Personalized Pain Medicine

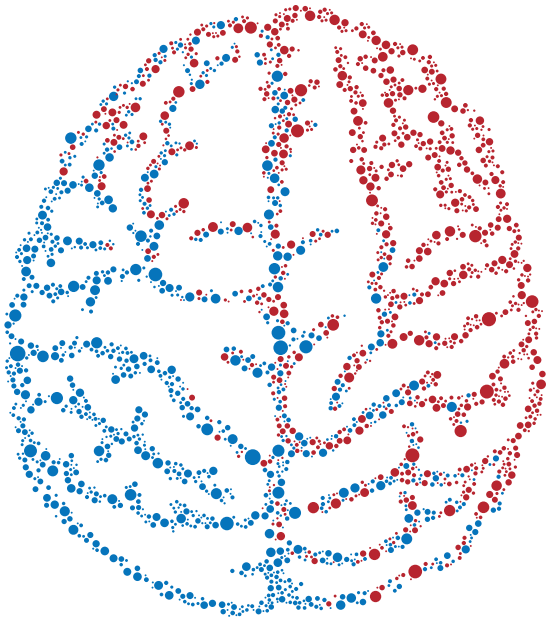
Using Electroencephalography and Machine Learning

Gram, Mikkel

DOI (link to publication from Publisher):
10.5278/vbn.phd.med.00041

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record



PERSONALIZED PAIN MEDICINE

USING ELECTROENCEPHALOGRAPHY AND MACHINE LEARNING

BY
MIKKEL GRAM

DISSERTATION SUBMITTED 2015

AALBORG UNIVERSITY
DENMARK

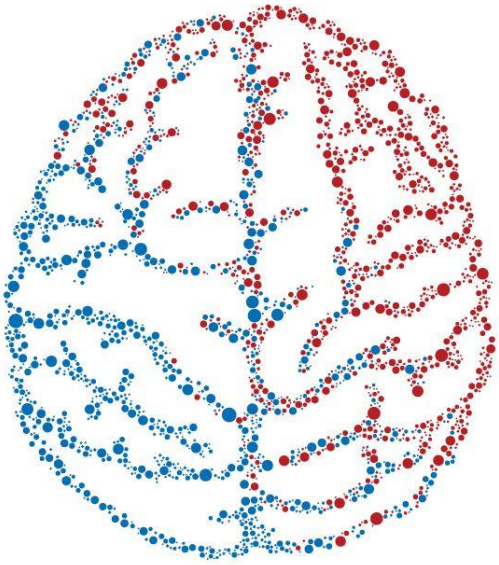
Personalized Pain Medicine

Using Electroencephalography and Machine Learning

by

Mikkel Gram

Dissertation submitted 2015



AALBORG UNIVERSITY
DENMARK

Thesis submitted: December 18, 2015

PhD supervisor: Prof. Asbjørn Mohr Drewes, MD, PhD, DMSc
Aalborg University Hospital
Aalborg University, Denmark

PhD committee: Associate Professor Sten Rasmussen, MD, PhD (chair.)
Aalborg University Hospital
Aalborg University, Denmark

Associate Professor Lone Nikolajsen, MD, PhD, DMSc
Aarhus University Hospital
Aarhus University, Denmark

Professor André Mouraux, MD, PhD
Université catholique de Louvain Institute of Neuroscience (IoNS), Belgium

PhD Series: Faculty of Medicine, Aalborg University

ISSN (online): 2246-1302
ISBN (online): 978-87-7112-439-2

Published by:
Aalborg University Press
Skjernvej 4A, 2nd floor
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2015

CV

Current position and work address:
2012 – 2015 PhD student at Department of Gastroenterology and Hepatology, Aalborg University Hospital, Hobrovej 18-22, 9100 Aalborg

Education:
2012 M.Sc. in Biomedical Engineering and Informatics

Published papers:

1. Brokjær A, Olesen AE, Kreilgaard M, Graversen C, **Gram M**, Christrup LL, Dahan A, Drewes AM: Objective markers of the analgesic response to morphine in experimental pain research. J. Pharmacol. Toxicol. Methods 2015; 73:7–14
2. **Gram M**, Graversen C, Olesen AE, Drewes AM: Machine learning on encephalographic activity may predict opioid analgesia. Eur. J. Pain 2015; 19:1552–61
3. Haas S, Brock C, Krogh K, **Gram M**, Lundby L, Drewes AM, Laurberg S: Abnormal neuronal response to rectal and anal stimuli in patients with idiopathic fecal incontinence. Neurogastroenterol. Motil. 2015; 27:954–62
4. Nilsson M, Sandberg TH, Poulsen JL, **Gram M**, Frøkjaer JB, Østergaard LR, Krogh K, Brock C, Drewes AM: Quantification and variability in colonic volume with a novel magnetic resonance imaging method. Neurogastroenterol. Motil. 2015; 27:1755–63
5. Sandberg TH, Nilsson M, Poulsen JL, **Gram M**, Frøkjær JB, Østergaard LR, Drewes AM: A novel semi-automatic segmentation method for volumetric assessment of the colon based on magnetic resonance imaging. Abdom. Imaging 2015; 40:2232–41
6. **Gram M**, Graversen C, Olesen SS, Drewes AM: Dynamic spectral indices of the electroencephalogram provide new insights into tonic pain. Clin. Neurophysiol. 2014; 126:763–71

7. Haas S, Brock C, Krogh K, **Gram M**, Nissen TD, Lundby L, Laurberg S, Drewes AM: Cortical evoked potentials in response to rapid balloon distension of the rectum and anal canal. Neurogastroenterol. Motil. 2014; 26:862–73
8. **Gram M**, Graversen C, Nielsen AK, Arendt-Nielsen T, Mørch CD, Andresen T, Drewes AM: A novel approach to pharmaco-EEG for investigating analgesics: assessment of spectral indices in single-sweep evoked brain potentials. Br. J. Clin. Pharmacol. 2013; 76:951–63
9. Hansen TM, Nilsson M, **Gram M**, Frøkjær JB: Morphological and functional evaluation of chronic pancreatitis with magnetic resonance imaging. World J. Gastroenterol. 2013; 19:7241–6
10. Frøkjær JB, Olesen SS, **Gram M**, Yavarian Y, Bouwense SAW, Wilder-Smith OHG, Drewes AM: Altered brain microstructure assessed by diffusion tensor imaging in patients with chronic pancreatitis. Gut 2011; 60:1554–62

This thesis is based on the following papers:

I. Gram, M., Graversen, C., Olesen, S. S., & Drewes, A. M. Dynamic spectral indices of the electroencephalogram provide new insights into tonic pain. *Clinical Neurophysiology, 2014, 126:763–71*

II. Gram, M., Graversen, C., Olesen, A. E., & Drewes, A. M. Machine learning on encephalographic activity may predict opioid analgesia. *European Journal of Pain, 2015, 19:1552–61*

III. Gram, M., Erlenwein, J., Petzke, F., Falla, D., Przemeck, M., Emons, M., Reuster, M., Olesen, S. S., & Drewes, A. M. Prediction of Post-Operative Opioid Analgesia using Clinical-Experimental Parameters and Electroencephalography. *Submitted to Anesthesiology*

Abbreviations

CNS	Central Nervous System
EEG	Electroencephalography
JMI	Joint Mutual Information
MPSS	Mainz' Pain Staging System
NRS	Numerical Rating Scale
PLI	Phase-Lag Index
QUIPS	Quality Improvement in Postoperative Pain Management
QST	Quantitative Sensory Testing
SVM	Support Vector Machine

English summary

Insufficient analgesia after surgery increases the risk of developing chronic postoperative pain, affecting the quality of life and is associated with large economic losses. Opioids are the drugs of choice for moderate to severe pain and are used by many chronic pain patients in Denmark. However, the analgesic effect varies widely between individuals, resulting in inadequate response for approximately one-third of patients. Biomarkers of opioid analgesia are therefore needed in order to personalize and optimize the treatment.

The objective of this PhD thesis was to investigate electroencephalography (EEG) as a biomarker for analgesic response to opioids. This included establishing changes in the biomarker during pain, the reliability of the biomarker and investigating if the signal contains information which could help determine if a given patient will respond to treatment.

Data from one experimental study and one clinical observational study formed the basis for this PhD thesis. The first study was a placebo-controlled experiment in healthy volunteers. EEG during rest and cold pain (hand immersed in 2°C water for two minutes) was recorded prior and 60 minutes after administration of 30 mg oral morphine. Paper I explored reliability and dynamics of spectral indices during rest and cold pain. Paper II investigated whether EEG could be used to predict the analgesic response to morphine. The second study was an observational survey of patients undergoing total hip replacement surgery and formed the basis for paper III. Patients were examined prior to surgery where EEG was also recorded. Pain was assessed in the first 24 hours after surgery to determine analgesic response from combined oxycodone and piritramide treatment.

The main finding from paper I was that the spectral indices derived from EEG are reliable over time, which is essential for a biomarker to be clinically viable. Furthermore, the theta (4 – 8 Hz) band correlated both overall and dynamically with pain ratings, suggesting close involvement in the pain experience. In paper II it was possible to predict which subjects would respond to morphine with an accuracy of 72% using EEG spectral indices during cold pain as baseline measurement. Resting EEG on the other hand, could not predict the morphine response. Paper III further confirmed these results by predicting response to oxycodone and piritramide after surgery with an accuracy of 65%, again this was only possible using EEG during cold pain.

In conclusion, EEG appears suitable as an objective biomarker for analgesic response to opioids. It is both reliable over time and contains information which can assist in determining if a given opioid is suitable for a patient. This represents the

first step towards a personalized medicine algorithm based on EEG. Future clinical studies should investigate the EEG further in more clinical studies. Another venue of interest would be to combine the EEG with other biomarkers, such as genetics, quantitative sensory testing and psychological questionnaires.

Dansk resumé

Utilstrækkelig smertebehandling efter operationer øger risikoen for udvikling af kronisk smerte, nedsætter livskvaliteten for patienten og er forbundet med store økonomiske omkostninger for samfundet. Opioider er den foretrukne medicin til at behandle moderat til svære smerter og bliver brugt at mange patienter med kronisk smerte i Danmark. Der er imidlertidig stor forskel mellem individer på den smertestillende effekt af opioider, hvilket resulterer i utilstrækkelig smertebehandling i omkring en tredjedel af patienterne. Derfor er der behov for biomarkører for opioiders effekt for at målrette behandlingen mod den enkelte patient.

Formålet med dette ph.d.-projekt var at undersøge elektroencefalografi (EEG) som en biomarkør for smertestillende effekt af opioider. Dette inkluderede etablering af biomarkørens ændringer under smerte, dens pålidelighed samt undersøgelse af om signalet indeholder information som kan bruges til at forudsige om en patient vil respondere på behandlingen.

Data fra et eksperimentelt og et klinisk studie danner basis for ph.d.-projektet. Det første studie var et placebo-kontrolleret eksperiment på raske frivillige. Der blev optaget EEG i hvile og under kuldesmerter (hånden nedsunket i 2˚C vand i 2 minutter) før og 60 minutter efter oral indtagelse af 30 mg morfin. Arbejde I undersøgte pålidelighed og dynamik af de spektrale indeks under hvile og kuldesmerter. Arbejde II forsøgte at forudsige responset på morfinbehandling ud fra EEG. Arbejde III er baseret på det andet studie som var et klinisk observations studie af patienter der skulle opereres for slidgigt i hoften med en hofteprotese. Patienterne blev undersøgt før operationen, hvor bl.a. EEG blev optaget. Patienternes smerte blev vurderet de første 24 timer efter operationen for at bestemme responset på kombinationsbehandling med oxycodon og piritramid.

Hovedresultaterne fra arbejde I var at de spektrale indeks fra EEG er pålidelige, hvilket er en essentiel kvalitet for en biomarkør til klinisk brug. Der blev også påvist en både overordnet og dynamisk sammenhæng mellem theta båndet og forsøgspersonernes smerte, hvilket indikerer en tæt forbindelse mellem theta (4 – 8 Hz) båndet og oplevelsen af smerte. I arbejde II var det muligt at forudsige hvilke patienter der ville respondere til morfin med 72% nøjagtighed på baggrund af EEG under kuldesmerter optaget før morfinindgift. EEG i hvile var derimod ikke i stand til at forudsige effekten af morfin. Arbejde III bekræftede disse resultater yderligere ved at forudsige effekten af behandling med oxycodon og piritramid med 65% nøjagtighed og dette var igen kun muligt ved brug af EEG under kuldesmerter.

EEG virker således brugbar som en biomarkør for smertestillende effekt af opioider. Målingen er både pålidelig og indeholder information som kan bestemme sandsynligheden for at et givet opioid har effekt ved den enkelte patient. Dette er det første skridt på vejen mod en algoritme til personlig smertebehandling på basis af EEG. EEG bør undersøges i nærmere i flere kliniske studier. En yderligere mulighed kunne være at kombinere EEG med andre biomarkører som genetik, kvantitativ sensorisk undersøgelse af smertesystemet og psykologiske spørgeskemaer.

Acknowledgements

I am thankful for the help and collaboration I have received from many people during my scientific work and would here like to mention a few.

First and foremost I would like to thank my supervisor Professor Asbjørn Mohr Drewes, MD, PhD, DMSc for his suggestions, incredibly fast responses and thorough comments, which has been indispensable.

I would like to thank my co-authors: Søren Schou Olesen, Anne Estrup Olesen, Carina Graversen, Joachim Erlenwein, Frank Petzke, Deborah Falla, Michael Przemeck, Miriam Emons and Michael Reuster for their contributions to papers I-III. Special thanks goes to Associate Professor Anne Estrup Olesen, PhD for her work on study I, Associate Professor Søren Schou Olesen, MD, PhD for all his help on statistics and editing, and Joachim Erlenwein, MD for his help in putting together the puzzle in study II.

I would like to thank all my colleagues at Mech-Sense, Department of Gastroenterology and Hepatology, and Department of Radiology for providing a great working environment. Especially thanks to the research nurses Isabelle Larsen and Birgit Kock-Henriksen who worked tirelessly in the laboratory to provide me with data.

I would like to thank all healthy volunteers and patients who enrolled in the studies.

I would like to acknowledge Innovation Fund Denmark - Individuals, Disease and Society for financial support. This work was supported by Lundbeck Foundation, Oticon Foundation and Otto Mønsted Foundation. All contributions have all been of great value.

I would like to thank Kirsten Wenneberg Pedersen for her always cheery and smiling demeanor.

I would like to thank all of my friends and family supporting me, in particular Shilpa Razdan for her wonderful support throughout the project and for pushing me forward when I needed it in the end as well as Jens T. Olesen for continuous encouragement and exceeding my expectations in interest for my work.

Table of contents

Chapter 1. Introduction

Proper control of pain still remains a problem in treatment of postoperative pain[1,2]. Inadequate pain treatment adversely affects patients immediately following surgery[3]. Acute pain felt immediately after surgery also increases the likelihood of developing persistent long-term chronic pain, more than the procedure itself[3,4]. This is connected to a loss of function and quality of life[5] as well as large direct and indirect economic costs for society[6].

Opioids are widely used as treatment for moderate-to-severe pain[7], with up to 13% of Danish chronic pain patients in treatment[8,9]. Despite it wide use, the response to opioid analgesia is heterogeneous with approximately 30% of patients not responding to the treatment[10]. The exact mechanism that determines opioid responsiveness still remains elusive, but gender, age and genetic variation are known factors[10–13]. Switching to another opioid can be effective in treatment of patients who initially do not respond. Thus, pain control was improved from 74% to 96% when switching non-responding patients from morphine to another opioid[14]. Personalized treatment of pain therefore shows great in optimizing treatment, by enabling the initial choice of the correct opioid. However, to date there is no accurate method for identification of opioid responders[15]. Several attempts have been made to predict the analgesic efficacy using quantitative sensory models (QST), but results have been conflicting between QST modalities and study populations[16,17]. Thus, there is a need for further investigations before personalized pain treatment is a reality[15,18].

Pain is a conscious perception arising in the brain[19]. Therefore, individual differences in brain function can also influence analgesic efficacy for the individual and a better understanding of these differences is needed to pave the road for personalized pain treatment[18]. Magnetic resonance imaging and positron emission tomography can assess brain function, but require specialized personnel and incur high costs both in acquisition and maintenance. Electroencephalography (EEG) is a more clinically feasible method due to lower costs and simplicity of the required equipment. EEG has previously been used to assess individual differences in brain function to predict the experience of pain for the individual[20]. However, it is possible that a single predictor will not enable personalized treatment, and instead several factors will have to be assessed simultaneously[18]. Machine learning enables such assessments and it has been shown previously to enable otherwise obscured findings when assessing the EEG by looking at several features at once[21].

1.1. Hypothesis

The hypothesizes for the PhD research project were: 1) EEG is a reliable measurement which is stable over time, 2) EEG contains information which can be used to accurately predict the outcome of opioid treatment and 3) the use of machine learning will improve prediction by taking into account several predictors at once on an invidualized level.

1.2. Aims

This PhD thesis took the following three-step approach for investigating personalized pain medicine using EEG and machine learning:

- To investigate reliability and dynamics of the EEG during rest and cold pain in healthy volunteers to assess robustness of the biomarker.
- Investigate if prediction of morphine analgesia is possible in healthy volunteers.
- Predict postoperative opioid analgesia in patients undergoing total hip replacement surgery.

This resulted in the following aims:

I. To investigate the reliability of EEG over time in order to establish if the measurements would be sufficiently robust to use in personalized medicine (Paper I).
II. Analyze dynamics of the EEG and pain ratings during cold pain to gain further insight into pain processing during cold pain.
III. To establish whether the EEG during rest or tonic cold pain is the most suited for prediction of opioid analgesia (Paper I, II and III).
IV. To compare a machine learning analysis approach to conventional group-based statistics in order to investigate if more sensitive prediction can be achieved (Paper II and III).
V. To explore if combination of several features in machine learning enables separation of responders and non-responders to morphine in healthy volunteers (Paper II).
VI. To investigate if machine learning can be used to predict the analgesic efficacy in patients with postoperative pain (Paper III).

Chapter 2. The Pain System

The international association for the study of pain defines pain as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage". Thus, pain serves as a warning signal for the brain to act or risk harm to the body. Pain is most commonly triggered by activation of peripheral nociceptors which relay information about the noxious stimuli to the central nervous system (CNS)[22]. However, pain can also be felt in the absence of noxious stimuli[23].

2.1. Primary afferents

Primary afferents are free nerve endings which respond to noxious stimuli and relay information to the CNS regarding the intensity and location of the stimulus[24]. Primary afferents are divided into three fiber categories depending on size and conduction velocity:

- Aβ: Large, highly myelinated fibers which conduct signals quickly to the CNS. Mostly conducts signals regarding light touch and tactile information
- Aδ: Thin, myelinated fibers, which conduct signals slower than Aβ-fibers. Responds to noxious chemical, mechanical and thermal stimulations
- C: Non-myelinated fibers, responding to the same stimuli as Aδ-fibers, but with slower conducting velocity due to the lack of myelin sheets

Thus pain signals are primarily conducted to the brain by Aδ- and C-fibers. Figure 1 shows the perceived pain following noxious stimulation, where the signals conducted by the faster Aδ-fibers will arrive first to the brain and will subsequently be felt as a sharp, pricking pain (termed "first pain"). This sensation is followed shortly after by the second pain, when the signals from the slower conducting C-fibers arrive. Second pain is felt as a more diffuse and dull pain.

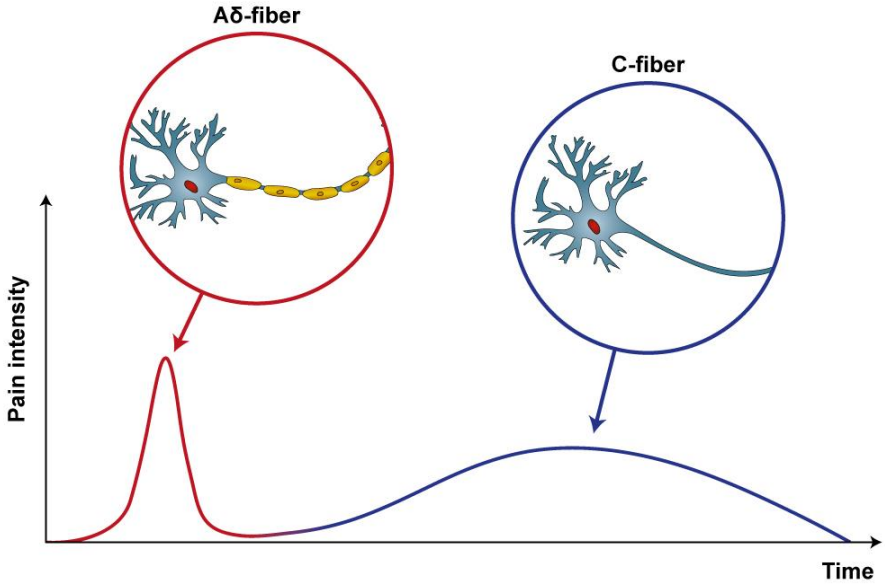


Figure 1: Pain following a stimulus is experienced first as a sharp pain (signals carried by the fast, myelinated Aδ-fibers) followed by a more dull and burning sensation (signals carried by slow, non-myelinated C-fibers).

2.2. Spinal cord processing

The primary afferents terminate in the dorsal horn where the signal is relayed to the secondary neurons. The dorsal horn is organized in layers (laminae I-VI). The Aδ- and C-fibers mostly terminate in the superficial layers (I-II), but some travel to the deeper layers. The information is relayed from the primary afferents via neurotransmitters, which can either inhibit or facilitate the neuronal activity[25]. Various neurotransmitters exist one such is glutamate which is important throughout the nervous system[26].

Once the nociceptive signal has entered the spinal cord, it is relayed via ascending pathways to the thalamus and the brainstem which in turn activates the higher cortical centers[26].

2.3. Cerebral processing

Once the nociceptive signal reaches the brain, an array of cortical centers are activated to process the signal. The exact areas involved depends on the noxious stimulus[26]. The most commonly activated centers include: the primary somatosensory cortex (S1), the secondary somatosensory cortex (S2), insula, anterior cingulate cortex and the prefrontal cortex[27]. These structures have been observed consistently across studies, and elicit responses which are correlated to the intensity of pain, and have therefore been named the "pain matrix" for its apparent participation in generating the perception of pain[28–30].

However, never studies have cast doubt on the concept of the pain matrix as solely responsible for generating the perception of pain from nociception, by demonstrating that the activity in the pain matrix also was modulated by the saliency of the noxious stimulus rather than the perceived pain intensity as well as the stimulus intensity[31]. This indicates that even though the cortical centers represented in the pain matrix are activated by painful stimuli, their response is not specific to the perception of pain and stimulus saliency seem just as dominant for the brains response[19].

Chapter 3. Pain treatment

Pain decreases quality of life significantly for patients, and is a common reason for patients to seek medical care[5]. The World Health Organization therefore devised a three-step ladder (see figure 2) for treatment of cancer pain, which has since been adopted for all pain treatments[32,33]. The ladder serves as a guideline for analgesic therapy with the goal to achieve successful pain control using the least potent analgesic first.

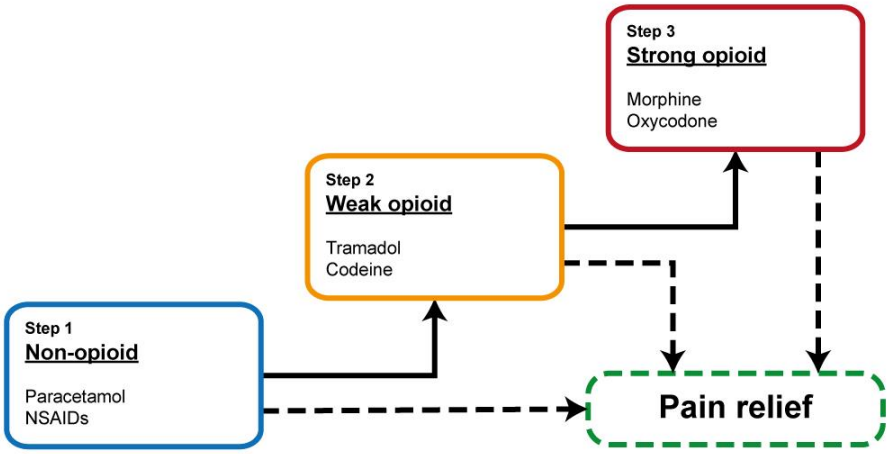


Figure 2: The WHO three-step ladder for pain treatment to guide physicians when prescribing analgesic therapy

According to the ladder, pain patients first presenting with pain should be treated with non-opioid analgesics. If the pain persists, a weak opioid such as tramadol can be added to treatment at the second step of the ladder. Should the weak opioids still be insufficient to achieve adequate pain relief, the last step of the ladder allows for substituting the weak opioid with a strong opioid such as the current golden standard in opioid therapy which is morphine[34].

Despite the existence of guidelines large differences exists in the agents prescribed as well as the quantity, putting into question how the guidelines are interpreted or to what extent they are followed[32]. Furthermore, doctors in Northern Europe are much more likely to prescribe opioids than in Southern Europe, with close to 0% of pain patients in Italy being on strong opioids such as morphine, while the figure for the United Kingdom was 12%[5]. This discrepancy in treatment approach probably contributes to the unsatisfactory pain control for around 40% of patients[5]

3.1. Opioids

Opioids is a group of analgesic drugs which binds to the opioid receptors found throughout the nervous system[35]. More specifically, four receptor types exist; μ (mu), δ (delta) and κ (kappa) receptors as well as the opioid-like receptor 1[35]. Opioids exert their effect on the receptors with either an *agonistic* or *antagonistic* effect. For example are both morphine and oxycodone thought to exert their analgesic effect by binding to the μ-receptor as an agonist, in CNS and to some degree in the peripheral nervous system[36].

Opioids exert large parts of their effect in the brain, and therefore the electrical activity recorded as EEG is affected[37,38]. However there is not complete agreement in the literature, as different methods for analysis complicate comparison of results and results on the EEG vary even for a single opioid[39]. The most common effect is a slowing of the EEG seen as an increase in the delta (1 – 4 Hz) band[40–42]. However, stronger opioids with higher receptor-affinity have been shown to affect the EEG in a more widespread way. Buprenorphine and fentanyl are both strong opioids with an affinity 25-100[43] and 75-100[36] times greater than morphine. These were studied to investigate the spectral EEG activity in response to painful electrical stimulation when administered through a transdermal patch[44]. Buprenorphine increased activity more, and in a more widespread frequency range than fentanyl (See figure 3), indicating a large difference in the EEG following treatment with the two opioids, despite being administered in equipotent doses. Therefore, impact on the EEG seems not defined by the opioid potency alone, but possibly by the receptors and mechanisms by which the opioids exerts their effect[44].

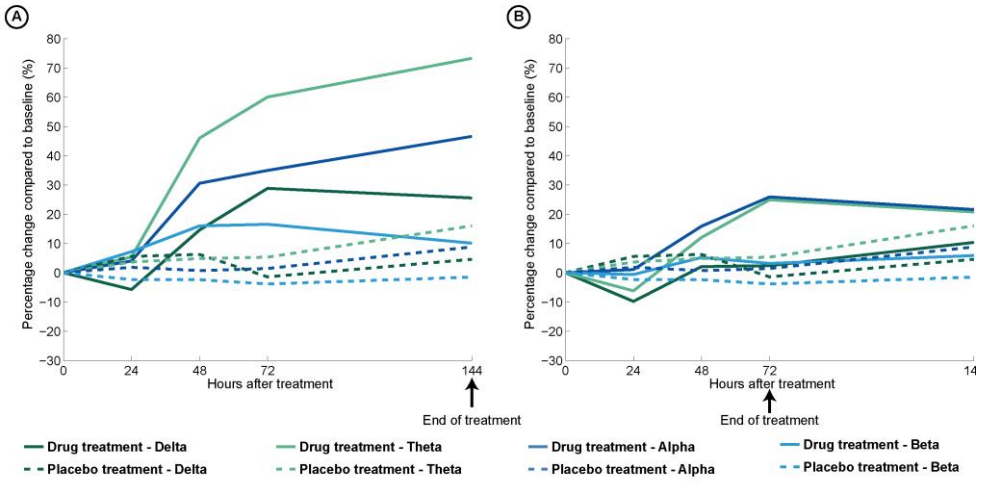


Figure 3: Relative changes in frequency bands over time in evoked brain potentials following electrical stimulation of the median nerve (wrist) after administration of a) buprenorphine and b) fentanyl. Both doses were equipotent in terms of morphine equivalent units and delivered via a transdermal patch.

3.2. Opioid responsiveness

The opioid response is heterogeneous with large parts of the population not responding. For morphine around 30% of patients are not non-responsive to morphine treatment[10]. There is no difference between opioids in response rate, but there seem to be large variation for the individual as to what opioid provide a satisfactory analgesic effect[34,45]. Currently, opioid selection is done based on the doctors experience on a trial-and-error basis, where non-responding patients can be switched to another opioid if the first choice is ineffective[46]. This approach to treatment does not account for the individual patient and often leads to unsatisfactory results[46], which is further reflected in 40% patients experiencing unsatisfactory pain management[5]. For example, one study raised the response rate from 74% to 96%, when patients not responding to morphine treatment was switched to another opioid[14]. This indicates that there is a large potential within pain treatment to select the right opioid for each individual patient[15,18].

3.3. Personalized medicine

Personalized medicine is the concept of using biomarkers to optimize medication type and dosage for the individual patient[15]. The concept of personalized medicine is explained in figure 4 together with the current method based on trial and error. Using the personalized medicine model it is more likely that patients will receive the right treatment initially and subsequently attain pain relief, as opposed to the current model where non-responders have to go through more treatments to achieve response. As a consequence suffering in these patients is prolonged, they experience more adverse effects and risks of development of chronic pain and drug abuse is increased.

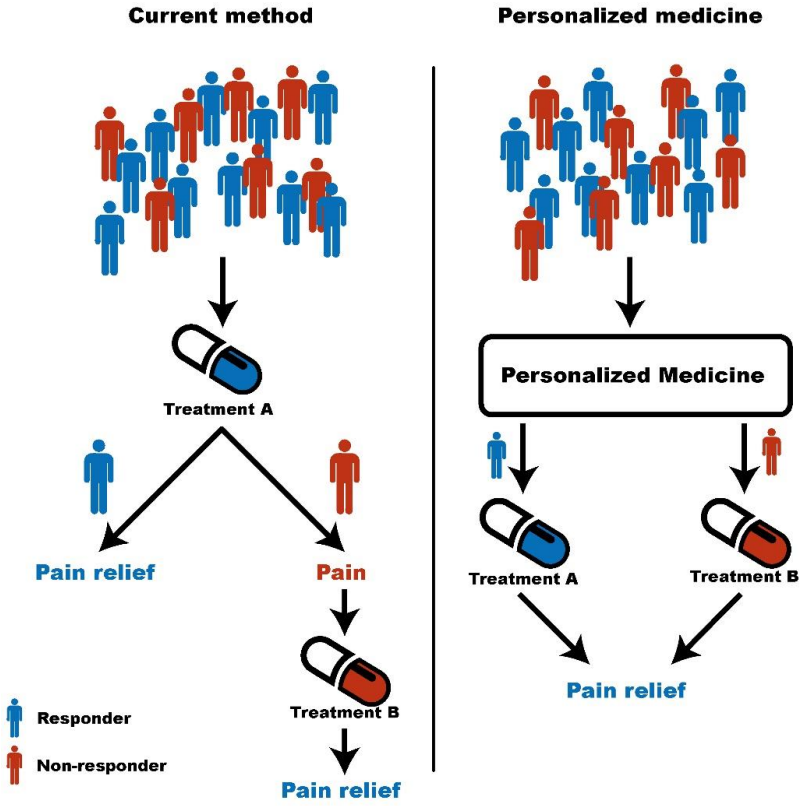


Figure 4: Schematic illustration of the current method for pain treatment and the principle behind personalized medicine. The current method has patients all treated with the same drug, of which some will respond and some will not. After it has been established that the non-responders are not properly treated they will be switched to an alternative analgesic to properly manage their pain. The personalized medicine approach will start out by classifying patients into different treatments depending on the appropriate drug of treatment, thus minimizing suffering and possible development of abuse.

The challenge with personalized pain medicine is to find reliable biomarkers for opioid responders and non-responders in order to identify potentially problematic patients. Several studies have attempted to identify simple parameters to distinguish between responders and non-responders using methods such as QST and different questionnaires such as pain catastrophizing, but results have been conflicting[16,17]. This indicates that there is still a need for additional parameters and/or new methods to combine parameters in order to achieve a reliable prediction of response[16]. Response to opioids is also affected by e.g., genetics, gender and age[11–13]. Since pain is a perception generated in the brain[19], differences in brain function could also reveal who will respond to different analgesics. EEG can assess individual changes in brain function and could therefore be a valuable tool for discovering biomarkers[20]. This study attempts to investigate if EEG can be used as a new biomarker for opioid analgesic response (Paper II and III). EEG is an objective method, which is also economically feasible and can be performed at bedside. The next chapter will explain EEG in further detail.

Chapter 4. Clinical parameters

Clinical data is usually connected to a degree of variability since the condition of individual patients are always different. To assess these differences and whether the patient cohort is representative, clinical parameters are collected. Some clinical parameters are common in most studies, such as age, gender and body-mass index, while others are entirely specific to the condition of interest. Paper III focuses on patients undergoing total hip replacement surgery, and in this chapter the clinical parameters specific to this patient group will be briefly explained.

4.1. Chronic Pain Assessments

Two assessments of chronic pain were used in paper III; one assessment by the clinician and one by the patient. The chronic pain grade was developed for patients to assess the severity of their chronic pain both in terms of pain intensity and how limiting the pain was for everyday life[47]. Chronic pain was divided into four pain grades which briefly can be explained as:

- Grade I: low pain intensity, low disability
- Grade II: high pain intensity, low disability
- Grade III: high disability – moderately limiting
- Grade IV: high disability – severely limiting

Patients who only experience low disability from their chronic pain belong in grade I and II as it is based on purely on pain intensity. Grade III and IV are based on the disability the patient is experiencing and is therefore only assessing how limiting the pain is[47].

The Mainz Pain Staging System (MPSS) was also used and is an assessment for clinicians for pain[48,49]. The clinician used the score to divide patients into three stages of pain chronicity.

4.2. Passive Hip Rotation

A simple test was developed to assess magnitude of evoked pain during passive rotation of the hip. The leg and knee was brought into a flexed position and rotated until the patient experienced pain. The leg was then brought into this position and held for 30 seconds. Then patients were asked to rate their overall pain on an 11-point numerical rating scale (NRS).

4.3. Response score

Since there is no validated definition for opioid response in a postoperative setting, we developed a "response score" in paper III to stratify patients into opioid responders and non-responders. The response score was determined based on the "Quality Improvement in Postoperative Pain Management" Questionnaire (QUIPS). QUIPS is a validated outcome measurement for postoperative quality control. A point-based response score was developed in order to stratify patients into opioid responders and non-responders. Points could be awarded up to a maximum of 10, and questions and instructions for awarding points can be found in Table 1. Cut-offs were based on typical intervention thresholds for postoperative pain management (e.g. pain on movement or maximal pain of 5)[50,51]. Patients with a response score of 5 or greater were considered opioid responders, a score of 4 or lower would indicate a non-responder.

Table 1: Explanation of the questions in the response score, and how points are distributed depending on the answers given for each question. Some questions are rated on an 11-point numerical rating scale, while others are simple yes or no questions.

Questionnaire entry	Answer	Points awarded
Maximal pain	0 – 4	2
	5 - 10	0
Pain on movement	0 – 4	2
	5 - 10	0
Minimum pain	0 – 2	2
	3 - 10	0
Pain on mobilization	Yes	1
	No	0
Pain when coughing	Yes	1
	No	0
Pain when waking up	Yes	1
	No	0
Pain affecting mood	Yes	1
	No	0

Chapter 5. Quantitative Sensory Testing

Clinical pain are often associated with other factors such as anxiety, cognitive and autonomic responses and fear[52]. Perception of pain is also influenced by gender and cultural background making evaluation difficult[53]. QST uses experimental pain models to assess the pain system by invoking pain under controlled circumstances and are advantageous since they can eliminate many of the aforementioned problems to some degree dependent on experimental setup[54,55]. QST models consist of a stimulus which is assessed with subjective methods such as a NRS or by objective measurements such as EEG[55]. The information which can be derived from the QST tests are largely dependent on the evoking stimulus, of which a large selection is available[55]. This chapter will describe the QST models utilized within papers I, II and III.

5.1. Cold Pain

Cold stimulations are carried to the brain primarily by the Aδ- (cold sensation) and C-fibers (cold pain)[56]. The most common modality for cold pain stimulation is the cold pressor test, which is also used in papers I, II and III. The test is performed by submerging the hand in cold water (temperature range 1 – 7 ˚C) and maintaining it there for a specified amount of time (e.g. two minutes)[57]. Patient will then rate their overall pain on a NRS or another pain rating scale, or in some cases rating will be continuous during the stimulation such as in paper I. The stimulation constitutes a tonic stimuli which has been shown to better mimic clinical pain due to the length of stimulus and the intense unpleasantness felt during stimulation[58]. Furthermore, tonic cold pain stimulation has been shown to be rather sensitive towards detecting opioid analgesia[55]. However, the cold pressor test has been shown to be somewhat variable depending on the individual and experimental methods[59]. The introduction of new specialized equipment for maintaining water temperature and circulation in the water is thought to mitigate some of this variability[57], but some inter-individual variability is still present.

Papers I and II used cold pain stimulation for determination of opioid response, since it sensitive to opioid analgesia[55]. Papers I, II and III included the method as a QST measure and recorded EEG during the evoked pain. Papers I and II used specialized temperature-controlled equipment with continuous circulation of the water. However, this equipment was not available for paper III which used a thermal container with cold water inside for the test. To ensure a consistent

temperature of the water was used, temperature was measured before pain stimulation each time.

5.2. Pressure Pain

Pressure can induce pain in the skin by use of pressure algometry[60]. This painful sensation is mediated to the brain through Aδ- and C-fibers[60]. An algometer is usually operated by the examiner, who increase pressure at a constant rate until the patient indicates that their pain threshold has been reached and the pressure required is the quantifiable pressure pain threshold or predefined suprathreshold levels. However, it can be difficult to increase the rate at a constant pace, and it is therefore recommended that the same examiner always perform the test in order to reduce variability[61]. Paper III included the pressure pain in order to assess the pain system of patients preoperatively.

5.3. Heat Pain

Heat pain is another kind of thermal stimulation which is also mediated by Aδ- and C-fibers. Heat can be applied both rapidly and slowly, depending on which kind of stimulation is of interest. Rapid stimulation will quickly activate Aδ-fibers which will mediate the sharp feeling of "first pain" within less than 0.5 seconds. "Second pain" is felt slower as it is mediated by the C-fibers. As such, the fibers that are activated can to some degree be controlled with the rate of temperature increase[55]. Paper III included slowly increasing (1˚C/s) heat pain stimulation as a QST measure to assess the preoperative state of the pain system in order to activate both Aδ- and C-fibers.

5.4. Conditioned Pain Modulation

Pain is modulated by inhibitory and facilitatory mechanisms which are influenced by a number of factors. Conditioned pain modulation (CPM) represents the relationship between the descending inhibition and facilitation. CPM is assessed by application of a conditioning noxious stimulus which will inhibit the pain induced by another test stimulus. CPM can be performed using a wide range of conditioning stimulation modalities[62], most common among them is the cold pressor test[63].

In paper III the CPM effect was measured since it has been shown to be decreased in patients with experimental and clinical pain[64–66]. CPM was assessed using the cold pressor test as a conditioning stimulus, and the slowly increasing heat stimulation as a test stimulus.

Chapter 6. Electroencephalography

EEG records the summation of electrical activity of the brain which is generated by the simultaneous firing of neurons and is well-known within pain and pharmacology research[27]. EEG can be recorded both locally (as local field potentials), on the surface of the cortex or on the surface of the scalp[67]. Due to the invasive nature of the first two methods, surface EEG is most commonly used as it is more clinically feasible to record and is therefore also chosen for this PhD thesis[67,68].

The primary advantage of EEG is the high temporal resolution and hence EEG allows for analysis of sequential brain activation, frequency information and direction of information flow due to pain or analgesics. The major disadvantage of surface EEG is the relatively poor spatial resolution due to all measurements being made at the scalp surface as brain signals recorded on the scalp are distorted due to volume conduction. Compared to imaging methods, the poor spatial resolution limits the ability of EEG in pinpointing the exact brain centers involved. On the other hand, the temporal resolution of imaging solutions is very poor (on second time-scale) and thus it is impossible to study sequential brain activity and ultimately separate pain specific from nonspecific activity as the pain stimulus reaches the brain within milliseconds of being presented. Furthermore, the costs and logistics involved in imaging solutions are severe limitations, making EEG a more suitable choice for developing a clinical bed-side system for prediction of opioid analgesic effect[69].

6.1. Generation of the EEG

The main contributor of the surface EEG is the summation of excitatory and inhibitory neurons in larger populations of neurons, which transmit to the surface via volume conduction[68]. The signal decays over distance, and therefore the largest potentials measured at the scalp surface is generated by the cortex, although deeper sources can also be identified[67,68] Since the generated electric field for the individual neuron is small, both the structure in which neurons are organized[67] as well as the synchronization of neuron firing is of great importance in order to generate a measurable potential at the scalp surface[39,67]. In the cortex, neurons are arranged perpendicular to the surface of the cortex, forming layers of neurons in palisade[67]. Activation can occur in synchrony and within defined layers, enabling generation of quite large potentials[67].

EEG can be recorded in two main categories: *spontaneous EEG* such as during rest or a prolonged stimulus and *event-related potentials* which are the direct response to a phasic and external stimulus[68]. This PhD thesis focuses on the spontaneous EEG during rest and cold pain.

6.2. Recording Surface EEG

Surface EEG has amplitudes in the order of microvolts (μV) and therefore needs to be amplified in order to be accurately measured[67]. Since the potentials are miniscale, differential amplification between two electrodes (an active electrode and the reference electrode) is necessary. This amplifies the difference in voltage between the two electrodes and has the advantage of cancelling out common signals between electrodes[67].

The frequency bandwidth of interest for EEG signals ranges from around 0.1 -100 Hz[67]. Filtering is often necessary due to noise such as:

- Slow voltage drift (slowly changing voltages).
- Muscle artifacts.
- Mains noise (50 or 60 Hz depending on country).
- Artifact from cables and other sources,

The voltage drift noise is caused by slowly changing properties of the electrode gel, electrodes or skin resistance[67]. These signals can be removed using a high-pass filter with a cut-off frequency of around 0.1 – 1 Hz[67]. Papers I, II and III all used 1 Hz as the cut-off frequency.
Muscle artifacts are generally high-frequency signals outside the frequency spectrum of the EEG, and is commonly removed using a low-pass filter set as low as possible depending on the spectrum of interest[67]. 30 Hz is a common cut-off, or 70/80 Hz if the gamma band is of interest. Papers I and II used 70 Hz as a cut-off while paper III used 32 Hz since the gamma band was excluded.
Mains noise can arise from the electrical net if the desired spectrum arises above 50 or 60 Hz (i.e. if the gamma band is of interest)[67]. To combat this problem a notch filter at the specified frequency is utilized. Since papers I and II included the gamma band, a notch filter designed for 50 Hz was employed.
Electrodes can be contaminated with noise which cannot be removed with filters. Signals from such electrodes are then rejected and an average signal is interpolated from neighboring electrodes to replace it. This is done manually and to preserve integrity of the recording it is essential that a limited number of channels is interpolated. If too many channels are affected by noise the recording will have to be rejected.

Since there are activity differences between brain structures, it is often desirable to record EEG from multiple locations spread over the scalp, to give a full

representation of the neuronal activity[67,68]. These electrodes are placed according to the 10-20 system which is universally accepted[67]. Number of electrodes can vary widely (up to 512 electrodes or more) depending on the application. A high number of electrodes is desirable for full representation of neuronal activity. However, as the number of electrodes increases more time is required for mounting, problems with high impedance between scalp and electrode increases and general expense of the system is also increased. Therefore a trade-off in the number of electrodes is required and depends entirely on the application. Papers I and II used a 62-electrode recording cap, as the study was experimental and included relatively few subjects. Paper III was a clinical study and therefore a 34-electrode cap was used to decrease mounting times. A clinical algorithm for personalized pain medicine will likely need to reduce this number even further, when the electrodes of interest have been established.

6.3. Evoked Brain Potentials

Evoked brain potentials is the time-locked response to an external stimulus [39,68]. Since the response to is highly dependent on the sequential activation of various brain centers, the evoked brain potentials have a specific morphology, as can be observed in figure 5 as the average potential[39]. In order to obtain the time-locked properties of the evoked brain potentials it is necessary for the evoking stimulus to be short and rise rapidly in intensity[67]. The potential is characterized by negative and positive peaks which is commonly quantified by their peaks (estimates amount of synchronously activated neurons) and latency (estimates delay conduction from the periphery and in activation caused by delayed cortico-cortical connections)[39]. The most common peaks are described in Table 2.

Table 2: The different peaks of evoked brain potentials, their latency and from which areas of the brain they originate.

Peaks	Latency [ms]	Typical brain area	General interpretation
Early[70]	20 - 60	Primary somato-sensory cortex	Non-pain specific somatosensory input from touch.
Intermediary[40]	60 – 120	Operculum Limbic system	Reflects nociceptive activation of supraspinal structures.
Late[71]	120 - 350	Operculum Limbic system	Reflects discomfort or Emotional aspects of pain.

Due to low signal-to-noise ratio of recorded evoked brain potentials, stimulation is normally repeated a number of times while recording the EEG and subsequently averaging all trials to get the average response to the stimuli[72]. When the technique was first developed EEG would often suffer from poor signal to noise ratio. Although much has been improved since in the recording techniques, post-processing and filtering, the averaging method is still widely used since the averaging procedure makes the evoked potential stand out from the background EEG in the signal[44]. However, there are many drawbacks to averaging, since it only preserves components which are time-locked to the stimulus[73]. Since this is not the case for all nociceptive input, averaging effectively removes useful information from the signal along with noise[73]. The reduction in variability by use of averaging is shown in figure 5. Use of more advanced time-frequency analysis on each EEG signal from each repeated stimulation (termed single-sweep analysis) has been shown to increase reliability[74], sensitivity to analgesics[44] and neuronal changes due to disease[75].

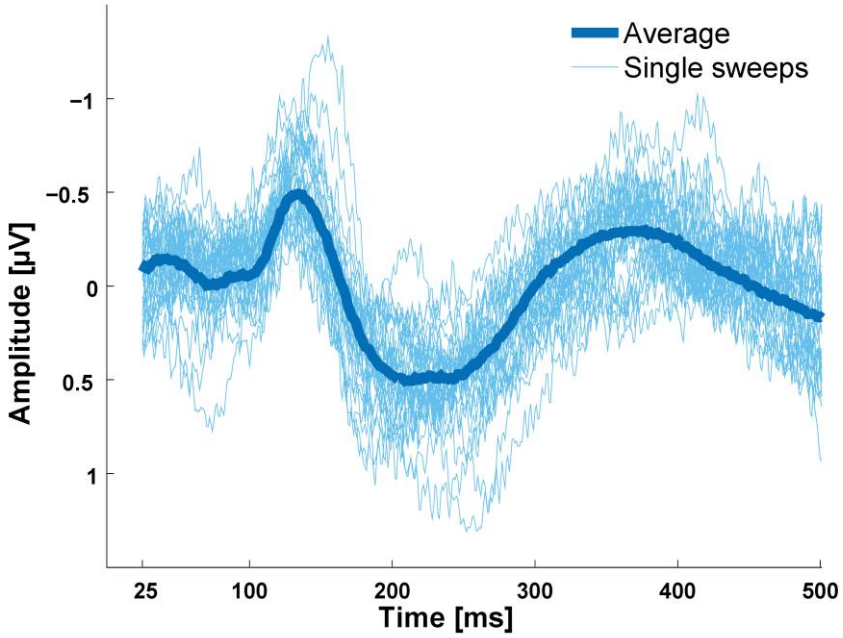


Figure 5: Single-sweep evoked brain potentials (green) and the resulting average potential (black). The loss of variation from the underlying single-sweeps by applying the averaging procedure is apparent.

The emergence of fast and affordable computing has enabled analysis of each individual trial, also termed single-sweep analysis. This analyses the signal including non-time locked components, and has been shown to be superior in detecting the effect of analgesic drugs[44], and in detecting differences in cerebral activity in patients with fecal incontinence[75].

6.4. Continuous EEG

Continuous EEG is not time-locked to a specific stimulus like evoked brain potentials. It is rather a recording of EEG over a specified time period. Continuous EEG lacks the easily discernible peaks of evoked brain potentials and as such it is usually divided into frequency bands to describe different processes in the brain. The exact limits of the frequency bands differ, but they are roughly defined as:

- Delta (δ): 1 – 4 Hz
- Theta (θ): 4 – 8 Hz
- Alpha (α): 8 – 12 Hz

- Beta (β): 12 – 32 Hz
- Gamma (γ): 32 – 80 Hz

Figure **6** illustrates an EEG signal and the division of the signal into frequency bands.

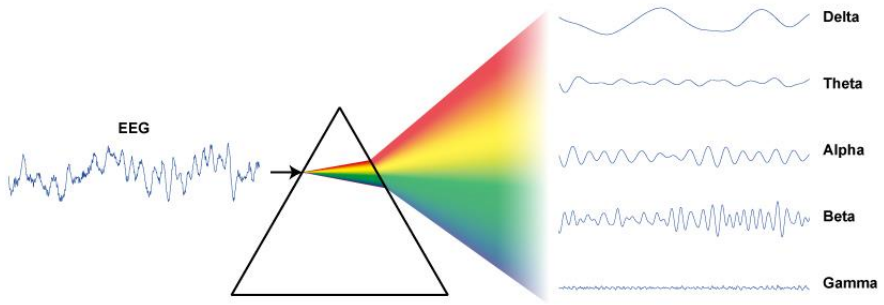


Figure 6: Illustration of how the spectrum of the EEG signal can be divided into the predefined frequency bands; delta (1 – 4 Hz), theta (4 – 8 Hz), alpha (8 – 12 Hz), beta (12 – 32 Hz) and gamma (32 – 80 Hz).

Continuous EEG is susceptible to artifacts from eye blinking, eye movements and muscle artifacts, which occur at amplitudes more than ten times that of the EEG signal[76]. Therefore, various methods have been developed to remove these artifacts using independent component analysis which can separate the signal into components that appear to be independent[76].

6.4.1. Resting state EEG

The most common recording of continuous EEG is made with the subject in a resting state. In this state subjects are instructed to be calm while either lying or sitting with little or no verbal communication. Resting EEG has been used extensively in combination with pharmacological agents to detect differences in the cortical activity following drug administration[39].

Resting EEG has a certain variability over time. To account for this variability it is recommended to used recordings of at least 120 seconds to get a steady recording[77].

Resting EEG has also been shown to drift gradually over time possibly due to shifts in attention state[78,79]. To maintain a constant attention level, special protocols have been developed to maintain vigilance, for example through simple arithmetic

tasks[80,81]. However, it has not been investigated if a similar drift occurs during these vigilance-controlled conditions.

6.4.2. Tonic Pain EEG

Tonic pain EEG is a variation of continuous EEG where the subject is undergoing a painful tonic stimulation. Tonic pain is thought to be a more physiologically meaningful stimulation that better mimics the perception of chronic pain due to the unpleasantness and time available for the brain to adopt to the painful event[58]. Furthermore, tonic pain models such as the cold pressor test are more sensitive to opioid analgesia possibly due to the deep pain evoked by the stimulation and higher level of unpleasantness[55]. Feeling of unpleasantness is connected to the limbic network which is an area where opioids are known to modulate pain response[27].

Thermal pain is a common tonic pain model, both heat[82,83] and cold[79,84,85] have been extensively used. Therefore, papers I, II and III employed the cold pressor test as a tonic pain stimulus. Tonic ischemic pain can also be induced by limiting the blood supply by use of a blood pressure cuff and has been used in clinical studies[86] and in combination with EEG before[87]. In addition, the relationship between ischemic pain and time is largely linear as opposed to the cold pressor test which experiences a ceiling effect[58].

Chapter 7. Feature extraction

EEG signals consist of millions of data points which in itself is impractical to assess. Feature extraction is the process of extracting overall characteristics from the many data points, in order to describe the signal with fewer numbers. A myriad of feature extraction methods exist, such as spectral analysis, mean dominant frequency, 95% spectral edge, sample entropy and connectivity[81,88–91]. For this thesis spectral analysis was used since it is well tested and has been proven useful in many studies. Also it does not discard any part of the spectrum and should therefore provide a more complete assessment of the signal.

7.1. Spectral analysis

Spectral analysis investigates the energy of the signal within the standard frequency bands. The Fourier transform has for many years been the golden standard for frequency analysis within EEG since the first introduction in 1932[92] although the method has several limitations[93]. Fourier analysis requires relatively long signal segments for detailed frequency resolution, and makes assumptions about the signal stationarity which is not fulfilled for EEG data[94,95]. Even so it is a commonly used method, and the variability of EEG measures derived using the Fourier transform is heavily dependent on the recording length[96]. Thus, longer recordings are more robust, and commonly a recording length of at least 20 seconds have been recommended[97]. However, newer evidence suggests that much longer recording lengths of at least 120 seconds should be used in clinical and pharmacological studies, to reduce influence from noise[96]. Several other methods exist for spectral analysis, the most common are listed in Table 3.

Table 3: Overview of common time-frequency algorithms for analysis

Name	Description	Advantages	Drawbacks
Fourier transform[98]	Classic frequency analysis.	Low computational cost	Time-frequency resolution
Wigner-Ville distribution[99]	Simple instance of Cohen's class	High temporal resolution. Well-suited for non-stationary signals such as EEG	Influenced by cross-term interference. Density estimate can result in negative values.
Matching pursuit[100–102]	Uses a redundant dictionary of atoms for decomposing the signal into a sparse representation	Time-frequency resolution adapted to the signal	Does not represent all frequencies, making comparisons to other studies difficult. Computationally costly
Wavelet analysis[103,104]	Multi-resolution analysis. Higher temporal resolution for high frequencies and higher spectral resolution for low frequencies	EEG signals have properties suitable for multi-resolution analysis.	Requires a mother wavelet function, which is an a priori assumption that can affect outcome

An alternative is the wavelet transform which takes a different approach to spectral analysis, by decomposing the signal into coefficients through use of a mother wavelet function[103]. Mother wavelet functions can be selected from a library or developed independently to satisfy the mother wavelet criteria; zero mean value, finite energy and relatively little low frequency content compared to high frequency content[95]. During analysis the mother wavelet is scaled, changing its center frequency and thus achieving affinity for different frequencies. The process also

illustrated in figure 7 is repeated for scales matching the frequency range of interest[103].

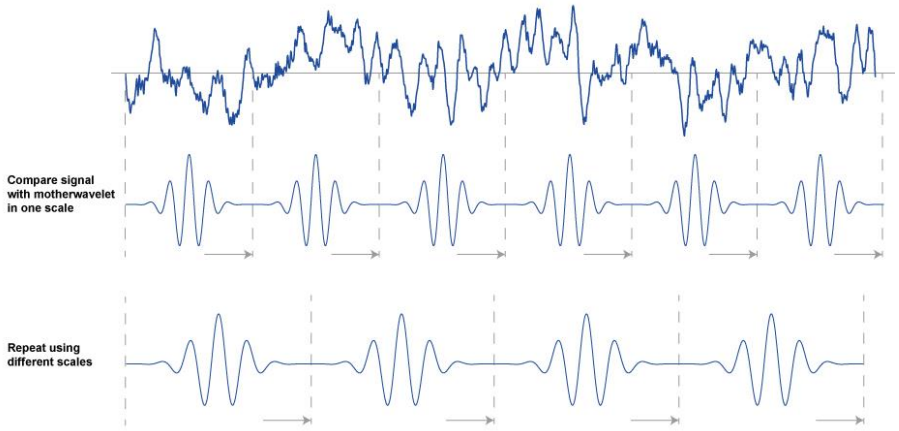


Figure 7: Decomposition of a signal using a mother wavelet function at different scales. The mother wavelet is scaled to a certain length, and compared to the signal, moving the wavelet along the signal to result in a continuous comparison. The mother wavelet is then scaled again and the process is repeated until all desired scales have been used.

The wavelet coefficients are then used to establish the frequency content within each band. Two variants exist to measure the content; power and amplitude. The only difference being that the amplitude uses the absolute of the wavelet coefficients, while the power uses the square of the absolute values. The two measures are more or less equivalent, however power tends to put over-emphasis on short-high amplitude bursts which is characteristic for eye movement artifacts and other sources of noise. For this reason, the amplitude was chosen as the measure of frequency content for this thesis (Papers I, II and III).

The average frequency content within each frequency band is calculated, and the values are termed as the absolute EEG indices. However, due to the high inter-individual variability introduced due to factors such as scalp thickness, electrode-gel conductance etc., the relative EEG indices have also been utilized as a feature in EEG research[105]. Relative indices are expressed as the percentual contribution of each band to the total EEG amplitude, calculated by dividing each absolute EEG index by the sum of all absolute indices. Relative EEG indices accounts for more inter-individual differences and is shown to show a higher correlation with brain perfusion as measured using positron emission topography[106]. However, relative EEG indices suffer from interactions between frequency bands since it is a relative measure. Thus, if one frequency band increases in amplitude, other frequency bands must decrease accordingly. This can complicate studies which compare conditions, such as drug studies and many pain studies. When several frequency bands are

changing at the same time, this can make an increase in a frequency band appear as a decrease simply because the other bands are counterbalancing these changes[79]. Both measures provide complimentary information, and neither is more correct than the other[106,107]. The measures are best used in unison, where full advantage can be taken of the information (Paper I). This is especially true for studies which compare EEG under different conditions for the same subject. Paper II and III only used the relative spectral content since it is less susceptible to inter-individual differences.

7.2. Functional connectivity

Another way to assess EEG is through functional connectivity which has been used extensively with structural and functional magnetic resonance imaging[108,109]. The common spectral analysis will assess EEG amplitude in different frequency bands, and is as such a measure of brain rhythmicity[110]. Functional connectivity tries to assess the flow of information between electrodes by considering the brain as a complex network of inter-connected nodes (electrodes)[108]. Many different methods exist for assessing functional connectivity, but most rely on the phase differences between electrodes to assess which electrodes are communicating and which direction information is flowing as exemplified on figure 8[111,112]. This is typically achieved by calculating the phase-relationships between signals. Electrodes with similar phase-relationships are considered to be exchanging information[111–113].

The methods have been used to show decreased cortical connectivity after remifentanil[110], distinguishing between levels of consciousness in comatose patients[114] and detection of decreased information flow in diabetes patients[115].

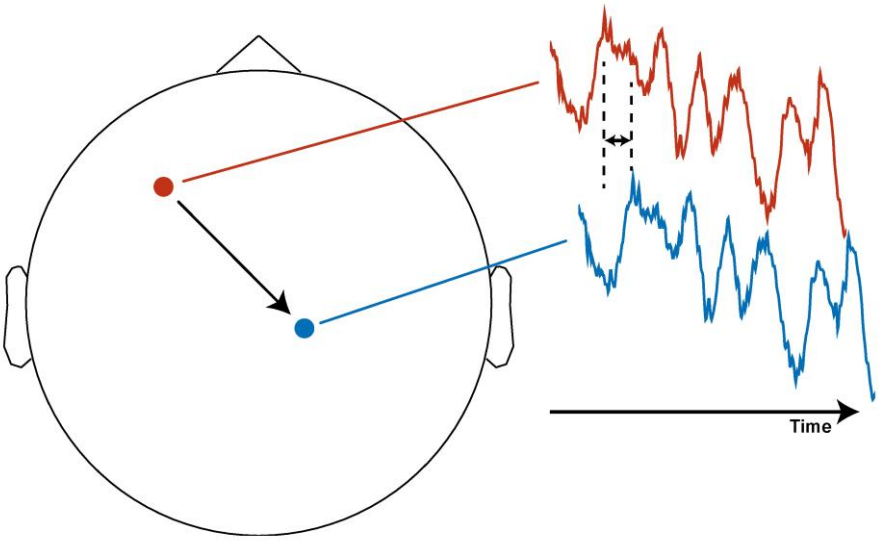


Figure 8: Illustration the principle behind functional connectivity. The signal from the receiving electrode (blue) is phase-shifted compared to the signal from the driving electrode (red). This indicates a strong connection between the two electrodes. The driving and receiving electrode can be distinguished by which signal occurs earlier if directionality is investigated.

This PhD thesis included functional connectivity in paper III to attempt assessment using a slightly more advanced parameter. The phase-lag index (PLI) is quite resistant to volume conduction since it discards most phase-relations between electrodes with no differences in phase (likely caused by volume conduction) and was therefore chosen for analysis[111]. Furthermore, PLI has already been shown to be reliable[108].

Chapter 8. Machine learning

Machine learning derives from pattern recognition and is a common term for methods which learn from and make predictions based on learning data. Machine learning methods work by constructing a model based on a set of training data belonging to two or more discrete groups, which can subsequently be used for predicting outcomes for future samples. It enables simultaneous assessment of multiple variables on an *individualized* basis. This is in contrast to conventional statistical methods, which often analyze single variables between *groups* of patients and rely on *a priori* assumptions.

Within machine learning there are three main branches of methods; supervised, non-supervised and reinforced learning;

- Supervised learning methods require information regarding the groups in order to construct a model optimized for a specific problem. Suitable for problems where a general rule is needed to determine an outcome based on a given data sample
- Non-supervised learning does not require any information regarding which samples belong to each group. Non-supervised learning will instead try to find structure in the training data and is thus able to find hidden patterns in data.
- Reinforced learning takes place when the model must perform a certain task and repeatedly learn to do the task better based on feedback it receives. An example of this could be a computer learning to play a computer game against an opponent.

Since the object of this project is to differentiate between two pre-defined groups, supervised learning was chosen as the appropriate method. Several different methods exist within supervised learning, among others; linear discriminant analysis, artificial neural networks, support vector machine and Bayesian classifiers.

8.1. Feature selection

Until little more than a decade ago, few areas of research investigated more than 40 features[116]. This has changed dramatically in recent years, where some studies investigate numbers of features in the range of tens of thousands[117]. Now all possibly relevant features are collected and subsequently subjected to a statistical feature selection process to establish which features are relevant to the classification problem. Feature selection is the process of selecting a subset of relevant features from a larger pool of variables. This is done for several reasons, among others;

- Making results easier to interpret
- Decrease training time for the machine learning models
- Increasing generalization by reducing over-fitting

The main goal of feature selection is to identify and remove features which are either redundant or irrelevant as they can be removed without significant loss of information[118].

Feature selection methods can be either *wrappers* or *filters*[117]. *Wrappers* are specific to the chosen machine learning classifier and work by searching through all features using the accuracy of this classifier. This however can be very expensive computationally (especially with large numbers of features), and might result in solutions which are overly specific to the classifier[117]. *Filter* methods work independently of the classifier and instead define a measure of *relevance* to approximate the value of a feature within the current classification problem[117]. This makes feature selection possible using only few assumptions. *Filter* methods are relatively computationally inexpensive while generally being less susceptible to over-fitting than *wrappers*[117]. As such *filter* methods were chosen as the feature selection approach for this study.

Filter methods are all defined primarily by their measure of relevance which tries to describe the utility of including a given feature in the reduced feature set [117]. However, there are countless ways to define relevance, and as a result a large array of different feature selection methods exist[117]. The question then becomes which method to choose over the other available methods. A recent study reviewed 17 methods for feature selection under three criteria[117]:

- Inclusion of a conditional redundancy term
- Balancing the relevance and redundancy terms
- Use of low dimensional approximation, important for small sample sizes

Only three out of 17 methods were found to fulfill all criteria; joint mutual information (JMI), mutual information maximization and conditional mutual information maximization[117]. JMI was recommended as giving the best tradeoff between accuracy and stability and thus was selected for this study[119].

8.2. Support vector machine

Support vector machines (SVM) were first introduced in 1963 by Vapnik and Lerner seeking a method to find the optimal (defined by the largest margin) hyperplane to separate two distinct classes of data[120]. The basic principle is explained in figure 9. Here it is illustrated that the data can be separated by a number of hyperplanes. However, to find the hyperplane with the largest margin, the SVM identifies points closest to the border (support vectors), and uses them in the calculation as shown in figure 9.

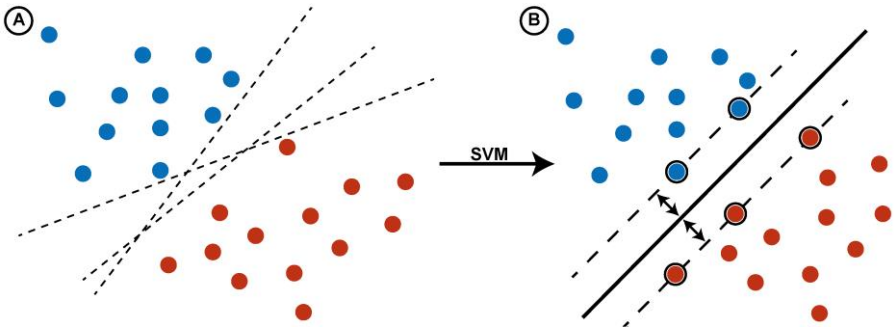


Figure 9: Illustration of the basic working principle of the SVM to separate two classes of data. A) The classes can be separated by a number of hyperplanes. B) The hyperplane with the largest margin (arrow) is shown with the support vectors used to calculate it

8.2.1. Soft margin SVM

The first implementation of the SVM assumed that data would be perfectly separable[120]. However, real world data and especially human biometric data as used in this study is rarely perfectly separable and will include a certain amount of overlap between classes. To accommodate this, the soft-margin SVM was proposed which allowed for misclassifications and which has since been the most common implementation[121]. The soft-margin SVM assigns a penalty for misclassified samples. The optimization of the hyperplane then becomes a compromise between a large margin and a small error penalty. The cost variable *C* was introduced to control the tradeoff between margin and error penalty and also serves as a regularization parameter for the SVM, which can be adjusted to increase generalization of the classifier[122]. This type of classifier has now widely replaced the original classifier, and is also used in papers I, II and III.

8.2.2. Kernel functions

The original SVM was developed as a linear classifier[120]. However, not all real-life problems will present with data that is linearly separable. To allow for non-linear classification, kernels were introduced[123]. This is an interesting implementation because the hyperplane is still linear, but instead the data is transformed into a high-dimensional feature-space where a linear hyperplane can separate the data, even though the separation might be non-linear in the original feature-space[123]. Figure 10 shows how the transformation into a different features space can make a non-linearly separable dataset become linearly separable.

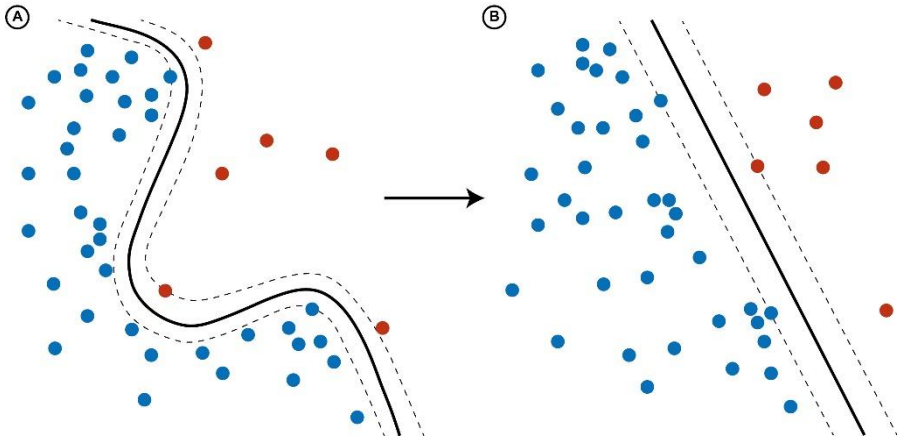


Figure 10: Transformation of a feature-set into a different feature space enables linear separation using a support vector machine. A) The features in the normal feature-space, where the decision boundary appears non-linear. B) The features in the transformed feature-space where the decision boundary is linear.

However, kernel functions can present with the problem of over-fitting and subsequent loss of generalization[124]. This is unfortunate since the results of the SVM are pointless unless the results are generalizable to the general population. The problem is demonstrated in figure 11, where an overly complicated decision hyperplane makes the performance poor when new data is introduced. For this reason it is recommended to use caution when using non-linear kernels to avoid over-fitting, as many times the generalization becomes inferior compared to SVMs with linear kernels[124]. Therefore, this project utilized a linear kernel to avoid over-fitting (Paper II and III).

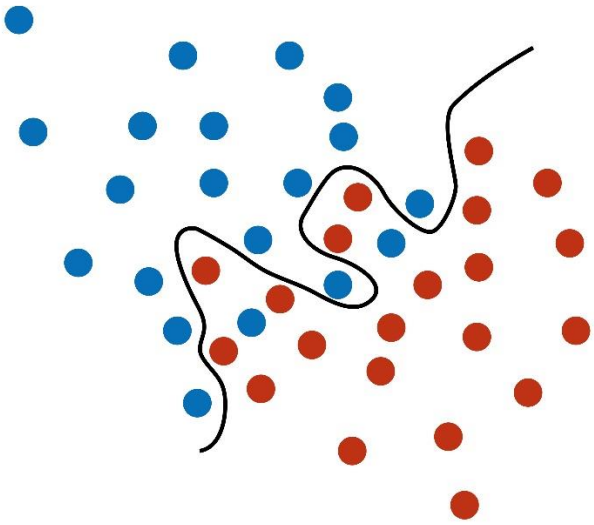


Figure 11: Over-fitting of the support vector machine using an overly complicated decision rule which will provide poor generalization and sub-optimal results when applied to new data.

8.2.3. Cross-validation

The idea behind machine learning is to train a model based on a set of training data which can then correctly classify new samples of the same type. This poses challenges for studies trying to validate a classifier when the number of samples is limited. Firstly, the classifier's ability to generalize new data also depends on the number of samples available for training, meaning that as many samples as possible should be included for training. Secondly, samples included for training cannot be used simultaneously for validating the classifier, as this can result in over-fitting to the data and subsequent poor ability to generalize to new data. The optimal solution is to build the model using training data from one study, and subsequently testing the model using completely separate data. However, this poses requires high numbers of samples which is not always possible or ethical to include in clinical studies.

Cross-validation attempts to handle this problem by dividing the data into training and test subsets which can be validated without possibility for over-fitting. This is repeated multiple times with different training and test subsets, and the accuracy of the classifier is finally calculated as the average accuracy for all iterations[125]. This has the advantage of maximizing the number of samples that can be used for both training and testing. Several types of cross-validation exist, most common among them are:

Leave-one-out: This type of cross-validation uses a single sample for testing, while the rest of the dataset is left for training. This has the advantage of maximizing the amount of samples for training in each iteration while still using all samples for testing in one iteration. This makes it well suited for small datasets, where the number of training samples is crucial. Furthermore, the method is exhaustive, meaning that all possible combinations of training and datasets are included and thus the result will always be the same. However, the method requires additional computer processing power due to the high number of iterations when using this method.

K-fold cross-validation: This version of cross-validation divides the data into k equally-sized subsets. Each iteration is then made by using one subset for testing while the other subsets are used for training the classifier. A common value for k is 10, but it can be set to any value of 2 or above. When $k = N$, it is equivalent to the leave-one-out method. K-fold cross-validation generally uses less computing power due to the less number of classifications (if $k < N$), but is also non-exhaustive, meaning that results will differ depending on how the data is split.

Since the number of samples in this study was limited and the amount of computer processing power was not, leave-one-out cross-validation was chosen as the preferred method to avoid over-fitting (Paper II and III).

8.2.4. Dealing with imbalanced datasets

Part of the calculation for the soft-margin SVM (current golden standard) is to assign a penalty for incorrectly classified samples[121]. However, when datasets are overlapping significantly and group sizes are uneven it can result in inaccurate results[126–128]. More specifically, the classifier will become biased towards the majority class. This phenomenon is caused by the sheer number of overlapping samples from the majority class and will cause a greater penalty for the decision rule than the penalty for most or all of the samples from the minority class. In this case it is more optimal (incurring less penalty) for the classifier to simply classify all or most samples as belonging to the majority class[128]. This will result in higher accuracies than a classifier which will pay equal attention to both classes, but the heightened accuracy is essentially meaningless for prediction. The problem with imbalanced datasets is quite common within genetic profiling and medical diagnosis[126] and usually it is the few samples minority class which holds the greatest importance for correct classification[129]. Figure 12 illustrates the problem for two classes with skewed group sizes. A typical SVM would strive for the solution shown in figure 12A because it yields fewer overall misclassified samples. However, the classifier would simply classify all samples as belonging to the same class, making the solution useless. A solution more similar to the one shown in

figure 12B would be more desirable since it correctly classifies the minority class, despite the lower overall accuracy.

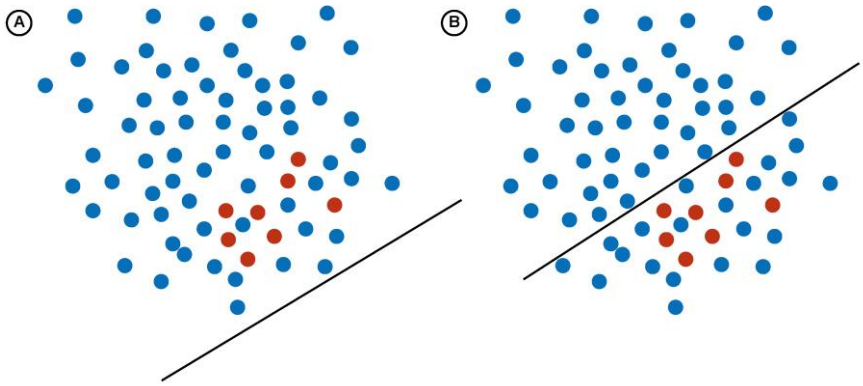


Figure 12: Skewed group sizes in overlapping data can have adverse effects on a machine learning classifiers ability to properly separate data. A) The more useful classification, which will have a lower overall accuracy due to the large number of misclassified samples in the majority class. B) The optimal solution for minimizing misclassified samples, but useless in practical terms since it always returns the same result.

This problem, which is universal in machine learning, and not just restricted to support vector machines, has been addressed in two different ways:

1. Balancing the original dataset
2. Adapting the classifier to the dataset

Balancing the dataset involves either under-sampling the majority class or oversampling the minority class until the classes are balanced[126,128]. Both methods have drawbacks, under-sampling being reported to perform sub-optimally[126] for the SVM, but in other cases reported as performing successfully when combined with specialized data elimination techniques[128]. Although it must be noted that information is removed using this method. Over-sampling on contrary does not remove information, but has instead been reported to shift bias towards the minority class[128].

Adapting the classifier completely depends on the classifier but is most commonly accomplished by assigning different costs for misclassifying the different samples[127,128]. Studies show that SVMs in particular have the ability to deal with imbalanced dataset without introducing noise, however there is a chance for the resulting classifier to over-fit the data[128]. This was implemented in paper III due to the slight imbalance in group sizes.

Chapter 9. Results

The key results from the paper I, II and III are found in this chapter. More detailed results can be found in the papers.

9.1. Key results Study I

- Recording of EEG (62 electrodes) during rest and cold pain was performed on healthy volunteers (N = 39) on two separate days and reliability between days were investigated.
- Spectral indices of EEG during cold pain and resting is reliable with coefficients of variation at 10% or below.
- Tonic cold pain induces widespread changes in the EEG over all frequency bands.
- The EEG activity in the theta, beta3 and gamma bands were correlated to the average pain intensity.
- Analysis of the EEG dynamics revealed that only theta was dynamically reflecting the pain response and therefore the more directly related to the perceived pain.

9.2. Key results Study II

- Healthy volunteers (N = 39) were included in a placebo-controlled morphine study.
- EEG was recorded (62 channels) during rest and cold pain before and 60 minutes after morphine administration where analgesic response was determined based on the reduction in pain ratings during cold pain.
- Tonic painful EEG combined with support vector machine classification enabled reliable classification of opioid responders at a rate of 72% in healthy volunteers.
- Conventional statistical methods could not differentiate between responders and non-responders.
- Resting state EEG did not provide equally high classification accuracy.

9.3. Key results study III

- Eighty-one patients undergoing total hip replacement surgery completed a clinical observational study.
- Prior to surgery QST (heat, pressure pain), clinical parameters and EEG during rest and cold pain was recorded.
- Patients were stratified into responders (N = 51) and non-responders (N = 30) to oxycodone and piritramide based on their pain ratings the first 24 hours after surgery.
- Conventional analysis showed that there was connection between pre-surgical pain state (chronic pain grade) and response to postoperative analgesic opioid treatment.
- Support vector machine classification on tonic painful EEG predicted response to postoperative opioid pain treatment in 65% of cases in patients undergoing full hip replacement.
- Resting EEG proved inferior to tonic painful EEG for prediction of analgesic effect.

Chapter 10. Discussion

This PhD thesis investigated EEG as a biomarker for personalized pain medicine in the following steps:

- To investigate reliability and dynamics of the EEG during rest and cold pain in healthy volunteers to assess robustness of the biomarker (Paper I).
- Investigate if prediction of morphine analgesia is possible in healthy volunteers (Paper II).
- Predict postoperative opioid analgesia in patients undergoing total hip replacement surgery (Paper III).

The first part of the discussion focuses on the methodological considerations for the three studies while the rest focuses on the findings and challenges regarding EEG as a biomarker and the current state of the art within personalized medicine for opioid analgesia.

10.1. Methodological considerations

Papers I, II and III investigated the delta, theta alpha and beta frequency bands while the gamma band was only included in papers I and II. The delta band was shown to increase during tonic pain in paper I which is in line with earlier studies[130]. Furthermore, it was the primarily selected feature for classification in paper II as well as paper III, demonstrating the relevance of this low-frequency band. Alpha and beta were both also part of the classification in paper II and are both important for the perception of pain[82,84,131] and opioids analgesic effects[44]. Furthermore, in paper I theta band activity seemed to be connected closer to perceived pain than any other frequency band.
Recently, large parts of evoked brain potentials previously thought to be directly connected to the experience of pain has been shown instead to be closely related to the saliency of the stimulus[31]. However, newer and more advanced analyses have revealed features in the gamma band to be more specifically associated the perception of pain[132,133]. This makes the gamma band interesting for pain research and was therefore included in paper I to study the reliability of the frequency bands as well that the connection to tonic pain. However, even though the gamma band was shown to be reliable and the average gamma activity correlated to the pain rating the gamma band activity was not dynamic during the cold pressor test. Since the pain ratings during the cold pressor test are highly dynamic, it is unlikely the gamma band was directly reflecting the painful experience. This contrasts other findings, but could be due to the nature of the stimulus which in previous studies

have focused on short, transient stimulations such as delivered by lasers compared to the tonic cold pain in paper I. When used for prediction in paper II, no gamma band features were selected for prediction of analgesic effect, indicating that the gamma band is of less importance for prediction of opioid analgesia. Due to the findings in papers I and II, and previous studies which suggest electromyographic artifacts are introduced into the EEG during the cold pressor test[81], the gamma band was not included in paper III, so as to avoid influence from the gamma on other frequency bands (due to the relative spectral activity).

Papers II and III attempted to predict opioid analgesia based on both resting EEG and EEG during tonic cold pain. Both studies showed that the EEG during cold pain was useful for prediction, while resting EEG was not. Paper I also revealed that resting EEG has slightly lower reliability than EEG during cold pain, possibly due to suppression of conscious and non-conscious processes during pain. Paper I additionally revealed unexpected drifting of the resting EEG during the 2 minute recording which could indicate drowsiness in the subject. Overall, it seems the painful stimuli bring out activity related to how the brain processes pain, which in turn is associated with the response to a given opioid. These findings indicate that EEG during pain is more suitable for prediction. However, only one stimulus (cold pain) was investigated so future studies could aim to investigate if other stimuli would yield better results.

Papers I, II and III all utilized the cold pressor test for the evocation of tonic pain. The test was as standardized as possible, however some discrepancies in methodology still remained. The water temperature in paper I and II was controlled with a specialized equipment, keeping the temperature at 2˚C while constantly circulating the water. This type of equipment is preferable for cold pressor studies, where the temperature of the water and lack of circulation can reduce reliability of the pain model[57]. However, for paper III this equipment was not available on-site and instead relied on cold water from the tap in a thermal container. To account for the added uncertainties from this approach, temperature of the water was measured, and was found to be consistent between groups. The primary confounder due to the cold pressor equipment in paper III seems to be the temperature (around 8˚C) which is comparatively higher than the 2˚C used in papers I and II, but also the normal range of the cold pressor test (1˚C to 7˚C) [57]. This also could contribute to explain why the overall accuracy in paper III was lower than in paper II. Comparison of resting EEG and EEG during cold pain revealed that the painful stimulation seemed to bring out the predictive features in the EEG. It therefore stands to reason that the lower stimulation intensity in paper III could fail to accentuate the predictive features sufficiently, thus resulting in a lower accuracy. This is supported by the comparison between resting EEG and EEG during cold pain revealed widespread differences in paper I, while the same differences were much lesser pronounced in paper III. This points to the need for an intense painful stimulus to properly individual differences in brain function related to opioid analgesic efficacy.

Paper II attempted to predict the analgesic effect of morphine on an experimental pain model (the cold pressor test) while paper III investigated the analgesia attained after total hip replacement surgery. Experimental pain models are advantageous since they are standardized with regards to the stimulus (nature, stimulation site, stimulus frequency and intensity)[134]. It also allows to exclude known confounding factors in clinical studies such as related symptoms, anxiety, immobility and pre-existing medication[55]. Paper II therefore avoided these confounders, and found an algorithm which could predict morphine analgesia in 72% of cases, indicating that the EEG can be utilized for prediction. However, personalized medicine needs to be able to deal with the real-life confounders which might impede classification accuracy. Therefore, paper III investigated patients undergoing total hip replacement surgery in order to predict their post-operative analgesics response using pre-operative EEG. The clinical aspect introduced several confounders into the study, firstly patients were already suffering from chronic pain which can affect the EEG[135,136]. Additionally, the patients were already using medication to treat their pain condition, which also has a known effect on the EEG[39]. Lastly, there is a source of variability introduced by the surgery in itself. Though all patients underwent the same surgical procedure, variations in complexity and duration of surgery will always vary, which could affect pain post-operatively. All these factors likely contribute to the relatively lower classification accuracy in paper III. More knowledge is needed about the confounders and their interaction with the EEG in relation to features which might predict analgesic efficacy. As more studies emerge on the subject, new algorithms will be able to account for these confounders in a more satisfying way.

Papers I and II analyzed the EEG by investigating the spectral content in predefined frequency bands. Paper III also included functional connectivity in the analysis to assess the communication between electrodes. Functional connectivity is a rapidly expanding field within EEG research, and is being heralded as the next step in analysis methods as it investigates the interactions between functional networks in the EEG activity[109]. More detailed information can be extracted by extending analysis from simple rhythmicity of individual electrodes to assessment of the flow of information between electrodes as well as direction[111]. Despite this, the feature reduction method did not result in selection of features from the functional connectivity features, but instead only selected a single feature from the spectral indices. Thus inclusion of these more advances features did not end up improving classification accuracy in paper III. Still, future studies should continue to investigate this method as it has proved promising in many studies[110,115,137]. Furthermore, functional connectivity parameters can be analyzed using graph theory, to describe overall characteristics of the network, which could be an interesting addition to the analysis[138,139]. Another interesting method could be source analysis to properly investigate the activity in specific brain centers related to analgesic response[140]. Such an analysis would not only help in developing the

personalized medicine algorithm, but also assist in understanding the underlying cause for heterogeneity in opioid response.

10.2. EEG as a biomarker for opioid response

Biomarkers are characteristics that can be objectively measured as an indicator for normal or pathological processes as well as pharmacological response to treatment[15]. Identification of opioid response phenotypes aim to identify sub-groups of patients, and tailor treatment for each sub-type[15]. Reliability of the biomarker is equally important as accuracy within personalized medicine but is unfortunately an area which is often overlooked within many areas of science[141,142]. EEG during cold pain appears more suitable for prediction of analgesic treatment outcome both in experimental pain models and in post-operative pain due to the higher prediction accuracy (Paper II and III). Furthermore, the spectral EEG content cold pain is reliable between days (Paper I) and provides similar results using the machine learning algorithm between days, both in terms of classification performance and in terms of the channels selected (Paper II).

EEG assesses the neuronal activity at the scalp and has a relatively poor spatial accuracy compared to imaging methods[143]. Neuroimaging has been suggested as an important approach on the way to personalizing pain treatment[18]. However, imaging methods would introduce substantial economic costs and logistical challenges, making it problematic to include as part of a personalized medicine approach. Though the spatial resolution limits the EEG in identifying the exact brain regions involved in a response, it is a low-cost method that could easily be brought to a patient bed-side. Furthermore, it has previously been shown to predict pain experience within individuals[20], proving its utility for investigations at the individual level.

Advances in the technical field is further improving the utility of EEG by moving away from the practice of establishing connection between electrode and scalp by electrode gel[144]. This practice requires scalp preparation and can introduce several problems, such as 1) loss of connection due to drying of electrode gel, 2) short-circuiting of electrodes due to excessive use of gel, 3) pain/discomfort in patients as scalp preparation usually involves abrasion of the outer skin layer[145]. Furthermore, the main time in EEG recordings is consumed in these preparations[145]. However, in recent years dry electrodes have been developed with no requirement for electrode gel or skin preparation[145,146]. Such equipment would greatly reduce mounting times, further making EEG more attractive for fast, clinical measurements.

10.3. Response stratification

When developing methods for personalized medicine it is important to have a gold standard for response in order to assess accuracy of the method and separate the opioid response from other factors that might influence analgesic effect. The most common threshold for clinical response to analgesic treatment is reduction of pain intensity by 30%. For QST models this threshold can be problematic as analgesic effect is generally lower[55]. This was also the case in paper II where the application of the traditional threshold for analgesic response to a clinically relevant dose (30mg) of morphine would have resulted in just one responder out of 39 subjects. Due to these limitations, the standard 30% threshold was not applied and a threshold was instead set at 5%. Since there is not definition of response for experimental pain models this was mostly based on methodological concerns to achieve balance between response groups, which is beneficial for machine learning methods when analyzing small data samples. Paper III investigated postoperative pain, for which the 30% threshold is not valid since there is no baseline pain measurement to compute pain reduction from. Therefore a response score based on several clinical pain measures which was employed in order to account for fact that both postoperative pain ratings and analgesic consumption will vary due to the postoperative patient controlled anesthesia (PCA)[147]. Ethical issues require that PCA is made available to patients after surgery in order to avoid insufficient analgesia[147]. This problem is one of the primary focus points for analgesic research, as it affects many analgesic studies[148,149]. A score integrating both analgesic consumption and perceived pain has been attempted[150], but has failed to achieve widespread acceptance and as such validity of the method remains unknown[16]. Thus, a score based on pain ratings was developed to assess the level of response for each patient, resulting in a response rate of 65% which is in line with this type of procedure. The lack of opioid consumption as a factor in the response score could be a confounder in paper III, however since stratification based on opioid consumption did not result in higher accuracies it is likely not a major confounder. However, methods for establishing robust criteria for analgesic response are severely lacking within not just postoperative analgesia, but also experimental pain models and should be of continued focus within analgesic research.

Paper III revealed a connection between preoperative pain severity and efficacy of postoperative pain management. Hence, higher severity of chronic pain grade, MPSS and higher levels of perceived pain during hip rotation was observed in non-responders to postoperative opioid analgesia. This is in line with previous literature which indicates that preoperative pain strongly predicts acute postoperative pain[2,151,152]. However, since the response score in paper III was determined by the postoperative pain ratings, it will by extension also be related to the preoperative pain condition. Although this is an important factor in postoperative pain treatment, it also complicates identification of correct algorithm for prediction of analgesia, since the preoperative pain condition affects both the EEG as well as the

postoperative pain experience. This makes it hard to discern if the features found in the EEG represents the preoperative pain condition or a specific phenotype for opioid response. There are indications that the findings could be connected to an opioid response phenotype, since the discriminative features found in paper II (where subjects had no prior pain) where somewhat similar to the patient population in paper III. However, it is not possible to establish if this is the case in paper III, but would instead require a clinical study on a patient population with no chronic pain condition.

Paper I and II investigated morphine while paper III investigated a combination treatment of oxycodone and piritramide. Different opioids may have different effects in the individual patients[34], and even equi-potent opioid doses of different opioids can have widely different effects on the EEG activity[44]. Therefore, results using different opioids cannot be directly compared. Furthermore, paper III is further complicated by the inclusion of two opioids for analgesic treatment, as this introduces several sub-groups (responders to both opioids, non-responders to both and responders to just one opioid) into the analysis, which are impossible uncover in the analysis. However as III was an observational study with no possibility for altering analgesic treatment this could not be avoided. Since this could act as a major confound in the study, future studies should try and perform clinical investigations using just one opioid at a time, to avoid similar issues.

10.4. Prediction of analgesic efficacy

Prediction of analgesic effect holds great potential to improve pain treatment by individualizing treatment for each patient[18]. Beyond the increased suffering, inadequate pain treatment after surgery increases the chance of developing chronic pain[153]. Furthermore, opioid use is associated with serious safety concerns regarding abuse and addiction which is resulting in an increasing number of deaths[154]. Personalized treatment is therefore equally about identifying the right treatment as well as sparing patients from risks associated with inadequate treatment[15,18].

To date studies have focused on the use of QST models for prediction of analgesic effect and some results have been promising[16]. However, no single QST method has so far emerged as a golden standard and results have been variable[16]. Paper III investigated QST measures (heat, cold and pressure pain) performed before surgery for predictive capability, but found no connection to the post-operative analgesic effect. This could be connected to central sensitization due to the pre-existing painful condition, which opioids have limited efficacy against[155].

Due to the pre-existing chronic pain condition patients in paper III were already on analgesic treatment prior to surgery. It is well know that analgesics affect the EEG[39] and therefore the preoperative medicine could have interfered with EEG recordings. However, as preoperative analgesic usage was relatively low the effect were likely minor. Even so, avoiding preoperative analgesic usage if possible is desirable for future studies.

Chronic pain is known to be associated with impaired descending inhibition (assessed by CPM paradigms), an effect which could possibly imply that the system is already maximally engaged[64,66]. Lowered CPM response is also associated with opioid consumption after surgery[156] and development of chronic pain[157,158]. This notion is supported by studies investigating the descending inhibition after the chronic pain has been relieved shows a return of function[65]. Paper III observed impaired descending inhibition in the osteoarthritis patients, meaning that their ability to regulate pain was compromised due to their pain condition. However, this was not a predictor of response to postoperative analgesia, but rather an overall condition for all patients.

Conventional group-wise statistics have limitation in use for personalized medicine since they focus on detecting differences between groups rather than assessing the individual subject which is of major clinical interest, making machine learning more suitable[159]. Furthermore, machine learning method are able to assess multiple features simultaneously which is important for high-dimensional data such as EEG. This has been seen before in studies on analgesic effect, which found the effect reflected in the EEG based on a SVM classifier which assessed all EEG frequency bands simultaneously[21]. The previous findings are further confirmed in paper II and III which both showed that traditional group-wise statistics were unable to assess the differences between responders and non-responders to opioid analgesia, whereas machine learning methods enabled prediction of analgesic effect.

The over-all accuracies for the predictions were 72% (paper II) and 65% (paper III). Positive predictive values were 70/75% (paper II) and 75% (paper III). Thus, if patients had been treated according to the machine learning results based on EEG prior to treatment, response rates would have been higher. In a personalized medicine scenario the remaining patients would then be treated with an alternative treatment although there was a higher degree of false negatives in paper III. However, in a personalized medicine scenario using EEG and machine learning to switch patients to an alternative treatment, the patients falsely classified as non-responders could still respond to the alternative treatment. Unfortunately the response rate of an alternative treatment could not be determined due to study design limitations. Future work could attempt such studies, but only after consensus has been made on a model for predicting analgesic opioid response.

The features selected in paper II were predominantly frontal and central electrodes of the delta band while only one frontal electrode from the delta band was selected in paper III. There is a level of similarity between the two studies in that they both selected features from frontal delta electrodes. Some differences are present like the number of selected electrodes, but given the differences between the studies (e.g. population, experimental pain vs. clinical pain, etc.) and the fact that the drugs were not consistent makes differences in the selected features very likely. This could potentially be an advantage as the number of needed electrodes would decrease which would reduce time required for making a measurement.

10.5. Clinical implications

Personalized treatment targeting the individual patient and identifying optimal treatment strategies holds great promise for improving current analgesic treatments and reducing the rising abuse of opioids seen worldwide[15,18]. Most studies investigating prediction of opioid analgesic effect has utilized subjective QST models. However, no single QST model has been discovered which can be recommended for prediction of analgesia and large variations exist application of the QST models[16]. Neuroimaging has previously proved useful with regards to understanding pain processing and could therefore also prove valuable for personalized medicine[18]. However, neuroimaging methods are infeasible logistically and economically in a clinical algorithm. EEG is an objective measurement and has lower costs than neuroimaging methods but can still assess neurological activity in the brain and has been shown in papers II and III to contain information which could help predict response to opioid treatment. Although accuracy is likely still too low for clinical application, this poses an important first step for personalized pain medicine with EEG. Furthermore, the machine learning approach also opens up the possibility for including other features for prediction in the same analysis, such as QST measures, clinical parameters or genetic profiling.

Chapter 11. Conclusions

This study marks the first investigations into objective biomarkers for prediction of opioid analgesia. We demonstrated the robustness of EEG as a biomarker and the ability to separate responders from non-responders in both healthy volunteers and patients scheduled to undergo total hip replacement surgery. The conclusions corresponding to each of the aims I-VI are presented below:

11.1. Aim I

To investigate the reliability of EEG over time to establish if the measurements would be sufficiently robust to use in personalized medicine (Paper I).

For any use of EEG in personalized medicine, it is imperative that the derived EEG parameters remain relatively stable over time. This was achieved by investigating the reliability of EEG parameters in paper I. Overall EEG parameters had good reliability and therefore seems appropriate for use in a personalized medicine algorithm. This was also underlined in paper II where the results were repeated when using EEG recorded on two separate days from the same subjects.

11.2. Aim II

Analyze dynamics of the EEG and pain ratings during cold pain to gain further insight into pain processing during cold pain (Paper I).

The analysis of the EEG revealed that the overall EEG activity was correlated to overall pain intensity from both the theta, beta3 and gamma bands. However, the dynamic analysis revealed that only the theta band was dynamic, and therefore likely to be more closely linked to the pain ratings, which during the cold pain are highly dynamic. Thus, dynamic analysis allowed for more specific information on pain processing during cold pain, and could be utilized in the future within pain research.

11.3. Aim III

To establish if the EEG during rest or tonic cold pain is more suited for prediction of opioid analgesia (Paper I, II and III).

EEG during tonic cold pain showed to be consistently superior to resting EEG, both in reliability and ability to discriminate between responders and non-responders. Thus tonic pain models seem to be more appropriate for future work using EEG in personalized medicine.

11.4. Aim IV

To compare a machine learning analysis approach to conventional group-based statistics in order to investigate if more sensitive prediction can be achieved (Paper II and III).

A relatively new analysis method within clinical research was introduced based on machine learning and advanced feature selection in order to assess the individual subject without *a priori* assumptions about the data. This method proved to be superior to group-based statistics for separating responders to opioid analgesia from non-responders both in paper II and III.

11.5. Aim V

To explore if combination of several features in machine learning enables separation of responders and non-responders to morphine in healthy volunteers (Paper II).

Machine learning enabled prediction of morphine analgesia based on EEG during cold pain recorded prior to drug administration. This prediction could be repeated using EEG during cold pain recorded on another day. The features selected for prediction suggest primarily the delta band contains information regarding the response to morphine.

11.6. Aim VI

To investigate if machine learning can be used to predict the analgesic efficacy in patients with postoperative pain (Paper III).

The machine learning approach enabled prediction of postoperative analgesic effect from oxycodone and piritramide in patients undergoing total hip replacement. The delta band was again selected as the most predictive feature.

Chapter 12. Future perspectives

The search for biomarkers associated with analgesic response is still ongoing and hold great potential for improving treatment of patients. This study has uncovered some interesting findings, but many questions still remain unanswered. In order to move the field forward, some areas of interest should be investigated:

Clinical studies should be conducted to further determine consistency of the results presented in this thesis. Additionally, many of the confounding factors should as well as possible be avoided. This could include investigations into a patient population undergoing surgery much like in this thesis, but with minor alterations. Firstly, treatment should be attempted using only one opioid. This should preferably be a widely used and well-known opioid (such as morphine) to increase the clinical utility of investigations. Secondly, chronic pain introduces variability into the EEG which can interfere with measurements. To combat this problem clinical studies could be performed on patient populations undergoing surgery, but without pain prior to surgery. Therefore we are currently investigating ~60 patients undergoing surgical funnel chest repair. The patients undergoing this surgery are generally healthy, pain-free and opioid naïve males. Since there is no pre-existing pain condition and medication, many of confounders can be avoided and hopefully a clearer signal can be found. This would help in determining which features in the EEG are connected purely to the analgesic response to opioids.

Of course personalized treatment should in the end also target chronic pain populations and therefore have to deal with these issues, and therefore we have conducted a multi-center study (ClinicalTrials.gov identifier: NCT02308306) on chronic pain patients at five centers throughout Europe and included 60 patients. The patients were all opioid naïve at the time of EEG recording, and follow-up was performed after 14 days to determine response to treatment. This study will investigate if any patterns can be found in the EEG and QST to predict analgesic effect in a chronic pain population.

This thesis focused purely on the analgesic effect of opioids as an outcome of treatment. However, adverse effect is also an interesting aspect in opioid treatment and can in some cases be so severe that patients deem the treatment worse that pain and discontinue treatment. This was not investigated in the current thesis in order to limit the scope of investigations, but should be considered for future studies. It is possible that the methods presented in this thesis could be used for this same purpose and would help cast light on a different but important aspect of treatment. When properly investigated results from prediction of analgesic response could be combined with prediction of adverse effects to present clinicians with a more complete view of the consequences of a given treatment.

The methods presented in this PhD thesis are relatively new within clinical research which mostly relies on conventional statistics. The prospective usefulness of these methods are not confined to pain research, but could in theory be used in many applications as a tool for decision support. We have started one such study for grading the severity of hepatic encephalopathy. As the severity of the disease increases, patients brains start to be affected to a degree which could make them a danger e.g. in traffic. However, it is difficult for clinicians to detect when the patient is impaired and special precautions should be taken. EEG has shown some promise in detecting severity of disease, but more accurate detection is needed. Our analysis will incorporate EEG (226 patients with different grades of hepatic encephalography and 137 healthy volunteers) into machine learning to allow for simultaneous assessment of several EEG parameters, and hopefully increasing sensitivity.

Machine learning can also be used for more simple measures than EEG, such as QST or clinical parameters. To illustrate this we analyzed 88 patients undergoing mastectomy for treatment of breast cancer for preoperative risk factors to develop postoperative chronic pain. Prior to surgery QST and clinical parameters were recorded and follow-up was made 6 months after surgery to establish who would go on to develop postoperative chronic pain. We analyzed the two groups and found simple two QST measures to be predict the development of chronic pain with a sensitivity and specificity of around 70 and 80%. These results, which are still under preparation for publishing, could pave the way for a simple test to establish which patients should be take special care of after surgery to prevent chronic pain developing. Furthermore, it shows the potential of machine learning within clinical research with simultaneous assessment of multiple parameters without any *a priori* assumptions.

Literature list

1. Dolin, S. J., Cashman, J. N. & Bland, J. M. Effectiveness of acute postoperative pain management: I. Evidence from published data. *Br. J. Anaesth.* **89,** 409–423 (2002).

2. Sommer, M. *et al.* The prevalence of postoperative pain in a sample of 1490 surgical inpatients. *Eur. J. Anaesthesiol.* **25,** 267–274 (2008).

3. VanDenKerkhof, E. G. *et al.* Impact of perioperative pain intensity, pain qualities, and opioid use on chronic pain after surgery: a prospective cohort study. *Reg Anesth Pain Med* **37,** 19–27 (2012).

4. Kehlet, H., Roumen, R. M., Reinpold, W. & Miserez, M. Invited commentary: Persistent pain after inguinal hernia repair: what do we know and what do we need to know? *Hernia* **17,** 293–7 (2013).

5. Breivik, H., Collett, B., Ventafridda, V., Cohen, R. & Gallacher, D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *Eur. J. Pain* **10,** 287–333 (2006).

6. Smith, B. H. & Torrance, N. in *ABC Pain* (2012).

7. Liu, S. S. & Wu, C. L. The effect of analgesic technique on postoperative patient-reported outcomes including analgesia: A systematic review. *Anesth. Analg.* **105,** 789–808 (2007).

8. Eriksen, J., Jensen, M. K., Sjøgren, P., Ekholm, O. & Rasmussen, N. K. Epidemiology of chronic non-malignant pain in Denmark. *Pain* **106,** 221–228 (2003).

9. Kurita, G. P., Sjøgren, P., Juel, K., Højsted, J. & Ekholm, O. The burden of chronic pain: a cross-sectional survey focussing on diseases, immigration, and opioid use. *Pain* **153,** 2332–8 (2012).

10. Maier, C., Hildebrandt, J., Klinger, R., Henrich-Eberl, C. & Lindena, G.

Morphine responsiveness, efficacy and tolerability in patients with chronic non-tumor associated pain - results of a double-blind placebo-controlled trial (MONTAS). *Pain* **97,** 223–33 (2002).

11. Nielsen, L. M. *et al.* Association Between Human Pain-Related Genotypes and Variability in Opioid Analgesia: An Updated Review. *Pain Pract.* **15,** 580–594 (2015).

12. Tremblay, J. & Hamet, P. Genetics of pain, opioids, and opioid responsiveness. *Metabolism* **59,** S5–S8 (2010).

13. Kalso, E., Edwards, J. E., Moore, R. A. & McQuay, H. J. Opioids in chronic non-cancer pain: systematic review of efficacy and safety. *Pain* **112,** 372–80 (2004).

14. Riley, J. *et al.* No pain relief from morphine? Individual variation in sensitivity to morphine and the need to switch to an alternative opioid in cancer patients. *Support. Care Cancer* **14,** 56–64 (2006).

15. Bruehl, S. *et al.* Personalized medicine and opioid analgesic prescribing for chronic pain: opportunities and challenges. *J. Pain* **14,** 103–13 (2013).

16. Grosen, K., Fischer, I. W. D., Olesen, a E. & Drewes, a M. Can quantitative sensory testing predict responses to analgesic treatment? *Eur. J. Pain* **17,** 1267–80 (2013).

17. Olesen, S. S. *et al.* Quantitative sensory testing predicts pregabalin efficacy in painful chronic pancreatitis. *PLoS One* **8,** e57963 (2013).

18. Ahmedzai, S. H. Personalized medicine--one size fits one: tailoring pain therapy to individuals' needs. *J. Pain Palliat. Care Pharmacother.* **27,** 83–5 (2013).

19. Legrain, V., Iannetti, G. D., Plaghki, L. & Mouraux, A. The pain matrix reloaded: A salience detection system for the body. *Prog. Neurobiol.* **93,** 111–124 (2011).

20. Schulz, E., Zherdin, A., Tiemann, L., Plant, C. & Ploner, M. Decoding an

Individual's Sensitivity to Pain from the Multivariate Analysis of EEG Data. *Cereb. Cortex* **22,** 1118–23 (2011).

21. Graversen, C. *et al.* The analgesic effect of pregabalin in patients with chronic pain is reflected by changes in pharmaco-EEG spectral indices. *Br. J. Clin. Pharmacol.* **73,** 363–72 (2012).

22. Sherrington, C. S. *The integrative action of the nervous system*. (1906).

23. Nikolajsen, L. & Jensen, T. S. in *Textb. Pain* 961– 971 (2006).

24. Zhu, Y. J. & Lu, T. J. A multi-scale view of skin thermal pain: from nociception to pain sensation. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **368,** 521–559 (2010).

25. Fürst, S. Transmitters involved in antinociception in the spinal cord. *Brain Res. Bull.* **48,** 129–141 (1999).

26. D'Mello, R. & Dickenson, A. H. Spinal cord mechanisms of pain. *Br. J. Anaesth.* **101,** 8–16 (2008).

27. Apkarian, a V., Bushnell, M. C., Treede, R.-D. & Zubieta, J.-K. Human brain mechanisms of pain perception and regulation in health and disease. *Eur. J. Pain* **9,** 463–84 (2005).

28. Ingvar, M. Pain and functional imaging. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **354,** 1347–1358 (1999).

29. Rainville, P. Brain mechanisms of pain affect and pain modulation. *Curr. Opin. Neurobiol.* **12,** 195–204 (2002).

30. Tracey, I. & Mantyh, P. W. The Cerebral Signature for Pain Perception and Its Modulation. *Neuron* **55,** 377–391 (2007).

31. Iannetti, G. D., Hughes, N. P., Lee, M. C. & Mouraux, a. Determinants of laser-evoked EEG responses: pain perception or stimulus saliency? *J. Neurophysiol.* **100,** 815–28 (2008).

32. Varrassi, G. *et al.* Pharmacological treatment of chronic pain – the need for CHANGE. *Curr. Med. Res. Opin.* **26,** 1231–1245 (2010).

33. Jadad, A. R. The WHO Analgesic Ladder for Cancer Pain Management. *JAMA* **274,** 1870 (1995).

34. Drewes, A. M. *et al.* Differences between opioids: Pharmacological, experimental, clinical and economical perspectives. *Br. J. Clin. Pharmacol.* **75,** 60–78 (2013).

35. Marvizon, J., Ma, Y.-Y., Charles, A., Walwyn, W. & CJ, E. in *Pharmacol. pain* 87–110 (2010).

36. Trescot, A. M., Datta, S., Lee, M. & Hansen, H. Opioid pharmacology. *Pain Physician* **11,** S133–S153 (2008).

37. Prichep, L. S., John, E. R., Howard, B., Merkin, H. & Hiesiger, E. M. Evaluation of the pain matrix using EEG source localization: a feasibility study. *Pain Med.* 1241–1248 (2011).

38. Saab, C. Visualizing the complex brain dynamics of chronic pain. *J. Neuroimmune Pharmacol.* **8,** 510–517 (2013).

39. Malver, L. P. *et al.* Electroencephalography and analgesics. *Br. J. Clin. Pharmacol.* **77,** 72–95 (2014).

40. Greenwald, M. K. & Roehrs, T. A. Mu-opioid self-administration vs passive administration in heroin abusers produces differential EEG activation. *Neuropsychopharmacology* **30,** 212–21 (2005).

41. Lötsch, J. *et al.* Effects of azapropazone on pain-related brain activity in human subjects. *Br. J. Clin. Pharmacol.* **40,** 545–52 (1995).

42. Bromm, B., Ganzel, R., Herrmann, W. M., Meier, W. & Scharein, E. The analgesic efficacy of flupirtine in comparison to pentazocine and placebo assessed by EEG and subjective pain ratings. *Postgrad. Med. J.* **63 Suppl 3,**

109–12 (1987).

43. Kress, H. G. Clinical update on the pharmacology, efficacy and safety of transdermal buprenorphine. *Eur. J. Pain* **13,** 219–30 (2009).

44. Gram, M. *et al.* A novel approach to pharmaco-EEG for investigating analgesics: assessment of spectral indices in single-sweep evoked brain potentials. *Br. J. Clin. Pharmacol.* **76,** 951–63 (2013).

45. Riley, J. L. & Hastie, B. A. Individual Differences in Opioid Efficacy for Chronic Noncancer Pain. *Clin. J. Pain* **24,** 509–520 (2008).

46. Noble, M. *et al.* Long-term opioid management for chronic noncancer pain. *Cochrane database Syst. Rev.* CD006605 (2010). doi:10.1002/14651858.CD006605.pub2

47. Von Korff, M., Ormel, J., Keefe, F. J. & Dworkin, S. F. Grading the severity of chronic pain. *Pain* **50,** 133–149 (1992).

48. Schmitt, N. & Gerbershagen, H. U. The mainz pain staging system (MPSS) for chronic pain. *Pain* **41,** 484 (1990).

49. Frettlöh, J., Maier, C., Gockel, H. & Hüppe, M. Validität des Mainzer Stadienmodells der Schmerzchronifizierung bei unterschiedlichen Schmerzdiagnosen. *Der Schmerz* **17,** 240–251 (2003).

50. Maier, C. *et al.* The quality of pain management in German hospitals. *Dtsch. Arztebl. Int.* **107,** 607–14 (2010).

51. Rothaug, J., Weiss, T. & Meissner, W. How simple can it get? Measuring pain with NRS items or binary items. *Clin. J. Pain* **29,** 224–32 (2013).

52. Melzack, R. The McGill Pain Questionnaire: Major properties and scoring methods. *Pain* **1,** 277–299 (1975).

53. Price, D. D. Psychological and neural mechanisms of the affective dimension of pain. *Science* **288,** 1769–1772 (2000).

54. Drewes, A. M., Gregersen, H. & Arendt-Nielsen, L. Experimental pain in gastroenterology: a reappraisal of human studies. *Scand. J. Gastroenterol.* **38,** 1115–1130 (2003).

55. Olesen, A., Andresen, T., Staahl, C. & Drewes, A. Human Experimental Pain Models for Assessing the Therapeutic Efficacy of Analgesic Drugs. *Pharmacol. Rev.* **64,** 722–779 (2012).

56. Fowler, C. J., Sitzoglou, K., Ali, Z. & Halonen, P. The conduction velocities of peripheral nerve fibres conveying sensations of warming and cooling. *J. Neurol. Neurosurg. Psychiatry* **51,** 1164–70 (1988).

57. Mitchell, L. A., MacDonald, R. A. R. & Brodie, E. E. Temperature and the cold pressor test. *J. Pain* **5,** 233–7 (2004).

58. Rainville, P., Feine, J. S., Bushnell, M. C. & Duncan, G. H. A psychophysical comparison of sensory and affective responses to four modalities of experimental pain. *Somatosens. Mot. Res.* **9,** 265–77 (1992).

59. Blasco, T. & Bayés, R. Unreliability of the Cold Pressor Test Method in pain studies. *Methods Find. Exp. Clin. Pharmacol.* **10,** 767–72 (1988).

60. Curatolo, M., Petersen-Felix, S. & Arendt-Nielsen, L. Sensory assessment of regional analgesia in humans: a review of methods and applications. *Anesthesiology* **93,** 1517–30 (2000).

61. Nussbaum, E. L. & Downes, L. Reliability of clinical pressure-pain algometric measurements obtained on consecutive days. *Phys. Ther.* **78,** 160–169 (1998).

62. Popescu, A., LeResche, L., Truelove, E. L. & Drangsholt, M. T. Gender differences in pain modulation by diffuse noxious inhibitory controls: A systematic review. *Pain* **150,** 309–318 (2010).

63. Pud, D., Granovsky, Y. & Yarnitsky, D. The methodology of experimentally induced diffuse noxious inhibitory control (DNIC)-like effect in humans. *Pain* **144,** 16–19 (2009).

64. Arendt-Nielsen, L. *et al.* Sensitization in patients with painful knee osteoarthritis. *Pain* **149,** 573–581 (2010).

65. Kosek, E. & Ordeberg, G. Lack of pressure pain modulation by heterotopic noxious conditioning stimulation in patients with painful osteoarthritis before, but not following, surgical pain relief. *Pain* **88,** 69–78 (2000).

66. Lewis, G. N., Rice, D. A. & McNair, P. J. Conditioned Pain Modulation in Populations With Chronic Pain: A Systematic Review and Meta-Analysis. *J. Pain* **13,** 936–944 (2012).

67. Drinkenburg, W. H. I. M. *Essentials and Applications of EEG Research in Preclinical and Clinical Pharmacology*. (2004).

68. Nunez, P. L. & Srinivasan, R. *Electric Fields of the Brain: The Neurophysics of EEG*. (2006).

69. Vespa, P. M., Nenov, V. & Nuwer, M. R. Continuous EEG monitoring in the intensive care unit: early findings and clinical efficacy. *J. Clin. Neurophysiol.* **16,** 1–13 (1999).

70. Bromm, B. & Lorenz, J. Neurophysiological evaluation of pain. *Electroencephalogr. Clin. Neurophysiol.* **107,** 227–253 (1998).

71. Zaslansky, R. *et al.* Pain-evoked potentials: What do they really measure? *Electroencephalogr. Clin. Neurophysiol. - Evoked Potentials* **100,** 384–391 (1996).

72. Dawson, G. A summation technique for detecting small signals in a large irregular background. *J. Physiol.* **115(1),** 2–3 (1951).

73. Luck, S. J. in *Event-Related Potentials A Methods Handb.* 17–32 (2004).

74. Haas, S. *et al.* Cortical evoked potentials in response to rapid balloon distension of the rectum and anal canal. *Neurogastroenterol. Motil.* **26,** 862–873 (2014).

75. Haas, S. *et al.* Abnormal neuronal response to rectal and anal stimuli in patients with idiopathic fecal incontinence. *Neurogastroenterol. Motil.* **27,** 954–962 (2015).

76. Croft, R. J. & Barry, R. J. Removal of ocular artifact from the EEG: a review. *Clin. Neurophysiol.* **30,** 5–19 (2000).

77. Brismar, T. The human EEG--physiological and clinical studies. *Physiol. Behav.* **92,** 141–7 (2007).

78. Oken, B. S. & Chiappa, K. H. Short-term variability in EEG frequency analysis. *Electroencephalogr. Clin. Neurophysiol.* **69,** 191–8 (1988).

79. Gram, M., Graversen, C., Olesen, S. S. & Drewes, A. M. Dynamic spectral indices of the electroencephalogram provide new insights into tonic pain. *Clin. Neurophysiol.* **126,** 763–771 (2014).

80. Jobert, M. *et al.* Guidelines for the recording and evaluation of pharmaco-EEG data in man: the International Pharmaco-EEG Society (IPEG). *Neuropsychobiology* **66,** 201–20 (2012).

81. Dowman, R., Rissacher, D. & Schuckers, S. EEG indices of tonic pain-related activity in the somatosensory cortices. *Clin. Neurophysiol.* **119,** 1201–12 (2008).

82. Nir, R.-R., Sinai, A., Raz, E., Sprecher, E. & Yarnitsky, D. Pain assessment by continuous EEG: association between subjective perception of tonic pain and peak frequency of alpha oscillations during stimulation and at rest. *Brain Res.* **1344,** 77–86 (2010).

83. Schulz, E. *et al.* Prefrontal Gamma Oscillations Encode Tonic Pain in Humans. *Cereb. Cortex* **25,** 4407–4414 (2015).

84. Backonja, M. *et al.* Tonic changes in alpha power during immersion of the hand in cold water. *Electroencephalogr. Clin. Neurophysiol.* **79,** 192–203 (1991).

85. Ferracuti, S., Seri, S., Mattia, D. & Cruccu, G. Quantitative EEG modifications during the Cold Water Pressor Test: hemispheric and hand differences. *Int. J. Psychophysiol.* **17,** 261–8 (1994).

86. Graven-Nielsen, T., Wodehouse, T., Langford, R. M., Arendt-Nielsen, L. & Kidd, B. L. Normalization of widespread hyperesthesia and facilitated spatial summation of deep-tissue pain in knee osteoarthritis patients after knee replacement. *Arthritis Rheum.* **64,** 2907–2916 (2012).

87. Reinert, A., Treede, R. D. & Bromm, B. The pain inhibiting pain effect: an electrophysiological study in humans. *Brain Res.* **862,** 103–110 (2000).

88. Marchetti, P. *et al.* Electroencephalography in Patients With Cirrhosis. *Gastroenterology* **141,** 1680–1689 (2011).

89. Noh, G.-J. *et al.* Electroencephalographic approximate entropy changes in healthy volunteers during remifentanil infusion. *Anesthesiology* **104,** 921–32 (2006).

90. Bruce, E., Bruce, M. & Vennelaganti, S. Sample entropy tracks changes in EEG power spectrum with sleep state and aging. *J. Clin. Neurophysiol.* **26,** 257–266 (2009).

91. Friston, K. J. Functional and Effective Connectivity: A Review. *Brain Connect.* **1,** 13–36 (2011).

92. Dietsch, G. Analyse von elektroenzephalogrammen des Menschen. *Pflügers Arch Ges Physiol* **230,** 106–112 (1932).

93. Tonner, P. H. & Bein, B. Classic electroencephalographic parameters: Median frequency, spectral edge frequency etc. *Best Pract. Res. Clin. Anaesthesiol.* **20,** 147–159 (2006).

94. Durka, P. J. From wavelets to adaptive approximations: time-frequency parametrization of EEG. *Biomed. Eng. Online* **2,** 1 (2003).

95. Samar, V. J., Bopardikar, A., Rao, R. & Swartz, K. Wavelet analysis of neuroelectric waveforms: a conceptual tutorial. *Brain Lang.* **66,** 7–60

(1999).

96. Maltez, J., Hyllienmark, L., Nikulin, V. V & Brismar, T. Time course and variability of power in different frequency bands of EEG during resting conditions. *Neurophysiol. Clin.* **34,** 195–202 (2004).

97. Gasser, T., Bacher, P. & Steinberg, H. Test-retest reliability of spectral parameters of the EEG. *Electroencephalogr. Clin. Neurophysiol.* **60,** 312–319 (1985).

98. Oppenheim, A. V., Schafer, R. W. & Buck, J. R. *Discrete-time signal processing*. (1989).

99. Durka, P. *Matching pursuit and unification in EEG analysis*. (2007).

100. Durka, P. J., Matysiak, A., Montes, E. M., Sosa, P. V. & Blinowska, K. J. Multichannel matching pursuit and EEG inverse solutions. *J. Neurosci. Methods* **148,** 49–59 (2005).

101. Krstulovic, S. & Gribonval, R. Mptk: Matching Pursuit Made Tractable. in *2006 IEEE Int. Conf. Acoust. Speed Signal Process. Proc.* **3,** III–496–III–499 (IEEE, 2006).

102. Mallat, S. & Zhang, Z. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.* **41,** 3397–3415 (1993).

103. Hubbard, B. B. *The World According to Wavelets*. (1996).

104. Akin, M. Comparison of wavelet transform and FFT methods in the analysis of EEG signals. *J. Med. Syst.* **26,** 241–7 (2002).

105. Law, S. K. Thickness and resistivity variations over the upper surface of the human skull. *Brain Topogr.* **6,** 99–109 (1993).

106. Leuchter, A. F., Uijtdehaage, S. H., Cook, I. A., O'Hara, R. & Mandelkern, M. Relationship between brain electrical activity and cortical perfusion in normal subjects. *Psychiatry Res.* **90,** 125–40 (1999).

107. Leuchter, a F. *et al.* Regional differences in brain electrical activity in dementia: use of spectral power and spectral ratio measures. *Electroencephalogr. Clin. Neurophysiol.* **87,** 385–93 (1993).

108. Hardmeier, M. *et al.* Reproducibility of Functional Connectivity and Graph Measures Based on the Phase Lag Index (PLI) and Weighted Phase Lag Index (wPLI) Derived from High Resolution EEG. *PLoS One* **9,** e108648 (2014).

109. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage* **52,** 1059–1069 (2010).

110. Khodayari-Rostamabad, A. *et al.* Disruption of Cortical Connectivity during Remifentanil Administration Is Associated with Cognitive Impairment but Not with Analgesia. *Anesthesiology* **122,** 140–149 (2014).

111. Nolte, G. *et al.* Robustly Estimating the Flow Direction of Information in Complex Physical Systems. *Phys. Rev. Lett.* **100,** 234101 (2008).

112. Varela, F., Lachaux, J. P., Rodriguez, E. & Martinerie, J. The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* **2,** 229–39 (2001).

113. Lachaux, J. P., Rodriguez, E., Martinerie, J. & Varela, F. J. Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* **8,** 194–208 (1999).

114. Lehembre, R. *et al.* Resting-state EEG study of comatose patients: A connectivity and frequency analysis to find differences between vegetative and minimally conscious states. *Funct. Neurol.* **27,** 41–47 (2012).

115. Cooray, G. K., Hyllienmark, L. & Brismar, T. Decreased cortical connectivity and information flow in type 1 diabetes. *Clin. Neurophysiol.* **122,** 1943–50 (2011).

116. Guyon, I. & Elisseeff, A. An Introduction to Variable and Feature Selection 1 Introduction. *J. Mach. Learn. Res.* **3,** 1157–1182 (2003).

117. Brown, G., Pocock, A., Zhao, M.-J. & Luján, M. Conditional Likelihood Maximisation: A Unifying Framework for Information Theoretic Feature Selection. *J. Mach. Learn. Res.* **13,** 27–66 (2012).

118. Bermingham, M. L. *et al.* Application of high-dimensional feature selection: evaluation for genomic prediction in man. *Sci. Rep.* **5,** 10312 (2015).

119. Brown, J. E., Chatterjee, N., Younger, J. & Mackey, S. Towards a physiology-based measure of pain: patterns of human brain activity distinguish painful from non-painful thermal stimulation. *PLoS One* **6,** e24124 (2011).

120. Vapnik, V. & Lerner, A. Pattern recognition using generalized portrait method. *Autom. Remote Control* **24,** 774–780 (1963).

121. Cortes, C. & Vapnik, V. Support-Vector Networks. *Mach. Learn.* **297,** 273–297 (1995).

122. Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F. & Arnaldi, B. A review of classification algorithms for EEG-based brain–computer interfaces. *J. Neural Eng.* **4,** R1–R13 (2007).

123. Boser, B. E., Guyon, I. M. & Vapnik, V. N. A training algorithm for optimal margin classifiers. in *Proc. fifth Annu. Work. Comput. Learn. theory - COLT '92* 144–152 (ACM Press, 1992). doi:10.1145/130385.130401

124. Ivanciuc, O. Applications of support vector machines in chemistry. *Rev. Comput. Chem.* **23,** 291–400 (2007).

125. Seni, G. & Elder, J. F. Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions. *Synth. Lect. Data Min. Knowl. Discov.* **2,** 1–126 (2010).

126. Akbani, R., Kwek, S. & Japkowicz, N. *Applying Support Vector Machines to Imbalanced Datasets*. *Mach. Learn. ECML 2004* **3201,** (Springer Berlin Heidelberg, 2004).

127. Thai-Nghe, N., Gantner, Z. & Schmidt-Thieme, L. Cost-sensitive learning methods for imbalanced data. in *2010 Int. Jt. Conf. Neural Networks* 1–8 (IEEE, 2010). doi:10.1109/IJCNN.2010.5596486

128. Wang, B. X. & Japkowicz, N. Boosting support vector machines for imbalanced data sets. *Knowl. Inf. Syst.* **25,** 1–20 (2009).

129. Sun, Y., Wong, A. K. C. & Kamel, M. S. Classification of Imbalanced Data: a Review. *Int. J. Pattern Recognit. Artif. Intell.* **23,** 687–719 (2009).

130. Chang, P. F., Arendt-Nielsen, L. & Chen, A. C. N. Dynamic changes and spatial correlation of EEG activities during cold pressor test in man. *Brain Res. Bull.* **57,** 667–75 (2002).

131. Nir, R.-R., Sinai, A., Moont, R., Harari, E. & Yarnitsky, D. Tonic pain and continuous EEG: prediction of subjective pain perception by alpha-1 power during stimulation and at rest. *Clin. Neurophysiol.* **123,** 605–12 (2012).

132. Zhang, Z. G., Hu, L., Hung, Y. S., Mouraux, A. & Iannetti, G. D. Gamma-Band Oscillations in the Primary Somatosensory Cortex--A Direct and Obligatory Correlate of Subjective Pain Intensity. *J. Neurosci.* **32,** 7429–7438 (2012).

133. Tiemann, L. *et al.* Differential neurophysiological correlates of bottom-up and top-down modulations of pain. *Pain* **156,** 289–296 (2015).

134. Staahl, C. & Drewes, A. M. Experimental human pain models: a review of standardised methods for preclinical testing of analgesics. *Basic Clin. Pharmacol. Toxicol.* **95,** 97–111 (2004).

135. Sarnthein, J., Stern, J., Aufenberg, C., Rousson, V. & Jeanmonod, D. Increased EEG power and slowed dominant frequency in patients with neurogenic pain. *Brain* **129,** 55–64 (2006).

136. Olesen, S. S. *et al.* Slowed EEG rhythmicity in patients with chronic pancreatitis: evidence of abnormal cerebral pain processing? *Eur. J. Gastroenterol. Hepatol.* **23,** 418–24 (2011).

137. Castellanos, N. Alteration and reorganization of functional networks: a new perspective in brain injury study. *Front. Hum. Neurosci.* **5,** 1–13 (2011).

138. Stam, C. J. & Reijneveld, J. C. Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomed. Phys.* **1,** 3 (2007).

139. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10,** 186–98 (2009).

140. Khodayari-Rostamabad, A. *et al.* A cortical source localization analysis of resting EEG data after remifentanil infusion. *Clin. Neurophysiol.* **126,** 898–905 (2014).

141. Atkinson, G. & Nevill, A. M. Measurement Error (Reliability) in Variables Relevant to Sports Medicine. *Sport. Med.* **26,** 217–238 (1998).

142. Bruton, A., Conway, J. H. & Holgate, S. T. Reliability: What is it, and how is it measured? *Physiotherapy* **86,** 94–99 (2000).

143. Edlinger, G., Wach, P. & Pfurtscheller, G. On the realization of an analytic high-resolution EEG. *IEEE Trans. Biomed. Eng.* **45,** 736–745 (1998).

144. Tallgren, P., Vanhatalo, S., Kaila, K. & Voipio, J. Evaluation of commercially available electrodes and gels for recording of slow EEG potentials. *Clin. Neurophysiol.* **116,** 799–806 (2005).

145. Lin, C. T. *et al.* Novel dry polymer foam electrodes for long-term EEG measurement. *IEEE Trans. Biomed. Eng.* **58,** 1200–1207 (2011).

146. Mullen, T. *et al.* Real-time Neuroimaging and Cognitive Monitoring Using Wearable Dry EEG. *IEEE Trans. Biomed. Eng.* **62,** 1–1 (2015).

147. Dai, F. *et al.* Integration of pain score and morphine consumption in analgesic clinical studies. *J. Pain* **14,** 767–777 (2013).

148. Turk, D. C. *et al.* Analyzing multiple endpoints in clinical trials of pain treatments: IMMPACT recommendations. Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials. *Pain* **139,** 485–93 (2008).

149. Dworkin, R. H. *et al.* Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J. Pain* **9,** 105–21 (2008).

150. Silverman, D. G., O'Connor, T. Z. & Brull, S. J. Integrated assessment of pain scores and rescue morphine use during studies of analgesic efficacy. *Anesth. Analg.* **77,** 168–170 (1993).

151. Gerbershagen, H. J. *et al.* Preoperative chronic pain in radical prostatectomy patients: preliminary evidence for enhanced susceptibility to surgically induced pain. *Eur. J. Anaesthesiol.* **27,** 448–454 (2010).

152. Pinto, P. R., McIntyre, T., Araújo-Soares, V., Costa, P. & Almeida, A. Differential Predictors of Acute Post-Surgical Pain Intensity After Abdominal Hysterectomy and Major Joint Arthroplasty. *Ann. Behav. Med.* **49,** 384–397 (2015).

153. Bisgaard, T., Rosenberg, J. & Kehlet, H. From acute to chronic pain after laparoscopic cholecystectomy: a prospective follow-up analysis. *Scand. J. Gastroenterol.* **40,** 1358–64 (2005).

154. Okie, S. A Flood of Opioids, a Rising Tide of Deaths. *N. Engl. J. Med.* **363,** 1981–1985 (2010).

155. Olesen, S. S. *et al.* Pharmacological pain management in chronic pancreatitis. *World J. Gastroenterol.* **19,** 7292–301 (2013).

156. Grosen, K., Vase, L., Pilegaard, H. K., Pfeiffer-Jensen, M. & Drewes, A. M. Conditioned pain modulation and situational pain catastrophizing as preoperative predictors of pain following chest wall surgery: a prospective observational cohort study. *PLoS One* **9,** e90185 (2014).

157. Yarnitsky, D. *et al.* Prediction of chronic post-operative pain: Pre-operative

DNIC testing identifies patients at risk. *Pain* **138,** 22–28 (2008).

158. Ip, H. Y. V., Abrishami, A., Peng, P. W. H., Wong, J. & Chung, F. Predictors of postoperative pain and analgesic consumption: a qualitative systematic review. *Anesthesiology* **111,** 657–677 (2009).

159. Khodayari-Rostamabad, A., Hasey, G. M., Maccrimmon, D. J., Reilly, J. P. & de Bruin, H. A pilot study to determine whether machine learning methodologies using pre-treatment electroencephalography can predict the symptomatic response to clozapine therapy. *Clin. Neurophysiol.* **121,** 1998–2006 (2010).

ISSN (online): 2246-1302
ISBN (online): 978-87-7112-439-2

AALBORG UNIVERSITY PRESS