. Khadijah Adam, N. Abdul Manan, and R. Basir, “General Pathways of Pain Sensation and the Major Neurotransmitters Involved in Pain Regulation,” *Int. J. Mol. Sci.*, vol. 19, no. 8, p. 2164, Jul. 2018, doi: 10.3390/ijms19082164.

[34] C. J. Woolf, “Pain modulation in the spinal cord,” *Front. Pain Res.*, vol. 3, Sep. 2022, doi: 10.3389/fpain.2022.984042.

[35] D. R. Kirkpatrick *et al.*, “Therapeutic Basis of Clinical Pain Modulation.,” *Clin. Transl. Sci.*, vol. 8, no. 6, pp. 848–856, Dec. 2015, doi: 10.1111/cts.12282.

[36] U. Bingel and I. Tracey, “Imaging CNS Modulation of Pain in Humans,” *Physiology*, vol. 23, no. 6, pp. 371–380, Dec. 2008, doi: 10.1152/physiol.00024.2008.

[37] M. H. Ossipov, K. Morimura, and F. Porreca, “Descending pain modulation and chronification of pain,” *Curr. Opin. Support. Palliat. Care*, vol. 8, no. 2, pp. 143–151, Jun. 2014, doi: 10.1097/SPC.0000000000000055.

[38] B. H. Smith *et al.*, “The IASP classification of chronic pain for ICD-11: applicability in primary care,” *Pain*, vol. 160, no. 1, pp. 83–87, Jan. 2019, doi: 10.1097/j.pain.0000000000001360.

[39] S. P. Cohen, L. Vase, and W. M. Hooten, “Chronic pain: an update on burden, best practices, and new advances,” *Lancet*, vol. 397, no. 10289, pp. 2082–2097, May 2021, doi: 10.1016/S0140-6736(21)00393-7.

[40] J. Barroso, P. Branco, and A. V. Apkarian, “Brain mechanisms of chronic pain: critical role of translational approach,” *Transl. Res.*, vol. 238, no. 5, pp. 76–89, Dec. 2021, doi: 10.1016/j.trsl.2021.06.004.

[41] J. Scholz *et al.*, “The IASP classification of chronic pain for ICD-11: chronic neuropathic pain.,” *Pain*, vol. 160, no. 1, pp. 53–59, Jan. 2019, doi: 10.1097/j.pain.0000000000001365.

[42] M. Costigan, J. Scholz, and C. J. Woolf, “Neuropathic pain: a maladaptive response of the nervous system to damage.,” *Annu. Rev. Neurosci.*, vol. 32, pp. 1–32, 2009, doi: 10.1146/annurev.neuro.051508.135531.

[43] L. Colloca *et al.*, “Neuropathic pain,” *Nat. Rev. Dis. Prim.*, vol. 3, no. 1, p. 17002, Dec. 2017, doi: 10.1038/nrdp.2017.2.

[44] S. A. Schug, H. C. S. Daly, and K. D. Slannard, “Pathophysiology of Pain,” in *Mechanisms of Vascular Disease: A Reference Book for Vascular Specialists*, R. Fitridge and M. Thompson, Eds. Adelaide: University of Adelaide, 2011.

[45] T. S. Jensen and N. B. Finnerup, “Allodynia and hyperalgesia in neuropathic pain: clinical manifestations and mechanisms.,” *Lancet. Neurol.*, vol. 13, no. 9, pp. 924–935, Sep. 2014, doi: 10.1016/S1474-4422(14)70102-4.

[46] N. Stanisic, B. Häggman-Henrikson, M. Kothari, Y. M. Costa, L. Avivi-Arber, and P. Svensson, “Pain’s Adverse Impact on Training-Induced Performance and

Neuroplasticity: A Systematic Review," *Brain Imaging Behav.*, vol. 16, no. 5, pp. 2281–2306, Oct. 2022, doi: 10.1007/s11682-021-00621-6.

[47] B. McCarberg and J. Peppin, "Pain Pathways and Nervous System Plasticity: Learning and Memory in Pain," *Pain Med.*, vol. 20, no. 12, pp. 2421–2437, Dec. 2019, doi: 10.1093/pm/pnz017.

[48] S. Hiraga, T. Itokazu, M. Nishibe, and T. Yamashita, "Neuroplasticity related to chronic pain and its modulation by microglia," *Inflamm. Regen.*, vol. 42, no. 1, p. 15, Dec. 2022, doi: 10.1186/s41232-022-00199-6.

[49] R. Melzack, T. J. Coderre, J. Katz, and A. L. Vaccarino, "Central neuroplasticity and pathological pain.," *Ann. N. Y. Acad. Sci.*, vol. 933, pp. 157–174, Mar. 2001, doi: 10.1111/j.1749-6632.2001.tb05822.x.

[50] L. Doan, T. Manders, and J. Wang, "Neuroplasticity Underlying the Comorbidity of Pain and Depression," *Neural Plast.*, vol. 2015, pp. 1–16, 2015, doi: 10.1155/2015/504691.

[51] P. Cortelli, G. Giannini, V. Favoni, S. Cevoli, and G. Pierangeli, "Nociception and autonomic nervous system," *Neurol. Sci.*, vol. 34, no. SUPPL. 1, 2013, doi: 10.1007/s10072-013-1391-z.

[52] D. J. Hohenschurz-Schmidt *et al.*, "Linking Pain Sensation to the Autonomic Nervous System: The Role of the Anterior Cingulate and Periaqueductal Gray Resting-State Networks," *Front. Neurosci.*, vol. 14, Feb. 2020, doi: 10.3389/fnins.2020.00147.

[53] J. A. Waxenbaum, V. Reddy, and M. Varacallo, *Anatomy, Autonomic Nervous System*. 2021. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/30969667

[54] Y. Hu, C. Converse, M. C. Lyons, and W. H. Hsu, "Neural control of sweat secretion: a review," *Br. J. Dermatol.*, vol. 178, no. 6, pp. 1246–1256, Jun. 2018, doi: 10.1111/bjd.15808.

[55] E. J. Dansie and D. C. Turk, "Assessment of patients with chronic pain," *Br. J. Anaesth.*, vol. 111, no. 1, pp. 19–25, Jul. 2013, doi: 10.1093/bja/aet124.

[56] O. Karcioglu, H. Topacoglu, O. Dikme, and O. Dikme, "A systematic review of the pain scales in adults: Which to use?," *Am. J. Emerg. Med.*, vol. 36, no. 4, pp. 707–714, Apr. 2018, doi: 10.1016/j.ajem.2018.01.008.

[57] A. Williamson and B. Hoggart, "Pain: a review of three commonly used pain rating scales," *J. Clin. Nurs.*, vol. 14, no. 7, pp. 798–804, Aug. 2005, doi: 10.1111/j.1365-2702.2005.01121.x.

[58] M. Haefeli and A. Elfering, “Pain assessment,” *Eur. Spine J.*, vol. 15, no. S1, pp. S17–S24, Jan. 2006, doi: 10.1007/s00586-005-1044-x.

[59] M. Saltychev, H. Vastamäki, R. Mattie, Z. McCormick, M. Vastamäki, and K. Laimi, “Psychometric Properties of the Pain Numeric Rating Scale When Applied to Multiple Body Regions among Professional Musicians,” *PLoS One*, vol. 11, no. 9, p. e0161874, Sep. 2016, doi: 10.1371/journal.pone.0161874.

[60] N. Sirintawat, K. Sawang, T. Chaiyasamut, and N. Wongsirichat, “Pain measurement in oral and maxillofacial surgery,” *J. Dent. Anesth. Pain Med.*, vol. 17, no. 4, p. 253, 2017, doi: 10.17245/jdapm.2017.17.4.253.

[61] G. A. Hawker, S. Mian, T. Kendzerska, and M. French, “Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF,” *Arthritis Care Res. (Hoboken).*, vol. 63, no. S11, pp. S240–S252, Nov. 2011, doi: 10.1002/acr.20543.

[62] L. Bosdet, K. Herron, and A. C. de C. Williams, “Exploration of Hospital Inpatients’ Use of the Verbal Rating Scale of Pain,” *Front. Pain Res.*, vol. 2, Aug. 2021, doi: 10.3389/fpain.2021.723520.

[63] A. Lazaridou, N. Elbaridi, R. R. Edwards, and C. B. Berde, “Pain Assessment,” in *Essentials of Pain Medicine*, Elsevier, 2018, pp. 39-46.e1. doi: 10.1016/B978-0-323-40196-8.00005-X.

[64] R. Melzack, “The McGill Pain Questionnaire: Major properties and scoring methods,” *Pain*, vol. 1, no. 3, pp. 277–299, Sep. 1975, doi: 10.1016/0304-3959(75)90044-5.

[65] S. D. Waldman, “Pain Assessment Tools for Adults,” in *Pain Review*, Elsevier, 2009, pp. 375–380. doi: 10.1016/B978-1-4160-5893-9.00222-7.

[66] K. Niere, “Measurement of headache,” in *Headache, Orofacial Pain and Bruxism*, Elsevier, 2009, pp. 153–165. doi: 10.1016/B978-0-443-10310-0.00013-7.

[67] R. L. Daut, C. S. Cleeland, and R. C. Flanery, “Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases,” *Pain*, vol. 17, no. 2, pp. 197–210, Oct. 1983, doi: 10.1016/0304-3959(83)90143-4.

[68] N. Poquet and C. Lin, “The Brief Pain Inventory (BPI),” *J. Physiother.*, vol. 62, no. 1, p. 52, Jan. 2016, doi: 10.1016/j.jphys.2015.07.001.

[69] N. Attal, D. Bouhassira, and R. Baron, “Diagnosis and assessment of neuropathic pain through questionnaires,” *LANCET Neurol.*, vol. 17, no. 5, pp. 456–466, 2018, doi: 10.1016/S1474-4422(18)30071-1.

[70] M. Bennett, “The LANSS Pain Scale: the Leeds assessment of neuropathic symptoms and signs,” *Pain*, vol. 92, no. 1, pp. 147–157, May 2001, doi: 10.1016/S0304-3959(00)00482-6.

[71] R. Freynhagen, R. Baron, U. Gockel, and T. R. Tölle, “pain DETECT : a new screening questionnaire to identify neuropathic components in patients with back pain,” *Curr. Med. Res. Opin.*, vol. 22, no. 10, pp. 1911–1920, Oct. 2006, doi: 10.1185/030079906X132488.

[72] Pfizer, “Pharmacy focus on pain - DN4 Questionnaire,” 2016. [Online]. Available: https://www.pfizerpro.ie/sites/default/files/dn.pdf

[73] B. S. Galer and M. P. Jensen, “Development and preliminary validation of a pain measure specific to neuropathic pain: The Neuropathic Pain Scale,” *Neurology*, vol. 48, no. 2, pp. 332–338, 1997, doi: 10.1212/WNL.48.2.332.

[74] E. Widerstrom-Noga, “The assessment and treatment of pain syndromes in neurorehabilitation,” in *Oxford Textbook of Neurorehabilitation*, 2015, pp. 314–327.

[75] M. L. Wong, L. Fleming, L. E. Robayo, and E. Widerström-Noga, “Utility of the Neuropathic Pain Symptom Inventory in people with spinal cord injury,” *Spinal Cord*, vol. 58, no. 1, pp. 35–42, Jan. 2020, doi: 10.1038/s41393-019-0338-5.

[76] D. L. Morton and A. K. P. Jones, “Brain imaging of pain : state of the art,” pp. 613–624, 2016.

[77] M. L. Loggia *et al.*, “Evidence for brain glial activation in chronic pain patients,” *Brain*, vol. 138, no. 3, pp. 604–615, Mar. 2015, doi: 10.1093/brain/awu377.

[78] A. D. Legatt, “Evoked Potentials,” in *Encyclopedia of the Neurological Sciences*, Elsevier, 2014, pp. 228–231. doi: 10.1016/B978-0-12-385157-4.00529-7.

[79] D. M. Zolezzi, L. M. Alonso-Valerdi, and D. I. Ibarra-Zarate, “Chronic neuropathic pain is more than a perception: Systems and methods for an integral characterization,” *Neurosci. Biobehav. Rev.*, vol. 136, p. 104599, May 2022, doi: 10.1016/j.neubiorev.2022.104599.

[80] S. D. Subramaniam, B. Doss, L. D. Chanderasekar, A. Madhavan, and A. M. Rosary, “Scope of physiological and behavioural pain assessment techniques in children

- a review.," *Healthc. Technol. Lett.*, vol. 5, no. 4, pp. 124–129, Aug. 2018, doi: 10.1049/htl.2017.0108.

[81] D. Magee, S. Bachtold, M. Brown, and P. Farquhar-Smith, "Cancer pain: where are we now?," *Pain Manag.*, vol. 9, no. 1, pp. 63–79, 2019, doi: 10.2217/pmt-2018-0031.

[82] R. M. Fink and E. Gallagher, "Cancer Pain Assessment and Measurement," *Semin. Oncol. Nurs.*, vol. 35, no. 3, pp. 229–234, 2019, doi: 10.1016/j.soncn.2019.04.003.

[83] M. Elgendi, "On the Analysis of Fingertip Photoplethysmogram Signals," *Curr. Cardiol. Rev.*, vol. 8, no. 1, pp. 14–25, Jun. 2012, doi: 10.2174/157340312801215782.

[84] T. K. L. Hui and R. S. Sherratt, "Coverage of Emotion Recognition for Common Wearable Biosensors," *Biosensors*, vol. 8, no. 2, p. 30, Mar. 2018, doi: 10.3390/bios8020030.

[85] F. Shaffer and J. P. Ginsberg, "An Overview of Heart Rate Variability Metrics and Norms," *Front. Public Heal.*, vol. 5, no. September, pp. 1–17, Sep. 2017, doi: 10.3389/fpubh.2017.00258.

[86] F. Li, L. Yang, H. Shi, and C. Liu, "Differences in photoplethysmography morphological features and feature time series between two opposite emotions: Happiness and sadness," *Artery Res.*, vol. 18, no. C, p. 7, 2017, doi: 10.1016/j.artres.2017.02.003.

[87] J.-J. Ye, K.-T. Lee, Y.-Y. Chou, H.-H. Sie, R.-N. Huang, and C.-C. Chuang, "Assessing Pain Intensity Using Photoplethysmography Signals in Chronic Myofascial Pain Syndrome.," *Pain Pract.*, vol. 18, no. 3, pp. 296–304, Mar. 2018, doi: 10.1111/papr.12601.

[88] C. C. Chuang, J. J. Ye, W. C. Lin, K. T. Lee, and Y. T. Tai, "Photoplethysmography variability as an alternative approach to obtain heart rate variability information in chronic pain patient," *J. Clin. Monit. Comput.*, vol. 29, no. 6, pp. 801–806, 2015, doi: 10.1007/s10877-015-9669-8.

[89] K. Hamunen, V. Kontinen, E. Hakala, P. Talke, M. Paloheimo, and E. Kalso, "Effect of pain on autonomic nervous system indices derived from photoplethysmography in healthy volunteers," *Br. J. Anaesth.*, vol. 108, no. 5, pp. 838–844, 2012, doi: 10.1093/bja/aes001.

[90] Y. La Yang, H. S. Seok, G. J. Noh, B. M. Choi, and H. Shin, "Postoperative pain assessment indices based on photoplethysmography waveform analysis," *Front. Physiol.*, vol. 9, no. AUG, pp. 1–11, 2018, doi: 10.3389/fphys.2018.01199.

[91] W. Boucsein, *Electrodermal Activity*. Boston, MA: Springer US, 2012. doi: 10.1007/978-1-4614-1126-0.

[92] G. Turpin and T. Grandfield, "Electrodermal Activity," in *Encyclopedia of Stress*, Elsevier, 2007, pp. 899–902. doi: 10.1016/B978-012373947-6.00139-2.

[93] N. A. S. Taylor, "Thermal Stress and Its Physiological Implications," in *Stress: Physiology, Biochemistry, and Pathology*, Elsevier, 2019, pp. 349–379. doi: 10.1016/B978-0-12-813146-6.00026-6.

[94] V. Carr, R. Minniti, and I. Pilowsky, "Electrodermal Activity in Patients With Chronic Pain: Implications for the Specificity of Physiological Indices in Relation to Psychopathology," *Psychophysiology*, vol. 22, no. 2, pp. 208–217, 1985, doi: 10.1111/j.1469-8986.1985.tb01588.x.

[95] D. S. Bari, H. Y. Y. Aldosky, C. Tronstad, H. Kalvøy, and Ø. G. Martinsen, "Electrodermal activity responses for quantitative assessment of felt pain," *J. Electr. Bioimpedance*, vol. 9, no. 1, pp. 52–58, Dec. 2018, doi: 10.2478/joeb-2018-0010.

[96] S. A. H. Aqajari *et al.*, "Pain Assessment Tool With Electrodermal Activity for Postoperative Patients: Method Validation Study," *JMIR mHealth uHealth*, vol. 9, no. 5, p. e25258, May 2021, doi: 10.2196/25258.

[97] H. F. Posada–Quintero, Y. Kong, and K. H. Chon, "Objective pain stimulation intensity and pain sensation assessment using machine learning classification and regression based on electrodermal activity," *Am. J. Physiol. Integr. Comp. Physiol.*, vol. 321, no. 2, pp. R186–R196, Aug. 2021, doi: 10.1152/ajpregu.00094.2021.

[98] H. T. Tran, Y. Kong, A. Talati, H. Posada-Quintero, K. H. Chon, and I. Chen, "The use of electrodermal activity in pulpal diagnosis and dental pain assessment," *Int. Endod. J.*, Nov. 2022, doi: 10.1111/iej.13868.

[99] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of EMG signal analysis: detection, processing, classification and applications," *Biol. Proced. Online*, vol. 8, no. 1, pp. 11–35, Dec. 2006, doi: 10.1251/bpo115.

[100] A. Kelati, E. Nigussie, I. Ben Dhaou, J. Plosila, and H. Tenhunen, "Real-Time Classification of Pain Level Using Zygomaticus and Corrugator EMG Features," *Electronics*, vol. 11, no. 11, p. 1671, May 2022, doi: 10.3390/electronics11111671.

[101] R. Mieronkoski *et al.*, "Developing a pain intensity prediction model using facial expression: A feasibility study with electromyography," *PLoS One*, vol. 15, no. 7, p. e0235545, Jul. 2020, doi: 10.1371/journal.pone.0235545.

[102] C. T. Candotti *et al.*, “Electromyography for Assessment of Pain in Low Back Muscles,” *Phys. Ther.*, vol. 88, no. 9, pp. 1061–1067, Sep. 2008, doi: 10.2522/ptj.20070146.

[103] N. H. Che Hamid, A. C. Zakaria, N. M. Mustafah, A. Azaman, and K. Johar, “Electromyography (EMG) for Assessment in Low Back Pain; Erector Spinae Muscle,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 834, no. 1, p. 012043, Apr. 2020, doi: 10.1088/1757-899X/834/1/012043.

[104] T. Daiana da Costa, M. de Fatima Fernandes Vara, C. Santos Cristino, T. Zoraski Zanella, G. Nunes Nogueira Neto, and P. Nohama, “Breathing Monitoring and Pattern Recognition with Wearable Sensors,” in *Wearable Devices - the Big Wave of Innovation*, IntechOpen, 2019. doi: 10.5772/intechopen.85460.

[105] H. Jafari, I. Courtois, O. Van den Bergh, J. W. S. Vlaeyen, and I. Van Diest, “Pain and respiration: a systematic review.,” *Pain*, vol. 158, no. 6, pp. 995–1006, Jun. 2017, doi: 10.1097/j.pain.0000000000000865.

[106] P. H. Charlton *et al.*, “Extraction of respiratory signals from the electrocardiogram and photoplethysmogram: technical and physiological determinants,” *Physiol. Meas.*, vol. 38, no. 5, pp. 669–690, May 2017, doi: 10.1088/1361-6579/aa670e.

[107] N. Seiger *et al.*, “1621 Heart Rates and Respiratory Rates are Associated with Manchester Pain Scores in Children Presented at the Emergency Department,” *Arch. Dis. Child.*, vol. 97, no. Suppl 2, pp. A459–A459, Oct. 2012, doi: 10.1136/archdischild-2012-302724.1621.

[108] R. Cao, S. A. H. Aqajari, E. Kasaeyan Naeini, and A. M. Rahmani, “Objective Pain Assessment Using Wrist-based PPG Signals: A Respiratory Rate Based Method,” in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Nov. 2021, pp. 1164–1167. doi: 10.1109/EMBC46164.2021.9630002.

[109] R. Sallis, “Developing healthcare systems to support exercise: exercise as the fifth vital sign,” *Br. J. Sports Med.*, vol. 45, no. 6, pp. 473–474, May 2011, doi: 10.1136/bjsm.2010.083469.

[110] C. Perruchoud, E. Buchser, L. M. Johanek, K. Aminian, A. Paraschiv-Ionescu, and R. S. Taylor, “Assessment of Physical Activity of Patients With Chronic Pain,” *Neuromodulation Technol. Neural Interface*, vol. 17, pp. 42–47, Jun. 2014, doi: 10.1111/ner.12036.

[111] J. A. Verbunt, I. P. J. Huijnen, and A. Köke, “Assessment of physical activity in daily life in patients with musculoskeletal pain,” *Eur. J. Pain*, vol. 13, no. 3, pp. 231–242, Mar. 2009, doi: 10.1016/j.ejpain.2008.04.006.

[112] K. V Patel, E. J. Dansie, and D. C. Turk, “Impact of chronic musculoskeletal pain on objectively measured daily physical activity: a review of current findings,” *Pain Manag.*, vol. 3, no. 6, pp. 467–474, Nov. 2013, doi: 10.2217/pmt.13.46.

[113] B. Leininger *et al.*, “Accelerometer-Determined Physical Activity and Clinical Low Back Pain Measures in Adolescents With Chronic or Subacute Recurrent Low Back Pain,” *J. Orthop. Sport. Phys. Ther.*, vol. 47, no. 10, pp. 769–774, Oct. 2017, doi: 10.2519/jospt.2017.7345.

[114] K. C. Wong, R. Y. Lee, and S. S. Yeung, “The association between back pain and trunk posture of workers in a special school for the severe handicaps,” *BMC Musculoskelet. Disord.*, vol. 10, no. 1, p. 43, Dec. 2009, doi: 10.1186/1471-2474-10-43.

[115] P. Werner, D. Lopez-Martinez, S. Walter, A. Al-Hamadi, S. Gruss, and R. Picard, “Automatic Recognition Methods Supporting Pain Assessment: A Survey,” *IEEE Trans. Affect. Comput.*, vol. X, no. July, pp. 1–1, 2019, doi: 10.1109/TAFFC.2019.2946774.

[116] D. Naranjo-Hernandez, J. Reina-Tosina, and L. M. Roa, “Sensor Technologies to Manage the Physiological Traits of Chronic Pain: A Review.,” *Sensors (Basel).*, vol. 20, no. 2, Jan. 2020, doi: 10.3390/s20020365.

[117] A. Leroux, R. Rzasa-Lynn, C. Crainiceanu, and T. Sharma, “Wearable Devices: Current Status and Opportunities in Pain Assessment and Management,” *Digit. Biomarkers*, vol. 5, no. 1, pp. 89–102, Apr. 2021, doi: 10.1159/000515576.

[118] J. Chen, M. Abbod, and J.-S. Shieh, “Pain and Stress Detection Using Wearable Sensors and Devices—A Review,” *Sensors*, vol. 21, no. 4, p. 1030, Feb. 2021, doi: 10.3390/s21041030.

[119] D. Castaneda, A. Esparza, G. Mohammad, C. Soltanpur, and H. Nazeran, “A review on wearable photoplethysmography sensors and their potential future applications in health care,” *Int. J. Biosens. Bioelectron.*, vol. 4, no. 4, pp. 100–106, 2018, doi: 10.15406/ijbsbe.2018.04.00125.

[120] C. Orphanidou, *Signal Quality Assessment in Physiological Monitoring*. Cham: Springer International Publishing, 2018. doi: 10.1007/978-3-319-68415-4.

[121] B. Bent, B. A. Goldstein, W. A. Kibbe, and J. P. Dunn, “Investigating sources of inaccuracy in wearable optical heart rate sensors,” *npj Digit. Med.*, vol. 3, no. 1, p. 18, Dec. 2020, doi: 10.1038/s41746-020-0226-6.

[122] S. Majumder, T. Mondal, and M. Deen, “Wearable Sensors for Remote Health Monitoring,” *Sensors*, vol. 17, no. 12, p. 130, Jan. 2017, doi: 10.3390/s17010130.

[123] D. R. Witt, R. A. Kellogg, M. P. Snyder, and J. Dunn, “Windows into human health through wearables data analytics,” *Curr. Opin. Biomed. Eng.*, vol. 9, no. 1, pp. 28–46, Mar. 2019, doi: 10.1016/j.cobme.2019.01.001.

[124] J. Moraes, M. Rocha, G. Vasconcelos, J. Vasconcelos Filho, V. de Albuquerque, and A. Alexandria, “Advances in Photopletysmography Signal Analysis for Biomedical Applications,” *Sensors*, vol. 18, no. 6, p. 1894, Jun. 2018, doi: 10.3390/s18061894.

[125] J. Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiol. Meas.*, vol. 28, no. 3, pp. R1–R39, Mar. 2007, doi: 10.1088/0967-3334/28/3/R01.

[126] F. B. Reguig, “Photoplethysmogram signal analysis for detecting vital physiological parameters: An evaluating study,” in *2016 International Symposium on Signal, Image, Video and Communications (ISIVC)*, 2016, pp. 167–173. doi: 10.1109/ISIVC.2016.7893981.

[127] P. A. Kyriacou and J. Allen, *Photoplethysmography: Technology, Signal Analysis and Applications*, 1st ed. Elsevier, 2021. doi: 10.1016/C2020-0-00098-8.

[128] B. W. Nelson and N. B. Allen, “Accuracy of Consumer Wearable Heart Rate Measurement During an Ecologically Valid 24-Hour Period: Intraindividual Validation Study,” *JMIR mHealth uHealth*, vol. 7, no. 3, p. e10828, Mar. 2019, doi: 10.2196/10828.

[129] M. Nardelli, N. Vanello, G. Galperti, A. Greco, and E. P. Scilingo, “Assessing the Quality of Heart Rate Variability Estimated from Wrist and Finger PPG: A Novel Approach Based on Cross-Mapping Method,” *Sensors*, vol. 20, no. 11, p. 3156, Jun. 2020, doi: 10.3390/s20113156.

[130] M. Zanon *et al.*, “A quality metric for heart rate variability from photoplethysmogram sensor data,” in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Jul. 2020, vol. 2020-July, pp. 706–709. doi: 10.1109/EMBC44109.2020.9175671.

[131] M. Lemay, M. Bertschi, J. Sola, P. Renevey, J. Parak, and I. Korhonen, “Application of Optical Heart Rate Monitoring,” in *Wearable Sensors*, Elsevier, 2014, pp. 105–129. doi: 10.1016/B978-0-12-418662-0.00023-4.

[132] N. Pinheiro *et al.*, “Can PPG be used for HRV analysis?,” in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2016, vol. 2016-Octob, pp. 2945–2949. doi: 10.1109/EMBC.2016.7591347.

[133] M. Elgendi, *PPG Signal Analysis*. Boca Raton : Taylor & Francis, [2018]: CRC Press, 2020. doi: 10.1201/9780429449581.

[134] Y. S. Can, N. Chalabianloo, D. Ekiz, and C. Ersoy, “Continuous Stress Detection Using Wearable Sensors in Real Life: Algorithmic Programming Contest Case Study,” *Sensors*, vol. 19, no. 8, p. 1849, Apr. 2019, doi: 10.3390/s19081849.

[135] M. Wang, C. Huang, H. Chen, and S. Ye, “Preprocessing PPG and ECG Signals to Estimate Blood Pressure Based on Noninvasive Wearable Device,” *DEStech Trans. Eng. Technol. Res.*, no. iceta, pp. 1103–1109, Apr. 2017, doi: 10.12783/dtetr/iceta2016/7140.

[136] P. M. Nabeel, S. Karthik, J. Joseph, and M. Sivaprakasam, “Experimental validation of dual PPG local pulse wave velocity probe,” *2017 IEEE Int. Symp. Med. Meas. Appl. MeMeA 2017 - Proc.*, pp. 408–413, 2017, doi: 10.1109/MeMeA.2017.7985911.

[137] S. Li, L. Liu, J. Wu, B. Tang, and D. Li, “Comparison and Noise Suppression of the Transmitted and Reflected Photoplethysmography Signals,” *Biomed Res. Int.*, vol. 2018, pp. 1–9, Sep. 2018, doi: 10.1155/2018/4523593.

[138] H. J. Baek and J. Shin, “Effect of Missing Inter-Beat Interval Data on Heart Rate Variability Analysis Using Wrist-Worn Wearables,” *J. Med. Syst.*, vol. 41, no. 10, p. 147, Oct. 2017, doi: 10.1007/s10916-017-0796-2.

[139] M. Elgendi, “Optimal Signal Quality Index for Photoplethysmogram Signals,” *Bioengineering*, vol. 3, no. 4, p. 21, Sep. 2016, doi: 10.3390/bioengineering3040021.

[140] Q. Yousef, M. B. I. Reaz, and M. A. M. Ali, “The Analysis of PPG Morphology: Investigating the Effects of Aging on Arterial Compliance,” *Meas. Sci. Rev.*, vol. 12, no. 6, pp. 266–271, Jan. 2012, doi: 10.2478/v10048-012-0036-3.

[141] M. Rinkevičius *et al.*, “Photoplethysmogram Signal Morphology-Based Stress Assessment,” Dec. 2019. doi: 10.22489/CinC.2019.126.

[142] M. Hickey, J. P. Phillips, and P. A. Kyriacou, “Investigation of peripheral photoplethysmographic morphology changes induced during a hand-elevation study,” *J. Clin. Monit. Comput.*, vol. 30, no. 5, pp. 727–736, Oct. 2016, doi: 10.1007/s10877-015-9761-0.

[143] T. Pereira *et al.*, “A Supervised Approach to Robust Photoplethysmography Quality Assessment,” *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 3, pp. 649–657, Mar. 2020, doi: 10.1109/JBHI.2019.2909065.

[144] J. A. Sukor, S. J. Redmond, and N. H. Lovell, “Signal quality measures for pulse oximetry through waveform morphology analysis,” *Physiol. Meas.*, vol. 32, no. 3, pp. 369–384, Mar. 2011, doi: 10.1088/0967-3334/32/3/008.

[145] N. Pradhan, S. Rajan, and A. Adler, “Evaluation of the signal quality of wrist-based photoplethysmography,” *Physiol. Meas.*, vol. 40, no. 6, p. 065008, Jul. 2019, doi: 10.1088/1361-6579/ab225a.

[146] D.-G. Jang *et al.*, “A Simple and Robust Method for Determining the Quality of Cardiovascular Signals Using the Signal Similarity,” *2018 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 478–481, Jul. 2018, doi: 10.1109/EMBC.2018.8512341.

[147] K. Vandecasteele *et al.*, “Artifact Detection of Wrist Photoplethysmograph Signals,” *Proc. 11th Int. Jt. Conf. Biomed. Eng. Syst. Technol.*, vol. 4, no. January, pp. 182–189, 2018, doi: 10.5220/0006594301820189.

[148] C. Fischer, B. Domer, T. Wibmer, and T. Penzel, “An Algorithm for Real-Time Pulse Waveform Segmentation and Artifact Detection in Photoplethysmograms,” *IEEE J. Biomed. Heal. Informatics*, vol. 21, no. 2, pp. 372–381, Mar. 2017, doi: 10.1109/JBHI.2016.2518202.

[149] G. B. Papini, P. Fonseca, X. L. Aubert, S. Overeem, J. W. M. Bergmans, and R. Vullings, “Photoplethysmography beat detection and pulse morphology quality assessment for signal reliability estimation,” in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2017, pp. 117–120. doi: 10.1109/EMBC.2017.8036776.

[150] C. Orphanidou, T. Bonnici, P. Charlton, D. Clifton, D. Vallance, and L. Tarassenko, “Signal Quality Indices for the Electrocardiogram and Photoplethysmogram: Derivation and Applications to Wireless Monitoring,” *IEEE J. Biomed. Heal. Informatics*, vol. 19, no. 3, pp. 1–1, 2014, doi: 10.1109/JBHI.2014.2338351.

[151] W. Karlen, K. Kobayashi, J. M. Ansermino, and G. A. Dumont, "Photoplethysmogram signal quality estimation using repeated Gaussian filters and cross-correlation," *Physiol. Meas.*, vol. 33, no. 10, pp. 1617–1629, Oct. 2012, doi: 10.1088/0967-3334/33/10/1617.

[152] Q. Li and G. D. Clifford, "Dynamic time warping and machine learning for signal quality assessment of pulsatile signals," *Physiol. Meas.*, vol. 33, no. 9, pp. 1491–1501, Sep. 2012, doi: 10.1088/0967-3334/33/9/1491.

[153] V. Hartmann, H. Liu, F. Chen, Q. Qiu, S. Hughes, and D. Zheng, "Quantitative Comparison of Photoplethysmographic Waveform Characteristics: Effect of Measurement Site," *Front. Physiol.*, vol. 10, Mar. 2019, doi: 10.3389/fphys.2019.00198.

[154] C. Orphanidou, *Signal Quality Assessment in Physiological Monitoring*. Cham: Springer International Publishing, 2018. doi: 10.1007/978-3-319-68415-4.

[155] J. L. Fleiss, "Measuring nominal scale agreement among many raters.," *Psychol. Bull.*, vol. 76, no. 5, pp. 378–382, 1971, doi: 10.1037/h0031619.

[156] G. E. P. Box and D. R. Cox, "An Analysis of Transformations Revisited, Rebutted," *J. Am. Stat. Assoc.*, vol. 77, no. 377, pp. 209–210, Mar. 1982, doi: 10.1080/01621459.1982.10477788.

[157] J. Goldberger, S. Roweis, G. Hinton, and R. Salakhutdinov, "Neighbourhood Components Analysis," *Adv. Neural Inf. Process. Syst.*, vol. 17, pp. 513–520, Jul. 2005, doi: 10.1109/TCSVT.2013.2242640.

[158] The Mathworks Inc., "Matlab 2021b." [Online]. Available: www.mathworks.com

[159] E. A. Freeman and G. G. Moisen, "A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa," *Ecol. Modell.*, vol. 217, no. 1–2, pp. 48–58, Sep. 2008, doi: 10.1016/j.ecolmodel.2008.05.015.

[160] W.-Y. Lin, V. Verma, M.-Y. Lee, and C.-S. Lai, "Activity Monitoring with a Wrist-Worn, Accelerometer-Based Device," *Micromachines*, vol. 9, no. 9, p. 450, Sep. 2018, doi: 10.3390/mi9090450.

[161] J. R. Landis and G. G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, vol. 33, no. 1, p. 159, Mar. 1977, doi: 10.2307/2529310.

[162] F. Peng, Z. Zhang, X. Gou, H. Liu, and W. Wang, "Motion artifact removal from photoplethysmographic signals by combining temporally constrained independent

component analysis and adaptive filter," *Biomed. Eng. Online*, vol. 13, no. 1, p. 50, 2014, doi: 10.1186/1475-925X-13-50.

[163] D. Pollreisz and N. TaheriNejad, "Detection and Removal of Motion Artifacts in PPG Signals," *Mob. Networks Appl.*, Aug. 2019, doi: 10.1007/s11036-019-01323-6.

[164] Y. Zhang *et al.*, "Motion Artifact Reduction for Wrist-Worn Photoplethysmograph Sensors Based on Different Wavelengths," *Sensors*, vol. 19, no. 3, p. 673, Feb. 2019, doi: 10.3390/s19030673.

[165] J. Lee, M. Kim, H. Park, and I. Y. Kim, "Motion Artifact Reduction in Wearable Photoplethysmography Based on Multi-Channel Sensors with Multiple Wavelengths," *Sensors*, vol. 20, no. 5, p. 1493, Mar. 2020, doi: 10.3390/s20051493.

[166] M. D. C. Peláez, M. T. L. Albalate, A. H. Sanz, M. A. Vallés, and E. Gil, "Photoplethysmographic Waveform Versus Heart Rate Variability to Identify Low-Stress States: Attention Test," *IEEE J. Biomed. Heal. Informatics*, vol. 23, no. 5, pp. 1940–1951, 2019, doi: 10.1109/JBHI.2018.2882142.

[167] S. Ahmed, T. A. Bhuiyan, and M. Nii, "PPG Signal Morphology-Based Method for Distinguishing Stress and Non-Stress Conditions," *J. Adv. Comput. Intell. Intell. Informatics*, vol. 26, no. 1, pp. 58–66, Jan. 2022, doi: 10.20965/jaciii.2022.p0058.

[168] M. Elgendi *et al.*, "The use of photoplethysmography for assessing hypertension," *npj Digit. Med.*, vol. 2, no. 1, p. 60, Dec. 2019, doi: 10.1038/s41746-019-0136-7.

[169] S. Haddad, A. Boukhayma, and A. Caizzone, "Continuous PPG-Based Blood Pressure Monitoring Using Multi-Linear Regression," *IEEE J. Biomed. Heal. Informatics*, vol. 26, no. 5, pp. 2096–2105, May 2022, doi: 10.1109/JBHI.2021.3128229.

[170] X. Xing, Z. Ma, M. Zhang, Y. Zhou, W. Dong, and M. Song, "An Unobtrusive and Calibration-free Blood Pressure Estimation Method using Photoplethysmography and Biometrics," *Sci. Rep.*, vol. 9, no. 1, p. 8611, Dec. 2019, doi: 10.1038/s41598-019-45175-2.

[171] K. Guk *et al.*, "Evolution of wearable devices with real-time disease monitoring for personalized healthcare," *Nanomaterials*, vol. 9, no. 6, pp. 1–23, 2019, doi: 10.3390/nano9060813.

[172] S. Cheng *et al.*, "Recent Progress in Intelligent Wearable Sensors for Health Monitoring and Wound Healing Based on Biofluids," *Front. Bioeng. Biotechnol.*, vol. 9, Nov. 2021, doi: 10.3389/fbioe.2021.765987.

[173] H. Luo and B. Gao, “Development of smart wearable sensors for life healthcare,” *Eng. Regen.*, vol. 2, pp. 163–170, 2021, doi: 10.1016/j.engreg.2021.10.001.

[174] M. Smuck, C. A. Odonkor, J. K. Wilt, N. Schmidt, and M. A. Swiernik, “The emerging clinical role of wearables: factors for successful implementation in healthcare,” *npj Digit. Med.*, vol. 4, no. 1, p. 45, Dec. 2021, doi: 10.1038/s41746-021-00418-3.

[175] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, “A review of wearable sensors and systems with application in rehabilitation,” *J. Neuroeng. Rehabil.*, vol. 9, no. 1, p. 21, Dec. 2012, doi: 10.1186/1743-0003-9-21.

[176] D. R. Seshadri *et al.*, “Wearable sensors for monitoring the physiological and biochemical profile of the athlete,” *npj Digit. Med.*, vol. 2, no. 1, p. 72, Dec. 2019, doi: 10.1038/s41746-019-0150-9.

[177] J. Lin *et al.*, “Wearable sensors and devices for real-time cardiovascular disease monitoring,” *Cell Reports Phys. Sci.*, vol. 2, no. 8, p. 100541, Aug. 2021, doi: 10.1016/j.xcrp.2021.100541.

[178] P. Schmidt, A. Reiss, R. Dürichen, and K. Van Laerhoven, “Wearable-Based Affect Recognition—A Review,” *Sensors*, vol. 19, no. 19, p. 4079, Sep. 2019, doi: 10.3390/s19194079.

[179] S. Huthart, M. Elgendi, D. Zheng, G. Stansby, and J. Allen, “Advancing PPG Signal Quality and Know-How Through Knowledge Translation—From Experts to Student and Researcher,” *Front. Digit. Heal.*, vol. 2, Dec. 2020, doi: 10.3389/fdgth.2020.619692.

[180] K. Pilt, K. Meigas, R. Ferenets, K. Temitski, and M. Viigimaa, “Photoplethysmographic signal waveform index for detection of increased arterial stiffness,” *Physiol. Meas.*, vol. 35, no. 10, pp. 2027–2036, Oct. 2014, doi: 10.1088/0967-3334/35/10/2027.

[181] M. Proenca, G. Bonnier, D. Ferrario, C. Verjus, and M. Lemay, “PPG-Based Blood Pressure Monitoring by Pulse Wave Analysis: Calibration Parameters are Stable for Three Months,” in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2019, pp. 5560–5563. doi: 10.1109/EMBC.2019.8857740.

[182] E.-S. Väliaho *et al.*, “Wrist Band Photoplethysmography Autocorrelation Analysis Enables Detection of Atrial Fibrillation Without Pulse Detection,” *Front. Physiol.*, vol. 12, May 2021, doi: 10.3389/fphys.2021.654555.

[183] A. Awang and N. A. Nayan, “Early detection of mental disorder signs using photoplethysmogram : A review,” *J. Discret. Math. Sci. Cryptogr.*, vol. 24, no. 8, pp. 2171–2180, Nov. 2021, doi: 10.1080/09720529.2021.2009191.

[184] J. Fine *et al.*, “Sources of Inaccuracy in Photoplethysmography for Continuous Cardiovascular Monitoring,” *Biosensors*, vol. 11, no. 4, p. 126, Apr. 2021, doi: 10.3390/bios11040126.

[185] I. Pi, I. Pi, and W. Wu, “External factors that affect the photoplethysmography waveforms,” *SN Appl. Sci.*, vol. 4, no. 1, p. 21, Jan. 2022, doi: 10.1007/s42452-021-04906-9.

[186] F. Scardulla, L. D’Acquisto, R. Colombarini, S. Hu, S. Pasta, and D. Bellavia, “A Study on the Effect of Contact Pressure during Physical Activity on Photoplethysmographic Heart Rate Measurements,” *Sensors*, vol. 20, no. 18, p. 5052, Sep. 2020, doi: 10.3390/s20185052.

[187] S. Dehghanojamahalleh and M. Kaya, “Sex-Related Differences in Photoplethysmography Signals Measured From Finger and Toe,” *IEEE J. Transl. Eng. Heal. Med.*, vol. 7, pp. 1–7, 2019, doi: 10.1109/JTEHM.2019.2938506.

[188] D. E. Giza, G. Iliescu, S. Hassan, K. Marmagkiolis, and C. Iliescu, “Cancer as a Risk Factor for Cardiovascular Disease,” *Curr. Oncol. Rep.*, vol. 19, no. 6, p. 39, Jun. 2017, doi: 10.1007/s11912-017-0601-x.

[189] C.-H. Shih, P.-C. Chou, T.-L. Chou, and T.-W. Huang, “Measurement of Cancer-Related Fatigue Based on Heart Rate Variability: Observational Study,” *J. Med. Internet Res.*, vol. 23, no. 7, p. e25791, Jul. 2021, doi: 10.2196/25791.

[190] C. Arab *et al.*, “Heart rate variability measure in breast cancer patients and survivors: A systematic review,” *Psychoneuroendocrinology*, vol. 68, pp. 57–68, Jun. 2016, doi: 10.1016/j.psyneuen.2016.02.018.

[191] A. Mayampurath, S. L. Volchenboum, and L. N. Sanchez-Pinto, “Using photoplethysmography data to estimate heart rate variability and its association with organ dysfunction in pediatric oncology patients,” *npj Digit. Med.*, vol. 1, no. 1, p. 29, Dec. 2018, doi: 10.1038/s41746-018-0038-0.

[192] C.-T. Chen *et al.*, “Pulse-waveform and laser-doppler indices for identifying colorectal-cancer patients,” *Biomed. Eng. Appl. Basis Commun.*, vol. 33, no. 01, p. 2150005, Feb. 2021, doi: 10.4015/S1016237221500058.

[193] Y. Bryce *et al.*, “Abnormal photoplethysmography waveforms are associated with chemotherapy induced neuropathy,” *Vasa*, vol. 51, no. 2, pp. 85–92, Mar. 2022, doi: 10.1024/0301-1526/a000987.

[194] T. J. Akl, M. A. Wilson, M. N. Ericson, and G. L. Coté, “Quantifying tissue mechanical properties using photoplethysmography,” *Biomed. Opt. Express*, vol. 5, no. 7, p. 2362, Jul. 2014, doi: 10.1364/BOE.5.002362.

[195] J. M. Ahn, “New Aging Index Using Signal Features of Both Photoplethysmograms and Acceleration Plethysmograms,” *Healthc. Inform. Res.*, vol. 23, no. 1, p. 53, 2017, doi: 10.4258/hir.2017.23.1.53.

[196] R. C. Neath and M. S. Johnson, “Discrimination and Classification,” in *International Encyclopedia of Education*, Elsevier, 2010, pp. 135–141. doi: 10.1016/B978-0-08-044894-7.01312-9.

[197] “Generalized Linear Mixed Models,” in *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*, 2455 Teller Road, Thousand Oaks, California 91320: SAGE Publications, Inc., 2018. doi: 10.4135/9781506326139.n286.

[198] A. C. Rencher and G. B. Schaalje, *Linear Models in statistics*, 2nd editio. Hoboken, NJ, 2008.

[199] H. Bozdogan, “Akaike’s Information Criterion and Recent Developments in Information Complexity,” *J. Math. Psychol.*, vol. 44, no. 1, pp. 62–91, Mar. 2000, doi: 10.1006/jmps.1999.1277.

[200] R. Anderson, P. Jönsson, and M. Sandsten, “Effects of Age, BMI, Anxiety and Stress on the Parameters of a Stochastic Model for Heart Rate Variability Including Respiratory Information,” in *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies*, 2018, pp. 17–25. doi: 10.5220/0006512900170025.

[201] D. Jarchi and A. Casson, “Description of a Database Containing Wrist PPG Signals Recorded during Physical Exercise with Both Accelerometer and Gyroscope Measures of Motion,” *Data*, vol. 2, no. 1, p. 1, Dec. 2016, doi: 10.3390/data2010001.

[202] P. H. Charlton, P. Kyriacou, J. Mant, and J. Alastruey, “Acquiring Wearable Photoplethysmography Data in Daily Life: The PPG Diary Pilot Study,” in *7th International Electronic Conference on Sensors and Applications*, Nov. 2020, p. 80. doi: 10.3390/ecsa-7-08233.

[203] Y. Hou *et al.*, “Cardiac risk stratification in cancer patients: A longitudinal patient–patient network analysis,” *PLOS Med.*, vol. 18, no. 8, p. e1003736, Aug. 2021, doi: 10.1371/journal.pmed.1003736.

[204] F. Kolar and B. Ostadal, “Sex differences in cardiovascular function,” *Acta Physiol.*, vol. 207, no. 4, pp. 584–587, Apr. 2013, doi: 10.1111/apha.12057.

[205] A. Shcherbina *et al.*, “Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort,” *J. Pers. Med.*, vol. 7, no. 2, p. 3, May 2017, doi: 10.3390/jpm7020003.

[206] H. Dao and R. A. Kazin, “Gender differences in skin: A review of the literature,” *Gend. Med.*, vol. 4, no. 4, pp. 308–328, Dec. 2007, doi: 10.1016/S1550-8579(07)80061-1.

[207] S. C. Millasseau, R. P. Kelly, J. M. Ritter, and P. J. Chowienczyk, “Determination of age-related increases in large artery stiffness by digital pulse contour analysis,” *Clin. Sci.*, vol. 103, no. 4, pp. 371–377, Oct. 2002, doi: 10.1042/cs1030371.

[208] J. Allen and A. Murray, “Age-related changes in the characteristics of the photoplethysmographic pulse shape at various body sites,” *Physiol. Meas.*, vol. 24, no. 2, pp. 297–307, May 2003, doi: 10.1088/0967-3334/24/2/306.

[209] D. Zheng, Y. Yao, I. Morrison, and S. Greenwald, “Photoplethysmographic assessment of arterial stiffness and endothelial function,” in *Photoplethysmography*, Elsevier, 2022, pp. 235–276. doi: 10.1016/B978-0-12-823374-0.00007-4.

[210] D. Djeldjli, F. Bousefsaf, C. Maaoui, F. Bereksi-Reguig, and A. Pruski, “Remote estimation of pulse wave features related to arterial stiffness and blood pressure using a camera,” *Biomed. Signal Process. Control*, vol. 64, p. 102242, Feb. 2021, doi: 10.1016/j.bspc.2020.102242.

[211] H. Sanuki, R. Fukui, T. Inajima, and S. Warisawa, “Cuff-less Calibration-free Blood Pressure Estimation under Ambulatory Environment using Pulse Wave Velocity and Photoplethysmogram Signals,” in *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies*, 2017, pp. 42–48. doi: 10.5220/0006112500420048.

[212] J. Park, H. S. Seok, S.-S. Kim, and H. Shin, “Photoplethysmogram Analysis and Applications: An Integrative Review,” *Front. Physiol.*, vol. 12, Mar. 2022, doi: 10.3389/fphys.2021.808451.

[213] V. Hartmann, H. Liu, F. Chen, W. Hong, S. Hughes, and D. Zheng, “Toward Accurate Extraction of Respiratory Frequency From the Photoplethysmogram: Effect of Measurement Site,” *Front. Physiol.*, vol. 10, Jun. 2019, doi: 10.3389/fphys.2019.00732.

[214] B. Liu *et al.*, “The Assessment of Autonomic Nervous System Activity Based on Photoplethysmography in Healthy Young Men,” *Front. Physiol.*, vol. 12, Sep. 2021, doi: 10.3389/fphys.2021.733264.

[215] F. Rijo-Ferreira and J. S. Takahashi, “Genomics of circadian rhythms in health and disease,” *Genome Med.*, vol. 11, no. 1, p. 82, Dec. 2019, doi: 10.1186/s13073-019-0704-0.

[216] M. Chesnut *et al.*, “Stress Markers for Mental States and Biotypes of Depression and Anxiety: A Scoping Review and Preliminary Illustrative Analysis,” *Chronic Stress*, vol. 5, p. 247054702110003, Jan. 2021, doi: 10.1177/24705470211000338.

[217] A. Ahmed, J. Ramesh, S. Ganguly, R. Aburukba, A. Sagahyroon, and F. Aloul, “Investigating the Feasibility of Assessing Depression Severity and Valence-Arousal with Wearable Sensors Using Discrete Wavelet Transforms and Machine Learning,” *Information*, vol. 13, no. 9, p. 406, Aug. 2022, doi: 10.3390/info13090406.

[218] V. Khullar, R. G. Tiwari, A. K. Agarwal, and S. Dutta, “Physiological Signals Based Anxiety Detection Using Ensemble Machine Learning,” 2022, pp. 597–608. doi: 10.1007/978-981-16-4284-5_53.

[219] C. J. Levin, J. M. Wai, J. D. Jones, and S. D. Comer, “Changes in cardiac vagal tone as measured by heart rate variability during naloxone-induced opioid withdrawal,” *Drug Alcohol Depend.*, vol. 204, p. 107538, Nov. 2019, doi: 10.1016/j.drugalcdep.2019.06.040.

[220] M. S. Mahmud, H. Fang, H. Wang, S. Carreiro, and E. Boyer, “Automatic Detection of Opioid Intake Using Wearable Biosensor,” in *2018 International Conference on Computing, Networking and Communications (ICNC)*, Mar. 2018, pp. 784–788. doi: 10.1109/ICCNC.2018.8390334.

[221] L. Garavaglia, D. Gulich, M. M. Defeo, J. Thomas Mailland, and I. M. Irurzun, “The effect of age on the heart rate variability of healthy subjects,” *PLoS One*, vol. 16, no. 10, p. e0255894, Oct. 2021, doi: 10.1371/journal.pone.0255894.

[222] S. Rukavina, S. Gruss, H. Hoffmann, J.-W. Tan, S. Walter, and H. C. Traue, “Affective Computing and the Impact of Gender and Age,” *PLoS One*, vol. 11, no. 3, p. e0150584, Mar. 2016, doi: 10.1371/journal.pone.0150584.

[223] S. Moscato *et al.*, “Virtual Reality in Home Palliative Care: Brief Report on the Effect on Cancer-Related Symptomatology,” *Front. Psychol.*, vol. 12, Sep. 2021, doi: 10.3389/fpsyg.2021.709154.

[224] R. Refinetti, G. Cornélissen, and F. Halberg, “Procedures for numerical analysis of circadian rhythms,” *Biol. Rhythm Res.*, vol. 38, no. 4, pp. 275–325, Aug. 2007, doi: 10.1080/09291010600903692.

[225] J. Kim, B. Ku, J.-H. Bae, G.-C. Han, and J. U. Kim, “Contrast in the circadian behaviors of an electrodermal activity and bioimpedance spectroscopy,” *Chronobiol. Int.*, vol. 35, no. 10, pp. 1413–1422, Oct. 2018, doi: 10.1080/07420528.2018.1486852.

[226] A. E. Warfield, J. F. Prather, and W. D. Todd, “Systems and Circuits Linking Chronic Pain and Circadian Rhythms,” *Front. Neurosci.*, vol. 15, Jul. 2021, doi: 10.3389/fnins.2021.705173.

[227] J. A. Strong, J.-M. Zhang, and H.-G. Schaible, “The Sympathetic Nervous System and Pain,” in *The Oxford Handbook of the Neurobiology of Pain*, Oxford University Press, 2018, pp. 156–178. doi: 10.1093/oxfordhb/9780190860509.013.6.

[228] W. H. Walker, J. C. Walton, A. C. DeVries, and R. J. Nelson, “Circadian rhythm disruption and mental health,” *Transl. Psychiatry*, vol. 10, no. 1, p. 28, Jan. 2020, doi: 10.1038/s41398-020-0694-0.

[229] K. Khanade and F. Sasangohar, “Efficacy of Using Heart Rate Measurements as an Indicator to Monitor Anxiety Disorders: A Scoping Literature Review,” *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 61, no. 1, pp. 1783–1787, Sep. 2017, doi: 10.1177/1541931213601927.

[230] G. P. Trotman, J. J. C. S. Veldhuijzen van Zanten, J. Davies, C. Möller, A. T. Ginty, and S. E. Williams, “Associations between heart rate, perceived heart rate, and anxiety during acute psychological stress,” *Anxiety, Stress. Coping*, vol. 32, no. 6, pp. 711–727, Nov. 2019, doi: 10.1080/10615806.2019.1648794.

[231] B. Helgadóttir, Y. Forsell, and Ö. Ekblom, “Physical Activity Patterns of People Affected by Depressive and Anxiety Disorders as Measured by Accelerometers: A Cross-Sectional Study,” *PLoS One*, vol. 10, no. 1, p. e0115894, Jan. 2015, doi: 10.1371/journal.pone.0115894.

[232] M. Sarchiapone *et al.*, “The association between electrodermal activity (EDA), depression and suicidal behaviour: A systematic review and narrative synthesis,” *BMC Psychiatry*, vol. 18, no. 1, p. 22, Dec. 2018, doi: 10.1186/s12888-017-1551-4.

[233] B. P. Chapman, B. T. Gullapalli, T. Rahman, D. Smelson, E. W. Boyer, and S. Carreiro, “Impact of individual and treatment characteristics on wearable sensor-based digital biomarkers of opioid use,” *npj Digit. Med.*, vol. 5, no. 1, p. 123, Aug. 2022, doi: 10.1038/s41746-022-00664-z.

[234] B. T. Gullapalli, S. Carreiro, B. P. Chapman, D. Ganesan, J. Sjoquist, and T. Rahman, “OpiTrack: A Wearable-based Clinical Opioid Use Tracker with Temporal Convolutional Attention Networks,” *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 5, no. 3, pp. 1–29, Sep. 2021, doi: 10.1145/3478107.

[235] S. Wu, M. Chen, J. Wang, B. Shi, and Y. Zhou, “Association of Short-Term Heart Rate Variability With Breast Tumor Stage,” *Front. Physiol.*, vol. 12, Sep. 2021, doi: 10.3389/fphys.2021.678428.

[236] L. K. Sumitra and M. Swetha, “Gender Variations in Electrodermal Activity among Medical Students in Response to Cold Pressor Test,” *Int. J. Physiol.*, vol. 7, no. 1, p. 135, 2019, doi: 10.5958/2320-608X.2019.00029.5.

[237] E. St. John Smith, “Advances in understanding nociception and neuropathic pain,” *J. Neurol.*, vol. 265, no. 2, pp. 231–238, Feb. 2018, doi: 10.1007/s00415-017-8641-6.

[238] R.-D. Treede *et al.*, “A classification of chronic pain for ICD-11,” *Pain*, vol. 156, no. 6, pp. 1003–1007, Jun. 2015, doi: 10.1097/j.pain.0000000000000160.

[239] D. De Ridder, D. Adhia, and S. Vanneste, “The anatomy of pain and suffering in the brain and its clinical implications,” *Neurosci. Biobehav. Rev.*, vol. 130, pp. 125–146, Nov. 2021, doi: 10.1016/j.neubiorev.2021.08.013.

[240] W. Raffaeli *et al.*, “Chronic Pain: What Does It Mean? A Review on the Use of the Term Chronic Pain in Clinical Practice,” *J. Pain Res.*, vol. Volume 14, pp. 827–835, Mar. 2021, doi: 10.2147/JPR.S303186.

[241] D. Arslan, “Interactions Between The Painful Disorders and The Autonomic Nervous System,” *Ağrı - J. Turkish Soc. Algol.*, 2022, doi: 10.14744/agri.2021.43078.

[242] H. Plácido da Silva, N. M. Garcia, and E. T. Solovey, “Editorial: Biomedical Signals for Human-Computer Interaction,” *Front. Comput. Sci.*, vol. 3, Nov. 2021, doi: 10.3389/fcomp.2021.799952.

[243] Y. Lin *et al.*, “Experimental Exploration of Objective Human Pain Assessment Using Multimodal Sensing Signals,” *Front. Neurosci.*, vol. 16, Feb. 2022, doi: 10.3389/fnins.2022.831627.

[244] L. Wang, Y. Guo, B. Dalip, Y. Xiao, R. D. Urman, and Y. Lin, “An experimental study of objective pain measurement using pupillary response based on genetic algorithm and artificial neural network,” *Appl. Intell.*, vol. 52, no. 2, pp. 1145–1156, Jan. 2022, doi: 10.1007/s10489-021-02458-4.

[245] M. Ibrahim, “Polyneuropathies,” in *Physical Rehabilitation*, Elsevier, 2007, pp. 514–537. doi: 10.1016/B978-072160361-2.50022-3.

[246] A. Greco, G. Valenza, A. Lanata, E. Scilingo, and L. Citi, “cvxEDA: a Convex Optimization Approach to Electrodermal Activity Processing,” *IEEE Trans. Biomed. Eng.*, pp. 1–1, 2016, doi: 10.1109/TBME.2015.2474131.

[247] A. Schäfer and K. W. Kratky, “Estimation of breathing rate from respiratory sinus arrhythmia: Comparison of various methods,” *Ann. Biomed. Eng.*, vol. 36, no. 3, pp. 476–485, 2008, doi: 10.1007/s10439-007-9428-1.

[248] A. Bánhalmi, J. Borbás, M. Fidrich, V. Bilicki, Z. Gingl, and L. Rudas, “Analysis of a Pulse Rate Variability Measurement Using a Smartphone Camera,” *J. Healthc. Eng.*, vol. 2018, pp. 1–15, 2018, doi: 10.1155/2018/4038034.

[249] Z. Visnovcova, M. Mestanik, M. Gala, A. Mestanikova, and I. Tonhajzerova, “The complexity of electrodermal activity is altered in mental cognitive stressors,” *Comput. Biol. Med.*, vol. 79, pp. 123–129, Dec. 2016, doi: 10.1016/j.compbiomed.2016.10.014.

[250] L. M. Tracy, L. Ioannou, K. S. Baker, S. J. Gibson, N. Georgiou-Karistianis, and M. J. Giummarra, “Meta-analytic evidence for decreased heart rate variability in chronic pain implicating parasympathetic nervous system dysregulation,” *Pain*, vol. 157, no. 1, pp. 7–29, Jan. 2016, doi: 10.1097/j.pain.0000000000000360.

[251] Z. He, “The control mechanisms of heart rate dynamics in a new heart rate nonlinear time series model,” *Sci. Rep.*, vol. 10, no. 1, p. 4814, Mar. 2020, doi: 10.1038/s41598-020-61562-6.

[252] M. Rinkevičius *et al.*, “Photoplethysmogram Signal Morphology-Based Stress Assessment,” Dec. 2019. doi: 10.22489/CinC.2019.126.

[253] T. D. Yeater *et al.*, “Chronic Pain is Associated With Reduced Sympathetic Nervous System Reactivity During Simple and Complex Walking Tasks: Potential Cerebral Mechanisms,” *Chronic Stress*, vol. 5, p. 247054702110302, Jan. 2021, doi: 10.1177/24705470211030273.

[254] S.-Y. Lin *et al.*, “Differences in Physiological Signals Due to Age and Exercise Habits of Subjects during Cycling Exercise,” *Sensors*, vol. 21, no. 21, p. 7220, Oct. 2021, doi: 10.3390/s21217220.

[255] D. S. Bari, H. Y. Yacoob Aldosky, and Ø. G. Martinsen, “Simultaneous measurement of electrodermal activity components correlated with age-related differences,” *J. Biol. Phys.*, vol. 46, no. 2, pp. 177–188, Jun. 2020, doi: 10.1007/s10867-020-09547-4.

[256] H. F. Posada-Quintero, J. P. Florian, A. D. Orjuela-Cañón, T. Aljama-Corrales, S. Charleston-Villalobos, and K. H. Chon, “Power Spectral Density Analysis of Electrodermal Activity for Sympathetic Function Assessment,” *Ann. Biomed. Eng.*, vol. 44, no. 10, pp. 3124–3135, 2016, doi: 10.1007/s10439-016-1606-6.

[257] R. B. Azhiri, M. Esmaeili, and M. Nourani, “EMG-Based Feature Extraction and Classification foProsthetic Hand Control,” *Epic Ser. Comput.*, vol. 83, pp. 136–145, 2022, doi: 10.29007/zflb.

[258] M. Radovic, M. Ghalwash, N. Filipovic, and Z. Obradovic, “Minimum redundancy maximum relevance feature selection approach for temporal gene expression data,” *BMC Bioinformatics*, vol. 18, no. 1, p. 9, Dec. 2017, doi: 10.1186/s12859-016-1423-9.

[259] The Mathworks Inc., “Matlab 2022b.”

[260] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[261] A. Caraceni and M. Shkodra, “Cancer pain assessment and classification,” *Cancers (Basel).*, vol. 11, no. 4, 2019, doi: 10.3390/cancers11040510.

[262] H. L. Edwards, M. R. Mulvey, and M. I. Bennett, “Cancer-Related Neuropathic Pain,” *Cancers (Basel).*, vol. 11, no. 3, Mar. 2019, doi: 10.3390/cancers11030373.

[263] K. H. Kumar and P. Elavarasi, “Definition of pain and classification of pain disorders,” *J. Adv. Clin. Res. Insights*, vol. 3, no. June, pp. 87–90, 2016, doi: 10.15713/ins.jcri.112.

[264] “WHO guidelines on the pharmacological treatment of persisting pain in children with medical illnesses,” 2012.

[265] J. K. Fink RM, Gates RA, “Pain assessment in the palliative care setting,” in Ferrell BR, Paice JA, eds. Oxford Textbook of Palliative Nursing, 5th Edition Oxford, UK: Oxford University Press, 2019.

[266] N. J. Neufeld, S. M. Elnahal, and R. H. Alvarez, “Cancer pain: A review of epidemiology, clinical quality and value impact,” *Futur. Oncol.*, vol. 13, no. 9, pp. 833–841, 2017, doi: 10.2217/fon-2016-0423.

[267] M. H. J. van den Beuken-van Everdingen, J. M. de Rijke, A. G. Kessels, H. C. Schouten, M. van Kleef, and J. Patijn, “Prevalence of pain in patients with cancer: A systematic review of the past 40 years,” *Ann. Oncol.*, vol. 18, no. 9, pp. 1437–1449, 2007, doi: 10.1093/annonc/mdm056.

[268] H. Breivik *et al.*, “Cancer-related pain: A pan-European survey of prevalence, treatment, and patient attitudes,” *Ann. Oncol.*, vol. 20, no. 8, pp. 1420–1433, 2009, doi: 10.1093/annonc/mdp001.

[269] B. J. L. Marshall, T. H. Cartwright, C. a Berry, S. a Stowell, and S. C. Miller, “Implementation of a Performance Improvement Initiative in Colorectal Cancer Care,” *Am. Soc. Clin. Oncol.*, vol. 8, no. 5, pp. 309–313, 2012.

[270] M. C. Madariaga Muñoz, F. Villegas Estévez, A. J. Jiménez López, A. Cabezón Álvarez, and B. Soler López, “Evaluation of Quality of Life and Satisfaction of Patients with Neuropathic Pain and Breakthrough Pain: Economic Impact Based on Quality of Life,” *Pain Res. Treat.*, vol. 2018, pp. 1–8, Sep. 2018, doi: 10.1155/2018/5394021.

[271] A. P. Abernethy and J. L. Wheeler, “A health economic model of breakthrough pain,” *Am. J. Manag. Care*, vol. 14, no. 5, pp. 129–40, 2008.

[272] D. B. Gordon *et al.*, “American Pain Society recommendations for improving the quality of acute and cancer pain management: American Pain Society quality of care task force,” *Arch. Intern. Med.*, vol. 165, no. 14, pp. 1574–1580, 2005, doi: 10.1001/archinte.165.14.1574.

[273] M. I. Bennett *et al.*, “Standards for the management of cancer-related pain across Europe—A position paper from the EFIC Task Force on Cancer Pain,” *Eur. J. Pain (United Kingdom)*, vol. 23, no. 4, pp. 660–668, 2019, doi: 10.1002/ejp.1346.

[274] K. McCracken, “The challenges of cancer pain assessment.,” *Ulster Med. J.*, vol. 84, no. 1, pp. 55–557, 2015.

[275] “IASP Terminology,” *IASP Task Force on Taxonomy*, 1994. https://www.iasp-pain.org/Education/Content.aspx?ItemNumber=1698 (accessed May 29, 2020).

[276] S. D. Subramaniam, B. Doss, L. D. Chanderasekar, A. Madhavan, and A. M. Rosary, “Scope of physiological and behavioural pain assessment techniques in children

– A review," *Healthc. Technol. Lett.*, vol. 5, no. 4, pp. 124–129, 2018, doi: 10.1049/htl.2017.0108.

[277] H. Breivik *et al.*, "Assessment of pain," *Br. J. Anaesth.*, vol. 101, no. 1, pp. 17–24, Jul. 2008, doi: 10.1093/bja/aen103.

[278] P. Cortelli and G. Pierangeli, "Chronic pain-autonomic interactions," *Neurol. Sci.*, vol. 24, no. S2, pp. s68–s70, May 2003, doi: 10.1007/s100720300045.

[279] H. Breivik, W. I. Campbell, and M. K. Nocholas, *Clinical Pain Management - Practice and Procedures*. 2008. [Online]. Available: http://www.ghbook.ir/index.php?name=های رسانه و فرهنگ نوین&option=com_dbook&task=readonline&book_id=13650&page=73&chkhashk=ED9C9491B4&Itemid=218&lang=fa&tmpl=component

[280] M. N. Baliki *et al.*, "Chronic pain and the emotional brain: specific brain activity associated with spontaneous fluctuations of intensity of chronic back pain.," *J. Neurosci.*, vol. 26, no. 47, pp. 12165–12173, Nov. 2006, doi: 10.1523/JNEUROSCI.3576-06.2006.

[281] A. V. Apkarian, M. C. Bushnell, R. D. Treede, and J. K. Zubieta, "Human brain mechanisms of pain perception and regulation in health and disease," *Eur. J. Pain*, vol. 9, no. 4, p. 463, 2005, doi: 10.1016/j.ejpain.2004.11.001.

[282] D. Mischkowski, E. E. Palacios-Barrios, L. Banker, T. C. Dildine, and L. Y. Atlas, "Pain or nociception? Subjective experience mediates the effects of acute noxious heat on autonomic responses," *Pain*, vol. 159, no. 4, pp. 699–711, 2018, doi: 10.1097/j.pain.0000000000001132.

[283] W. B. Cannon, Bodily changes in pain, hunger, fear and rage. 1915.

[284] P. Gouverneur, F. Li, W. M. Adamczyk, T. M. Szikszay, K. Luedtke, and M. Grzegorzek, "Comparison of Feature Extraction Methods for Physiological Signals for Heat-Based Pain Recognition," *Sensors*, vol. 21, no. 14, p. 4838, Jul. 2021, doi: 10.3390/s21144838.

[285] J. Lee and S. K. Yoo, "The Design of Feature Selection Classifier based on Physiological Signal for Emotion Detection," *J. Inst. Electron. Eng. Korea*, vol. 50, no. 11, pp. 206–216, Nov. 2013, doi: 10.5573/ieek.2013.50.11.206.

[286] D. M. Hallman and E. Lyskov, "Autonomic Regulation in Musculoskeletal Pain," in *Pain in Perspective*, 2012, pp. 38–39. [Online]. Available: https://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics

[287] B. Rim, N.-J. Sung, S. Min, and M. Hong, “Deep Learning in Physiological Signal Data: A Survey,” *Sensors*, vol. 20, no. 4, p. 969, Feb. 2020, doi: 10.3390/s20040969.

[288] D. Moher *et al.*, “Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement,” *PLoS Med.*, vol. 6, no. 7, 2009, doi: 10.1371/journal.pmed.1000097.

[289] A. Liberati *et al.*, “The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration,” *PLoS Med.*, vol. 6, no. 7, 2009, doi: 10.1371/journal.pmed.1000100.

[290] A. Cook, D. Smith, and A. Booth, “Beyond PICO: the SPIDER tool for qualitative evidence synthesis,” *Qual. Health Res.*, vol. 22, no. 1435, 2012.

[291] “Mendeley - Reference Management Software.” https://www.mendeley.com/?interaction_required=true

[292] “Rayyan QCRI - the Systematic Reviews web app.” https://rayyan.qcri.org/

[293] P. F. Whiting *et al.*, “QUADAS-2: A Revised Tool for the Quality Assessent of Diagnostic Accuracy Studies,” *Ann. Intern. Med.*, no. 4, 2011.

[294] National Institute of Health, “Quality Assessment Tool for Before-After (Pre-Post) Studies With No Control Group,” 2014. https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools (accessed Jan. 22, 2021).

[295] J.-B. Delmotte *et al.*, “Electrochemical Skin Conductance as a Marker of Painful Oxaliplatin-Induced Peripheral Neuropathy,” *Neurol. Res. Int.*, vol. 2018, pp. 1–9, Sep. 2018, doi: 10.1155/2018/1254602.

[296] H. Yesil, S. Eyigor, M. Kayıkcıoglu, R. Uslu, M. Inbat, and B. Ozbay, “Is neuropathic pain associated with cardiac sympathovagal activity changes in patients with breast cancer?,” *Neurol. Res.*, vol. 40, no. 4, pp. 297–302, Apr. 2018, doi: 10.1080/01616412.2018.1438225.

[297] L. Guasti *et al.*, “Pain perception, blood pressure levels, and peripheral benzodiazepine receptors in patients followed for differentiated thyroid carcinoma: A longitudinal study in hypothyroidism and during hormone treatment,” *Clin. J. Pain*, vol. 23, no. 6, pp. 518–523, 2007, doi: 10.1097/AJP.0b013e3180735e5e.

[298] L. K. Badr, H. Puzantian, M. Abboud, A. Abdallah, and R. Shahine, “Assessing procedural pain in children with cancer in Beirut, Lebanon,” *J. Pediatr. Oncol. Nurs.*, vol. 23, no. 6, pp. 311–320, 2006, doi: 10.1177/1043454206291699.

[299] A. T. Ferrell-Torry and O. J. Glick, “The use of therapeutic massage as a nursing intervention to modify anxiety and the perception of cancer pain.,” *Cancer Nurs.*, vol. 16, no. 2, pp. 93–101, Apr. 1993.

[300] S. Uchida, Y. Kadoi, and S. Saito, “Effect of low dose remifentanil on postoperative pain relief and heart rate variability in post-anaesthesia care unit [Anestezi sonrası bakım ünitesinde düşük dozda remifentanilin postoperatif ağrıyı gidermede ve kalp atım hızı değişkenliği üzerindeki etk,” *Turk Anesteziyoloji ve Reanimasyon Dern. Derg.*, vol. 45, no. 5, pp. 297–302, 2017, doi: 10.5152/TJAR.2017.34341.

[301] S. H. Yu and G. H. Seol, “Lavandula angustifolia Mill. Oil and Its Active Constituent Linalyl Acetate Alleviate Pain and Urinary Residual Sense after Colorectal Cancer Surgery: A Randomised Controlled Trial,” *Evidence-Based Complement. Altern. Med.*, vol. 2017, pp. 1–7, 2017, doi: 10.1155/2017/3954181.

[302] E. Masel, P. Huber, T. Engler, and H. H. Herbert Watzke, “Heart rate variability during treatment of breakthrough pain in patients with advanced cancer: a pilot study,” *J. Pain Res.*, vol. Volume 9, pp. 1215–1220, Dec. 2016, doi: 10.2147/JPR.S120343.

[303] P. Węgorowski *et al.*, “The effect of pre-emptive analgesia on the level of postoperative pain in women undergoing surgery for breast neoplasm,” *Współczesna Onkol.*, vol. 2, no. 2, pp. 158–164, 2016, doi: 10.5114/wo.2016.60071.

[304] E. G. Boland *et al.*, “Central Pain Processing in Chronic Chemotherapy-Induced Peripheral Neuropathy: A Functional Magnetic Resonance Imaging Study,” *PLoS One*, vol. 9, no. 5, p. e96474, May 2014, doi: 10.1371/journal.pone.0096474.

[305] F. Burrai, V. Micheluzzi, and V. Bugani, “Effects of Live Sax Music on Various Physiological Parameters, Pain Level, and Mood Level in Cancer Patients,” *Holist. Nurs. Pract.*, vol. 28, no. 5, pp. 301–311, Sep. 2014, doi: 10.1097/HNP.0000000000000041.

[306] H. Nahman-Averbuch *et al.*, “Associations between autonomic dysfunction and pain in chemotherapy-induced polyneuropathy,” *Eur. J. Pain*, vol. 18, no. 1, pp. 47–55, Jan. 2014, doi: 10.1002/j.1532-2149.2013.00349.x.

[307] S.-W. Jane, D. J. Wilkie, B. B. Gallucci, R. D. Beaton, and H.-Y. Huang, “Effects of a Full-Body Massage on Pain Intensity, Anxiety, and Physiological Relaxation in Taiwanese Patients with Metastatic Bone Pain: A Pilot Study,” *J. Pain Symptom Manage.*, vol. 37, no. 4, pp. 754–763, Apr. 2009, doi: 10.1016/j.jpainsymman.2008.04.021.

[308] A. Buvanendran, A. Ali, T. R. Stoub, J. S. Kroin, and K. J. Tuman, “Brain activity associated with chronic cancer pain.,” *Pain Physician*, vol. 13, no. 5, pp. E337-42, 2010, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/20859325

[309] S. Uchida, Y. Kadoi, and S. Saito, “Effect of Low Dose Remifentanil on Postoperative Pain Relief and Heart Rate Variability in Post-Anaesthesia Care Unit,” *Turkish J. Anesth. Reanim.*, vol. 45, no. 5, pp. 297–302, Oct. 2017, doi: 10.5152/TJAR.2017.34341.

[310] E.-H. Jang, B.-J. Park, S.-H. Kim, and J.-H. Sohn, “Three differential emotion classification by machine learning algorithms using physiological signals: Discriminantion of emotions by machine learning algorithms,” in *ICAART 2012 - Proceedings of the 4th International Conference on Agents and Artificial Intelligence*, 2012, vol. 1, pp. 528–531.

[311] F. Al Machot, A. Elmachot, M. Ali, E. Al Machot, and K. Kyamakya, “A deep-learning model for subject-independent human emotion recognition using electrodermal activity sensors,” *Sensors (Switzerland)*, vol. 19, no. 7, pp. 1–14, 2019, doi: 10.3390/s19071659.

[312] S. A. Taylor, N. Jaques, E. Nosakhare, A. Sano, and R. Picard, “Personalized Multitask Learning for Predicting Tomorrow's Mood, Stress, and Health,” *IEEE Trans. Affect. Comput.*, vol. 14, no. 8, 2017, doi: 10.1109/TAFFC.2017.2784832.

[313] H. F. Posada-Quintero *et al.*, “Using electrodermal activity to validate multilevel pain stimulation in healthy volunteers evoked by thermal grills,” *Am. J. Physiol. Integr. Comp. Physiol.*, vol. 319, no. 3, pp. R366–R375, Sep. 2020, doi: 10.1152/ajpregu.00102.2020.

[314] P. Thiam, P. Bellmann, H. A. Kestler, and F. Schwenker, “Exploring Deep Physiological Models for Nociceptive Pain Recognition.,” *Sensors (Basel).*, vol. 19, no. 20, Oct. 2019, doi: 10.3390/s19204503.

[315] H. O’Leary, K. M. Smart, N. A. Moloney, and C. M. Doody, “Nervous System Sensitization as a Predictor of Outcome in the Treatment of Peripheral Musculoskeletal Conditions: A Systematic Review.,” *Pain Pract.*, vol. 17, no. 2, pp. 249–266, Feb. 2017, doi: 10.1111/papr.12484.

[316] D. Cho *et al.*, “Detection of stress levels from biosignals measured in virtual reality environments using a kernel-based extreme learning machine,” *Sensors (Switzerland)*, vol. 17, no. 10, 2017, doi: 10.3390/s17102435.

[317] A. W. Burton, T. Chai, and L. S. Smith, “Cancer pain assessment,” *Curr. Opin. Support. Palliat. Care*, vol. 8, no. 2, pp. 112–116, Jun. 2014, doi: 10.1097/SPC.0000000000000047.

[318] K. G. Korotkov, “Gender Differences in the Activity of the Autonomic Nervous Systems of Healthy and Hypertensive Patients in Russia,” *J. Appl. Biotechnol. Bioeng.*, vol. 3, no. 6, pp. 459–463, 2017, doi: 10.15406/jabb.2017.03.00084.

[319] H. A. Abhishekh *et al.*, “Influence of age and gender on autonomic regulation of heart,” *J. Clin. Monit. Comput.*, vol. 27, no. 3, pp. 259–264, 2013, doi: 10.1007/s10877-012-9424-3.

[320] M. Havelka, J. D. Lucanin, and D. Lucanin, “Biopsychosocial model--the integrated approach to health and disease.,” *Coll. Antropol.*, vol. 33, no. 1, pp. 303–10, Mar. 2009, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/19408642

[321] S. Singer, “Psychosocial Impact of Cancer,” 2018, pp. 1–11. doi: 10.1007/978-3-319-64310-6_1.

[322] K. Jordan *et al.*, “European Society for Medical Oncology (ESMO) position paper on supportive and palliative care,” *Ann. Oncol.*, vol. 29, no. 1, pp. 36–43, Jan. 2018, doi: 10.1093/annonc/mdx757.

[323] P. R. Tutelman *et al.*, “The Implementation Effectiveness of a Freely Available Pediatric Cancer Pain Assessment App: A Pilot Implementation Study,” *JMIR Cancer*, vol. 4, no. 2, p. e10280, 2018, doi: 10.2196/10280.

[324] J. Yang *et al.*, “Development and testing of a mobile app for pain management among cancer patients discharged from hospital treatment: Randomized controlled trial,” *JMIR mHealth uHealth*, vol. 7, no. 5, 2019, doi: 10.2196/12542.

[325] V. LeBaron *et al.*, “Exploring the Feasibility and Acceptability of the ``Behavioral and Environmental Sensing and Intervention for Cancer{''} to Improve Pain Management: Protocol for a Descriptive Pilot Study,” *JMIR Res. Protoc.*, vol. 8, no. 12, 2019, doi: 10.2196/16178.

[326] C. H. Lee and H.-J. Yoon, “Medical big data: promise and challenges,” *Kidney Res. Clin. Pract.*, vol. 36, no. 1, pp. 3–11, Mar. 2017, doi: 10.23876/j.krcp.2017.36.1.3.

[327] A. Khalaf *et al.*, “Analysis of multimodal physiological signals within and between individuals to predict psychological challenge vs. threat,” *Expert Syst. Appl.*, vol. 140, p. 112890, Feb. 2020, doi: 10.1016/j.eswa.2019.112890.

[328] J. Kim, M. Yadav, T. Chaspari, and C. R. Ahn, “Environmental Distress and Physiological Signals: Examination of the Saliency Detection Method,” *J. Comput. Civ. Eng.*, vol. 34, no. 6, p. 04020046, Nov. 2020, doi: 10.1061/(ASCE)CP.1943-5487.0000926.

[329] W. Li, Z. Zhang, and A. Song, “Physiological-signal-based emotion recognition: An odyssey from methodology to philosophy,” *Meas. J. Int. Meas. Confed.*, vol. 172, no. November 2020, p. 108747, 2021, doi: 10.1016/j.measurement.2020.108747.

[330] A. Johnson *et al.*, “Use of mHealth Apps and Wearable Technology to Assess Changes and Predict Pain During Treatment of Acute Pain in Sickle Cell Disease (Preprint),” *JMIR mHealth uHealth*, vol. 7, 2019, doi: 10.2196/13671.

[331] A. Badura, A. Masłowska, A. Myśliwiec, and E. Piętka, “Multimodal Signal Analysis for Pain Recognition in Physiotherapy Using Wavelet Scattering Transform,” *Sensors*, vol. 21, no. 4, p. 1311, Feb. 2021, doi: 10.3390/s21041311.

[332] C. Moro *et al.*, “Edmonton symptom assessment scale: Italian validation in two palliative care settings,” *Support. Care Cancer*, vol. 14, no. 1, pp. 30–37, Jan. 2006, doi: 10.1007/s00520-005-0834-3.

[333] H. F. Posada-Quintero *et al.*, “Time-varying analysis of electrodermal activity during exercise,” *PLoS One*, vol. 13, no. 6, pp. 1–12, 2018, doi: 10.1371/journal.pone.0198328.

[334] R. Zangróniz, A. Martínez-Rodrigo, J. M. Pastor, M. T. López, and A. Fernández-Caballero, “Electrodermal activity sensor for classification of calm/distress condition,” *Sensors (Switzerland)*, vol. 17, no. 10, pp. 1–14, 2017, doi: 10.3390/s17102324.

[335] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, “The WEKA data mining software,” *ACM SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, Nov. 2009, doi: 10.1145/1656274.1656278.

[336] A. Timm, S. Knecht, M. Florian, H. Pickenbrock, B. Studer, and T. Schmidt-Wilcke, “Frequency and nature of pain in patients undergoing neurorehabilitation,” *Clin. Rehabil.*, vol. 35, no. 1, pp. 145–153, Jan. 2021, doi: 10.1177/0269215520956784.

[337] C. A. Porro *et al.*, “Diagnosing and assessing painin neurorehabilitation:from translational research to the clinical setting: Evidence and recommendations from the Italian consensus conference on pain in neurorehabilitation,” *Eur. J. Phys. Rehabil. Med.*, vol. 52, no. 5, pp. 717–729, 2016.

[338] L. M. Benrud-Larson and S. T. Wegener, “Chronic pain in neurorehabilitation populations: Prevalence, severity and impact,” *NeuroRehabilitation*, vol. 14, no. 3, pp. 127–137, Jun. 2000, doi: 10.3233/NRE-2000-14302.

[339] S. Paolucci *et al.*, “Assessing and treating pain associated with stroke, multiple sclerosis, cerebral palsy, spinal cord injury and spasticity: Evidence and recommendations from the Italian Consensus Conference on Pain in Neurorehabilitation,” *Eur. J. Phys. Rehabil. Med.*, vol. 52, no. 6, pp. 827–840, 2016.

[340] D. Seixas, P. Foley, J. Palace, D. Lima, I. Ramos, and I. Tracey, “Pain in multiple sclerosis: A systematic review of neuroimaging studies,” *NeuroImage Clin.*, vol. 5, pp. 322–331, 2014, doi: 10.1016/j.nicl.2014.06.014.

[341] C. H. Marck *et al.*, “Pain in People with Multiple Sclerosis: Associations with Modifiable Lifestyle Factors, Fatigue, Depression, Anxiety, and Mental Health Quality of Life,” *Front. Neurol.*, vol. 8, Sep. 2017, doi: 10.3389/fneur.2017.00461.

[342] A. B. Sullivan, J. Scheman, A. LoPresti, and H. Prayor-Patterson, “Interdisciplinary Treatment of Patients with Multiple Sclerosis and Chronic Pain,” *Int. J. MS Care*, vol. 14, no. 4, pp. 216–220, Dec. 2012, doi: 10.7224/1537-2073-14.4.216.

[343] C. Clarke, R. Howard, M. Rossor, and S. Shorvon, *Neurology: A Queen Square textbook: Second edition*. 2016. doi: 10.1002/9781118486160.

[344] S. D. Subramaniam, B. Doss, L. D. Chanderasekar, A. Madhavan, and A. M. Rosary, “Scope of physiological and behavioural pain assessment techniques in children – A review,” *Healthc. Technol. Lett.*, vol. 5, no. 4, pp. 124–129, 2018, doi: 10.1049/htl.2017.0108.

[345] A. Mozzi, M. Meregaglia, C. Lazzaro, V. Tornatore, M. Belfiglio, and G. Fattore, “A comparison of EuroQol 5-Dimension health-related utilities using Italian, UK, and US preference weights in a patient sample,” *Clin. Outcomes Res.*, p. 267, Jun. 2016, doi: 10.2147/CEOR.S98226.

[346] I. Gilron, R. Baron, and T. Jensen, “Neuropathic Pain: Principles of Diagnosis and Treatment,” *MAYO Clin. Proc.*, vol. 90, no. 4, pp. 532–545, Apr. 2015, doi: 10.1016/j.mayocp.2015.01.018.

[347] Z. Dou and L. Yang, “The Application of Functional Magnetic Resonance Imaging in Neuropathic Pain,” in *Medical Imaging - Principles and Applications [Working Title]*, IntechOpen, 2019. doi: 10.5772/intechopen.89200.

[348] M. Teixeira *et al.*, "Beta Electroencephalographic Oscillation Is a Potential GABAergic Biomarker of Chronic Peripheral Neuropathic Pain," *Front. Neurosci.*, vol. 15, Feb. 2021, doi: 10.3389/fnins.2021.594536.

[349] A. Johnson *et al.*, "Use of Mobile Health Apps and Wearable Technology to Assess Changes and Predict Pain During Treatment of Acute Pain in Sickle Cell Disease: Feasibility Study.," *JMIR mHealth uHealth*, vol. 7, no. 12, p. e13671, Dec. 2019, doi: 10.2196/13671.

[350] S. Moscato, S. Orlandi, A. Giannelli, R. Ostan, and L. Chiari, "Automatic pain assessment on cancer patients using physiological signals recorded in real-world contexts," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Jul. 2022, pp. 1931–1934. doi: 10.1109/EMBC48229.2022.9871990.

[351] I. EMPATICA, "Empatica Sensors Empatica E4 TechSpecs," pp. 4–8, 2015, Accessed: Jan. 31, 2019. [Online]. Available: www.empatica.com/docs

[352] J. F. Kurtzke, "Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS)," *Neurology*, vol. 33, no. 11, pp. 1444–1444, Nov. 1983, doi: 10.1212/WNL.33.11.1444.

[353] F. La Porta, M. Franceschini, S. Caselli, P. Cavallini, S. Susassi, and A. Tennant, "Unified Balance Scale: An activity-based, bed to community, and aetiology-independent measure of balance calibrated with Rasch analysis," *J. Rehabil. Med.*, vol. 43, no. 5, pp. 435–444, 2011, doi: 10.2340/16501977-0797.

[354] K. R. Jones, C. P. Vojir, E. Hutt, and R. Fink, "Determining mild, moderate, and severe pain equivalency across pain-intensity tools in nursing home residents," *J. Rehabil. Res. Dev.*, vol. 44, no. 2, p. 305, 2007, doi: 10.1682/JRRD.2006.05.0051.

[355] S. Taylor, N. Jaques, Weixuan Chen, S. Fedor, A. Sano, and R. Picard, "Automatic identification of artifacts in electrodermal activity data," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2015, vol. 2015-Novem, pp. 1934–1937. doi: 10.1109/EMBC.2015.7318762.

[356] Empatica Inc, "Empatica E4 user manual." https://empatica.app.box.com/v/E4-User-Manual (accessed Mar. 14, 2021).

[357] M. Elgendi, I. Norton, M. Brearley, D. Abbott, and D. Schuurmans, “Detection of a and b waves in the acceleration photoplethysmogram,” *Biomed. Eng. Online*, vol. 13, no. 1, p. 139, 2014, doi: 10.1186/1475-925X-13-139.

[358] J. H. Migueles *et al.*, “Comparability of accelerometer signal aggregation metrics across placements and dominant wrist cut points for the assessment of physical activity in adults,” *Sci. Rep.*, vol. 9, no. 1, p. 18235, Dec. 2019, doi: 10.1038/s41598-019-54267-y.

[359] J. Fridolfsson *et al.*, “Effects of Frequency Filtering on Intensity and Noise in Accelerometer-Based Physical Activity Measurements,” *Sensors*, vol. 19, no. 9, p. 2186, May 2019, doi: 10.3390/s19092186.