UNIVERSITY OF TWENTE.

Biomedical Engineering

Integrative analysis towards biomarker discovery on individual level in chronic pain patients

E.M.H. Verhoeven

M.Sc. Thesis
May 2021

Supervisors:
Dr. ir. J.R. Buitenweg
B. van den Berg M.Sc.
T. Berfelo M.Sc.
Committee:
Dr. ir. J.R. Buitenweg
B. van den Berg M.Sc.
T. Berfelo M.Sc.
Dr. ir. F.H.C. de Jongh

Biomedical Signals and Systems
Faculty of Electrical Engineering,
Mathematics and Computer Science,
Biomedical Engineering
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Preface

After almost eight years of education at the University of Twente, including a bachelor technical medicine and a master biomedical engineering (after switching from the technical medicine master), I can finally present this thesis. The last ten months I have been working on this thesis at the biomedical signals and systems (BSS) research group at the University of Twente. Starting this graduation project, I was enthusiastic to work on signal analysis of physiological data. Next to the conducted conventional analysis, I learned about machine learning techniques for the first time in my educational program. Additionally, pain research was a whole new field for me since my previous internships included respiratory medicine and vascular medicine. However, after several weeks, I became more familiar with the subject and learned how much there is to gain by studying chronic pain. Furthermore, I did get the opportunity for academic development during my master assignment. Therefore, I want to start thanking my graduation committee.

First of all I want to express my gratitude to my daily supervisors, Boudewijn en Tom. I would like to thank you for the many Monday afternoons where we discussed questions and where I could receive advise from you. Especially in these times, during the COVID-19 pandemic, these hours were valuable for me to keep in touch with the research groups at the University and the St. Antonius hospital. I also valued our mountain bike meeting in Rijssen. Boudewijn, thank you for the technical and academic advise. I learned from your positive approach to issues. Tom, thank you for helping me to see the bigger picture of the research in the clinic. It made me very enthusiastic about contributing to the Nocicept-project.

Jan, I would like to thank you for the discussions at the committee meetings. After those meetings, I always got a new insights and new inspiration for improving my study. I valued your effort in enabling a work day at the University last summer. Before, I only knew my supervisors and other students on the topic from behind a computer screen. At that point, it gave me new energy to continue my work on the pain project.

Frans, I would like to thank you for your honest opinion during my graduation and my internship at Medisch Spectrum Twente. Discussions with you always let me see issues from another point-of-view.

Next, I would like to thank the students and staff from the BSS research group and the st. Antonius hospital. Unfortunately, due to the current pandemic, it was not possible to meet a lot in person. However, I always could ask my questions if I needed to and everyone was willing to help. In addition, I would like to thank the BSS research group for enabling a place at the research group. I would also like to thank the research group at the st. Antonius hospital, for making available subject data for my study.

Addtionally,I would like to thank my friends from technical medicine, biomedical engineering and v.v. Drienerlo. Spending time with you in the weekends, ensured me of fresh energy on Monday mornings.

Finally, my family and Jildert. At first I would like to thank my parents, my sister,

Jildert and his family the infinite support. You were always supporting my decisions and helped me finding my own road to where I am at now. I also would like my grandparents. Your enthusiasm gave me energy during my educational path. Thank you for all the support these years.

Abstract

Current chronic pain treatments focus on pain relief and symptom control and not on underlying neurophysiological mechanisms due to lacking diagnostic tools. The recently developed nociceptive detection threshold - evoked potential (NDT-EP) experiment is adequate in studying brain signals around the detection threshold, and aims to contribute in improved observation of nociceptive processing in chronic pain patients. Previous studies focused on statistic differences between groups. However, studies exploring features and combinations of features from the NDT-EP experiment on individual level, are lacking. Therefore, this study aims to observe altered nociceptive processing on individual level by using features from the NDT-EP experiment in chronic pain patients and pain-free subjects.

Features were obtained per measurement. The diagnostic performance of single features was explored by receiver operating characteristic analysis. The performance of features in combinations was studied using a random forest classification model.

Results show large sets of significant nociceptive detection threshold (NDT) features (4-12 out of 13) in contrast to evoked potential features (0-9 out of 11) for all comparisons between groups. More significant area under the curve values were retrieved from the P2 latency (0-9 out of 11) opposing those values at the N1 latency (0-4 out of 11). Considering feature performance in combinations, five features with the highest predictor estimates were mainly NDT, and a few P2 features.

This study identified important biomarkers from the NDT-EP experiment for observation of varying nociceptive processing on individual level between different groups. The NDT features showed the largest differentiating performance and were therefore considered most important, followed by P2 and N1 features. In feature combinations, predictor importance estimates showed a similar outcome. However, adaptions should be applied to prevent overfitting in random forest classification.

Abstract (Dutch)

Huidige behandelingen van chronische pijn focussen zich op pijnstilling en het controleren van symptomen, maar niet op onderliggende neurofysiologische mechanismen door ontbrekende diagnostiek. Recent is een experiment ontwikkeld die stimulus gerelateerde hersenresponsen (EP) meet rond de nociceptieve detectie drempel (NDT), het NDT-EP experiment. Het doel van dit experiment is om bij te dragen aan verbeterde observatie van nociceptieve signaalverwerking in patiënten met chronische pijn. Voorgaande onderzoeken concentreerden zich voornamelijk op statistische verschillen tussen groepen. Echter, studies die zich focussen op het exploreren van kenmerken en combinaties van kenmerken van het NDT-EP experiment ontbreken. Daarom is het doel van deze studie om veranderde nociceptieve signaalverwerking op individueel niveau te observeren, door het gebruik van kenmerken van het NDT-EP experiment in patiënten met chronische pijn en pijnvrije proefpersonen.

Kenmerken waren verkregen per meting. De diagnostische prestatie van losse kenmerken werden geëxploreerd door middel van 'receiver operating characteristic' analyse. De prestatie van kenmerken in combinaties werd bestudeerd door de toepassing van een 'random forest' classificatie model.

Resultaten laten voor alle vergelijkingen tussen groepen, grote sets van significante NDT kenmerken zien (4-12 van de 13) in tegenstelling tot EP kenmerken (0-9 van de 11). Daarnaast werden meer significante waarden van de kenmerken voor de oppervlakte onder de curve verkregen van de P2 latentie (0-9 van de 11) ten opzichte van de N1 latentie (0-4 van de 11). Bij de analyse van kenmerken in combinaties, waren vijf kenmerken met de hoogste voorspellende waarden voornamelijk NDT kenmerken, gevolgd door een paar P2 kenmerken.

Deze studie identificeerde belangrijke kenmerken van het NDT-EP experiment voor observatie van variërende nociceptieve signaal verwerking op individueel niveau tussen verschillende groepen. De NDT kenmerken lieten de grootste differentiërende prestaties zien en werden hierdoor beschouwd als belangrijkste kenmerken, gevolgd door P2 en N1 kenmerken. In combinaties van kenmerken waren de uitkomsen vergelijkbaar. Echter, aanpassingen zijn nodig om overfitting te voorkomen in random forest classificatie.

Contents

List of Abbreviations

ACC	anterior cingulate cortex
ARAS	ascending reticular activating system
AUC	area under the curve
BMI	body mass index
BSS	biomedical signals and systems
CSI	central sensitization inventory
D	detection
DM	diabetes mellitus
DMnp	pain-free diabetes mellitus
DP	double pulse
DPN	diabetic neuropathic pain
DPNp	painful diabetic poly neuropathy
DPSN	distal symmetrical polyneuropathy
EEG	electroencephalogram
EP(s)	evoked potential(s)
ER	the closest to (0,1) criteria
FBSS	failed back surgery syndrome
FN	false negative
FP	false positive
GFP	global field power
GL(M)M	generalized linear (mixed) model
HC(s)	healthy control(s)
IASP	international association for the study of pain
IC	intensive care
ICD-11	eleventh revision of the international classification of diseases
IES	intra-epidermal electrocutaneous stimulation
IPI	inter-pulse interval
IQR	interquartile range
L(M)M	linear (mixed) model
ML	machine learning
MTT	multiple threshold tracking
MTT-EP	multiple threshold tracking - evoked potential
NDT(s)	nociceptive detection threshold(s)
NDT-EP	nociceptive detection threshold - evoked potential
NoP	number of pulses
NRM	nucleus raphe magnus
NRS	numeric rating scale
OOB	out of bag
PAG	periaquaductal grey

PW	pulse-width
QST	quantitative sensory testing
RF	random forest
ROC	receiver operating characteristic
SE	standard error
SFN	small fiber neuropathy
SNR	signal-to-noise ratio
SP	single pulse
SP1	pulse amplitude of the first pulse
SP2_10	second pulse with 10 ms IPI
SP2_40	second pulse with 40 ms IPI
T1D	type 1 diabetes
T2D	type 2 diabetes
TENS	transcutaneous electrical nerve stimulation
TES	transcutanious electrical stimulation
TN	true negative
TP	true positive
TRL	trial
TRPV-1	acid sensitive vanilloid receptor 1
VAS	visual analog scale
VDS	verbal descriptor scale
VPL	ventral posterolateral
VRS	verbal rating scale
YLDS	years lived with disability

Chapter 1

Introduction

1.1 Problem specification

Chronic pain is a major globally occurring healthcare problem and is defined as a "persistent or recurrent pain lasting longer than 3 months" by Treede et al. (2015) [1]. Chronic pain is responsible for a negative impact on daily life and affects 19% of the European adults [2; 3]. Most people suffering from chronic pain are not capable of performing many daily activities [2]. As a result, chronic pain patients often suffer from low self-esteem, anxiety, deconditioning or depression and therefore chronic pain is accountable for the clinical and social impact on many patients' lives [2; 4; 5; 6]. Next to this clinical and social impact, causes the extensive number of chronic pain patients a substantial economic burden. The effect of less efficient employment and increasing costs of social compensations impact social-economic circumstances [7; 8]. In addition, the direct health costs of chronic pain have a 3% larger share of the gross domestic product, compared to cardial or oncological disorders [4]. Treatment of chronic pain might reduce these burdens. Current treatment of chronic pain symptoms consists of multidisciplinary cooperation between different fields: The (bio)social (e.g., physiotherapy or occupational therapy), biological (e.g., pharmacological treatment, deep brain stimulation) and psychological (e.g., multidisciplinary pain management or cognitive-behavioural therapy) field [5]. However, these conventional treatments of chronic pain focus on symptom control and pain relief and not on underlying neurophysiological mechanisms of nociceptive processing [5].

Treatment based on the underlying neurophysiological mechanisms is currently not feasible. This is caused by lacking objective observational techniques of nociceptive processing in patients suffering from chronic pain. Traditional and current used monitoring techniques in terms of chronic pain, are mainly questionnaires (e.g., numeric rating scale (NRS) or central sensitization inventory (CSI)) and physical examinations [9; 10; 11]. An enhanced observational technique for chronic pain including the underlying mechanisms, can therefore contribute to objective observation of pain mechanisms. This enables future diagnostics and treatments to focus on these mechanisms in chronic pain patients.

In previous research, a method was proposed for exploration of nociceptive processing by the use of psychophysical thresholds, the nociceptive detection thresholds (NDTs) [12]. In this procedure, intra-epidermal electrocutaneous stimulation (IES) is facilitating the activation of superficial nociceptive fibers, responsible for pain sensation. Subsequently, the NDTs, based on the stimulus settings and the subject's response, can be calculated by the multiple threshold tracking (MTT) algorithm [13]. The NDTs are partly based on the subject's responses and can therefore be subjective. Accordingly, evoked potentials (EPs), derived from the electroencephalogram (EEG) are added to the method for increased

objectivity. This combination of techniques regarding the NDTs and the EPs is known as the nociceptive detection threshold - evoked potential (NDT-EP) experiment.

The biomedical signals and systems (BSS) research group from the University of Twente already conducted the NDT-EP experiment in different subject groups in collaboration with the St. Antonius hospital in Nieuwegein. Those subject groups were including healthy controls (HCs) and patients suffering from expected central or peripheral sensitization. Berfelo et al. (2019) observed the behaviour of NDTs and EPs in HCs and in patients suffering from failed back surgery syndrome (FBSS) (expected to suffer from central alterations) [14]. In patients suffering from FBSS, back pain is experienced after lumbar surgery where the outcome contradicts with the patient's and the surgeon's intention [15]. Gefferie et al. (2020) also observed the behavior of NDTs and EPs. However, these subjects included patients suffering from painful diabetic poly neuropathy (DPNp) and pain-free diabetes mellitus (DMnp) (expected to suffer from peripheral alterations and perhaps additional central alterations) [16]. Patients suffering from diabetes mellitus (DM) develop chronic glycemic alterations which can result in neuropathy [17; 18].

Previous research focused on differences in brain responses around the NDT on group level. However, exploration of features concerning NDTs and EPs on individual level of subjects of a specific research group (e.g., FBSS or the healthy control group (HC)) has not been performed yet. More insight into these features and their differentiating abilities on individual level, is expected to contribute to observation of neurophysiological mechanisms responsible for chronic pain. Long-term benefits of improved knowledge concerning neurophysiological mechanisms are likely to result in objective diagnostic methods. Subsequently, future therapies can also be based on underlying neurophysiological mechanisms instead of the conventional focus on pain relief or symptom control.

1.2 Research objective

The aim of this study is to observe altered nociceptive processing, using the features and combinations of features from the NDT-EP experiment on individual level. This is established with subjects suffering chronic pain (i.e., FBSS and DPNp) and pain-free subjects (i.e., HC and DMnp). When these features are considered relevant in differentiating the researched diagnoses, they can be referred to as biomarkers.

Moreover, the quality in terms of duration and diagnostic performance of the current NDT-EP experiment is taken into consideration for possible adaptions of the experiment. These adaptions might additionally contribute to the observation of neurophysiological mechanisms in terms of nociceptive processing.

1.3 Thesis outline

This thesis about observation of possible features important in nociceptive processing, is divided into seven chapters. After this introducing chapter (Chapter 1), the background will be discussed (Chapter 2). The Background provides more specific technical and medical information to the reader in order to fill in on possible missing knowledge and is completed with the implications. In this section, the reader is provided with a small overview of the most important information and includes the primary and secondary research questions. Moreover, Chapter 3 shows the methods. Thereafter, in Chapter 4 an overview of the results is given. These results are being discussed in Chapter 5. Next part of this thesis, concerning the conclusion of this study is assembled in Chapter 6. Finally,

Chapter 7 consists of the future perspectives. In this chapter, the possible improvements regarding the NDT-EP procedure will be discussed next to general recommendations.

Chapter 2

Background

This chapter presents an extensive background of chronic pain research. The first few sections (Section 2.1, 2.2 and 2.3) present general background information to become acquainted with pain, physiology and the pathophysiology. Section 2.4 outlines several current measurement methods of pain, including the method currently used by the BSS research group at the University of Twente, Enschede. Next, a machine learning classification method is discussed in Section 2.5. The final section of this chapter is Section 2.6. This section describes the the most important rationales of researching this topic, concluding with the research questions.

2.1 Pain

According to the International Association for the Study of Pain (IASP), pain is defined as "An unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage" [19]. An important concept in the definition of pain is the requirement of consciousness and the subjectivity of the experience. The latter makes pain a hard to measure concept [20].

In the field of pain research, various types of pain are described. The IASP differentiates nociceptive pain and neuropathic (peripheral or central) pain. Here, nociceptive pain is defined as "pain that arises from actual or threatened damage to non-neural tissue and is due to the activation of nociceptors". On the other hand, neuropathic pain is known as "pain caused by a lesion or disease of the somatosensory nervous system" [21].

Recently, a new type of pain was added to the definitions of pain. This third pain mechanism, proposed by Kosek et al. (2016) is currently known as nociplastic pain. The IASP council adopted this pain type to their taxonomy in 2017 [22; 23]. The corresponding definition of nociplastic pain is currently defined as "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" [21].

The last type of pain is known as 'mixed pain'. This type of pain is not included in the taxonomy of IASP, though it is a quite familiar term in the clinical field [22]. After the introduction of nociplastic pain, Freynhagen et al. (2019) defined mixed pain as "a complex overlap of the different known pain types (nociceptive, neuropathic, nociplastic) in any combination, acting simultaneously and/or concurrently to cause pain in the same body area" [24].

2.2 Nociceptive regulation

The definition of pain indicates the presence of 'experience' in the perception of pain. This experience differentiates the term nociception from pain [20]. Nociception is the neural encoding of a noxious stimulus through transduction, transmission and modulation for perception in the brain [20; 21; 25]. A noxious stimulus is known as actual or impending tissue damage [21]. The actual pain awareness varies since it depends on multiple factors (e.g., past pain and life experiences, gender, culture, individual health and genetics) [25].

2.2.1 Transduction

After activation of nociceptive peripheral receptors (nociceptors), pain can be experienced [26]. The process of the generation of a pain signal after a stimulus by a nociceptor is called transduction [27]. Nociceptors are free endings of peripheral nerve fibers known as mechanoreceptors sensitive to mechanical deformation and polymodal nociceptors sensitive to multiple noxious stimuli (thermal, mechanical and chemical) [20; 26]. The two types of polymodal nociceptors are Aδ and C nerve fibers [20; 27]. Aδ fibers are finely myelinated fibers, causing a pricking, fast and well localized pain sensation. These fibers have a 2-5 μm diameter and the conduction velocity is 5-15 m/s. The unmyelinated C fibers create a diffuse, dull, aching and slow pain sensation with a conduction velocity of 0.5-2 m/s and a diameter less than 2 μm [20]. Majority of C and Aδ fibers terminate in the dorsal horn. Next to the C and Aδ fibers, Aβ fibers terminate in the dorsal horn. Aβ fibers are fast conducting myelinated fibers, sensitive to touch and vibrations. All three types of fibers, enter through the dorsal root via ventrolateral bundle (Figure 2.1) [20]. Besides, about 30% of the C fibers enter via the ventral root [20; 27].

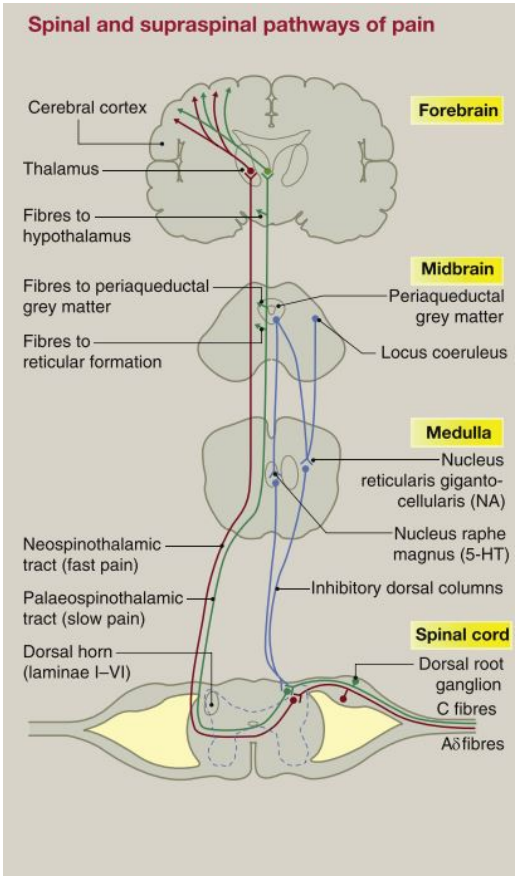


Figure 2.1: Pain pathway, obtained from Steeds et al. (2016) [20]

2.2.2 Transmission

After transmission of the action potentials along the peripheral nerve fibers, the signal reaches the dorsal horn, being located in the spinal chord. Here a connection between peripheral and second-order neurons is established and modulation is able to occur. The different fibers terminate in different laminae of the dorsal horn as is shown in Figure 2.2. In lamina II, the C fibers terminate. This is in the substantia gelatinosa. In lamina I and V, Aδ fibers synapse with second order neurons and the large Aβ fibers have collateral branches that terminate in laminae III to V. These second order neurons are categorized in three types. First of all, nociceptive specific neurons responsible for processing stimuli with high thresholds. Secondly, wide

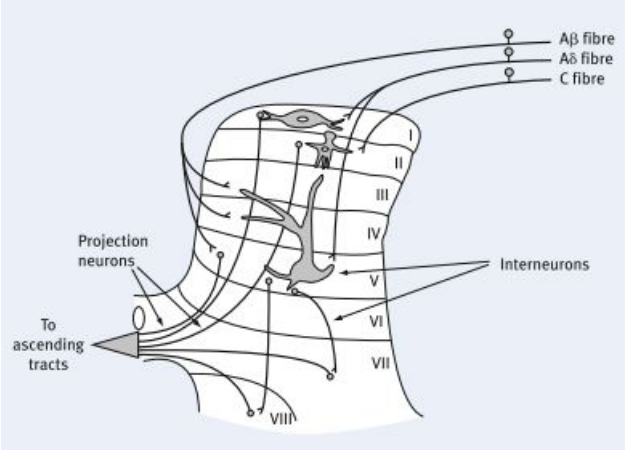


Figure 2.2: Nociceptive fibers in the dorsal horn. Originally Welchew et al. (2003), reproduced by (with consent of B. P. Rice) and obtained from Birnbaum (2010) [27; 28].

dynamic range neurons, responding to an extensive amount of noxious stimuli. And third the second-order low-threshold neurons, being able to process non-noxes [20; 27].

2.2.3 Perception

The next factor in nociceptive processing is pain perception [25]. This starts with second-order neurons functioning as ascending pathways. They propagate contralateral through the spine via the spinothalamic and spinoreticular tract to the thalamus. [20; 27; 30]. In Figure 2.1, the ascending pathways can be recognized by their green and red color and Figure 2.3 gives an overview of the ascending pathways in the brain [20; 30].

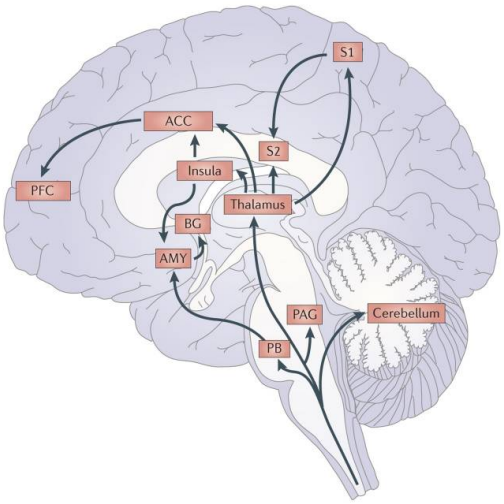


Figure 2.3: Ascending brain pathway, obtained from Dydyk et al. (2020) [29]

The spinothalamic tract leads Aδ neurons originating from lamina I and V to VII and propagates through the brain stem to the thalamus in the ventral posterolateral (VPL) and the posterior nucleus [27; 30; 31; 32]. There they synapse to axons that continue to the somesthetic region of the brain. The pain conducted through the spinothalamic tract can be experienced as sharp or pin-pricking [31].

The spinoretecular tract leads the neurons conducting complex pain properties originating from lamina VII and VIII of the dorsal horn up to the thalamus and hypothalamus through the reticular formation of the brain stem [20; 31]. The reticular formation is part of the cerebral cortex, the thalamus and the ascending reticular activating system (ARAS). The latter is responsible for arousal and attention [31].

2.2.4 Modulation

As previously mentioned, many processes are established in the dorsal horn. This is called, modulation. The mechanisms responsible for modulation are (1) segmental (spinal)

inhibition including the effects of endogenous opioids, (2) inhibition from the descending pathway and (3) the effect of non-nociceptive input [20; 27]. The latter is also known as the 'gate-control' theory first mentioned by Melzack et al. (1965) [33].

Firstly, Melzack et al. (1965) stated that activation of inhibitory interneurons in lamina II can be established by Aβ fibers. The consequence is the suppression of transmission of C fibers resulting in a pain reduction. Practically this means that when a sensitive area is rubbed, this results in pain relief [20; 33]. This theory is the foundation for transcutaneous electrical nerve stimulation (TENS) [27; 33].

Secondly, at segmental inhibition, local inhibition is established. This can be done by a collection of opioids or neuroreceptors [27].

The final modulation mechanism influencing pain perception is inhibition by descending pathways [27]. The inhibition causes the pain to reduce. Conclusive brain stem areas providing this reduction are the nucleus raphe magnus (NRM) and the periaquaductal grey (PAG). By stimulation of raphe nuclei inhibition is provided causing analgesia. The released serotonin overrules noradrenaline and therefore prevents pain transmission. The PAG receives information from (1) the spinothalamic tract, (2) the cortex, (3) the thalamus and (4) the hypothalamus. Next to the serotonin, they stimulate the NRM causing a blockage in the pain transmission [27].

2.3 Pathophysiology of pain

2.3.1 Chronic pain

Chronic pain is defined as a "persistent or recurrent pain lasting longer than 3 months" by Treede et al. (2015) [1]. However, there are researchers that use other classification factors causing the prevalence to vary [2]. However, the disorder chronic pain itself was subdivided in seven categories, in order to fit the classification under one category in the eleventh revision of the international classification of diseases (ICD-11) [1].

Normally, pain is essential in survival. A noxious stimulus can be detected, whereby the person is warned. However, one can label pain as chronic pain when long after recovery of an acute injury, pain still persists. Temporary sensitization occurs in order to guard the damaged area. This temporary sensitization, causes allodynia and hyperalgesia to occur [32]. At the IASP, allodynia is defined as "pain due to a stimulus that does not normally provoke pain" [21]. This means that when for example a light touch is applied to 'healthy' skin, this can be experienced as pain [34]. Hyperalgesia implies "increased pain from a stimulus that normally provokes pain" [21]. Here, patients experience an increased response to a pain stimulus [34].

2.3.2 Central sensitization

The first publications, implying an altered central sensitization, were written by Woolf et al. in 1983 [35]. Nowadays central sensitization is documented in the IASP terminology as an "increased responsiveness of nociceptive neurons in the central nervous system to their normal or subthreshold afferent input" [21].

There are various mechanisms that contribute to the occurrence of central sensitization. However the most apparent rationale of central sensitization, is caused by an increased sensitivity of glutamate (neurotransmitter) transmission in the central nervous system. After afferent noxious stimuli, the peripheral C fibers synapse, using glutamate release, to AMPA and NMDA receptors. These receptors are located on pain transmission fibers and prolong to the brain [36; 37]. As a result, less neurotransmitter is necessary in order

to establish successful pain transmission. Therefore, less stimuli are needed for pain experience. This process is constitutional for central sensitization in patients suffering from chronic pain [37].

This sensitivity shift can be recognized by the appearance of hyperalgesia and allodynia [38]. This altered sensitivity is caused by the neural plasticity. These alterations contribute substantially to a variety of pain syndromes. Thus, central sensitization is considered to be an important factor in treatments of clinical pain syndromes [39].

2.3.3 Failed back surgery syndrome

The 2016 Global Burden of Disease study, acknowledge low back pain as the predominant cause considering years lived with disability (YLDs) [40]. Chronic low back pain has a prevalence of 37.1% and the lifetime prevalence includes 85.5% in the German population [41]. This prevalence increases linearly between adults of the age of 30 and 60 [42]. These large numbers, together with the impact of individuals cause for chronic low back pain it to be an extensive economic and social burden [43].

Together with the number of spine surgery procedures in patients suffering from chronic low back pain, the number of patients with FBSS has increased [43]. Waguespack et al. (2002) used for FBSS the following definition: "FBSS implies that the outcome of a lumbar spine surgery did not meet the expectations established by the patient and the surgeon before surgery" [15]. There is a variety in diagnoses for patients who suffer from FBSS. The study from Waguespack et al. (2002) found out that in their population, about half of the population suffers from foraminal stenosis or painful disc(s) [15]. There are multiple treatment options for this disorder such as neuromodulation or medical management. However, it is important to maintain a multidisciplinary approach for management of FBSS [44].

Despite all research, there is still lacking knowledge on the exact pain processing including in patients suffering from FBSS [45].

Recently, Berfelo et al. (2019) researched the behavior of NDTs and EPs in healthy subjects (including the replicability with healthy subjects from the University of Twente) and FBSS patients [46; 13]. Corresponding results were observed between healthy patients at the University of Twente and in the St. Antonius Hospital. Additionally, altered behavior of the EPs and the NDTs was recognized in the FBSS patients . Therefore it was concluded that the NDT-EP experiment (previously the multiple threshold tracking - evoked potential (MTT-EP) experiment) can be used to identify patients suffering from for central sensitized chronic pain syndromes [46].

2.3.4 Diabetes Mellitus

DM is a globally occurring chronic hyperglycemic syndrome [17]. The prevalence of this disorder rises since the prevalence of obesity rises and the physical inactivity increases. Additionally, population growth, urbanization and the age are rising. In 2000 the overall prevalence was about 2.8% with a total amount of patients around 171 million worldwide. Wild et al. (2004) estimated the growth for 2030. It was expected, a prevalence of 4.4% with 366 million cases worldwide will occur [47].

There are two types of DM. Type 1 diabetes (T1D) and type 2 diabetes (T2D), respectively. T1D is characterized by an absolute insulin secretion deficiency. Patients suffering from T2D, experience insulin resistance and a poor insulin secretion response [48]. These types of DM are associated with a series of short- and long-term consequences. Short-term consequences can be polydipsia, blurred vision, polyuria, weight loss with possible

polyphagia. Possible long-term consequences mainly occur due to neuropathy and can result in retinopathy with possible loss of vision, renal failure due to nephropathy, possible ulcerating feet, increased cardiovascular disorders and more [48; 18]. Neuropathy can occur in both groups of DM patients. As a matter of fact, over 90% of the DM patients familiar with a neuropathy, suffer from distal symmetrical polyneuropathy (DSPN). DPSN or other neuropathies may cause various symptoms of pain, such as mononeuropathies, neuropathies caused by entrapments and radiculopathies [18]. There are various known theories about diabetic neuropathic pain (DPN). However, the total truth of the pathogenesis is still unknown [49].

The last couple of years, various research groups explored pain mechanisms in patients suffering from DM. A team that focuses on this topic is the BSS research group at the University of Twente in Enschede together with St. Antonius hospital in Nieuwegein [16]. Another example of a research group investigating pain processing is a group in Brussels. They research the nociceptive processing by applying a thermal stimulus [50]

2.4 Measures of pain

To attain the best possible treatment for the pain patient, pain should be measured [51]. Various measurement methods are currently used in order to measure acute and chronic pain [51; 52]. Three frequently used tests measuring the pain intensity are: (1) The visual analog scale (VAS), (2) the NRS and (3) verbal rating scale (VRS)/verbal descriptor scale (VDS) [51]. The VAS uses a line of 10 cm with on the left end 'no pain' and on the right end 'pain as bad as it could be' [9; 53; 54]. Another and the most common measure of pain is the NRS, consisting of a scale from 0 tot 10. In this scale 0 is considered 'no pain' and the 'worst imaginable pain' can be indicated with 10 [54; 55]. Finally, the VDS or VRS is organized as a queue of terms (e.g. 'none', 'mild', 'moderate', 'severe') [51; 53; 54]. Despite the clinical relevance of those methods, there is still much room for improvement [51].

Previous questionnaires mainly focus on pain in general. However, a more focused measure on central sensitization, is the CSI. [56; 10]. This questionnaire consists of two parts. The first part measures 25 symptoms with five possible answers (e.g. never, rarely, sometimes, often, always) and the second part contains a list of ten possible central disorders [56].

2.4.1 Physiological measures of nociceptive pain

Quantative measurements of nociception

Quantitative sensory testing (QST) is considered as a more adequate approach for measuring stimulus intensity, establishing certain sensory perceptions [57]. Examples of these stimuli are touch, pressure, thermal, vibrating and pain stimuli [58]. QST provides better characteristic control in order to successfully evaluate hyperalgesia or allodynia [57]. Therefore, QST has currently an increased availability and is becoming a more potential neurophysiologic tool [58]. Underneath, several approaches of QST are discussed.

Previous work

Due to the increasing interest and availability of different QST methods in the neurophysiological field, various pain studies focused on this method of testing [58].

Various researchers focus on the assessment of pain sensitivity. Several stimulus methods can be conducted in order to get more insight in this sensitivity. This is studied with various stimuli in various diseases. For example, Coronado et al. (2014) investigated pain responses in patients suffering from unilateral shoulder pain [59]. Here, pressure and heat stimuli were applied and aimed was to examine pain responses in the extremities of patients suffering from shoulder pain. Additionally they were interested in the intra-individual of different states of sensitization. The study concluded that sensitization was the response of pain, stimulated by pressure and heat [59].

Another example of a study, researched pain sensitivity by analysing thresholds from pressure pain and heat pain and heat temporal summation. This was done in the study from Eckenrode et al. (2019) in healthy and disordered patients suffering from chronic achilles tendinopathy [60]. They concluded that patients suffering from achilles tendinopathey indeed seemed experience central and peripheral sensitization [60].

Currently, a research group in Brussels, Belgium is working on thermal stimulation for assessment of pain processing. This group investigated in one of their latest studies from Courtin et al. (2020) various derivatives from psychometric functions including the slope and the threshold DM patients and compared them with healthy controls [50]. The type of stimuli they used was heath and cold detection. In this research, the DM and the control group were exposed to cold and warm stimuli. These stimuli were being generated with a CO_2 laser, controlling the temperature and a Peltier thermode. It was suggested that patients suffering from neuropathic changes due to DM, have a varying psychometric function. This was most apparent in cold stimuli. Here, on group level was shown a less steeper slope in patients suffering from DM compared to the controls. Additionally, the detection thresholds were lower at cold stimuli at the DM group when being compared to the controls and higher at the laser heat stimuli [50].

Nociceptive stimulation

As previously mentioned in this section, there are several ways to perform QST. Most frequent application of nociceptive stimulation consists of laser and mechanical stimulation [61]. However, Steenbergen et al. (2012) introduced an improved system, able to induce tactile sensations by transcutanious electrical stimulation (TES) and nociceptive sensations by IES [61; 62].

Earlier, electrical stimulation of Aβ fibers was established by using surface electrodes in various studies (e.g. Inui et al. (2003)) [63]. This study was also one of the studies applying IES to superficial Aδ fibers by using small needles [63].

The electrode was originally developed by Steenbergen et al. (2008) [62]. This bimodal electrode applies nociceptive stimuli by the five needle electrodes stimulating Aδ fibers and tactile stimulation by the four flat electrodes stimulating the Aβ fibers [62]. However, over the years, the electrode changed into an electrode without flat electrodes, but with five needles perpetrating the skin for 0.2 mm for stimulation of the Aδ fibers. The electrode is shown in Figure 2.4 [64]. This electrode has already been used in several studies from the BSS group and the St. Antonius hospital in Nieuwegein (e.g. Doll et al. (2016), van den Berg et al. (2018), Berfelo et al. (2019) and Gefferie et al. (2020)) [46; 13; 62; 64; 16]. All studies used the electrode for observation of the nociceptive processing.

Psychometric curve

One of the most important psychophysical measures from the NDT-EP experiment used by various studies (e.g. Doll et al. (2016)) is the psychometric curve. The psychometric

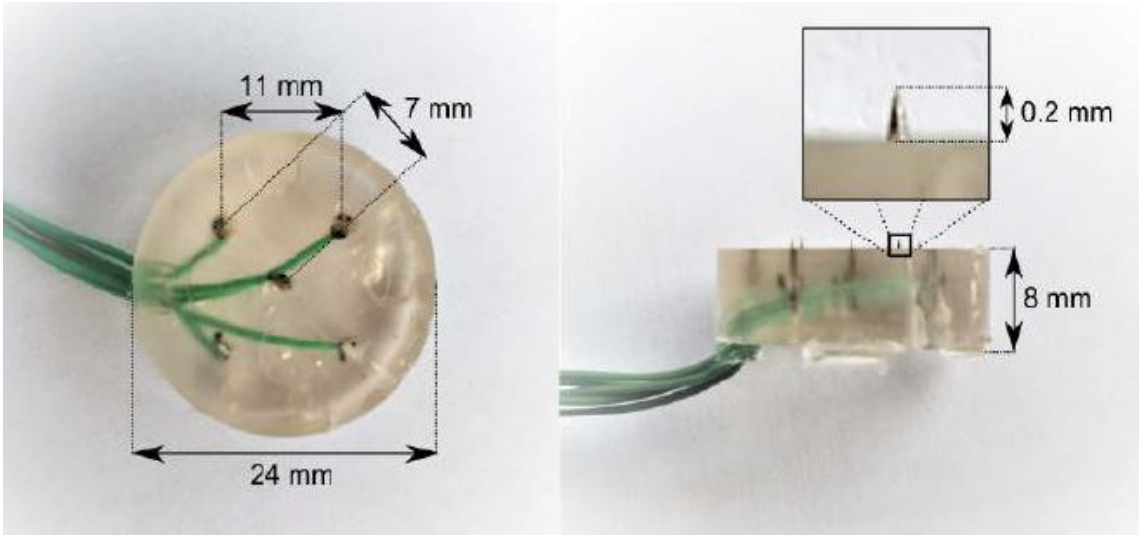


Figure 2.4: The current used electrode, adapted from van den Berg et al. (2018) [64]. Originally developed by Steenbergen et al. (2012) [61].

curve determines the detection probability while the responses and stimuli of the subject are used. This curve is shown in Figure 2.5. The left graph of the figure shows an experimental setting where for three different settings stimuli are given to a subject. A closed unit represents a detected stimulus and an open unit an undetected stimulus. These stimulus response pairs are utilized for generation of the psychometric curve (the right graph in Figure 2.5) [13].

This psychometric curve was created using logistic regression by determination of the probability p with amplitude x. Equation 2.1 with as threshold parameter $\alpha(t)$ and slope parameter $\beta(t)$ was used for the generation of the psychometric curve [12].

$$p(x) = (1 + exp(\beta(\alpha - x)))^{-1} \tag{2.1}$$

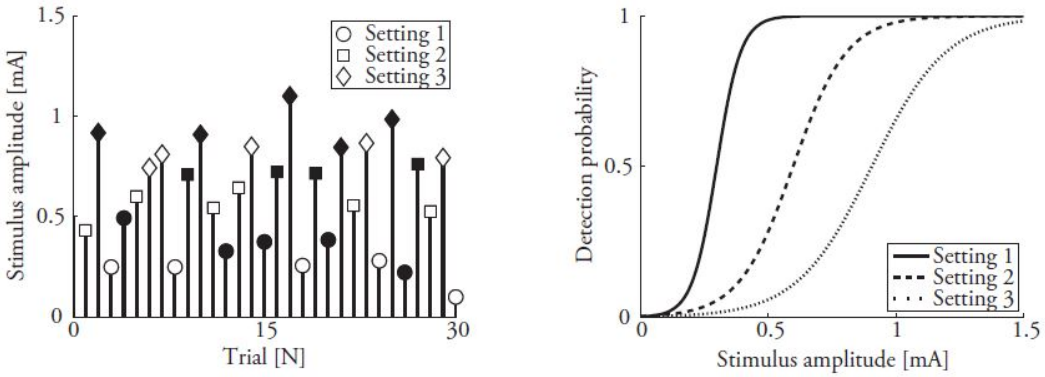


Figure 2.5: Left: Stimuli with three different settings are shown in a random order. Open figures are non-detected stimuli and closed are detected stimuli. Right: Psychophysical function, generated by using the stimulus response pairs from the experiment as shown in the left figure. Adapted from Doll et al. (2016) [13].

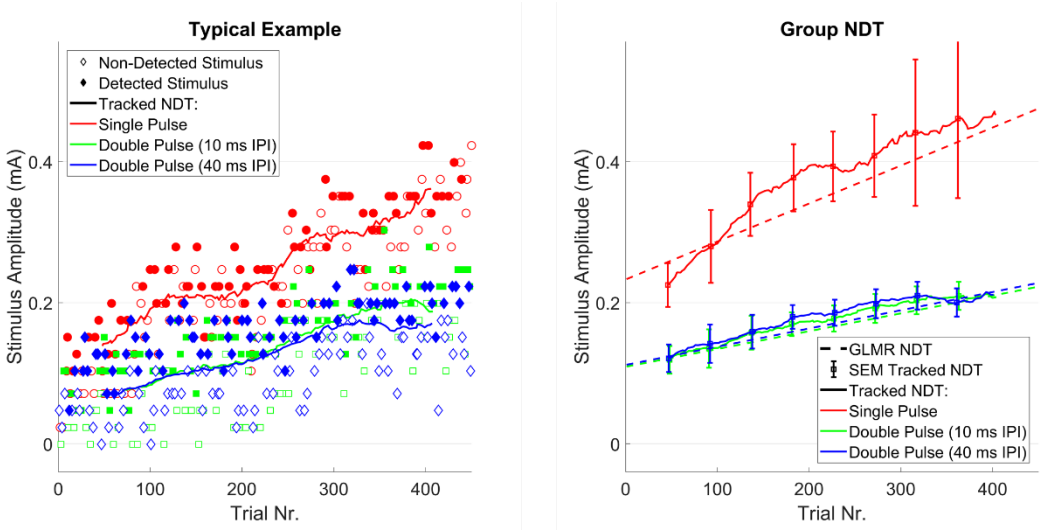


Figure 2.6: An example of a typical experiment measurement, including the calculated group NDTs. Obtained from van den Berg et al. (2021) [65].

Multiple theshold tracking

Previous research from the BSS research group from the University of Twente Doll et al. (2014) introduced the MTT procedure, currently known as the NDT procedure [13]. Doll et al. advised the use of a random staircase procedure instead of another more simple method. Despite the habituation present in both methods, the random staircase procedure realized a higher precision [13]. This staircase procedure triggers an increase in case the stimulus was considered not-perceived and decreased when it was perceived [12]. In fact, Doll et al. included the developed MTT procedure with different stimuli sets by varying settings. These stimuli sets contrast on properties regarding the number of pulses (NoP), inter-pulse interval (IPI) and pulse-width (PW) [13]. An example of a typical experiment and the calculated group NDTs is shown in Figure 2.6.

2.4.2 Neurophysiological measures of nociceptive pain

The data used for this research consists partly of EEG data. Therefore, the neurophysiological measures used in this research focuses on the electrical aspects in this field. EEG and EP are discussed underneath and other neurophysiological measures (e.g. magnetic resonance imaging) are considered outside the scope of this study.

Electroencephalography

The first neurophysiological recordings of electrical activity were achieved in the late 1800s. This was done on animals by Richard Caton. It was not until 1924 that Hans Berger applied EEG onto human subjects [66]. Nowadays EEG is widely used in many research areas as, for instance, the field of cognitive sciences, psychology, physiology, neuroscience and medicine [67]. One of the most decisive applications of EEG in medicine, is in the field of epileptic evaluation. Classification of epilepsy and seizures is an important function of the technique [68]. Additionally, in disorders such as metabolic or sleep disorders, neurodegenerative diseases or brain death EEG is clinically relevant [69]. Besides, in other area's such as pain research being investigated in this research, EEG is a well-known technique. This can be seen by various studies, such as the research of Slater et al.

(2010) [70]. They used time-lock EEG recordings in newborn infants and showed electric potentials, able to be categorized as a nociceptive-specific potential [70].

In order to generate an EEG recording, bioelectric currents transported by ions are translated into electric currents transported by electrons. A sensor collects the information. This information migrates through an amplifier to the actuator, presenting a final user interface display of the biological information. The information, visualized in the EEG recording starts with singular current dipoles, being too small for interpretation. These currents arise from pyramidal cells. However, after activity synchronization, multiple parallel dipoles are able to display the information in an EEG recording [69].

A regularly used method for obtaining an EEG recording, is by using silver-silverchloride electrodes along with a conducting gel. A reference signal respresenting the mean of all electrodes can be used to comprehend voltage differences. These voltage differences are used for the creation of a montage, realizing an overview of combinations between different EEG channels [69].

Currently, EEG is a part of the NDT-EP experiment was adapted and expanded over the years. More recent, EEG was used as an supplemental and more objective technique to explore the actual responses from the subjects.

Evoked potentials

EPs arise after application of a stimulus, and therefore functions as the response of the stimulus. There are different types of sensory stimulus inputs, namely visual, auditory and somatosensory. An example of this particular response can be seen in Figure 2.7 [69; 71]. In this study, where the NDT-EP experiment is conducted, electrical stimuli are applied to peripheral nociceptive nerves. This nociceptive EP provides information about the whole migration of the signal starting at the site of stimulus until it arrives contralateral at the somatosensory cortex and the anterior cingulate cortex (ACC) (see Figure 2.3) [20; 29; 31]. Therefore, the EP represents the status of the myelinated peripheral and central nerves [71].

There are different methods for describing the latency of the present peaks in the EP waveform. The first method is shown in Figure 2.7 as P2 and N2. P1 and N1 are located before P2 and N2, but they were not visible in Figure 2.7. P1 defines the first positive peak in the EP, N1 the first negative peak and so on. An additional method for the description is by using the time in ms after the stimulus. For instance, in Figure 2.7, P1 can be called P100 in this second way of defining the peaks, since this peak is present about 100 ms after the application of the stimulus [69; 71]. In this research, mainly the first notation is used.

2.4.3 NDT-EP experiment

The experimental procedure

For the NDT-EP experiment, the subject is seated in a chair facing the wall to be able to focus on one point on the wall [65]. Intra-epidermal electrocutaneous stimuli are applied to the subject by using the IES-5 electrodes by Steenbergen et al (2008) being previously described and shown in Figure 2.4 [62]. These electrodes are placed on both hands consecutively. A TENS electrode, sized 50 by 90 mm serving as an anode, is located distally from the stimulus electrode. A constant current stimulator (NociTRACK AmbuStim, University of Twente, Enschede, the Netherlands) was used for generation of cathodic, rectangular shaped pulses and manually controlled by the subject. The information for

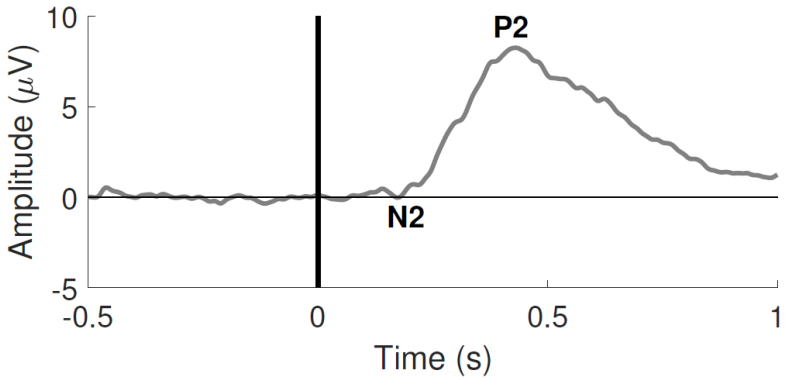


Figure 2.7: An example of an evoked potential at CPz-A1,A2 with the P2 and N2 parameters. Retrieved from van den Berg et al. (2018) [64]

this pulse generation is sent from a laptop to the NociTRACK stimulator through Bluetooth in the familiarization phase and through a trigger cable during the main experiment. Additionally, the response data is returned to this laptop. Besides, trigger codes originating from the NociTRACK stimulator are received by the EEG amplifier (ANT Neuro 72-channel Refa EEG amplifier) and sent to a second laptop by an optical cable. This information is sent together with the cortical activity from the ANT Neuro Waveguard 64 EEG cap through this EEG amplifier to the same laptop with ANT Neuro Asalab software for EEG recording [46; 65].

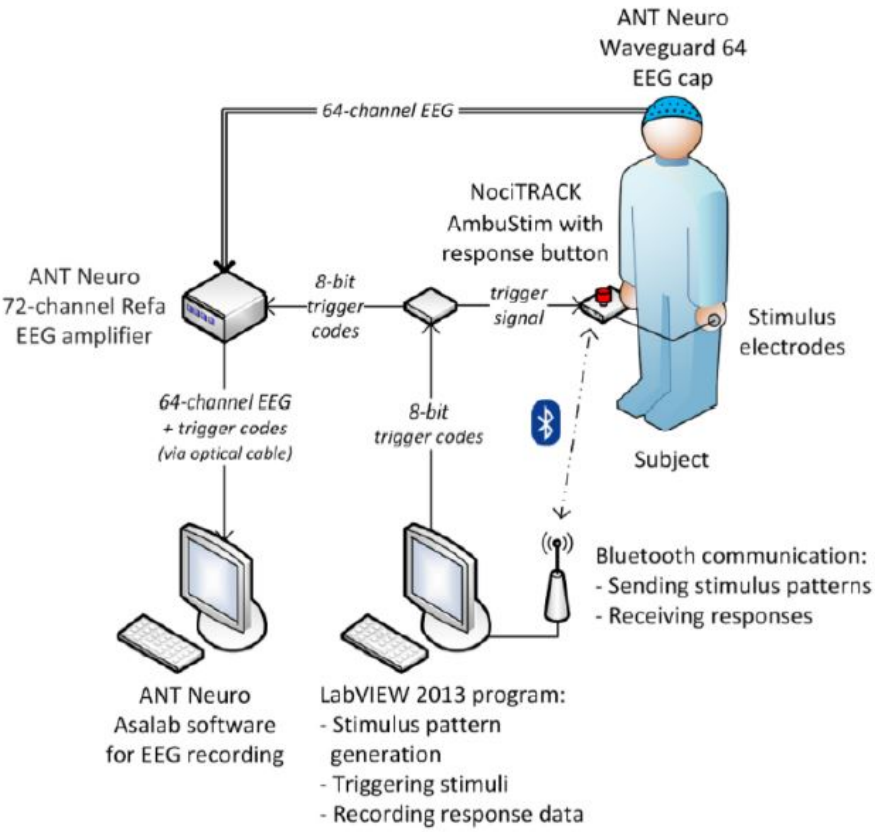


Figure 2.8: The adapted NDT-EP setup, retreived from Schooneman et al. (2015) [72].

Previous results

After the development of the MTT experiment by R. Doll et al. (2016), EEG was added to the research protocol and was entitled the MTT-EP experiment (currently the NDT-EP experiment). B. van den Berg et al. (2018) worked on a linear mixed model and researched possible improved analysis of the EP using a linear mixed model during the NDT determination. Using a linear mixed model it was found that the amplitude of the EP was considerably influenced by conscious detection of the electrical stimulus. Additionally, a smaller influence of the amplitude on the EP was observed at stimulus parameters [13; 64].

While using the previous developed NDT-EP experiment, Berfelo et al. (2019) compared the results of the NDT-EP experiment of the HCs at the St. Antonius hospital with HCs measured at the BSS research group at the University of Twente [46]. It was concluded that the results of both studies, at the University and the clinic were similar and can therefore be replicated. Additionally, the HCs were compared to patients suffering from FBSS. In those patients, the NDTs and EPs of the FBBS patients showed altered behaviour, while being compared to pain-free subjects. The NDTs were higher in patients suffering from FBSS and they also increased faster compared to the pain-free subject group. In chronic pain patients, higher stimuli were required for stimulus detection. The psychometric curves also showed a lower steepness in chronic pain patients. It was concluded that the NDT-EP experiment is likely to provide insight into altered neurophysiological mechanisms in subjects being central sensitized and suffering from chronic pain. Therefore, it was recommended to do further research of this altered behaviour in the NDTs and EPs of chronic pain patients [46].

On this note S. Gefferie et al. (2020) explored similar parameters from the NDT-EP experiment in DM patients with and without chronic pain and compared the results with a small fiber neuropathy (SFN) model [16]. The latter was established by the administration of lidocaine to pain-free subjects. Again, this research showed some alterations in the psycho-physiological parameters from the NDT-EP experiment. It was shown that subjects receiving the lidocaine treatment, showed lowered EP amplitudes, however, similar detection probabilities. Between DM patients with and without chronic pain and between the DM group and pain-free controls, EP amplitudes and detection probabilities differed. However, in the previous comparisons, EPs were considered similar. It was recommended by this paper to further investigate the influence of other possible confounders as for instance, demographic factors [16].

2.5 Classification learning

Over the last couple of years, many machine learning techniques are applied in the statistical field. Machine learning (ML) uses input data instead of a programmed script in order to accomplish a specific task. This process improves conducting the specific task after more experience is gained. This is the training part of the process [73]. Afterwards test data can be used for validation of the performance of the model. An example of a validation technique is k-fold cross-validation and is described in Section 2.5.4 [74].

2.5.1 Current applications of Random Forest Classification

Random forest (RF) classification is a popular machine learning technique and is functioning as a predictive classification model. RF classifiers are beneficial for processing data containing a large number of features. This classification technique is often applied

in studies, aiming to decrease features in order to increase the classification performance of the model [75]. Hereby, RF classifiers are applied in a large number of medical and non-medical fields, for example in gene selection or patients suffering from Alzheimer's disease [76; 77].

Recently, RF classification has been successfully applied in central sensitized subjects. The University of Groningen researched RF classifiers for prediction of high (CSI score: 40 to 100) or low (CSI score: < 40) central sensitization in patients suffering from chronic low back pain. Shapley values were calculated for predictor importance quantification. It was concluded that the accurate (81%) RF classification results could indicate that patients suffering from chronic low back pain including a high or low CSI score, showed different gait outcomes [78].

2.5.2 Decision trees

Decision trees are the foundation of random forest (RF) classifiers which are being discussed underneath. However, they can also function as an independent tree-structured classifier. Nevertheless, an important difference between both classification models is the learning algorithm. Though, the structure of those decision trees are specified, since they occur in RF classification [79].

Decision trees work top-down. At the top, a specific parameter is discussed and is divided into one or more values of that parameter. This tree propagates and each division is entitled, 'node'. This continues until the final node, the 'leaf'. An example of a decision tree is shown in Figure 2.9 [79].

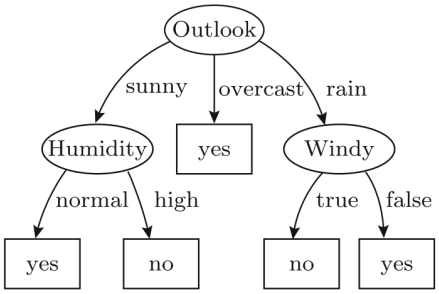


Figure 2.9: An example of a decision tree with a binary classification questioning the golf possibility depending on the weather. Adapted from Johannes et al. (2016) [79].

2.5.3 Bagging

RFs were described by Breiman et al. (2001) and is an important method based on 'bagging' [80; 81]. When bagging is applied to RF classification, various decision trees are combined into a RF [81]. This fundamental part of RF classifiers, bagging origins from bootstrap aggregation and can be applied to several classification or regression techniques. Bootstrap aggregation is a method for predicting improvement by reducing the variance in unstable procedures and it reduces the chance of overfitting on a particular data set [82; 83]. Bagging uses subsets of training data points to create a model. In this method, the repeated occurrence of a particular data point is not uncommon. After bagging, the calculated mean of the models is considered, the final prediction. Data points that were not used in a certain subset of data can be referred to as out of bag (OOB) data. This data can be used for calculating performance of the classification model. The definition of

the bagged estimate ($\hat{f}_{bag}$) can be referred to as in Equation 2.2. Here, the number of the bootstrap sample is defined by $b = 1, 2, ..., B$. The prediction of each bootstrap is entitled $\hat{f}^{*b}$ [82].

$$\hat{f}_{bag}(x) = \frac{1}{B}\sum_{b=1}^{B}\hat{f}^{*b}(x) \tag{2.2}$$

Delta error

There are multiple techniques available for determination of important features of the RF classification model. One of these methods being used in this research is the delta error. This method uses randomization every subset of data, being created by bagging. After generation of a bootstrap, the accuracy is calculated by using the OOB data samples as test data. Thereafter, these data samples from the OOB data are randomized of one feature. Subsequently, the accuracy was calculated again. The accuracy decrease is being calculated for all bootstraps and averaged. This average defines the importance of a certain feature [82].

2.5.4 K-fold cross-validation

Cross-validation is an overall used method for predicting evaluating performances of a machine learning model. It is based on the division of data where a large part of the data is used for training the model and the resulting data for the validation [74]. The term 'k' in k-fold cross-validation, used in this research, describes the number of established folds in the data. One part is the test data set and the rest is training data [84]. Figure 2.10 represents an example of a 5-fold cross-validation.

Figure 2.10: Example of k-fold (in this figure $k = 5$) cross-validation. Every blue box represents train data set and every green box a test data set. Every row is a fold and includes the whole data set. In every fold, another part of the data is considered test data. Adapted from Fedotenkova et al. (2017) [85].

Afterwards, the sum of all folds k over iterations i can be calculated and divided by the number of folds k as in Equation 2.3 [82].

$$Evaluation_{Total} = \frac{1}{k} \cdot \sum_{i=k}^{k}(Evaluation_i) \tag{2.3}$$

2.6 Implications

Chronic pain is a globally occurring health complication affecting 19% of the European population [2]. This large number of chronic pain patients, together with the large economic and social burden, implies the requirement of chronic pain treatments based on neurophysiological mechanisms in nociceptive processing. However, conventional treatments focus on pain relieve and symptom control.

To enable treatment protocols based on neurophysiological mechanisms, the observation of these mechanisms in nociceptive processing needs to be improved. The BSS research group from the University of Twente took a first step in the measurement of altered nociceptive processing in chronic pain with their NDT-EP experiment. This was explored in the Nocicept project [86].

Previously, statistical differences of the NDT and EP features (obtained by the NDT-EP experiment) between pain-free subjects and chronic pain patients were observed on group level. However, for identification of chronic pain in a single subject, individual features need to be explored to improve this observation.

2.6.1 Primary research objective

The objective of this study is to explore the relevance of different features from the NDT-EP experiment in order to observe pain processing in different subject groups. In order to define the feature performance, the features will be used to see to what extent they confirm a diagnose established by a doctor. In that event, techniques in the of nature of receiver operating characteristic (ROC) analysis will be performed. By measurement determination (e.g., the area under the curve) a proclamation can be made about the features' relevance. The three groups that are being compared in order to answer the research question are (1) pain-free controls versus chronic pain patients, (2) pain-free controls versus patients suffering from DM without chronic pain and (3) DM patients with and without chronic pain.

Primary research question:

Which biomarkers from the NDT-EP experiment are relevant for the identification of nociceptive processing on individual level in order to observe variations in neurophysiological mechanisms in subjects suffering from chronic pain (i.e., FBSS and DM patients with chronic neuropathic pain) compared to pain-free controls (i.e., healthy controls and DM patients without pain)?

2.6.2 Secondary research objectives

Secondary research objectives from this study are focusing on the feature performances in feature combinations and on the quality of the NDT-EP experiment. The aim of the first question is to explore which features are relevant in possible feature combinations. This can be studied by using a machine learning classification technique. The second research question focuses on possible improvements of the measuring method itself. Recommendations will be presented by using results from primary research question and the first secondary research question. Therefore, this part of the research is being discussed in Future perspectives (see Section 7.1, Chapter 7).

Secondary research questions:

- Which biomarkers from the NDT-EP experiment are relevant in possible biomarker combinations for the identification of nociceptive processing on individual level in order to observe variations in neurophysiological mechanisms in subjects suffering from chronic pain (i.e., FBSS and DM patients with chronic neuropathic pain) compared to pain-free controls (i.e., healthy controls and DM patients without pain)?
- How can the NDT-EP experiment be adapted into a more clinical applicable experiment to improve diagnostic performance and reduce the experiment duration?

Chapter 3

Methods

3.1 Study design

This explorative, retrospective study was a collaboration between the BSS research group at the University of Twente (Enschede, The Netherlands) and the department of anesthesiology, intensive care (IC) and pain medicine from the St. Antonius Hospital (Nieuwegein, The Netherlands). Approval of this research was provided by the Medical research Ethics Committees United (MEC-U: NL66.163.100.18).

3.2 Study population

The data used in this study is acquired at the St. Antonius Hospital. HCs, patients suffering from FBSS and patients suffering from DM were included [46; 16]. All exclusion criteria for each subject group can be found the research protocols. The exact groups are shown in Table 3.1. These subject groups are classified by the conditions: DMnp, DPNp, HCs and FBSS.

Table 3.1: The number of subjects and measurements in each subject group from the St. Antonius Hospital. The explored conditions are: diabetes mellitus without pain (DMnp), painful diabetic poly neuropathy (DPNp), healthy control (HC) and failed back surgery syndrome (FBSS).

Subject group	Subjects before exclusion	Excluded subjects	Subjects after exclusion	Measurements
DMnp	22	2	20	40
DPNp	16	2	14	28
HC	42	5	37	74
FBSS	20	4	16	32

3.3 The experimental procedure

Before the start of the experiment, the preparation procedure was completed, including sterilization of electrodes, calibration of the NociTRACK stimulator, signing of the informed consent, completion of questionnaires (i.e., case report form, NRS, CSI), removal of mobile phones from the room and, if necessary, neurological evaluation. Thereafter, the

Table 3.2: The stimulus protocol during an NDT-EP experiment with number of pulses (NoP) describing the amount of applied stimuli and the inter pulse interval (IPI), representing the time between two pulses. Adapted from Berfelo et al. (2020) [87].

NoP	**IPI**	**Experiment Randomized max. 1.5 mA**
1	**n.a.**	150
2	**10 ms**	150
2	**40 ms**	150
Total		450 stimuli

EEG cap, the stimulating electrode and TENS electrode were applied at one hand, while the stimulator was held in the other hand [87].

The experiment was initiated by a familiarization phase in which a series of stimuli with increasing amplitude was applied. The subject was asked to release the response button as soon as a perceived stimulus was clearly noticed. Then they were instructed to do the same when a stimulus was slightly noticed. For a good familiarization, this procedure was repeated twice. This data was used for the estimation of the sensation threshold [87].

For the actual experiment, the subject was asked to release and repress the NociTRACK stimulator when a stimulus sensation is noticed. The stimuli for the familiarization and the actual experiment are described in Table 3.2 [87]. Finally, this whole procedure was repeated on the other hand [87].

3.4 Data Processing

Figure 3.1 represents a general overview of the analysis pathway in this study. This figure includes the data collection and all pre-processing and analysis steps.

The data was processed and analyzed in Matlab R2017a (The MathWorks, Inc., Natick, Massachusetts, United States). Matlab R2020b (The MathWorks, Inc., Natick, Massachusetts, United States) was used for the second part of the research considering the fast developing machine learning techniques [88; 89].

3.4.1 Latency determination

The paragraph evoked potentials from subsection 2.4.2, describes different notations for evoked potentials. In this study, the latencies will be referred to as P1, P2, N1 and N2. However, only the N1 and P2 are observed. The P1 was indistinguishable in the data. As a result of low frequency filter settings used in this research, an N2 peak was absent. Therefore, the N2 was considered a filter artefact and was not taken into consideration in this research.

Both P2 and N1 were derived from the global field power (GFP) for each research group from different previous studies from the St. Antonius Hospital (see Table 3.1). N1 was quantified as the maximum GFP between 150 and 200 ms and P2 ranged between 350 and 500 ms. Thereafter, the means of both latencies were calculated with a weight factor based on measurement size per subject group.

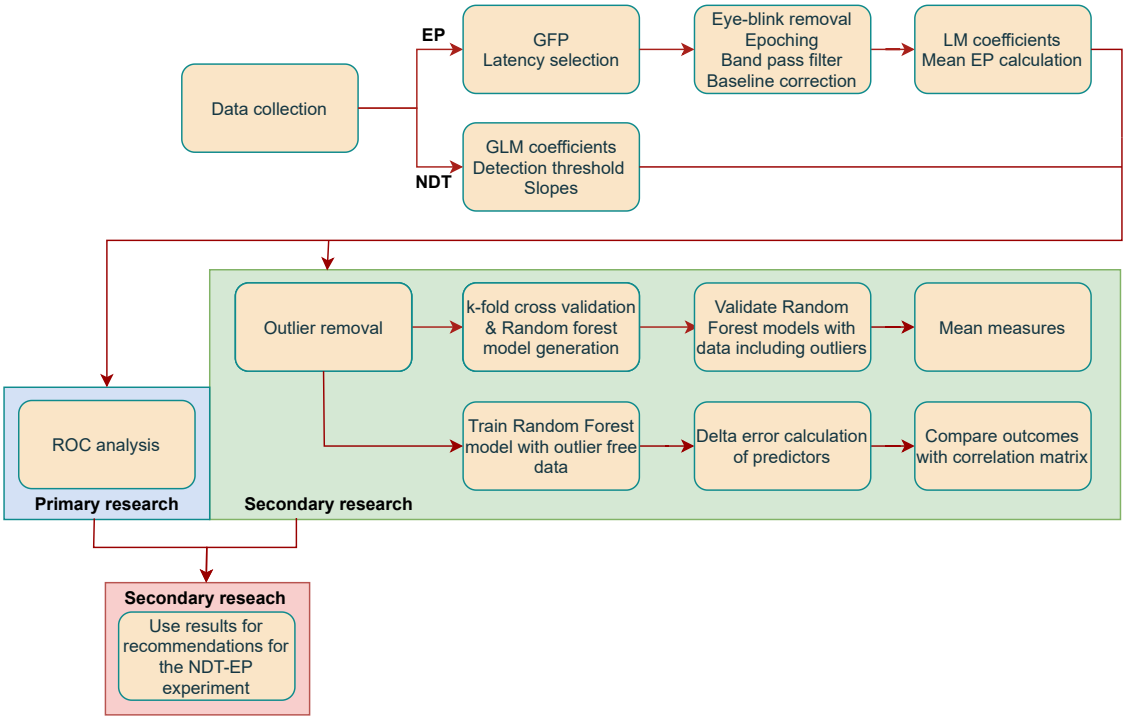


Figure 3.1: Flowchart representing an overview of the methods. The first part describes data collection and processing including two pathways. One for the evoked potential (EP) and one for the nociceptive detection threshold (NDT). The blue box contains the receiver operating characteristic (ROC) analysis for the primary research question. The green box represents the secondary research question in terms of combined features. The red box is a recommendation for the nociceptive detection threshold - evoked potential (NDT-EP) experiment.

3.4.2 Channel selection

In accordance with previous research in terms of the NDT-EP experiment, the early components from the EP originated contralateral for N1 (in this research, T7-F4). For P2 these components were derived centrally (in this research, CPz-A1,A2) [46; 64].

3.4.3 Nociceptive detection threshold analysis

The NDT is determined from the psychometric curve. Here, the detection probability is defined by 0.5. For parameter estimation of the psychometric curve, a generalized linear model (GLM) was used. Subsequently, these parameters from the psychometric curve were used for computation of the detection threshold. The predicted response of the NDT was written in the Wilkinson's notation and shown in Equation 3.1.

$$D \sim 1 + TRL + SP1 + SP2\_10 + SP2\_40 \tag{3.1}$$

In this equation the outcome measure, known as the response is referred to as detection (D). Habituation is expressed in the GLM by means of the number of received stimuli (TRL). The stimulus characteristics are described with regard to the pulse amplitudes of the first pulse (SP1), of the second pulse with 10 ms IPI (SP_10) and the second pulse with 40 ms IPI (SP_40). Since the analysis required individual GLMs, no random effects were taken into consideration.

3.4.4 Evoked potential analysis

The data was preprocessed using the Matlab Toolbox FieldTrip, specialized in EEG analysis [90]. Independent component analysis was used in order to remove eye-blinks. Additionally, electromyography and movement artifacts were visually selected and removed from the data. Thereafter, epoching of the evoked potential was established using a window of -0.5 s up to 1.0 s around the stimulus. A band pass filter was applied ranging from 0.01 to 40 Hz and a baseline correction from -0.5 s up to the stimulus onset was conducted.

For visualisation of the different EPs a linear model (LM) was used. This LM, likewise in Wilkinson's notation, is shown in Equation 3.2.

$$EP \sim 1 + D * TRL + SP1 + SP2\_10 + SP\_40 \tag{3.2}$$

Similar to the GLM (Equation 3.1), this LM contains the amplitudes for single pulse stimuli and double pulse stimuli with 10 and 40 ms IPI, respectively. Nevertheless, Equation 3.2 contains an extra interaction next to the main effects, known as $D * TRL$. This interaction represents possible variations after change in trial number in brain activity proportionate to stimulus detection.

3.5 Individual feature observation

The data was obtained from from previous studies described in Table 3.1. Outcome measures needed to be defined for the ROC curves and were classified by two conditions. Each set of conditions will be referred to as a comparison (see Table 3.3).

Table 3.3: Groups of subjects being compared in this research.

Comparison
FBSS versus HC
DPNp versus HC
DMnp versus HC
DMnp versus DPNp

3.5.1 ROC curves

The features explored in this research are: (1) the direct coefficients from the GLM or the LM (e.g., amplitudes), (2) the information derived from these coefficients (e.g., slopes) and (3) the information directly obtained from the NDT-EP experiment (e.g., means). A summary of these features is available in the Appendix in Section 8.1. For exploration of the feature classification performance in terms of nociceptive processing, ROC curves were used. These curves inform about the sensitivity as a function of 1-specificity of a binary classifier. The sensitivity (see Equation 3.3) is calculated with the number of true positive (TP) and false negative (FN) classifications at a certain threshold. The specificity (see Equation 3.4) is determined by using the true negative (TN) and false positive (FP) classifications. In this research, the previous discussed subject categories from each comparison in Table 3.3, were considered outcome measures.

$$Sensitivity = \frac{TP}{TP + FN} \tag{3.3}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3.4}$$

For this part of the study, no outliers were removed. Matlab R2017a (The MathWorks, Inc., Natick, Massachusetts, United States) was used for producing ROC curves for each feature of the NDT-EP experiment (see Appendix) [88; 91]. For quantification of the differentiation quality of each feature, the area under the curve (AUC) was calculated by application of trapezoidal numerical integration. An AUC of 0.5 implies incompetent differentiation of a feature between the outcome measures. Consequently, determination of the p-value was necessary for assessment of the AUC significance. The required standard error (SE) for the p-value calculation, was determined by the method of Hanley and McNeil (1982) [92].

Optimal cut-off value

Several additional parameters can provide insight into the differentiation quality of a certain parameter. The optimal point in an ROC curve can be calculated in multiple ways [93]. However, a method calculating the minimal distance to (0,1) seemed most sensitive in this data set. At (0,1) the sensitivity and specificity are maximal. The minimal distance to a potential optimal point can be referred to as 'the closest to (0,1) criteria' (ER) and is calculated in Equation 3.5. Se and Sp signify the sensitivity and specificity at a cut-off value (c) at the ROC curve [93].

$$ER(c) = \sqrt{(1 - Se(c))^2 + (1 - Sp(c))^2)} \tag{3.5}$$

This method was chosen applied sensitivity was considered most important since this decreased the possibility of false identification of the absence of pain in chronic pain patients. Next to the sensitivity, the specificity and the accuracy were calculated.

Cross correlation

Finally, the correlation coefficient between all different features in each outcome comparison was calculated. This analysis was performed in Matlab and visualized through a heatmap.

3.6 Combined feature classification performance

3.6.1 Outlier removal

The NDT and the EEG data was tested for normality by creating QQ-plots in Matlab. Predictors being considered not normally distributed were transformed by using a cube root transformation. Cube root enables transformation of negative values, adjacent to, for example, a log transformation.

Afterwards, the interquartile range (IQR) was used for outlier quantification in the NDT and EEG data for all four conditions. Any value outside the Q1 and Q3 boundaries ($< Q1 - 1.5 \cdot IQR$ and $> Q3 + 1.5 \cdot IQR$, representing 25% and 75% of the data, respectively) were removed from the data together with the corresponding features appropriate to that particular measurement from the same latency or NDT.

3.6.2 Random forest classifier

Delta error

A complete RF model was created per comparison outlined in Table 3.3. The OOB classification was plotted as a function of the number of trees to determine the stability of the model after a certain amount of trees. Additionally, the delta error was calculated to determine the predictive characteristics of each feature. ROC curves were presented for evaluation of the performance of the of the RF classification model.

Cross-validation

While the group size of all conditions is limited, 5-fold cross-validation has been applied. The data was randomly split in Matlab R2020b (The MathWorks, Inc., Natick, Massachusetts, United States) [89]. However, each two measurements from one subject were assigned to the same data set in order to decrease the possibility for overfitting on subject specific characteristics. Splitting the data this way resulted in five different training data sets (80% of the data) and five different validation data sets (20% of the data). Each training data set contained cleaned data while it was aimed to prevent model fitting based on outliers. On the other hand, the validation data contained data with the outliers, while in clinical setting, these outliers are also existent.

Accuracy calculation

After every fold from the 5-fold cross-validation in each comparison group, an RF classification model was trained with the training data set. The predictive values of the model were calculated using the test data set. Subsequently, for each fold, several measures were calculated (i.e., the accuracy, sensitivity and specificity). Afterwards, each mean was calculated in terms of the measures.

Chapter 4

Results

4.1 Subject characteristics

An overview of the group subject characteristics is represented in Table 4.1. The calculated p-values after one-way ANOVA (significance level <0.05) for the difference between all groups is, 0.001 for age and is <0.001 for BMI, NRS of the week before and during the experiment and the CSI score.

Table 4.1: Subject characteristics of pain experiencing failed back surgery syndrome (FBSS) subjects, diabetic poly neuropathy (DPNp) subjects, the pain-free diabetes mellitus (DMnp) subjects and the healthy control (HC) subject group. Note: Medication type differs per subject group.

	FBSS	**DPNp**	**DMnp**	**HC**
Number of subjects	16	13	20	37
Sex (M/F)	9/7	11/2	9/11	16/21
Age in years (mean ± SD)	50.1 ± 9.1	64.2 ± 7.6	53.4 ± 16.5	42.7 ± 18.8
BMI (mean ± SD)	25.8 ± 3.3	30.6 ± 4.7	26.8 ± 4.0	23.2 ± 3.2
NRS last week (mean ± SD)	7.3 ± 1.6	4.5 ± 1.6	0.2 ± 0.5	0.3 ± 0.9
NRS now (mean ± SD)	7.0 ± 2.1	3.2 ± 1.9	0.1 ± 0.3	0.0 ± 0.0
CSI score (mean ± SD)	44.6 ± 13.9	33.8 ± 13.9	17.6 ± 8.0	14.2 ± 8.8
Medication intake (Yes/No)	7/9	6/7	19/1	7/30
Disorder in years (mean ± SD)	4.6 ± 5.1	18.2 ± 10.3	16.0 ± 10.4	-
Pain in years (mean ± SD)	13.3 ± 13.2	9.5 ± 4.3	-	-

4.1.1 Failed back surgery syndrome

The FBSS group consists of 16 subjects whereof 9 are men and 7 women. The mean and standard deviation (SD) are 50.1 and 9.1, respectively. They range between 39 and 76 years of age. The body mass index (BMI) (25.8 ± 3.3) ranges between 20.4 and 31.2. The NRS score of last week (7.3 ± 1.6) and during the experiment (7.0 ± 2.1) range between 3 and 9 and 2 and 9, respectively. The scope of the CSI score (44.6 ± 13.9) is between 18 and 65. 7 out of 16 subjects took medicine before the experiment (mainly pain relieving and sedative medication). The duration of FBSS (4.6 ± 5.1) and pain (18.2 ± 10.3) in years range between 0.5 and 21.0 and 3.0 and 54.0, respectively.

4.1.2 Diabetic poly neuropathy

The number of subjects suffering from DPNp is 13, whence 11 males and 2 females with the age (64.2 ± 7.6) ranging between 45 and 72. The subjects' BMI (30.6 ± 4.7) range between 23.7 and 40.7. The scopes of the NRS last week (4.5 ± 1.6) and the NRS during the experiment (3.2 ± 1.9) range between 1 and 7, and 1 and 8, respectively. The CSI score (33.8 ± 13.9) extents from 12 to 57. In total, 6 subjects took medication before the experiment (predominantly DM related medication, followed by pain relieve medication and cardiovascular medication) and the duration (18.2 ± 10.3) of the DM diagnosis and the pain (9.5 ± 4.3) in years range between 4.0 and 41.0, and 3.5 and 17.0, respectively.

4.1.3 Diabetes mellitus (pain-free)

The DMnp group consists of 20 subjects. There are 9 females and 11 males between 20 and 73 years of age (53.4 ± 16.5). 19 subjects took medicine (mostly DM related, followed by cardiovascular medication) before the experiment. The BMI (26.8 ± 4.0) of the subjects are scoped between 21.4 and 33.7. The NRS of last week (0.2 ± 0.5) and the NRS during the experiment (0.1 ± 0.3) range between 0 and 3, and 0 and 2, respectively. Additionally, the CSI score (17.6 ± 8.0) is ranging between 5 and 33. The duration of DM has a mean of 16.0 and the SD is 10.4, ranging between 3 and 43 years.

4.1.4 Healthy controls

The HCs range between 18 and 73 years of age (42.7 ± 18.8) and contain 16 out of 37 males. 7 out of 37 subjects took medication (mainly, cardiovascular medication) before the experiment. The BMI (23.2 ± 3.2) ranges between 18.3 and 33.5. The NRS of last week (0.3 ± 0.9) and the CSI score (14.2 ± 8.8) range between 0 and 3, and 0 and 37, respectively. Finally, the NRS during the experiment (0.0 ± 0.0) is equal to 0 in all subjects.

4.2 Selected settings

The N1 for the subjects suffering from FBSS, derived from the GFP, is 147 ms and the P2 is 407 ms (the butterfly plots including the GFP are presented in section 8.2 in the appendix). In the DPNp research group, no N1 is identified and is therefore not taken into consideration for the weighted mean of the N1. However, the latency for P2 is, in this case, 486 ms. For the other diabetic subject group, N1 is specified at 168 ms and P2 at 383 ms. The N1 is 176 and 154 ms for both groups of HCs. Here, the latencies of P2 are 494 and 424 ms.

The final, weighted means of all groups together are set for N1 at a latency of 167 ms. For P2, this is 432 ms. The scalp topographies of all groups at all latencies are shown in the Appendix at in Section 8.6.

4.3 Single features

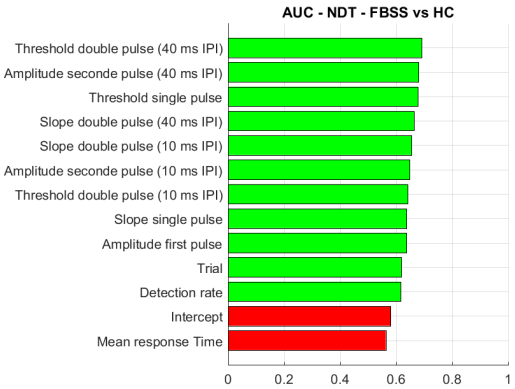
4.3.1 Area under the curve

The bar graphs for the values of the AUC for the NDT, N1 and P2 features, for all four comparisons (Table 3.3) are shown in Figure 4.1, 4.2 and 4.3, respectively. These figures present the AUCs for all feature of that particular measure (i.e., NDT, N1 and P2). Each

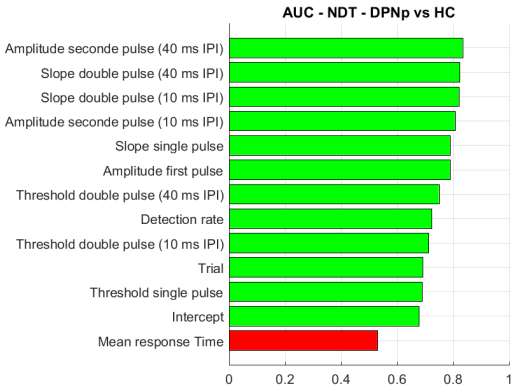
green bar in the figures represents a significant AUC value ($p < 0.05$) and each red bar is considered not significant and therefore includes a p-value ≥ 0.05.

Psychophysical features

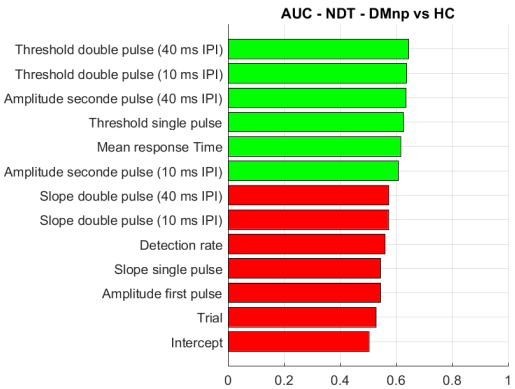
A striking result of all comparisons of the NDT measure in terms of the AUC, is the number of significant features. In Figures 4.1a, 4.1b and 4.1d, only one to three features are considered not significant. For the comparison of both pain-free outcome measures, DMnp and HC, Figure 4.1c shows a significant value for the AUCs of only six out of thirteen features. In addition, the comparisons DPNp versus HC and DMnp versus DPNp have the largest maximum values for the AUC of the features of 0.83 and 0.79, respectively. The maximum values for FBSS versus HC and DMnp versus HC, are smaller since they are 0.69 and 0.64, respectively.



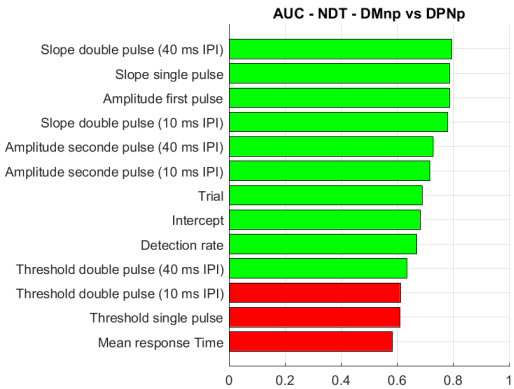
(a) NDT values for AUC with outcomes: FBSS and HC.



(b) NDT values for AUC with outcomes: DPNp and HC.



(c) NDT values for AUC with outcomes: DMnp and HC.

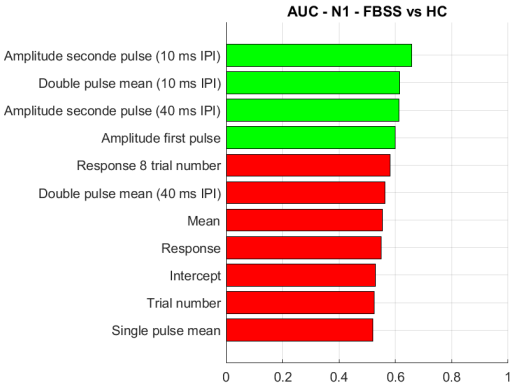


(d) NDT values for AUC with outcomes: DMnp and DPNp.

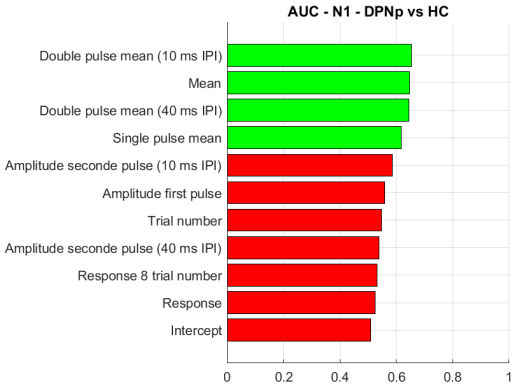
Figure 4.1: Bar graphs of nociceptive detection threshold (NDT) data of the subjects of four different comparisons. These conditions in the comparisons are failed back surgery syndrome (FBSS), painful diabetic poly neuropathy (DPNp), pain-free diabetic mellitus (DMnp) and healthy controls (HC). The y-axis show the different features, sorted on the area under the curve (AUC) value. The x-axis shows the value for the AUC.

N1 features

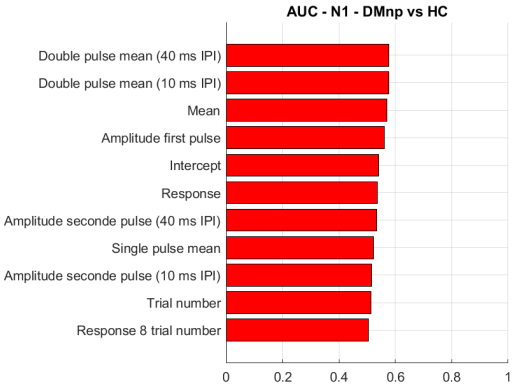
A compelling result in terms of the AUC of the latency at N1 in all comparisons are the limited significant AUCs. The amount of significant features (see Figure 4.2), ranges between zero and four out of eleven. DMnp vs HC (Figure 4.2c) does not show any significant AUCs of features for differentiating the outcome measures DMnp and HC. For DMnp vs DPNp in Figure 4.2d, only a single feature is considered significant. The other two figures (Figure 4.2a and 4.2b) only show four significant values for the AUC. Additionally, all four comparisons show small maximum AUC values (i.e., FBSS versus HC shows a maximum AUC of 0.66, followed by DPNp versus HC (0.65), DMNP versus DPNp (0.62) and DMnp versus HC (0.58)).



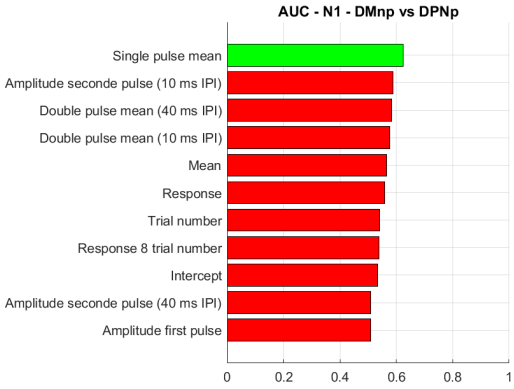
(a) N1 values for AUC with outcomes: FBSS and HC.



(b) N1 values for AUC with outcomes: DPNp and HC.



(c) N1 values for AUC with outcomes: DMnp and HC.



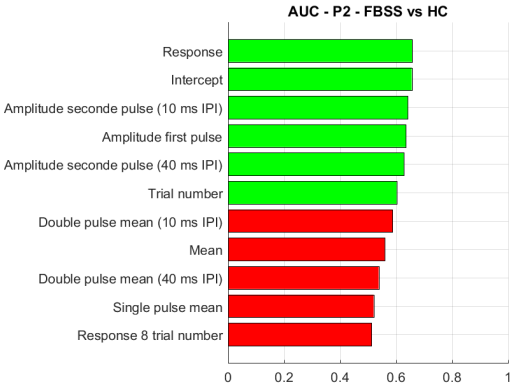
(d) N1 values for AUC with outcomes: DMnp and DPNp.

Figure 4.2: Bar graphs evoked potential (EP) data at the N1 latency of the subjects of four different comparisons. These conditions in the comparisons are failed back surgery syndrome (FBSS), painful diabetic poly neuropathy (DPNp), pain-free diabetic mellitus (DMnp) and healthy controls (HC). The y-axis show the different features, sorted on the area under the curve (AUC) value. The x-axis shows the value for the AUC.

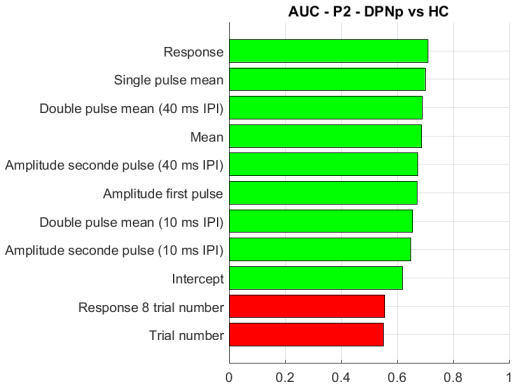
P2 features

Figure 4.3 is a representation of the AUC values of the ROC curves of all four comparisons being investigated in this study at the latency of P2. An apparent result is the number of

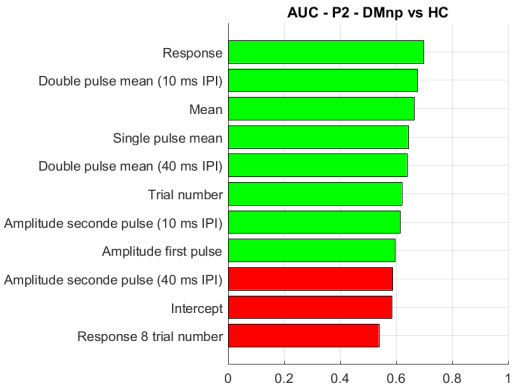
zero of significant AUC values for the features when comparing DMnp vs DPNp (Figure 4.3d). In contrast, DPNp vs HC (Figure 4.3b and DMnp vs HC (Figure 4.3c) show nine and eight out of eleven significant AUC values, respectively. FBSS vs HC (Figure 4.3a contains six out of eleven significant AUC values. In addition, the maximum AUC values are for FBSS versus HC, DPNp versus HC, DMnp versus HC and DMNP versus DPNp, 0.66, 0.71, 0.70 and 0.57, respectively.



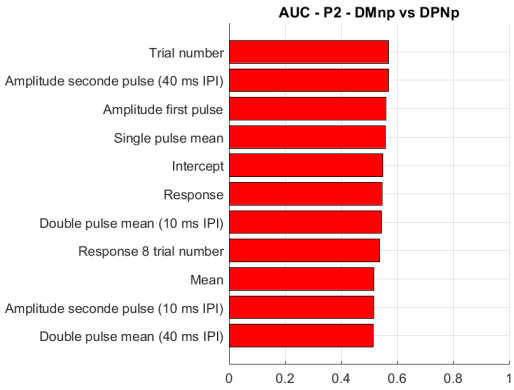
(a) P2 values for AUC with outcomes: FBSS and HC.



(b) P2 values for AUC with outcomes: DPNp and HC.



(c) P2 values for AUC with outcomes: DMnp and HC.



(d) P2 values for AUC with outcomes: DMnp and DPNp.

Figure 4.3: Bar graphs evoked potential (EP) data at the P2 latency of the subjects of four different comparisons. These conditions in the comparisons are failed back surgery syndrome (FBSS), painful diabetic poly neuropathy (DPNp), pain-free diabetic mellitus (DMnp) and healthy controls (HC). The y-axis show the different features, sorted on the area under the curve (AUC) value. The x-axis shows the value for the AUC.

4.3.2 Accuracy

The maximal values for the accuracy of individual features at the optimal point at each ROC curve is shown in Table 4.2. This table resumes the four features with the largest AUCs for each comparison. The accuracy of each feature is shown in the Appendix at Table 8.1.

Table 4.2: Features with the largest accuracy for each comparison of all comparisons (including: Pain experiencing failed back surgery syndrome (FBSS) subjects, diabetic poly neuropathy (DPNp) subjects, the pain-free diabetes mellitus (DMnp) and the healthy controls (HC))

	Feature	**Accuracy (%)**
FBSS versus HC	P2 Intercept	70
DPNp versus HC	NDT Slope double pulse (10 ms IPI)	80
DMnp versus HC	NDT Threshold double pulse (10 ms IPI)	70
DMnp versus DPNp	NDT Slope double pulse (40 ms IPI)	76

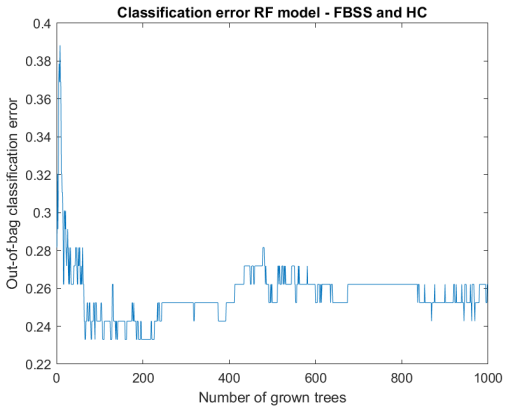
4.4 features in combinations

4.4.1 Out-Of-Bag Classification error

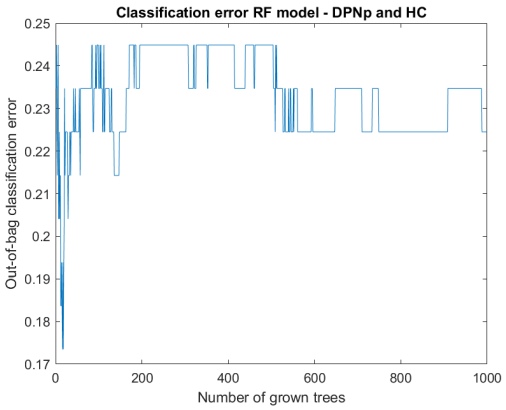
Figure 4.4 represents the number of trees used for the RF classification model as a function of the classification error. In the final RF model, 1000 trees were used. Stabilization of the graph is established in FBSS versus HC (Figure 4.4a) after about 600 trees. In the graph (Figure 4.4b) in terms of DPNp versus HC, stabilization occurs after 500 trees. Figure 4.4c shows DMnp versus HC, including stabilization after 200 trees. Figure 4.4d, representing DMnp versus DPNp is stable after about 150 to 300 trees.

4.4.2 Psychophysical and neurophysiological features

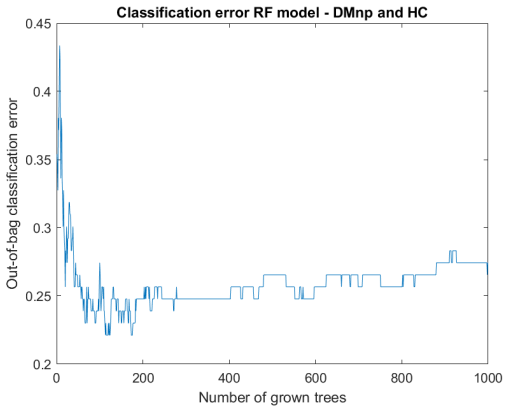
The predictor importance estimate for all four comparisons (Table 3.3), is shown in Figure 4.5. The uppermost features show highest importance estimates. For three out of four comparisons, the NDT features are predominant. In the predictor importance estimates from DMnp versus HC, the NDT features are followed by P2 features in the five most important predictor importance estimates.



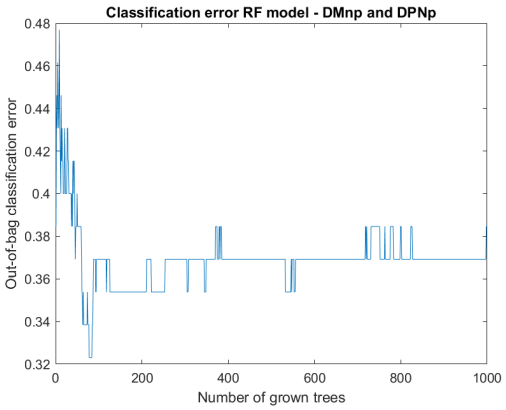
(a) Classification error of RF model in terms of FBSS versus HC.



(b) Classification error of RF model in terms of DPNp versus HC.



(c) Classification error of RF model in terms of DMnp versus HC.



(d) Classification error of RF model in terms of DMnp versus DPNp.

Figure 4.4: Each figure describes the random forest (RF) classification error against the number of grown trees. The conditions in the figures are failed back surgery syndrome (FBSS), painful diabetic poly neuropathy (DPNp), pain-free diabetic mellitus (DMnp) and healthy controls (HC).

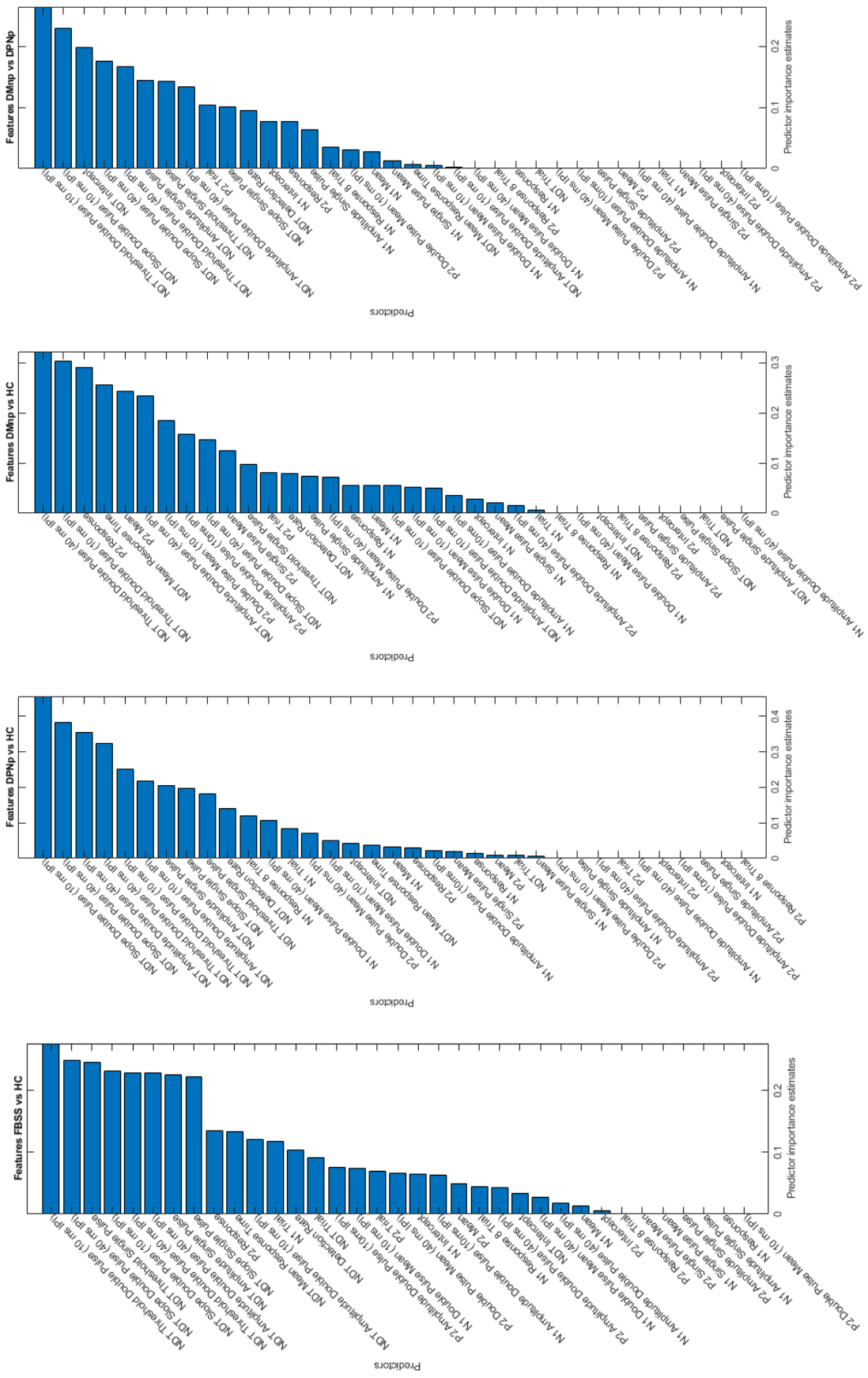


Figure 4.5: Feature (or predictor) importance estimates for the random forest (RF) classification model of all comparisons in Table 3.3. The diagnoses include: Failed back surgery syndrome (FBSS), diabetic poly neuropathy (DPNp), not painful diabetes mellitus (DMnp) and healthy controls (HC). The y-axis shows the (nociceptive detection threshold (NDT) features and evoked potential (EP) features (N1 and P2).

4.4.3 Correlation of the features

Figure 4.6, on the consecutive page, represent an overview of all correlation coefficients between all features from the NDT-EP experiment of the comparison FBSS versus HC. This figure includes cleaned data (without outliers and cube root transformed when data was considered not normally distributed). The color bar informs about the extent of correlation of between two features. Most high correlation coefficients are including NDT features relative to NDT features.

The correlation matrices for all comparisons show similar results. Therefore, only Figure 4.6 is present in this chapter. The correlation matrices for the remaining three comparisons (Figure 8.9, 8.10 and 8.11) are shown in the Appendix (Section 8.5).

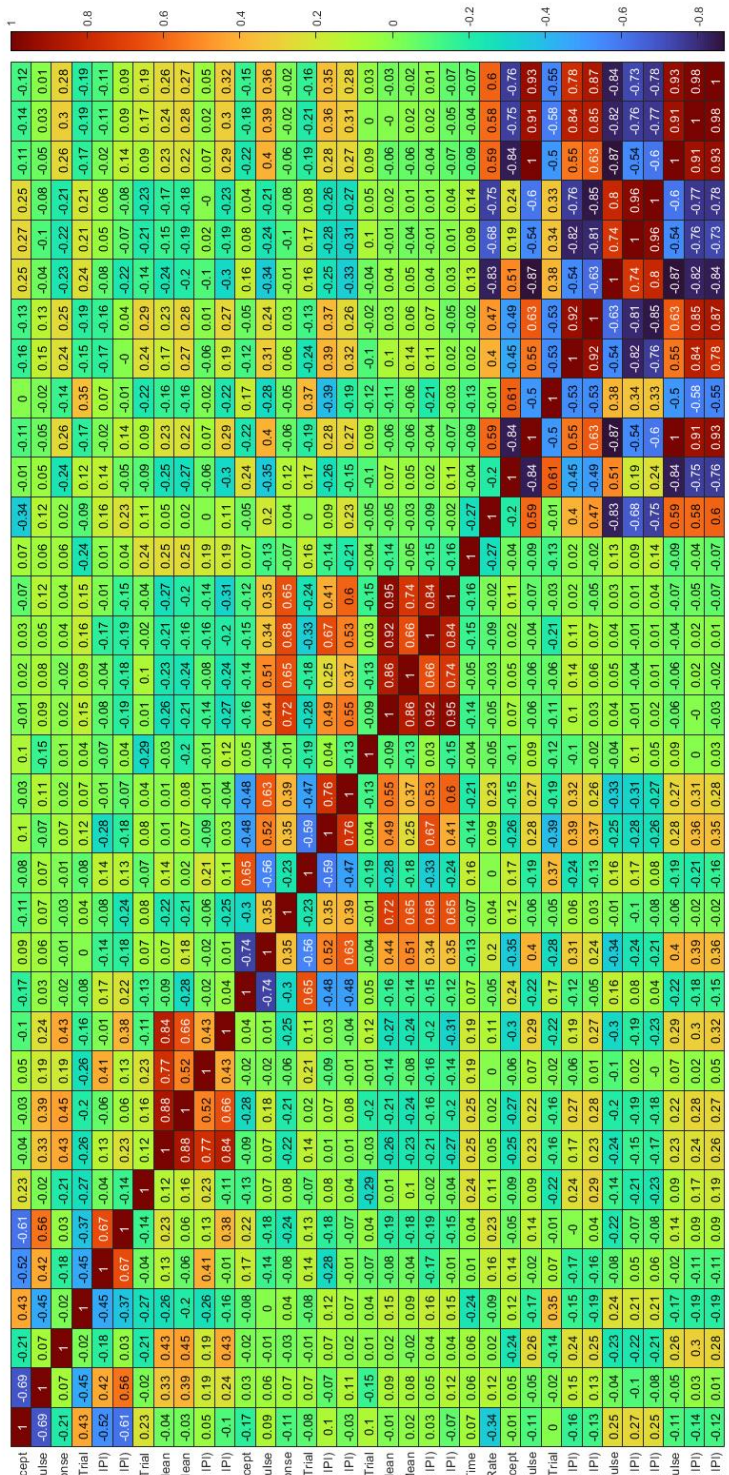
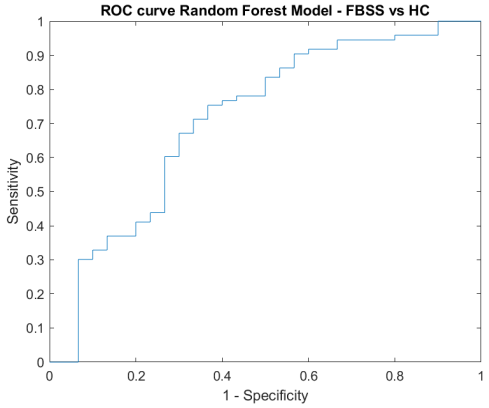


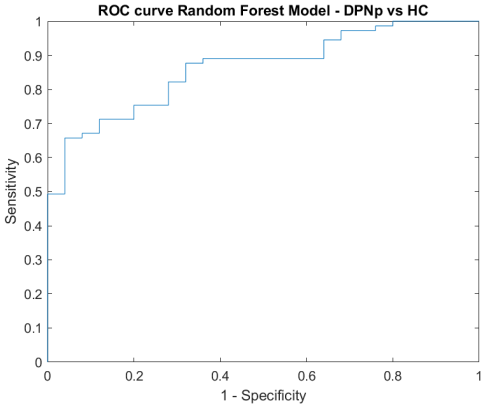
Figure 4.6: Correlation coefficients of failed back surgery syndrome (FBSS) versus healthy controls (HC).

4.4.4 ROC curves random forest classifiers

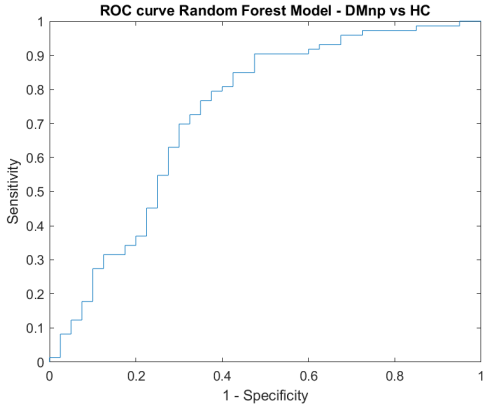
Several ROC curves of the RF classification model for the different comparisons are shown in Figure 4.7. These ROC curves are generated with the cleaned OOB train data. Figure 4.7a shows the ROC curve of the outcome measures FBSS and HC. The AUC for these outcome measures is 0.71. The comparison DPNp versus HC is visualized in Figure 4.7b. Here, the AUC is 0.86. Figure 4.7c shows the ROC in terms of DMnp versus HC, including an AUC of 0.73. The ROC curve of DMnp versus DPNp also includes an AUC of 0.73 and is shown in Figure 4.7d.



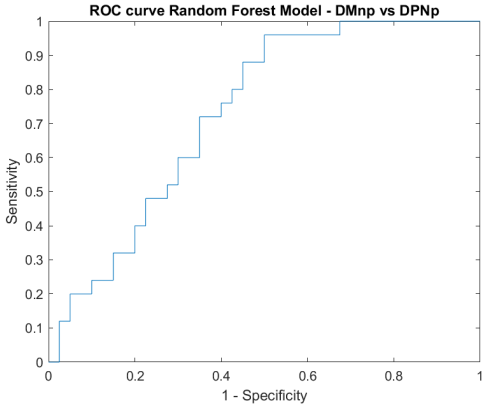
(a) ROC curve of FBSS versus HC. The AUC is 0.71.



(b) ROC curve of DPNp versus HC. The AUC is 0.86.



(c) ROC curve of DMnp versus HC. The AUC is 0.73.



(d) ROC curve of DMnp versus DPNp. The AUC is 0.73.

Figure 4.7: Each figure describes the receiver operating characteristic (ROC) curve of the random forest (RF) classification model. In these curves the sensitivity is presented as a function of 1 - specificity. The area under the curve (AUC) is shown in the subscription of each figure. The conditions in the figures are failed back surgery syndrome (FBSS), painful diabetic poly neuropathy (DPNp), pain-free diabetic mellitus (DMnp) and healthy control (HC).

4.4.5 Random forest model accuracy

Table 4.3 represents the accuracy, specificity and sensitivity of FBSS versus HC, DPNp versus HC, DMnp versus HC and DMnp verus DPNp, respectively. These measures in the table are determined after 5-fold cross validation including clean training data and outliers including validation data (however, when necessary, transformed for normal distribution).

Table 4.3: Accuracy, sensitivity and specificity specification of each of the researched comparisons after cross validation. The conditions in the table can be referred to as failed back surgery syndrome (FBSS), painful diabetic poly neuropathy (DPNp), pain-free diabetic mellitus (DMnp) and healthy control (HC).

Comparison	**Accuracy (%)**	**Sensitivity(%)**	**Specificity(%)**
FBSS versus HC	64	39	75
DPNp versus HC	73	50	81
DMnp versus HC	71	50	83
DMnp versus DPNp	67	85	45

Chapter 5

Discussion

In this research, various physiological and EEG features from the NDT-EP experiment were extracted for all four comparisons (Table 3.3). This study focused on the individual features and their performance in exploring neurophysiological mechanisms for observation of nociceptive processing. The research question answered in this part of the study is: '*Which biomarkers from the NDT-EP experiment are relevant for the identification of nociceptive processing on individual level in order to observe variations in neurophysiological mechanisms in subjects suffering from chronic pain (i.e., FBSS and DM patients with chronic neuropathic pain) compared to pain-free controls (i.e., healthy controls and DM patients without pain)?*'. This part will be discussed in Section 5.1. Important to note is that a feature can be considered a biomarker, when it consists of good differentiating abilities, resulting in a large importance of observation of nociceptive processing.

Moreover, the same features were researched to explore their combining ability in feature combinations to explore neurophysiological mechanisms for the observation of nociceptive processing. Here the research question is aimed to be answered is: '*Which biomarkers from the NDT-EP experiment are relevant in possible biomarker combinations for the identification of nociceptive processing on individual level in order to observe variations in neurophysiological mechanisms in subjects suffering from chronic pain (i.e., FBSS and DM patients with chronic neuropathic pain) compared to pain-free controls (i.e., healthy controls and DM patients without pain)?*'. This part is discussed in Section 5.2.

The results of the previous mentioned research questions were taken into consideration to answer the research question concerning possible improvements of the NDT-EP experiment. These future perspectives will be discussed in Chapter 7.

5.1 Part I - Individual features

The AUC illustrates the performance of a feature over all calculated thresholds. The NDT features (Figure 4.1) consisted of the largest set of significant AUC values and showed higher AUC values in contrast to the EEG features (N1 (Figure 4.2) and P2 (Figure 4.3)) for all comparisons. This could have been induced by a high sensitivity of EEG for artifacts and confounding. Factors responsible for this inconsistency can be procedural (e.g., time of the day) and participant (e.g., age and movement) related. Duncan et al. (2013) reviewed confounds for multiple neuro-imaging techniques and suggested manners to avoid confounding [94]. Thus, as a result of confounding factors for the EEG features, the lower and less significant AUC values implicated the importance of the NDT features. It can therefore be assumed that the NDT features were more competent in differentiating between groups, resulting in better observational abilties in nociceptive processing.

5.1.1 NDT features

It is interesting to note that features derived from double pulse stimuli resulted in larger values for the AUC, compared to features after single pulse stimulus settings. This indicated enhanced discriminating characteristics in terms of features after double pulse stimulus settings. These results were in accordance with prior research from Mouraux et al. (2014), where, similar to the NDT-EP experiment in this study, IESs were applied. They administered stimuli in various trains of two to four stimuli with 5 ms IPI. Single pulse stimuli provided a very low signal-to-noise ratio (SNR) next to a weak sensation due to a small amount of reached afferents in contrast to stimulus trains. Subsequently, a significant increase was found in perception intensity and EP magnitude after application of an increased number of pulses [95]. The latter could explain the larger values for the AUC after double pulse stimulus settings in this thesis. The inferior perception intensity, EP magnitude and superior SNR after single pulse stimuli presumably caused less variation in the data. Consequently, the double pulse features were responsible for the enhanced discriminating characteristics between groups, compared to single pulse features.

Another interesting finding was the small number of significant AUC values (i.e., 6 out of 13) in the comparison of DMnp versus HC, opposed to the other researched comparisons. These subject groups both contained pain-free subjects. In accordance, the NRS scores (during and the week prior to the experiment) of both groups were adjacent. It could be expected that this moderate pain experience compared to the other groups contributed to the small number of significant AUC values for the features. This small number was probably established by peripheral alterations as a result of nerve impairment in the diabetic subject group [18].

Features from the comparison between both subject groups suffering from DM, showed multiple significant AUC values (i.e., 10 out of 13). It was interesting that this number of significant and high AUC values was present in one encompassing diagnose. This large amount of significant AUC values was probably caused by central alterations. The CSI score from the DPNp subject group was almost twice as large as the DMnp group, indicating the presence of (starting) central sensitization.

The comparison between DPNp versus HC resulted into high values and a large amount (i.e., 12 out of 13) of significant values for the AUC. Similar to the previously discussed comparison between both DM groups, these differences between groups were a result of central sensitization. This was confirmed by the large differences between CSI scores. However, in the comparison of DPNp versus HC, even larger maximum AUC values were identified. This suggested the supplementary influence of peripheral alterations in the DPNp subject group in contrast to the HCs.

The most obvious finding to emerge from ROC analysis, in terms of the outcome measures FBSS and HC, was the large number of significant AUC values of the studied features. Similar to the comparison between both DM groups, one group did score a large CSI score (i.e., FBSS) in contrast to the other group (i.e., HC). The CSI score of the FBSS group was about three times as large as the HCs, indicating central sensitization. This presumably caused the high amount of significant AUC values. Surprisingly, the maximum of the AUCs were smaller compared to DPNp versus HC and DMnp versus DPNp. This was probably a result of the experimental procedure in combination with the affected pain-full and insensitive areas in patients suffering from FBSS. These patients often experience pain in their legs or pinched nerves in the spine [43]. The fact that the IES-5 electrode was localized on the subject's hand might bypass processing alterations in their legs. Therefore, placement of the electrode on a leg might increase the noticed differences between FBSS and HC, causing increased AUC values for NDT features.

5.1.2 EEG features

In almost all studied comparisons (see Table 3.3), P2 features showed a larger number of significant values and higher values for the AUC, in contrast to N1 features. This implied overall better distinctive properties with respect to the outcome measures at the P2 latency, compared to the features at the N1 latency. This finding could have attributed to the reflected functions in processing at the different latencies. Lee et al. (2009) suggested that N1 (at N160) indicates an early sensory processing phase. In this early phase, laser stimuli (perceived and unperceived) induced equal latency and amplitude [96; 97]. This presumably emphasized the absence of stimulus perception, at the by the N1 reflected deeper brain area [96]. In addition, the P2 (at P400) indicates a later phase of sensory processing where awareness and attention were established. Several brain areas (e.g., the anterior cingulate cortex and operoculo-insular regions) including areas responsible for awareness and attention (e.g., the anterior insular cortex), were activated between 230-400 ms [96; 97; 98]. The attention factor in this late phase, presumably indicated the task-related properties of P2 features. This probably caused for P2 to produce a high SNR compared to N1. The low SNR at N1 expectantly disadvantaged the differentiating performance of N1. Subsequently, the P2 showed higher values for the AUC of the ROC curves compared to N1. However, one should keep in mind that despite these benefits towards P2 features, the task-related properties contributed to the results. Due to the absence of stimulus perception at N1 in the study from Lee et al. (2009), it is hypothesized that N1 is more directly related to the sensory activity compared to P2 because of it's assumed task-related component [96]. Overall, the results of the P2 feature showed better differentiating abilities. However, this did not mean in advance that the quality of the information provided by the P2 features is higher.

Another remarkable result was the reduced latency for P2 in patients suffering from DMnp (i.e., 383 ms) in contrast to the mean of all subject groups (i.e., 432 ms). This result was similar to the results in the study from Mouraux et al. (2010). One of their aims was to test afferent fiber responses elected by IES after induced denervation by capsaicin application [99]. Capsaicin induced activation of the noxious heat and acid sensitive vanilloid receptor 1 (TRPV-1). In addition, hyperalgesia and pain sensations were expected to be influenced by capsaicin [99; 100]. Mouraux et al. stated that capsaicin did not affect the tactile Aβ fibers. They found reduced latencies for 75% of the researched negative and positive peaks in the capsaicin group in contrast to the control group [99]. These reduced values for P in the study from Mouraux et al., are in line with the results of this study. These reduced latencies might be induced by fast conducting Aβ fibers, while Mouraux et al. concluded that these fibers were unaffected [20; 99]. Patients suffering from diabetes are often familiar with damaged peripheral nerves [17]. Therefore, it can be suggested that nerve atrophy might have contributed to the reduced latencies of P1 in diabetic patients. The unaffected fast conducting tactile nerves were probably stimulated. However, this does not explain the increased P2 in the DPNp subject group. Nevertheless, exploration of the GFP did not result in one clear peak for P2. If an earlier peak would have been selected, the latency of this P2 for DPNp would be reduced compared to other research groups. The latter would be in accordance with the prior stated explanation.

Previous clarification for reduced latencies in the DPNp research group might have caused the low number of significant features in the comparison between DMnp versus DPNp. Features were determined by use of the mean latency of all research groups. This caused feature calculation at a shifted latency of the EP in DPNp, resulting in the fact that information of the EP after the actual event is obtained. On the contrary, the other two comparisons including both DM groups (i.e., DPNp versus HC and DMnp versus HC)

consisted of many significant features. This is due to the P2 latencies for the HC being close to the used mean P2 and the P2 latency for the DM group prior to the mean P2. Thus, these large amounts of significant features in both diabetic comparisons with the HCs might have been induced by the shifted position at the EP, before feature calculation, in the diabetic groups.

This study showed AUC values for the ROC curves for three types of EEG features: (1) features functioning as direct coefficients of the LM (e.g., amplitude), (2) features adapted from the LM coefficients (e.g., response) and (3) remaining features independent from the LM (e.g., the different means). There was no specific group that performed better than others. Therefore, in this study, all these groups of features helped differentiating one outcome measure from another. However, one feature at P2, entitled as 'Response' (see Appendix 8.1 for the definition), had the largest AUC in three out of four comparisons. This feature was adapted from consisting LM coefficients. This suggested that the application of the LM did add value to the differentiation of the outcome measures. However, there is no specific reason found yet, explaining the advanced performance of this feature compared to other features.

5.2 Part II - Features in combinations

In this second part of the research it was found that NDT features had an essential position in an RF classification model. The five features with the highest predictor importance estimates were mainly NDT features, followed by P2 features. This was in accordance with previously discussed results regarding better discriminating abilities of individual NDT features. Equal to the first part of the research, it can be implied that the many confounding factors might have induced noise in the EEG features. Therefore, NDT had a more essential role in the RF classification models. In addition to the importance of the NDT features, the features at the P2 latency occupied higher positions compared to N1 in the same predictor importance estimate figures. Also, due to the same reasons as previously described, the P2 proclaimed better predictive abilities compared to N1.

One cannot conclude that only a combination from the upper few NDT features were competent in creating the best possible feature combination. Multiple features were correlated, and therefore two different features might include equal information. This is most accurate for the NDT features. All correlation matrices, for all different comparisons, showed strong correlations within the NDT group. Only about four were correlated within the the P2 group and only a few of the N1 features showed a strong correlation. As previously mentioned, this strong correlation for the NDT features in particular, might have decreased the functionality of feature combinations in the RF classification model.

An unanticipated finding was the height of the accuracy of the combined features, compared to the accuracy of a single feature. In three out of four comparisons, the accuracy of a single feature was higher than a combination of features. This is in contrast to the expectation that combined features would contribute to an increased accuracy. There are three possible explanations for this inconsistency. At first, the optimization of the RF classifier could contribute to the reduced accuracy in the classification model. However, the used algorithm optimized on accuracy. Therefore, this could not have contributed to the low accuracy of the RF classifier. Another possible explanation was the outlier removal. For generation of the model, only clean data was used to prevent overfitting on outliers. The test data was not clean since outliers are present in the clinical environment. However, the outlier removal process may have caused the removal of too much data. If in one part of a particular feature group (NDT, N1 or P2) an outlier was found, all features from that

group were being removed. Together with the small amount of measurement data, this might have caused inconsistencies in the accuracy calculation. It was hypothesized that when more data is available, the outlier removal does not have this large impact on the accuracy calculation and a more accurate approach of the true accuracy might be obtained. A third and the most presumable explanation of this inconsistency was the overfitting disadvantages in machine learning algorithms. It has already been tried to prevent this by application of k-fold cross-validation and the choice of machine learning technique. RF classification is less sensitive to overfitting compared to, for instance, decision trees through bootstrapping [101]. However, aiming to train the data too extensive, and therefore using too many features, results in overfitting. This means that the test data will show bad classification. Decreasing the amount of features and extension of the data set will decrease the possibility of the occurrence of overfitting [82].

The 'outlier argument' from the previous paragraph was supported when the AUC values of the RF classifier were researched. This was caused by the fact that the ROC-curves for this part of the research were created by using OOB data. For the generation of this final RF model clean data was used. However, this caused the OOB data to be clean too. This might contribute to a better performance. In three out of four cases, the ROC curves of the comparisons showed increased values for the AUC. This might suggest a better performance for feature combinations compared to single features.

5.3 Strengths and Limitations

The strength of this study was the extensive analysis of features between different subject groups. In previous research, generalized linear mixed models (GLMMs) and linear mixed models (LMMs) were calculated on group level. In this study, features of individual measurements were calculated by the application of GLMs and LMs. Those features were used for predicting subject classification into a diagnosis. Therefore, this study gave insight into features' performances in differentiation per individual measurement. The classifying results of this study contributed excessively in exploration of important features. As a consequence, additional insight is acquired in terms of the observation of underlying neurophysiological mechanisms in nociceptive processing. Conventional treatments mainly concentrate on symptom control and pain relief. After extensive future research, the latter might contribute to improved diagnostics and treatments for chronic pain based on the underlying neurophysiological mechanisms.

The second part of the research focused on features in combinations after application of a RF classification model. First steps towards feature combinations were established for the observation of nociceptive processing. An advantage of this study includes the benefits that come with various machine learning techniques (e.g., RF classification). Single feature classification is based on linearity. The addition of more features to a classification model provides multi-dimensional and non-linear classification, expectantly resulting in an improved accuracy. As a consequence, this will be beneficial for the observation of nociceptive processing. In this research the first steps were made in biomarker combination quantification using machine learning. However, an improved accuracy had not been accomplished yet. Nevertheless, more research focusing on classification techniques, machine learning settings and overfitting properties, will probably result in an increased accuracy. Therefore, future research in machine learning techniques might be beneficial for the observation of neurophysiological mechanisms in nociceptive processing.

In this study there was a large distribution in demographic data (e.g., age, medication and BMI) between the subject groups. This inclusion bias might have induced or

emphasized differences between groups based on demographic characteristics. Therefore, the feature quality might have been based on more than only the neurophysiological properties. For instance, the mean age of the DPNp subject group was about 11 to 21 years higher compared to the other subject groups. This might have caused inconsistencies in especially the task-related activity from the NDT-EP experiment. Knyazev et al. (2015) correlated age-related changes to reduced attention tasks because of the decreased and random connectivity within the brain networks [102]. These neural changes probably affected the results based on more than only the consequence of a diagnose.

Another important disadvantage in terms of the demographic data is the varying BMIs in the different conditions. This demographic factor also presumably contributed to the outcome measure since Price et al. (2013) concluded decreased sensitivity on subcutaneous excess fat including areas. They applied thermal stimuli on the abdomen and forehead to explore pain thresholds and tolerance and detection thresholds [103]. Thus, it can be suggested that this decreased sensitivity affected the results next to the contribution of diagnose to the outcome of the research.

5.4 Implications

Both the singular and combined exploration of feature performance on individual level has not been conducted in previous studies. In these studies, only statistic differences on group level were researched. However, in order to identify characteristics of the individual subject and why this subject experiences chronic pain in contrast to other subjects, one should study the effects of features on individual level. A deeper understanding of these features can reveal which elements contribute to underlying neurophysiological mechanisms of nociceptive processing. The results of this study demonstrated possible indications of chronic pain. A model was developed, in terms of a singular feature and a combination of features. These models could classify chronic pain, based on specific values of features. This is the first step towards a larger model, able to observe pain and the pain experience, based on underlying neurophysiological mechanisms of nociceptive processing. When future research enables this observation of pain and pain experience, one can individually adapt therapies for chronic pain patients. However, the grounds of those therapies, should be part of a future study.

Chapter 6

Conclusion

This study identified important biomarkers from the NDT-EP experiment for improved observation of underlying neurophysiological mechanisms on individual level in nociceptive processing. These biomarkers were identified for subjects suffering from chronic pain and pain-free subjects. There was not one specific feature quantified as the ideal feature. However, NDT features were, in contrast to the EEG features, specified as the features with most differentiating abilities in all comparisons between the diagnoses. When considering the EEG features, P2 features showed an overall better differentiating performance compared to the N1 features. Also features, derived from double pulse stimuli showed overall better classifying abilities.

Besides, it could be concluded that this study identified features from the NDT-EP experiment, relevant in feature combinations for nociceptive processing on individual level. Similar to the first part of the research, did this part show a more prominent role in classification by feature combinations for the NDT features compared EEG features. It was also found that in terms of the EEG features, P2 features were more relevant compared to N1 features.

However, the accuracy did not improve when the RF classification model was aimed to differentiate between two outcome measures compared to single features. Nevertheless, several measures (e.g., feature correlation or prevention of overfitting) can be explored in future research for increasing the accuracy of this classification model.

Chapter 7

Future Perspectives

The first part of this chapter, Section 7.1, describes the recommendations in terms of the following research question: '*How can the NDT-EP experiment be adapted into a more clinical applicable experiment to improve diagnostic performance and reduce the experiment duration?*'. The second part of this chapter describes general recommendations for future studies (Section 7.2).

7.1 Recommendations for the NDT-EP experiment

This study successfully identified the importance of specific features of the NDT-EP experiment. This knowledge can be used for decreasing the experiment time of the NDT-EP experiment. The experiment time is currently experienced as too long by subjects. Moreover, reducing time is beneficial for clinical application. It was found that the NDT features were most significant when analyzing the performance of single features (the first part of this study). Also in the second part of the study, the most important predictor estimates consisted of NDT features. This might suggest that, measuring the NDT would be sufficient in differentiating between two diagnoses. Therefore, EEG could be excluded from the measurement protocol. This results in decreased preparation time due to the extraction of the EEG-cap application from the protocol. However, one should keep in mind that EEG is a more objective measure than the NDT data being determined by the subject's response. It is expected that patients, suffering from chronic pain, are likely to want to convince the physician of their pain experience. Therefore, it can be questioned whether the diagnostic performance is going to be increased by this measure, and it is recommended that first other possible performance increasing methods will be applied.

Therefore, there might be better ways in establishing more benefit in decreasing the duration time of the experiment. One important result from this study is the number of significant features after application of a double pulse stimulus compared to a single pulse stimulus setting. The research protocol currently consists of three settings (i.e., single pulse, double pulse with 10 ms IPI and double pulse with 40 ms IPI), each consisting of 150 trials. When one or two settings can be extracted, the experiment duration will be shortened by 150 or 300 trials per experiment, dependent on the amount of removed settings. However, one cannot simply take out one or two settings. Firstly, the informative capacity of all settings need to be addressed. This study already discussed the promising differentiating capacities of double pulse stimulus settings compared to single pulse settings. This might imply the possible removal of single pulse stimuli. Further research should be undertaken in order to explore the possibility of removing the least informative stimulus settings. One can start with an extensive assessment of correlation matrices. In

this research, already highly correlated features were shown between similar features with different stimulus settings.

Future research concerning integrative feature analysis in terms of the NDT-EP experiment should focus on optimization of feature combinations in predicting diagnose classification. This study investigated the single features and combinations on individual level and their classifying performances. It was expected that a combination of features would increase the accuracy. However, single features produced occasionally a higher accuracy than a combination of the RF classification model. This was probably induced by overfitting, many removed outliers and correlation between features. This study already applied several measures (e.g., k-fold cross validation and the choice of a certain classification technique) to reduce the effects of overfitting and the small number of outliers. However, additional measures can be applied in order to increase the accuracy of the RF classification model and therefore improve the diagnostic performance of the NDT-EP experiment. The most obvious measure is increasing the subject population. Another important recommendation is decreasing the amount of features in the RF classification model. Currently, 35 features (i.e., 13 NDT features, 11 N1 features, 11 P2 features) are added to the classification model. This large amount of features causes the RF classification model to be prone for overfitting. Reducing this amount of features is expected to improve the classification performance. Unimportant features, as reported by the predictor importance estimates, should be extracted from the model. Afterwards the performance of the model can be determined by calculation of the accuracy. Repetition of this process might be necessary in order to reach the most optimal performance. However, one should keep in mind that the correlation also contributes to the accuracy of the RF classification model. When only strong negative or positive correlated features are inserted into the model, the accuracy might be small since those features are likely to give similar information. Therefore, during feature extraction one should consider the correlation between the features next to the predictor importance estimates.

Another suggestion to improve the diagnostic performance of the NDT-EP experiment is by using different latencies for different subject groups in the analysis part of the research. The GFP showed different latencies for the N1 and P2 between the different diagnoses. In this research, the weighted mean was used. In line with previous research it was hypothesized that the DPNp subject group showed earlier latencies due to nerve degeneration and activation of the fast conducting tactile nerve fibers [99]. The application of the mean latency of P2 and N1 of all subject groups might result in, for instance, a too late latency for the DPNp subject group and too short for the HCs. Therefore, one could consider to determine the ideal latencies for each subject group apart and determine the EEG features by using a seperate determined latency per subject group.

Finally, in order to improve diagnostic performances of the NDT-EP experiment, one can consider the possibility of assigning different feature combinations to each comparison. Each comparison shows slightly differing orders in terms of feature classification performance. Therefore a specific feature set for a specific comparison of diagnoses can be considered.

7.2 General Recommendations

The first highly recommended factor that can decrease a current bias is concerning the inclusion of subjects in the different subject groups. The subject characteristics (see Table 4.1) show exceptional large differences in included means or numbers of, for example, age in HCs (e.g., younger hospital staff members) in contrast to DPNp subjects (older

because of late onset of DM). The latter also applies to the other subject groups (i.e., DMnp and FBSS). In addition, subjects suffering from DM (mainly type 2) are familiar with overweight, causing a high BMI in DM subject groups [17]. All these differences might influence neurophysiological signal processing resulting in an inaccurate feature performance. Therefore, future research will expectantly benefit from a smaller inclusion bias. This can be established by, for example, inclusion of matching subjects based on characteristics (i.e., sex, age, BMI).

Continuing on previous suggestion, one should keep in mind the medication used by the subjects. In this research, the medication intake has not been taken into consideration. However, especially pain relieving medication is expected to influence pain perception and therefore the results. Therefore, future research should investigate the possible influences of medication on the pain system and try to visualize these specific effects on the results of the NDT-EP experiment. Hereby, one can state whether results originate from influences of medication or they belong to a certain diagnose.

When more data from the disordered patient groups is available, one can consider to insert the data from the HCs measured at the University of Twente (van den Berg, 2018) [64]. In this study, however, only data from the St. Antonius Hospital has been used. Berfelo et al. (2019) researched similarities between the HC subject groups of the University of Twente and the St. Antonius Hospital. They were seemingly similar [46]. However, since the current HCs was already about twice as large compared to the disordered patients, this data was not taken into consideration yet. The influence of the HCs on, for example, the GFP for the latency determination would be too large compared to the other subject groups. Additionally, it seemed better to use data from one centre to decrease the chance on confounding as much as possible. It is not expected that the results would differ as much when data from van den Berg (2018) was used, as Berfelo (2019) showed agreement between the research groups.

Bibliography

[1] Antonia Barke Qasim Aziz Michael I. Bennett Rafael Benoliel Milton Cohen Stefan Evers Nanna B. Finnerup Michael B. First Maria Adele Giamberardino Stein Kaasa Eva Kosek Patricia Lavand'homme Michael Nicholas Shuu-Jiun Wang Rolf-Detlef Treede, Winfried Rief. A classification of chronic pain for ICD-11. *Pain*, 156(6):1003–1007, 2015.

[2] Vittorio Ventafridda Rob Cohen Derek Gallacher Harald Breivik, Beverly Collett. Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment. *European Journal of Pain*, 10:287–333, 2006.

[3] Kim J. Reid, Julie Harker, Malgorzata M. Bala, Carla Truyers, Eliane Kellen, Geertruida Elsiena Bekkering, and Jos Kleijnen. Epidemiology of chronic non-cancer pain in Europe: Narrative review of prevalence, pain treatments and pain impact. *Current Medical Research and Opinion*, 27(2):449–462, 2011.

[4] Nicholas Henschke, Steven J. Kamper, and Chris G. Maher. The epidemiology and economic consequences of pain. *Mayo Clinic Proceedings*, 90(1):139–147, 2015.

[5] Nicholas Hylands-White, Rui V. Duarte, and Jon H. Raphael. An overview of treatment approaches for chronic pain management. *Rheumatology International*, 37(1):29–42, 2017.

[6] Rebecca Dale and Brett Stacey. Multimodal Treatment of Chronic Pain. *Medical Clinics of North America*, 100(1):55–64, 2016.

[7] F. M. Blyth, L. M. March, K. N. Nicholas, and M. J. Cousins. Chronic pain, work performance and litigation. 103:41–47, 2003.

[8] Marianne Kjettrup Jensen, Per Sjøgren, Ola Ekholm, Niels Kristian Rasmussen, and Jørgen Eriksen. Identifying a long-term/chronic, non-cancer pain population using a one-dimensional verbal pain rating scale: An epidemiological study. *European Journal of Pain*, 8(2):145–152, 2004.

[9] M.P. Jensen, P. Karoly, and S. Braver. The measurement of clinical pain intensity: a comparison of six methods. *Pain*, 27:117–126, 1986.

[10] Randy Neblett, Howard Cohen, Yunhee Choi, Meredith M. Hartzell, Mark Williams, Tom G. Mayer, and Robert J. Gatchel. The central sensitization inventory (CSI): Establishing clinically significant values for identifying central sensitivity syndromes in an outpatient chronic pain sample. *Journal of Pain*, 14(5):438–445, 2013.

[11] Keizer D Köke AJA Schiere S Van Bommel M Van Coevorden RS Van de Vusse A Van den Donk M Van Es A Veldhoven CMM Verduijn MM. De Jong L, Janssen PGH. Pijn.

[12] Robert J. Doll, Jan R. Buitenweg, Hil G.E. Meijer, and Peter H. Veltink. Tracking of nociceptive thresholds using adaptive psychophysical methods. *Behavior Research Methods*, 46(1):55–66, 2014.

[13] R. Doll. *Psychophysical methods Nociceptive processing.* PhD thesis, University of Twente, Enschede, 2016.

[14] T Berfelo, B Van Den Berg, I P Krabbenbos, and J R Buitenweg. Observing electrical brain responses around the nociceptive detection threshold. 2019.

[15] Alexis Waguespack, Jerome Schofferman, Paul Slosar, and James Reynolds. Etiology of long-term failures of lumbar spine surgery. *Pain Medicine*, 3(1):18–22, 2002.

[16] Silvano Gefferie. Observing small fiber dysfunction using nociceptive stimulus detection and evoked potentials. 2020.

[17] P Kumar and M L Clark. *Kumar and Clark's Clinical Medicine.* Elsevier Health Sciences, 2012.

[18] Solomon Tesfaye, Andrew J.M. Boulton, and Anthony H. Dickenson. Mechanisms and management of diabetic painful distal symmetrical polyneuropathy. *Diabetes Care*, 36(9):2456–2465, 2013.

[19] IASP. IASP Announces Revised Definition of Pain, 2020.

[20] Charlotte E. Steeds. The anatomy and physiology of pain. *Surgery (United Kingdom)*, 34(2):55–59, 2016.

[21] IASP. IASP Terminology, 2017.

[22] Anne Priscille Trouvin and Serge Perrot. New concepts of pain. *Best Practice and Research: Clinical Rheumatology*, 33(3), 2019.

[23] Eva Kosek, Milton Cohen, Ralf Baron, Gerald F. Gebhart, Juan Antonio Mico, Andrew S.C. Rice, Winfried Rief, and A. Kathleen Sluka. Do we need a third mechanistic descriptor for chronic pain states? *Pain*, 157(7):1382–1386, 2016.

[24] Rainer Freynhagen, Harold Arevalo Parada, Carlos Alberto Calderon-Ospina, Juythel Chen, Dessy Rakhmawati Emril, Freddy J. Fernández-Villacorta, Hector Franco, Kok Yuen Ho, Argelia Lara-Solares, Carina Ching Fan Li, Alberto Mimenza Alvarado, Sasikaan Nimmaanrat, Maria Dolma Santos, and Daniel Ciampi de Andrade. Current understanding of the mixed pain concept: a brief narrative review. *Current Medical Research and Opinion*, 35(6):1011–1018, 2019.

[25] Deborah L. Ellison. Physiology of Pain. *Critical Care Nursing Clinics of North America*, 29(4):397–406, 2017.

[26] E Mtui, G Gruener, and M J T FitzGerald. Clinical Neuroanatomy and Neuroscience. chapter 34. Elsevier Health Sciences UK, sixth edition, 2011.

[27] Matthew Birnbaum. *Handbook of Pain Management*, volume 35. 2010.

[28] E.A. Welchew. Clinical Pain Management: Acute Pain. *British Journal of Anaesthesia*, 90(3):405–406, 2003.

[29] Alexander M. Dydyk and Amy Givler. *Central Pain Syndrome*. StatPearls Publishing, Treasure Island (FL), 2020.

[30] Karin N. Westlund and William D. Willis. Pain System. In *The Human Nervous System*, pages 1144–1186. Elsevier, third edit edition, 2012.

[31] Gregory D Cramer and Susan A Darby. *Spine, Spinal Cord and ANS*. Elsevier Inc., St. Louis, third edition, 2014.

[32] Allan I. Basbaum, Diana M. Bautista, Grégory Scherrer, and David Julius. Cellular and Molecular Mechanisms of Pain. *Cell*, 139(2):267–284, 2009.

[33] Ronald Melzack and Patrick D. Wall. Pain Mechanism: A New Theory. *Science*, 150(3699):971–979, 1965.

[34] Harold Merskey and Nikolai Bogduk. *Classification of Chronic Pain*, volume 23. Second edition, 1994.

[35] Clifford J. Woolf. Evidence for a central component of post-injury pain hypersensitivity, 1983.

[36] Ru Rong Ji, Tatsuro Kohno, Kimberly A. Moore, and Clifford J. Woolf. Central sensitization and LTP: Do pain and memory share similar mechanisms? *Trends in Neurosciences*, 26(12):696–705, 2003.

[37] E. A. Joosten. Pijn, nociceptie en inflammatie: inzichten in de mechanismen en mogelijke therapeutische aanpak. *Nederlands tijdschrift voor tandheelkunde*, 123(10):458–462, 2016.

[38] Clifford J. Woolf. Central sensitization: Implications for the diagnosis and treatment of pain. *Pain*, 152(SUPPL.3):S2–S15, 2011.

[39] Alban Latremoliere and Clifford J. Woolf. Central Sensitization: A Generator of Pain Hypersensitivity by Central Neural Plasticity. *Journal of Pain*, 10(9):895–926, 2009.

[40] Theo Vos. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, 386(9995):743–800, 2015.

[41] Carsten Oliver Schmidt, Heiner Raspe, Michael Pfingsten, Monika Hasenbring, Heinz Dieter Basler, Wolfgang Eich, and Thomas Kohlmann. Back pain in the German adult population: Prevalence, severity, and sociodemographic correlates in a multiregional survey. *Spine*, 32(18):2005–2011, 2007.

[42] Rodrigo Dalke Meucci, Anaclaudia Gastal Fassa, and Neice Muller Xavier Faria. Prevalence of chronic low back pain: Systematic review. *Revista de Saude Publica*, 49:1–10, 2015.

[43] Chin-wern Chan and Philip Peng. Failed back surgery syndrome. *Pain Medicine*, 12:577–606, 2011.

[44] Kliment Gatzinsky, Sam Eldabe, Jean Philippe Deneuville, Wim Duyvendak, Nicolas Naiditch, Jean Pierre Van Buyten, Philippe Rigoard, and Baogan Peng. Optimizing the Management and Outcomes of Failed Back Surgery Syndrome: A Proposal of a Standardized Multidisciplinary Team Care Pathway. *Pain Research and Management*, 2019(c):6–10, 2019.

[45] S. Blond, P. Mertens, R. David, M. Roulaud, and P. Rigoard. From "mechanical" to "neuropathic" back pain concept in FBSS patients. A systematic review based on factors leading to the chronification of pain (part C). *Neurochirurgie*, 61(S1):S45–S56, 2015.

[46] T. Berfelo. Electrical brain responses during processing of nociceptive stimuli around the detection threshold: an explorative study in pain-free subjects and failed back surgery syndrome pat ients. Technical Report February, University of Twente, 2019.

[47] Sarah Wild, Gojka Roglic, Anders Green, Richard Sicree, and Hilary King. Global Prevalence of Diabetes: Estimates for the year 2000 and projections for 2030. *Diabetes Care*, 27(5):1047–1053, 5 2004.

[48] American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*, 31(SUPPL. 1), 2008.

[49] Anne K Schreiber. Diabetic neuropathic pain: Physiopathology and treatment. *World Journal of Diabetes*, 6(3):432, 2015.

[50] Arthur S. Courtin, Sofia Maldonado Slootjes, Gilles Caty, Michel P. Hermans, Léon Plaghki, and André Mouraux. Assessing thermal sensitivity using transient heat and cold stimuli combined with a Bayesian adaptive method in a clinical setting: A proof of concept study. *European Journal of Pain (United Kingdom)*, (January):1–10, 2020.

[51] Ozgur Karcioglu, Hakan Topacoglu, Ozgur Dikme, and Ozlem Dikme. A systematic review of the pain scales in adults: Which to use? *American Journal of Emergency Medicine*, 36(4):707–714, 2018.

[52] Gillian A. Hawker, Samra Mian, Tetyana Kendzerska, and Melissa 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 and Research*, 63(SUPPL. 11):240–252, 2011.

[53] Carol S. Burckhardt and Kim D. Jones. Adult measures of pain: The McGill Pain Questionnaire (MPQ), Rheumatoid Arthritis Pain Scale (RAPS), Short-Form McGill Pain Questionnaire (SF-MPQ), Verbal Descriptive Scale (VDS), Visual Analog Scale (VAS), and West Haven-Yale Multidisciplinary Pain Inven. *Arthritis & Rheumatism*, 49(S5):S96–S104, 2003.

[54] Mark C. Bicket and Jianren Mao. Chronic Pain in Older Adults. *Anesthesiology Clinics*, 33(3):577–590, 2015.

[55] Catherine A Marco and Alan P Marco. Assessment of pain. pages 10–18, 2018.

[56] Descheemaeker F Cagnie B. van Wilgen P, Meeus M. Meetininstrumenten in de zorg - Central Sensitization Inventory, 2013.

[57] Peripheral Neuropathy Association. Quantitative sensory testing. *Quantative sensory testing: A consensus report from the Peripheral Neuropathy Assocication*, 43:1050–1052, 1993.

[58] Peter Siao and Didier P. Cros. Quantitative sensory testing. *Physical Medicine and Rehabilitation Clinics of North America*, 14(2):261–286, 2003.

[59] Rogelio A. Coronado, Corey B. Simon, Carolina Valencia, and Steven Z. George. Experimental Pain Responses Support Peripheral and Central Sensitization in Patients with Unilateral Shoulder Pain. 30(2):1–20, 2015.

[60] Brian J. Eckenrode, David M. Kietrys, and Scott K. Stackhouse. Pain Sensitivity in Chronic Achilles Tendinopathy. *International Journal of Sports Physical Therapy*, 14(6):945–956, 2019.

[61] Peter Steenbergen, Jan R. Buitenweg, Jörg Trojan, Esther M. van der Heide, Teun van den Heuvel, Herta Flor, and Peter H. Veltink. A system for inducing concurrent tactile and nociceptive sensations at the same site using electrocutaneous stimulation. *Behavior Research Methods*, 44(4):924–933, 2012.

[62] P. Steenbergen, J. R. Buitenweg, E. M. Van Der Heide, and P. H. Veltink. Characterization of a bimodal electrocutaneous stimulation device. *IFMBE Proceedings*, 22:230–234, 2008.

[63] K. Inui, T. D. Tran, Y. Qiu, X. Wang, M. Hoshiyama, and R. Kakigi. A comparative magnetoencephalographic study of cortical activations evoked by noxious and innocuous somatosensory stimulations. *Neuroscience*, 120(1):235–248, 2003.

[64] Boudewijn van den Berg. Stimulus related evoked potentials around the nociceptive detection threshold. 2018.

[65] Boudewijn van den Berg and Jan R. Buitenweg. Observation of Nociceptive Processing: Effect of Intra-Epidermal Electric Stimulus Properties on Detection Probability and Evoked Potentials. *Brain Topography*, 34(2):139–153, 2021.

[66] Hopp JL et al. Britton JW, Frey LC. Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants. American Epilepsy Society, Chicago, 2016.

[67] Shuli Huang and Huabo Xiao. The application of EEG related. *Procedia Environmental Sciences*, 10(PART B):1338–1342, 2011.

[68] Hai Chen and Mohamad Z. Koubeissi. Electroencephalography in Epilepsy Evaluation. *CONTINUUM Lifelong Learning in Neurology*, 25(2):431–453, 2019.

[69] Michel J.A.M. van Putten. *Essentials of Neurophysiology: Basic Concepts and Clinical Applications for Scientists and Engineers*. Springer Berlin Heidelberg, Enschede, 2010.

[70] Rebeccah Slater, Alan Worley, Lorenzo Fabrizi, Siân Roberts, Judith Meek, Stewart Boyd, and Maria Fitzgerald. Evoked potentials generated by noxious stimulation in the human infant brain. *European Journal of Pain*, 14(3):321–326, 2010.

[71] M.J.T. Fitzgerald, G. Gruener, and E. Mtui. *Clinical neuroanatomy and neuroscience*. Elsevier Saunders, six edition, 2012.

[72] Marc Schooneman. Measurement of evoked potentials during multiple threshold tracking of nociceptive electrocutaneous stimuli. 2015.

[73] Issam El Naqa and Martin J Murphy. Machine Learning in Radiation Oncology. *Machine Learning in Radiation Oncology*, pages 3–11, 2015.

[74] Yongli Zhang and Yuhong Yang. Cross-validation for selecting a model selection procedure. *Journal of Econometrics*, 187(1):95–112, 2015.

[75] Jaime Lynn Speiser, Michael E. Miller, Janet Tooze, and Edward Ip. A comparison of random forest variable selection methods for classification prediction modeling. *Expert Systems with Applications*, 134:93–101, 2019.

[76] Kohbalan Moorthy and Mohd Saberi Mohamad. Random forest for gene selection and microarray data classification. *Bioinformation*, 7(3):142–146, 2011.

[77] Alessia Sarica, Antonio Cerasa, and Aldo Quattrone. Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: A systematic review. *Frontiers in Aging Neuroscience*, 9(OCT):1–12, 2017.

[78] Reneman M. Ansuategui Echeita J. Kruitbosch H. Otten B. & Lamoth C. Zheng, X. Classification of patients with chronic low back pain and high or low central sensitization by gait outcomes using machine learning methods. 2021.

[79] F Johannes. Encyclopedia of Machine Learning and Data Mining. *Encyclopedia of Machine Learning and Data Mining*, pages 1–5, 2016.

[80] Leo Breiman. Random Forests. *Machine Learning*, 45:5–32, 2001.

[81] Tae Hwy Lee, Aman Ullah, and Ran Wang. Bootstrap Aggregating and Random Forest. *Advanced Studies in Theoretical and Applied Econometrics*, 52:389–429, 2020.

[82] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Second edition, 2001.

[83] Maheshakya Wijewardena. Implementation of Bootstrap Aggregating Algorithm for Scikit-learn Library. (November 2013), 2016.

[84] Sanjay Yadav and Sanyam Shukla. Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification. *Proceedings - 6th International Advanced Computing Conference, IACC 2016*, (Cv):78–83, 2016.

[85] Mariia Fedotenkova. Extraction of multivariate components in brain signals obtained during general anesthesia. (December 2016), 2017.

[86] NeuroCIMT Project 2: NOCICEPT, 2020.

[87] T Berfelo. Research protocol: Electrical brain responses during processing of nociceptive stimuli around detection threshold: an explorative study in pain patients - February 2020. (February):1–66, 2020.

[88] MATLAB. *version 9.2.0 (R2017a)*. The MathWorks Inc., Natick, Massachusetts, 2017.

[89] MATLAB. *version 9.9.0 (R2020b)*. The MathWorks Inc., Natick, Massachusetts, 2020.

[90] R. Oostenveld, E. Maris P. Fries, and J.-M. Schoffelen. Fieldtrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011, 2011.

[91] Victor Martínez-Cagigal. ROC Curve, 2021.

[92] J.A. Hanley; and B.J. McNeil. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Diagnostic Radiology*, 143(1):29–36, 1982.

[93] Ilker Unal. Defining an optimal cut-point value in ROC analysis: An alternative approach. *Computational and Mathematical Methods in Medicine*, 2017, 2017.

[94] Niall W. Duncan and Georg Northoff. Overview of potential procedural and participant-related confounds for neuroimaging of the resting state. *Journal of Psychiatry and Neuroscience*, 38(2):84–96, 2013.

[95] André Mouraux, Emilie Marot, and Valéry Legrain. Short trains of intra-epidermal electrical stimulation to elicit reliable behavioral and electrophysiological responses to the selective activation of nociceptors in humans. *Neuroscience Letters*, 561:69–73, 2014.

[96] Michael C. Lee, André Mouraux, and Gian Domenico Iannetti. Characterizing the cortical activity through which pain emerges from nociception. *Journal of Neuroscience*, 29(24):7909–7916, 2009.

[97] Valéry Legrain, Jean Michel Guérit, Raymond Bruyer, and Léon Plaghki. Attentional modulation of the nociceptive processing into the human brain: Selective spatial attention, probability of stimulus occurrence, and target detection effects on laser evoked potentials. *Pain*, 99(1-2):21–39, 2002.

[98] L. Garcia-Larrea, M. Frot, and M. Valeriani. Brain generators of laser-evoked potentials: From dipoles to functional significance. *Neurophysiologie Clinique*, 33(6):279–292, 2003.

[99] A. Mouraux, G. D. Iannetti, and L. Plaghki. Low intensity intra-epidermal electrical stimulation can activate Aδ-nociceptors selectively. *Pain*, 150(1):199–207, 2010.

[100] Michael J. Caterina, Tobias A. Rosen, Makoto Tominaga, Anthony J. Brake, and David Julius. A capsaicin-receptor homologue with a high threshold for noxious heat. *Nature*, 398(6726):436–441, 1999.

[101] Jehad Ali, Rehanullah Khan, Nasir Ahmad, and Imran Maqsood. Random Forests and Decision Trees. *International Journal of Computer Science Issues*, 9(5):272–278, 2012.

[102] Gennady G. Knyazev, Nina V. Volf, and Ludmila V. Belousova. Age-related differences in electroencephalogram connectivity and network topology. *Neurobiology of Aging*, 36(5):1849–1859, 2015.

[103] Rebecca C. Price, J. F. Asenjo, N. V. Christou, S. B. Backman, and P. Schweinhardt. The role of excess subcutaneous fat in pain and sensory sensitivity in obesity. *European Journal of Pain (United Kingdom)*, 17(9):1316–1326, 2013.

Chapter 8

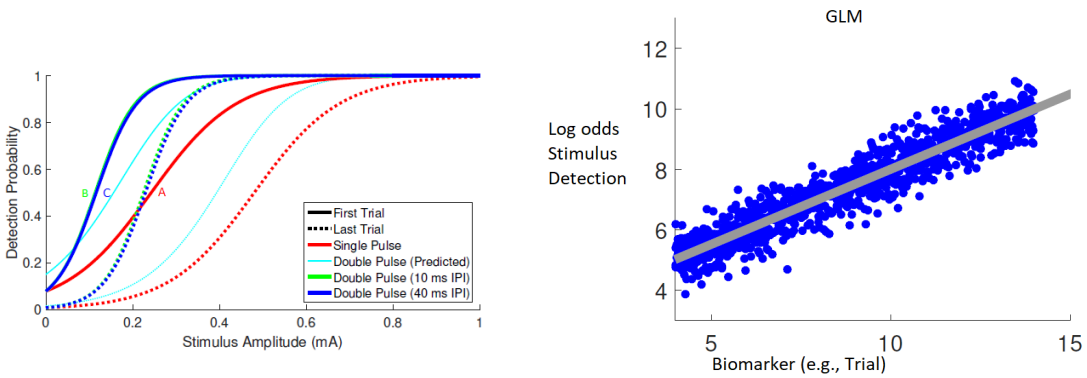
Appendix

8.1 Feature list

Underneath, definitions from the different NDT and EEG features are summarized. For both groups, the features are categorized in, (1) a direct feature from the GLM or LM, (2) a feature emerged from the GLM or LM coefficients and (3) the remaining features not linked to the GLM or LM.

8.1.1 NDT features

Figure 8.1 shows a psychometric curve (Figure 8.1a) and GLM scatter cloud (Figure 8.1b), respectively. The first figure shows what specific NDT features represent and the latter shows the determination of GLM coefficients. The grey line is created by the GLM through the data points (blue dots). The slope of this grey line is the value for the coefficient of the GLM.



(a) Psychometric curve with a certain trial presented by the full or the dotted line, the color represents the setting and the definition of each character (A, B and C) is explained below. Adapted from van den Berg et al. (2018) [64].

(b) GLM for a certain feature. The blue dots represent the data points. The grey line is the created GLM prediction. The slope of the grey line is the value for the feature coefficient. Adapted from van den Berg et al. (2018) [64].

Figure 8.1: NDT feature assistance figures.

GLM coefficients

All GLM coefficients are being calculated by determination of the slope of the GLM prediction in Figure 8.1b.

Amplitude first pulse: This amplitude in the regression model, represents the influence of the first pulse on the outcome measure (detection chance).
Amplitude second pulse (10 ms IPI): This amplitude in the regression model, represents the influence of the second pulse (10 ms after the first pulse) on the outcome measure (detection chance).
Amplitude second pulse (40 ms IPI): This amplitude in the regression model, represents the influence of the second pulse (40 ms after the first pulse) on the outcome measure (detection chance).
Intercept: Baseline activity.
Trial: The trial number. The NDT changes over time. Trial is the influence of amount of stimuli on the detection threshold representing the habituation. This feature is represented in Figure 8.1a as a dotted (last trial) or full line (first), respectively.

Adapted from the GLM coefficients

Slope single pulse: The slope of the psychometric curve on it's steepest point. This feature originates directly from from the regression coefficients. Thus, the slope from the single pulse is equal to the amplitude's coefficient from the first pulse. This feature is marked with 'A' in Figure 8.1a.
Slope double pulse (10 ms IPI): The slope of the psychometric curve after application of the second pulse (10 ms after the first pulse) on it's steepest point. This feature originates from the regression coefficients. Thus, the slope from the double pulse (10 ms IPI) is equal to the amplitude's coefficient from the first pulse and the coefficient of the second pulse after 10 ms IPI. This feature is marked with 'B' in Figure 8.1a.
Slope double pulse (40 ms IPI): The slope of the psychometric curve after application of the second pulse (40 ms after the first pulse) on it's steepest point. This feature originates from the regression coefficients. Thus, the slope from the double pulse (40 ms IPI) is equal to the amplitude's coefficient from the first pulse and the coefficient of the second pulse after 40 ms IPI. This feature is marked with 'C' in Figure 8.1a.
Threshold single pulse: The detection threshold for the single pulse. This feature is 'A' (at 0.5 detection probability) in Figure 8.1a.
Threshold double pulse (10 ms IPI): The detection threshold for a double pulse with 10 ms IPI. This feature is 'B' (at 0.5 detection probability) in Figure 8.1a.
Threshold double pulse (40 ms IPI): The detection threshold fora double pulse with 40 ms IPI This feature is 'C' (at 0.5 detection probability) in Figure 8.1a.

Remaining features

Detection rate: Detected percentage of all stimuli.
Mean response time: The mean time between the stimulus and the subject's response.

8.1.2 EEG features

Figure 8.2 is the explaining Figure for the feature coefficients. The blue dots in the figure represent the data points. The grey line is the created LM prediction. The slope of the grey line is the value for the feature coefficient.

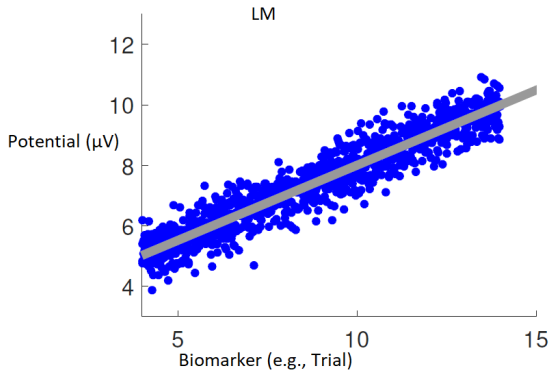


Figure 8.2: LM for a certain feature. The blue dots represent the data points. The grey line is the created LM prediction. The slope of the grey line is the value for the feature coefficient. Adapted from van den Berg et al. (2018) [64].

LM coefficients

All LM coefficients are being calculated by determination of the slope of the LM prediction in Figure 8.2.

Amplitude first pulse: This amplitude in the regression model, represents the influence of the first pulse on the outcome measure. This amplitude of the first pulse is the contribution from this amplitude to the neurophysiological activity.
Amplitude second pulse (10 ms IPI): This amplitude in the regression model, represents the influence of the second pulse after 10 ms IPI on the outcome measure. This amplitude is the contribution from this second pulse amplitude to the neurophysiological activity.
Amplitude second pulse (40 ms IPI): This amplitude in the regression model, represents the influence of the second pulse after 40 ms IPI on the outcome measure. This amplitude is the contribution from this second pulse amplitude to the neurophysiological activity.
Intercept: Mean baseline of the EEG.
Trial number: The brain activity change in μV per trial.

Adapted from the LM coefficients

Response: The height of the detection.
Response * Trial: Also entitled 'Response 8 trial', representing the interaction between response and trial. It is the task related activity, thus the activity considering the stimulus detection. It can be interpreted as the habituation of the task related component.

Remaining features

Double pulse mean (10 ms IPI): Is the mean EEG considering a double pulse with 10 ms IPI.
Double pulse mean (40 ms IPI): Is the mean EEG considering a double pulse with 40 ms IPI.
Mean: The mean value of the EEG.
Single pulse mean: Is the mean EEG considering the single pulse.

8.2 Butterfly plots

Underneath, five butterfly plots are presented of the four explored diagnoses. By using the GFP, presented in these butterfly plots the ideal values were determined for the researched latencies for the negative and positive peaks. Figure 8.3, shows the data from subjects suffering from DM without pain. Figure 8.4, represents the butterfly plot of the subjects suffering from DPN. The third butterfly plot, Figure 8.5, shows the plot of the patients suffering from FBSS. And the final two butterfly plots (Figure 8.6 and 8.7) are the representations of the butterfly data of the pain-free subjects.

Previous determined values for the N1 and P2 were used for calculation of the weighted means of the latencies for these peaks. Therefore, the N1 was set at 167 ms and the P2 was set at 435 ms for the EP analysis.

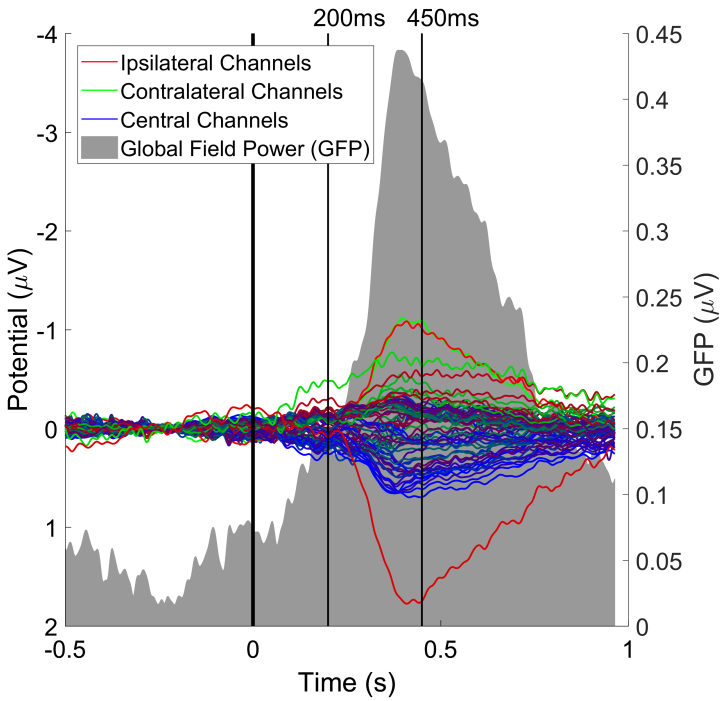


Figure 8.3: This butterfly plot shows the data from the subjects suffering from DM without pain (measurements = 40). The N1 was set at 168 ms and the P2 was set at 383 ms.

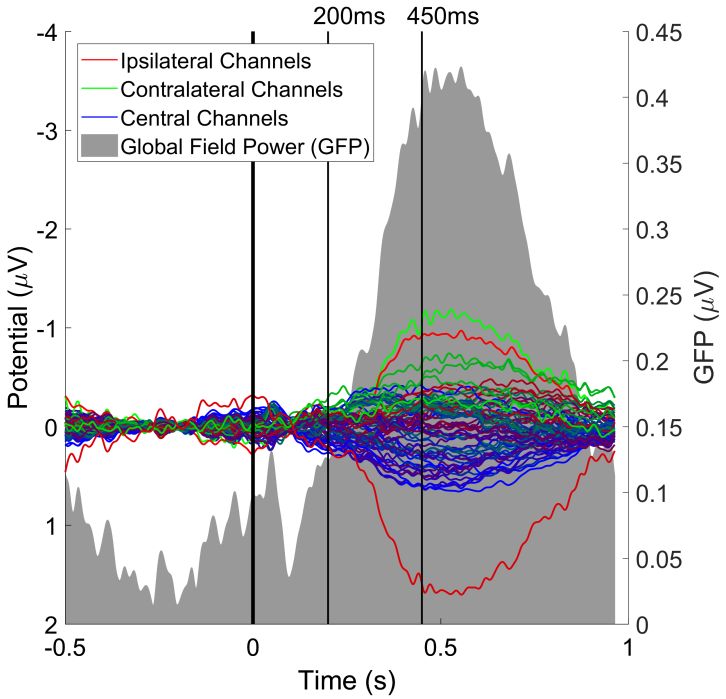


Figure 8.4: The butterfly plot of the subjects suffering from DPN (measurements = 26). The peak for N1 was not recognizable and the peak for P2 was set at 486 ms.

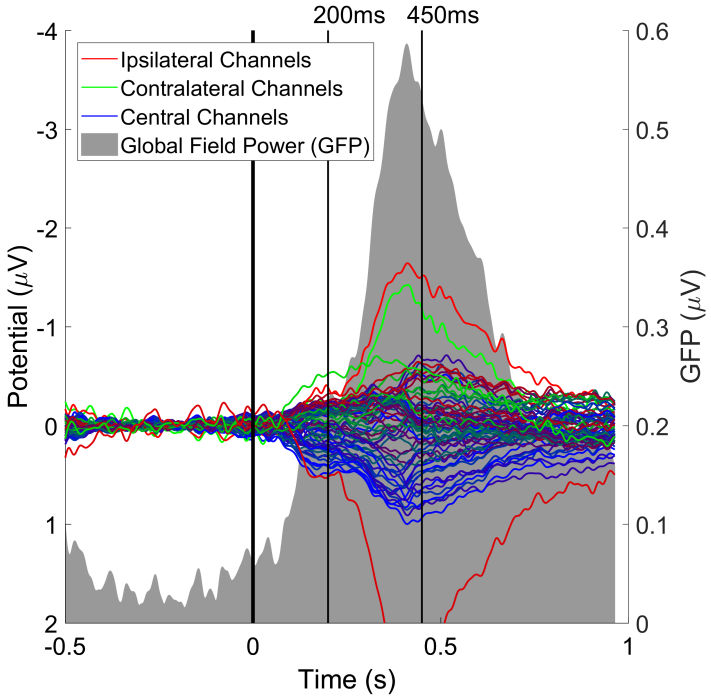


Figure 8.5: This butterfly plot is representing subjects suffering from FBSS (measurments = 32). The N1 was chosen at 174 ms and the P2 at 407 ms.

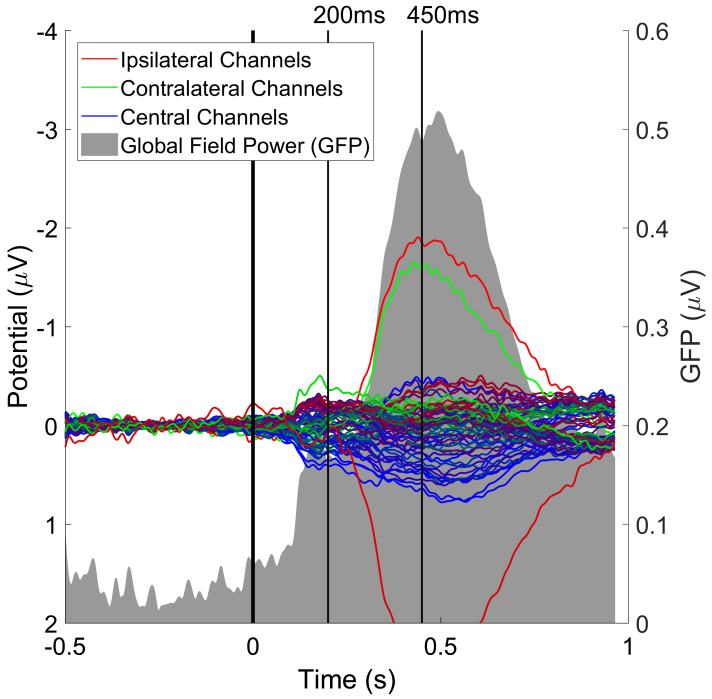


Figure 8.6: This butterfly plot shows the first half of measurements of pain-free subjects (measurements = 34). The N1 was set at 176 ms and the P2 at 494 ms.

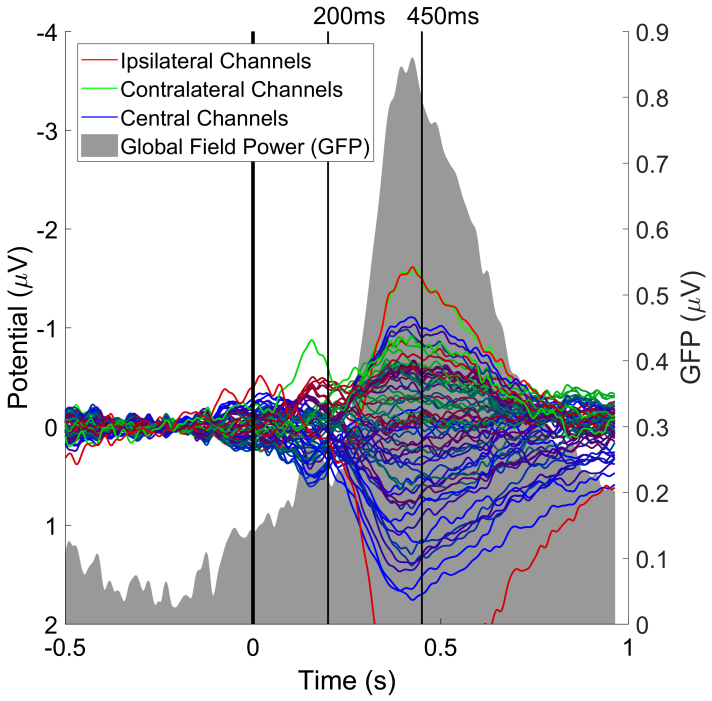


Figure 8.7: This butterfly plot represents the data from the second half of the pain-free subject measurements (measurements = 40). Here the N1 was set at 154 ms and the P2 was chosen at 424 ms.

8.3 Accuracy single features

Table 8.1 shows for all features of each comparison, the calculated accuracy.

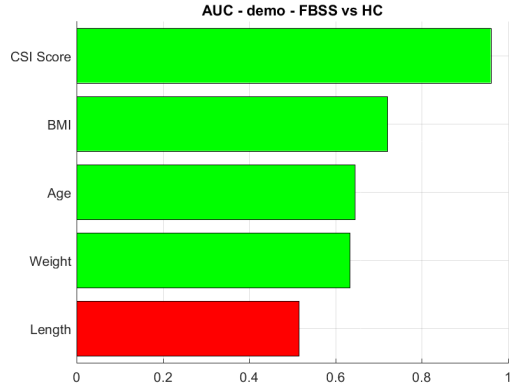
Table 8.1: Representation of the calculation of the accuracy (%) at each optimal point of the receiver operating characteristic (ROC) curve. This is done for four comparisons with diagnoses, failed back surgery syndrome (FBSS), the painful diabetic poly neuropathy (DPNp), the pain-less diabetic mellitus (DMnp) and the healthy controls (HC). In this table, the inter pulse interval is referred to as IPI. The maximum values are shown in red.

	FBSS vs HC	**DPNp vs HC**	**DMnp vs HC**	**DMnp vs DPNp**
N1 Double pulse mean (40 ms IPI)	58	67	58	61
N1 Double pulse mean (10 ms IPI)	59	61	59	56
N1 Single pulse mean	49	56	53	62
N1 Amplitude first pulse	56	56	57	58
N1 Amplitude seconde pulse (10 ms IPI)	57	55	51	59
N1 Amplitude seconde pulse (40 ms IPI)	64	53	54	56
N1 Intercept	59	62	61	61
N1 Mean	57	60	58	56
N1 Response	54	55	54	56
N1 Response * trial number	59	61	51	59
N1 Trial number	55	57	60	55
NDT Amplitude first pulse	63	76	58	74
NDT Amplitude seconde pulse (10 ms IPI)	66	74	62	67
NDT Amplitude seconde pulse (40 ms IPI)	66	75	64	71
NDT Detection rate	60	69	61	65
NDT Intercept	58	63	53	64
NDT Mean response Time	59	58	61	56
NDT Slope double pulse (10 ms IPI)	63	80	62	73
NDT Slope double pulse (40 ms IPI)	65	77	64	76
NDT Slope single pulse	63	76	58	74
NDT Threshold double pulse (10 ms IPI)	69	72	70	70
NDT Threshold double pulse (40 ms IPI)	66	69	67	68
NDT Threshold single pulse	65	72	57	61
NDT Trial	66	70	54	67
P2 Double pulse mean (40 ms IPI)	51	62	67	50
P2 Double pulse mean (10 ms IPI)	58	63	68	58
P2 Single pulse mean	54	68	68	55
P2 Amplitude first pulse	59	67	58	58
P2 Amplitude seconde pulse (10 ms IPI)	65	67	65	58
P2 Amplitude seconde pulse (40 ms IPI)	64	66	61	56
P2 Intercept	70	58	58	55
P2 Mean	57	64	68	56
P2 Response	64	68	67	62
P2 Response * trial number	59	68	58	70
P2 Trial number	58	65	61	64

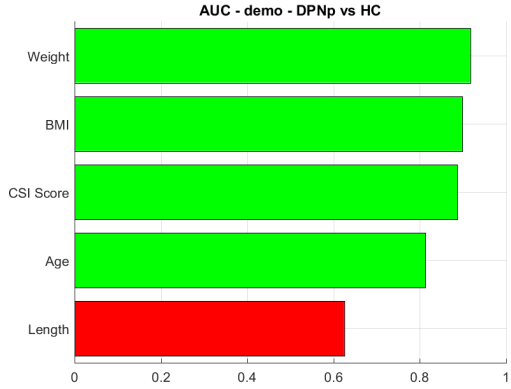
8.4 Demographic measures

Figure 8.8 describes the AUCs of the ROC curves for the different comparisons (i.e., FBSS vs HC (Figure 8.8a, DPNp vs HC (Figure 8.8b), DMnp vs HC (8.8c) and DMnp vs DPNp (Figure 8.8d)). This information is useful when in future classification models, demographic data is being inserted as additional features. Length is in all four comparisons not significant in classification of the two outcome measures. The only other not significant feature is the CSI score in DMnp versus HC. This can be explained by the absent central alterations in both outcome measures. The latter also explains why in other comparisons the CSI score is significant.

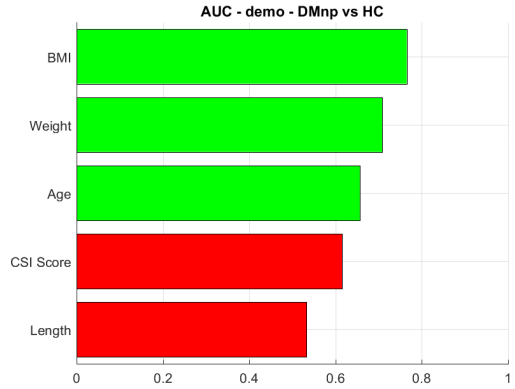
The fact that age is significant can be explained by the difference in subject characteristics. However, the prevalence in certain age groups of a specific condition might be higher and therefore differentiate in outcome. This is important for interpretation of the results and they can only be considered as an additional feature and not an individual feature (e.g., not all higher aged subject do suffer from a FBSS). This also applies for the BMI and weight. Those are also both considered significant in all figures.



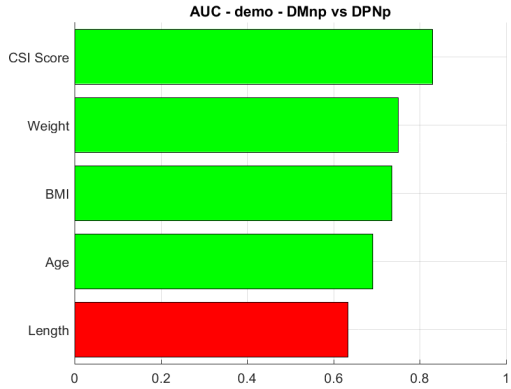
(a) Demographic values for AUC with outcomes: FBSS and HC (healthy controls).



(b) Demographic values for AUC with outcomes: DPNp and HC (healthy controls).



(c) Demographic values for AUC with outcomes: DMnp and HC (healthy controls).



(d) Demographic values for AUC with outcomes: DMnp and DPNp.

Figure 8.8: Bar graphs of demographic data of the subject of four different comparisons. The diagnoses include: Failed back surgery syndrome (FBSS), the painful diabetic poly neuropathy (DPNp), the pain-free diabetic mellitus (DMnp) and the healthy controls (HC).

8.5 Correlation coefficients data excluding outliers

The three remaining correlation coefficient matrices are shown in on the next pages. Figure 8.9, shows the correlation coefficients of DPNp versus HC. Next, the correlation coefficients of DMnp versus HC are shown in Figure 8.10, followed by the matrix of DMnp versus DPNp in Figure 8.11.

The correlation matrix of FBSS versus HC is shown in the results in Figure 4.6. Because of similar matrices, the remaining matrices are shown in this section of the Appendix.

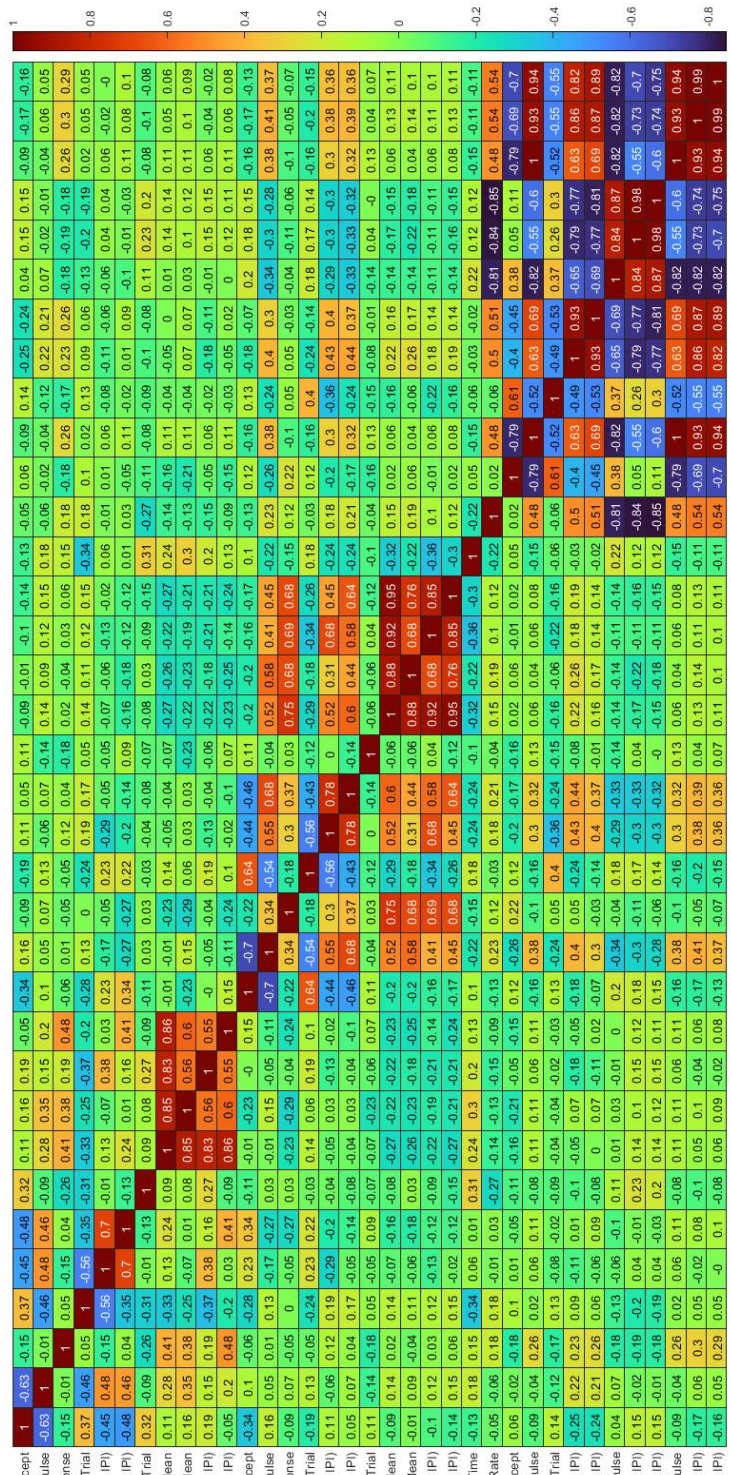


Figure 8.9: Correlation coefficients of painful Diabetic Poly Neuropathy (DPNp) versus Healthy Controls (HC).

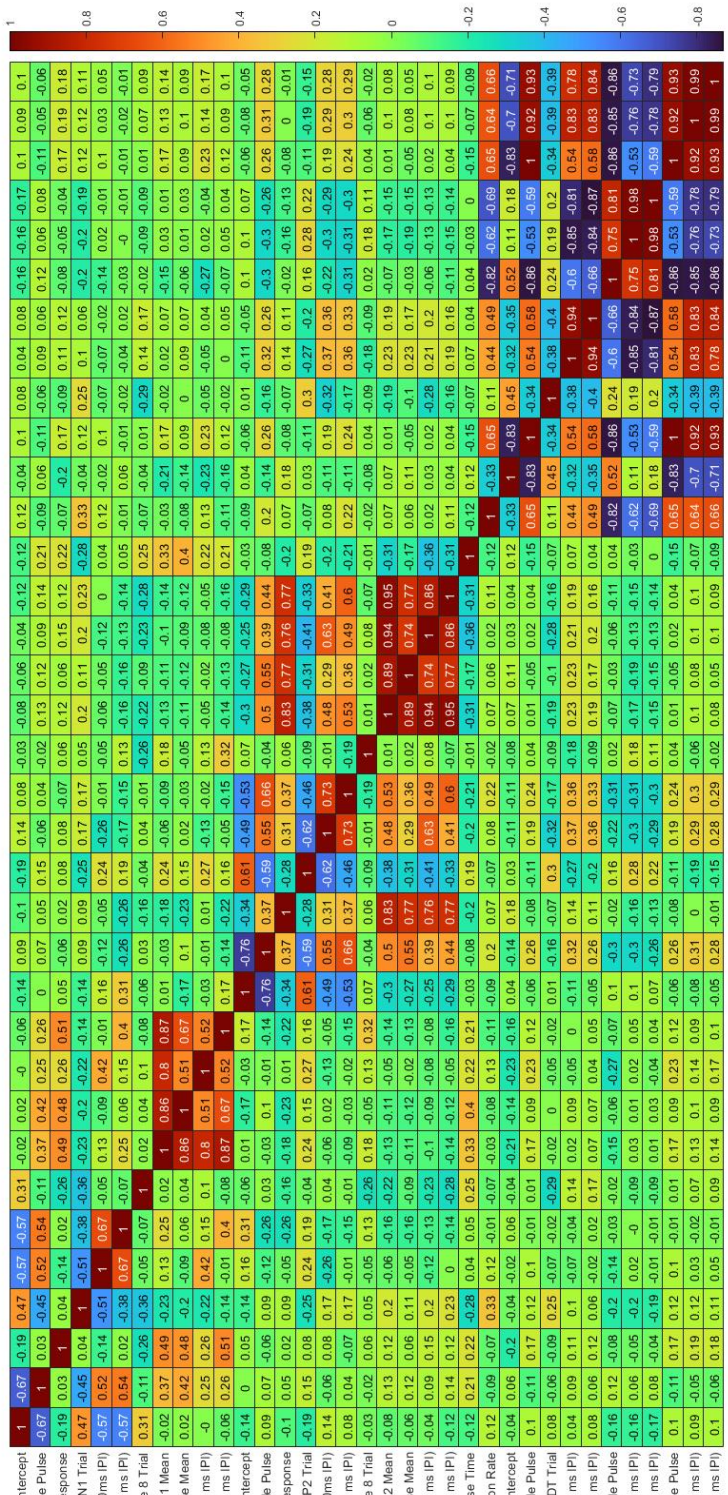


Figure 8.10: Correlation coefficients of not painful Diabetes Mellitus (DMnp) versus Healthy Controls (HC).

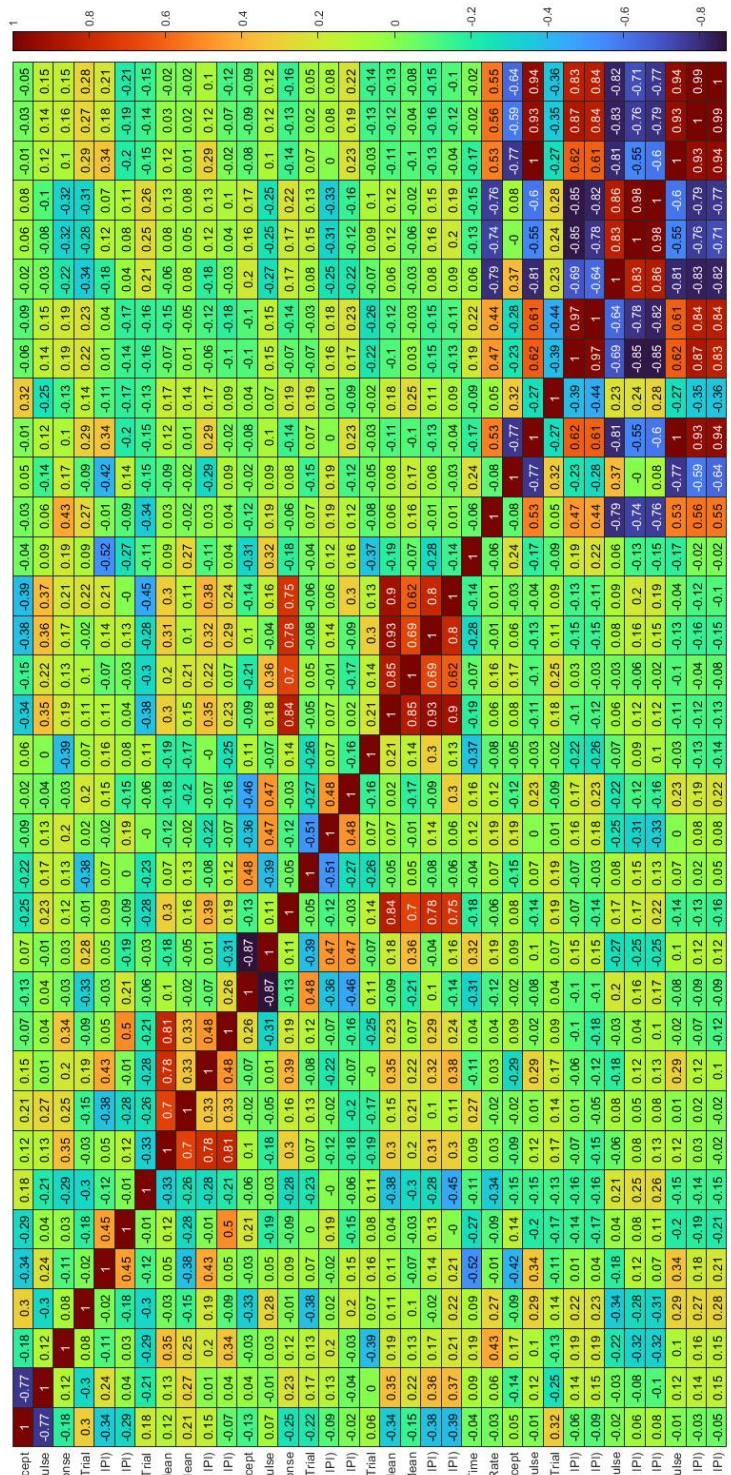
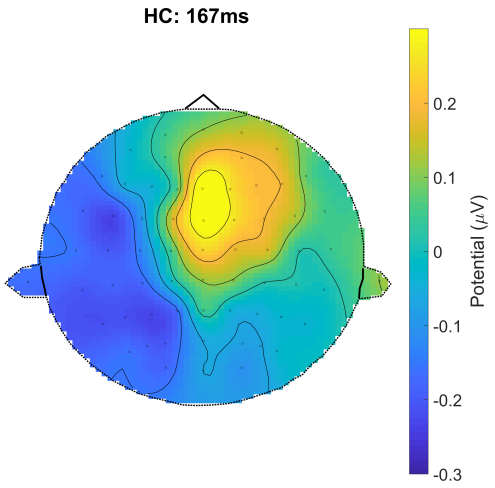


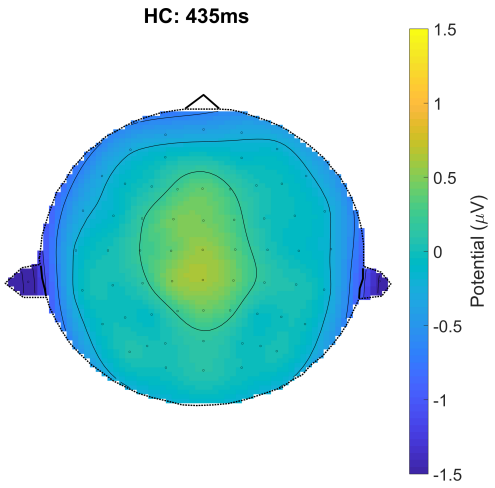
Figure 8.11: Correlation coefficients of Diabetes Mellitus (DMnp) versus Diabetic Poly Neuropathy (DPNp).

8.6 Topographies

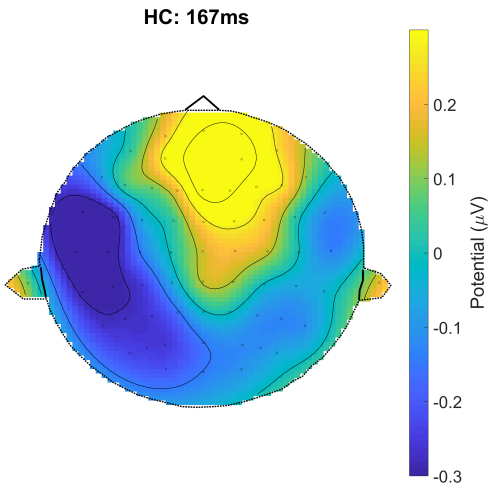
Underneath, the topographies are shown for each research group at N167 (N1) and P435 (P2). Figure 8.12, only contains the topographies of the first and the second half of the HCs. The remaing topographies of the subject groups are shown in Figure 8.13. It can be seen that for principal part of the topographies at N1, there is a large lateral potential difference at the left side. This is in correspondence with the researched channel choice of N1 (T7-F4). In addition, for P2, the potential difference is most visible at the centre. This is also in correspondence with the researched channel for P2 (CPz-A1,A2).



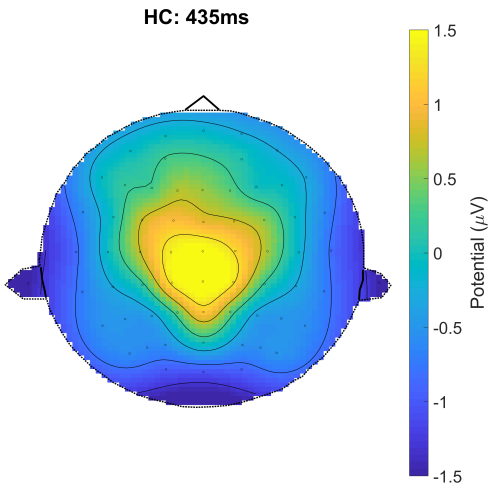
(a) First half of the HC research group at N1.



(b) First half of the HC group at P2.

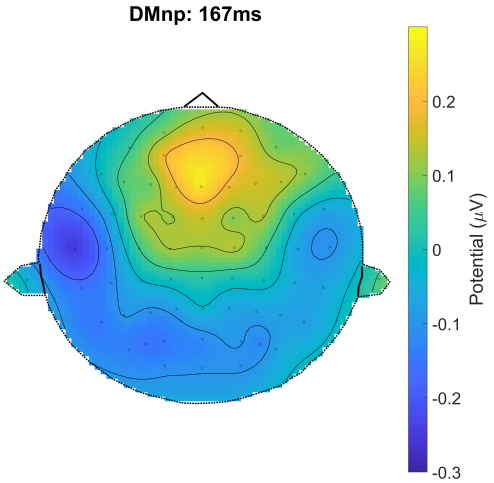


(c) Second half of the HC group at N1.

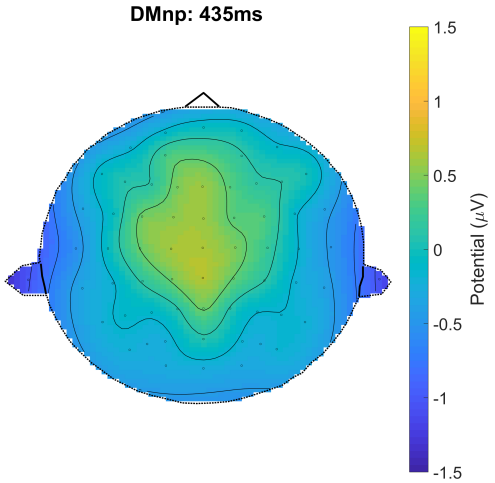


(d) Second half of the HC group at P2.

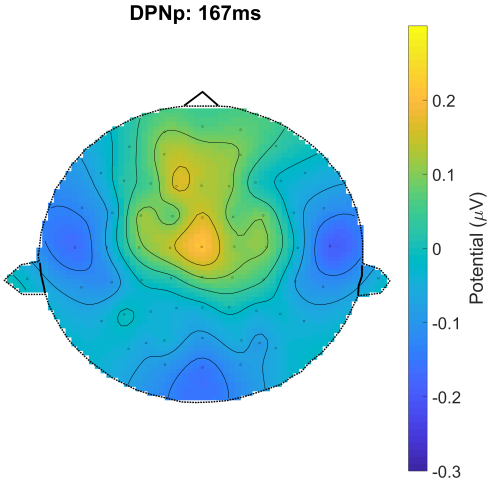
Figure 8.12: Topographies of the HCs at N1 and P2.



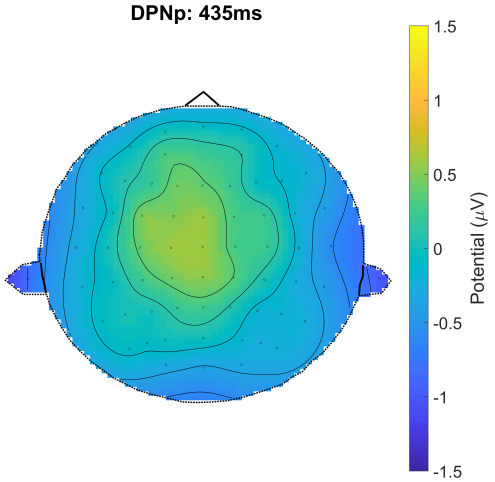
(a) DMnp group at N1.



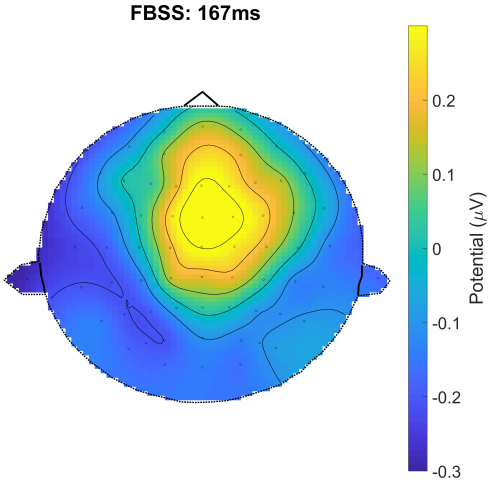
(b) DMnp group at P2.



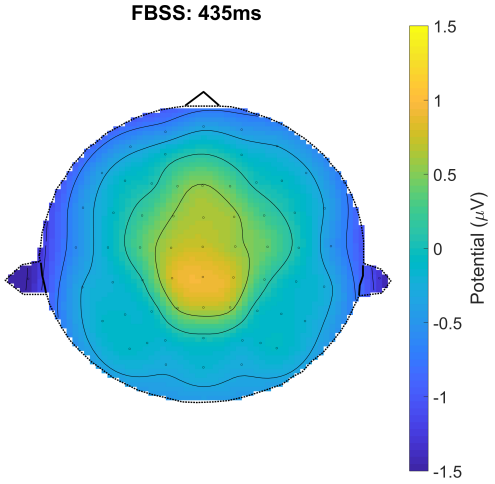
(c) DPNp group at N1.



(d) DPNp group at P2.



(e) FBSS group at N1.



(f) FBSS group at P2.

Figure 8.13: Topographies at N1 and P2 at groups of disordered subject groups.