Matti Helin

Changes in electromyographic amplitude as a measure of adequacy of sedation in critical care patients

Faculty of Electronics, Communications and Automation

Thesis submitted for examination for the degree of Master of Science in Technology.

24.5.2010

Thesis supervisor:

Prof. Mikko Sams

Thesis instructor:

D.Sc.(Tech.) Kimmo Uutela

Aalto University
School of Science
and Technology

AALTO UNIVERSITY
SCHOOL OF SCIENCE AND TECHNOLOGY

ABSTRACT OF THE
MASTER'S THESIS

Author: Matti Helin

Title: Changes in electromyographic amplitude as a measure of adequacy of sedation in critical care patients

Date: 24.5.2010 Language: English Number of pages:7+49

Faculty of Electronics, Communications and Automation

Bioinformation Technology

Professorship: Computational Technology Code: S-114

Supervisor: Prof. Mikko Sams

Instructor: D.Sc.(Tech.) Kimmo Uutela

Sedation is a treatment where consciousness level of critical care patients is lowered with sedative drugs in order to help patients to better withstand the strains of intensive care. Too deep or too light sedation can lengthen the time in critical care and raise the risk of complications. In current practice the level of sedation is controlled by different clinical assessment scales. As of now there is no automatic measure of sedation that has been validated for use in critical care. The most promising objective measure is recently developed EMG responsiveness.

The goal of this study is to examine frontal electromyographic (EMG) changes in critical care patients and to find out whether stimulus related responses of EMG can be used as a measure of depth of sedation in intensive care patients or not. Studied stimuli are a spoken command played back by earphones, a loud noise and two electric stimuli, one of which was painful. EMG stimulus related properties were analysed quantitatively and qualitatively and several different measures were developed based on those properties. The measures' applicability as monitor of sedation in intensive care patients were estimated statistically.

The results show that after a stimulus is given typically a rise in EMG power level occurs when the patient responds to the stimulus. Usually EMG level stays elevated for quite long. This elevation did not occur in cases where patient did not respond to the stimulus except when a painful stimulus was given. Also EMG power level before the stimulus tends to be higher in responsive patients compared to unresponsive ones.

Stimulus related changes are very suitable for measuring the depth of sedation at a given time. The best developed measure works significantly better than a continuous responsiveness measure. In some cases EMG power seemed to be more sensitive to patient reactions than nurse.

Keywords: Sedation, EMG, Response

AALTO-YLIOPISTO
TEKNILLINEN KORKEAKOULU

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Matti Helin

Työn nimi: Lihassähkökäyrän amplitudin muutokset tehohoitopotilaiden sedaation syvyyden mittarina

Päivämäärä: 24.5.2010 Kieli: Englanti Sivumäärä:7+49

Elektroniikan, tietoliikenteen ja automaation tiedekunta

Bioinfromaatioteknologia

Professuuri: Laskennallinen tekniikka Koodi: S-114

Valvoja: Prof. Mikko Sams

Ohjaaja: TkT Kimmo Uutela

Sedaatioksi kutsutaan hoitotoimenpidettä, jolla tehohoitopotilaiden tajunnan tasoa madalletaan sedatiivisilla aineilla, jotta potilas kestäisi paremmin tehohoidon aiheuttamat rasitukset. Liian syvä tai kevyt sedaatio voi pidentää hoidon kestoa ja lisätä komplikaatioriskiä. Sedaation syvyyttä säädellään nykyisin erilaisilla kliinikon arvioon perustuvilla asteikoilla. Automaattista tehohoitokäyttöön validoitua mittaria ei ole. Lupaavin objektiivinen mittari on viimeaikoina kehitetty lihassähkökäyrään (EMG) perustuva responsiivisuus.

Tämän työn tavoitteena on tutkia tehohoitopotilaiden otsan lihassähkökäyrän erilaisiin ärsykkeisiin liittyviä muutoksia sekä selvittää, voiko ärsykkeisiin liittyviä muutoksia käyttää tehopotilaiden sedaation syvyyden mittarina. Tarkasteltuja ärsykkeitä olivat kuulokkeilla toistetut puhuttelu ja voimakas kohina sekä kaksi sähköistä ärsykettä, joista toinen oli kivulias. Työssä analysoitiin laadullisesti ja määrällisesti EMG-tehon tyypillisiä ärsykkeisiin liittyviä ominaisuuksia, joiden pohjalta muodostettiin erilaisia mittareita. Näiden soveltuvuutta tehohoitopotilaan sedaation syvyyden mittariksi arvioitiin tilastollisin menetelmin.

Tulokset osoittavat että annettua ärsykettä seuraa tyypillisesti EMG-tehon tason nousu, kun potilas reagoi havaittavasti ärsykkeeseen. Yleensä EMG-teho myös jää tällöin korkeammalle tasolle pidemmäksi aikaa. Vastaavaa ilmiötä ei havaittu, kun potilas ei reagoinut annettuun ärsykkeeseen, paitsi silloin, kun annettiin kivulias ärsyke. Myös EMG-tehon lähtötaso on usein korkeampi niillä potilailla jotka reagoivat ärsykkeseen kuin niillä jotka eivät reagoi.

Ärsykkeeseen liittyvät muutokset soveltuvat hyvin tehohoitopotilaiden sedaation syvyyden arvioimiseen valitulla ajan hetkellä. Parhaiten toimiva mittari luokittelee potilaat reagoiviin ja reagoimattomiin merkittävästi paremmin kuin jatkuvaaikainen responsiivisuusmittari. EMG teho vaikutti joissain tapauksissa havaitsevan reaktioita ärsykkeisiin herkemmin kuin hoitaja.

Avainsanat: Sedaatio, EMG, vaste

Preface

This study was funded by and carried out at GE Healthcare Finland Oy.

I want to thank my supervisor Professor Mikko Sams for the suggestions on the subject of this thesis and for helping me out whenever required. Many thanks to my instructor D.Sc. Kimmo Uutela for his excellent guidance and tireless patience to help me throgh the writing process.

I want to thank Professor Pekka Meriläinen for providing me the opportunity to work on such an interesting research subject. Thanks to the whole Clinical Research team of GE Healthcare for the inspiring and pleasant athmosphere to work in. Special thanks goes to M.Sc. Petteri Lapinlampi for his help with several Matlab utilities and to MD, PhD Markku Paloheimo for the enlightening conversations on physiology related to my thesis and for giving me a large amount of reading material.

Finally I want to thank my parents for sparking my interest in technology and science at an early age and my wife for her support on my studies and for a gentle prod or two where needed.

Espoo, 24.5.2010

Matti Helin

Contents

Abbreviations

BIS	Bispectral Index
dB	Decibels
EEG	Electroencephalogram, electroencephalography
EMG	Electromyogram, electromyography
fEMG	Frontal electromyogram
ICU	Intensive care unit
MU	Motor unit
MUAP	Motor unit action potential
MUAPT	Motor unit action potential train
RASS	Richmond agitation-sedation scale
RE	Response Entropy
RMS	Root mean square
ROC	Receiver operating characteristics
SD	Standard deviation
SE	State Entropy
TOF	Train-of-four

1 Introduction

Intensive care patients are under significant amounts of anxiety and stress during their stay on intensive care unit (ICU). To increase patient comfort and to reduce risk of stress related symptoms and problems such as tachycardia and ventilator discompliance, during most of their stay in ICU patients are sedated with various drugs. On the other hand, oversedation increases the length of ventilation, which in turn is associated with higher morbidity and mortality of the patient. Therefore, an optimal level of sedation is a balance between patient comfort and safety.

Several methods to assess the level of sedation have been developed. Currently only scoring systems based on nurse or physician assessment have been proven to work in clinical setting. However these scoring systems are subjective and do not provide continuous information of level of sedation. So changes in patient's level of consciousness can go unnoticed for relatively long periods of time. Some objective measures have also been developed for assessing the level of sedation. Most of these methods are based on EEG signal analysis but other biosignals have also been used. So far these methods have shown only limited clinical usefulness in ICU setting [1].

The objective of this thesis is to examine properties of frontal electromyographic (fEMG) power signal related to different kinds of stimuli in ICU patients and to determine whether these properties can be used to derive an algorithm for measuring depth of sedation in the patients. The study includes two parts. In the first part general properties of stimulus related fEMG power are examined. In the second part of the study performance of several different fEMG based parameters as measure of depth of sedation are tested.

EMG power has been proven to be an effective measure of adequacy of sedation [2]. However a continuous depth of sedation algorithm based on EMG power seems to require relatively long history of data and therefore may have serious lag in the results. Also there has not been a reported study to assess the relationship between nurse observations and EMG responses to various stimuli. These phenomena will be addressed in this thesis.

2 Literary review

Purpose of this chapter is to describe underlying physiology behind frontal electromyography and review current techniques for measuring depth of sedation. Basics and current practices of sedation in intensive care units is discussed and hepatic encephalopathy is also reviewed shortly since all patients in the experimental part of the study suffered at least a medium degree encephalopathy.

Facial expressions are one of the most important ways of conveying emotions between humans. Both primal and more complex emotions are conveyed by facial muscle movements that are to a great extent automatic. Some of more primal expressions such as surprise or painful expression can be seen as responses to stimuli. Some of these reactions can also be seen even when a person is asleep. As muscle reactions are electrically active reactions they result in a potential changes that spread to the proximity of the muscle. If the muscle is close to the skin as facial muscles are, changes in potential are measurable with surface electrodes placed on skin. This kind of stimulus - muscle contraction - EMG potential measurement relationship is the basis of this study and will also be focus of this chapter.

2.1 Physiology

2.1.1 Muscle activation

Muscle tissue consists of muscle cells or muscle fibers and of nerve cells which innervate them. Muscle tissue has four special properties that together allow such functions as movement, maintaining of posture and facial expressions [3]. First, muscle cells are electrically excitable. Action potentials can propagate along cell's membrane. Muscle tissue can be stimulated either electrically or chemically. The most common stimulation method is neurotransmitters released by neurons of motor end plates. The second special property of muscle tissue is contractility. Excited muscle cells shorten and create tension pulling on its attachment points. Third, muscle tissue is extensible. Muscle tissue can be stretched beyond its original length without damaging it. Last, muscle tissue has great elasticity and can return to its original formation after contraction or extension.

There are three kinds of muscle tissue: skeletal, cardiac and smooth muscle tissue. Skeletal or striated muscle is a type of muscle that is mainly used to generate movement and maintain posture. Skeletal muscle contractions are generally voluntary but contractions can also be involuntary. Great deal of voluntary contractions are also subconscious by nature. For example maintaining posture or breathing do not require conscious effort. Breathing will continue even at sleep. Cardiac muscle tissue can only be found in the heart. Cardiac muscle is striated but its contractions are completely involuntary. Smooth muscle tissue also contracts involuntarily. It can mainly be found in internal structures such as intestines, blood vessels and airways but also attached to hair follicles. Activity of cardiac muscle and smooth muscle are

regulated by autonomous nervous system whereas skeletal muscles are controlled by somatic nervous systems.

Basic building block of skeletal muscle tissue is a muscle fiber. A single fiber is surrounded by cell membrane or sarcolemma. The fiber is permeated by thousands of minuscule invaginations of the sarcolemma called T tubules [3]. Other features distinct to muscle fibers are the actual contracting units or myofibrils and sarcoplasmic reticulum surrounding it.

A muscle contraction in skeletal muscle is initiated when connecting neurons firing rate is high enough. First acetylecholine is released from the neuron at neuromuscular junction. If the acetylecholine level at the junction is high enough an action potential is spread across the sarcolemma and through the T tubules. Action potential induces release of calcium from the sarcoplasmic reticulum, which in turn is essential for myofibril contraction. Since the action potential spreads through the whole muscle fiber the contraction of a single fiber is all or nothing type of action.

The smallest functional element in skeletal muscle contraction is called motor unit (MU) and consists of a single motoneuron and the individual muscle fibers it innervates. Number of muscle fibers in a motor unit depends on precision of the action of the muscle and can vary from as little as ten to couple of thousands muscle fibers per MU [3]. Action potentials of the muscle fibers in single MU are heavily correlated and can be measured as a single motor unit action potential (MUAP). During muscle activation several MUAPs are fired in succession in a motor unit action potential train (MUAPT). Firing frequency of a MUAPT is directly proportional to the strength of the muscle contraction until maximum contraction.

2.1.2 Upper facial muscles and facial expressions

A major function of facial muscles is expression of emotions. Emotions such as surprise and discomfort cause a complex and to a great extent automatic muscle contractions. This section describes the relevant facial muscles involved in emotional responses to abrupt stimuli and pain.

Musculus Frontalis muscle or frontal belly of *Occipitofrontalis* covers most of the forehead. It connects to skin superior to supraorbital margin and to epicranial aponeurosis (Figure 2.1). Movement of *M. Frontalis* draws scalp anteriorly, raises eyebrows and wrinkles skin of forehead horizontally [3].

Musculus Orbicularis Oculi or *Musculus Orbicularis Palbebrarum*, as it was called earlier, is a sphincter muscle located around the orbit (Figure 2.1). It's function is to close eyelids [3, 4].

Musculus Corrugator Supercilii is located nearly directly under the eyebrow (Figure 2.1). It draws eyebrows down and wrinkles the skin of forehead vertically in a frowning motion [3].

Cranial nerve VII (*nervus facialis*) is the motornerve that innervates the most facial muscles [5].

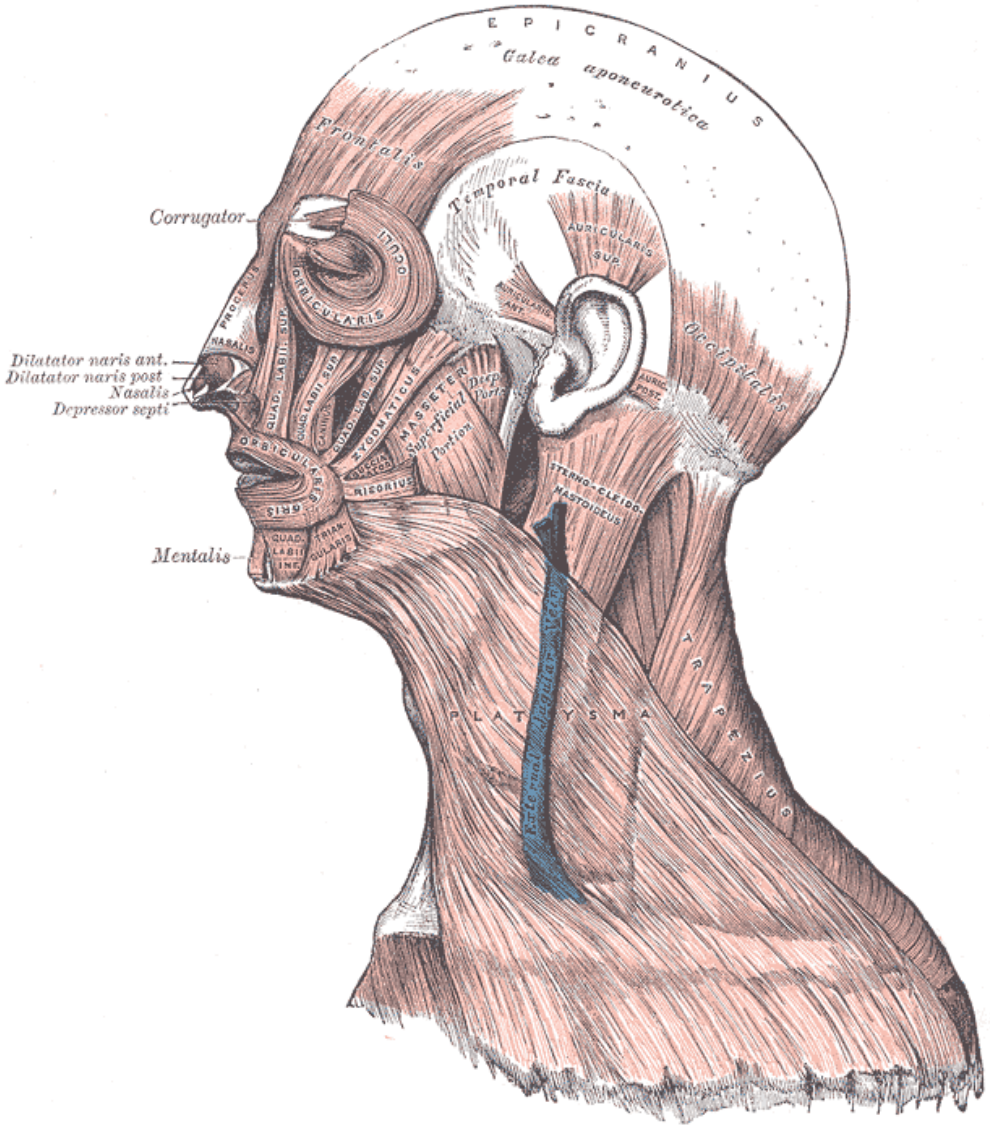

Figure 2.1: *Muscles of face and neck. Figure reprinted from [4].*

Activity of facial muscles is closely related to emotions. People react to pain and other stimuli with facial expressions. These expressions are to large extent involuntary and can be seen as stimulus related muscle contractions of facial muscles. Particularly well studied phenomenon is so called startle reflex or startle response, which occurs after an abrupt strong stimulus, usually a loud auditory stimulus or a noxious stimulus. Interestingly the startle reflex seems to generalise well in most of the muscles of the face and some of the upper body [6]. With the exception of blink response the latency of electromyographically measurable response to the stimulus seems to mostly depend on distance from caudal brainstem. Indeed startle reflex in man can be measured earliest from sternocleidomastoid that is very close to the brainstem but it also generalises well over the whole face [6].

2.1.3 Hepatic Encephalopathy

Hepatic encephalopathy is a form of a metabolic encephalopathy that is caused by acute or chronic liver failure. Clinical symptoms of hepatic encephalopathy vary greatly and it often is difficult to distinguish from other metabolic encephalopathies [7]. Hepatic encephalopathies are classified into two categories. *Fulminant hepatic encephalopathy* is a result of acute liver failure due to viral hepatitis or liver intoxication. Specifically in United States and Europe paracetamol ingestion accidentally or as a suicide attempt is a prominent cause. Prognosis for this form of hepatic encephalopathy is poor. Only about 20 % of patients survive without liver transplantation [7]. Death is often caused by cerebral oedema and can follow within days of onset of the symptoms. *Portal systemic encephalopathy* is due to chronic liver failure, mainly liver chirrosis caused by alcoholism. Prognosis for this condition depends on severity of underlying liver disease. If encephalopathy develops to coma stage only 35 % of patients will show a good recovery within a year [7].

Clinically hepatic encephalopathy is divided to four stages of severity. Psychological symptoms range from dysphoria, agitation, anxiety and concentration problems in stage 1 to disorientation, personality changes, disorientation and cognitive impairment in stage 2 and to progressive clouding of consciousness and eventual coma in stages 3 and 4 [7]. Neurological symptoms consist of gradually increasing asterixis in all stages, ataxia in stage 2, and muscular rigidity and hyperflexia in stage 3. Hepatic encephalopathy may also show rise in arterial ammonia level and changes in electroencephalogram (EEG). However even advanced hepatic encephalopathia may not show major changes in EEG [8].

The common changes in EEG signal are slowing of EEG and existence of *triphasic waves*. Slowing often parallels arterial ammonia level [8]. Sudden shifts between normal alpha frequency and slower variants are common [9]. Triphasic waves are slow waves with an initiating sharp transient. Although these waves do not occur in every patient and are not completely specific to hepatic encephalopathy they are highly indicative of it [9]. Triphasic waves are also indicate a poor prognosis [8]. Other EEG changes may also occur especially due to chronic hepatocerebral degeneration and interestingly due to disturbances after liver transplantation [8].

Only know effective treatment for fulminant hepatic encephalopathy is a liver transplantation. The symptoms may be alleviated and the condition improved by massive plasma exchange or certain medicines such as diuretics, but these treatments are not curative. For chronic hepatic encephalopathy there is no know therapy and hence the management is usually conservative. Due to liver dysfunction and cerebral changes not all medical treatments are suitable. For example use of benzodiazepines precipitates hepatic encephalopathy [10].

2.2 Electrophysiological Measurements

2.2.1 Electroencephalography

Electroencephalography (EEG) is a surface potential measurement of electrical activity in brain. The first concrete evidence of electrically measurable activity of brain was found by Richard Caton in late 19th century [8]. Half a century later in the year 1929 the first human electroencephalogram was recorded by Hans Berger. Since the early days of EEG there has been much development. Most notably introduction of frequency domain analysis to EEG and computerisation of EEG allow an increasing number of automated analysis methods that can be used. Although automatical EEG analysis methods cannot be expected completely replace expert evaluation of recordings automatic analysis can lighten the workload of experts. Automatic analysis methods can also bring EEG to fields, such as anesthesia monitoring, where EEG experts are not constantly available.

In principle EEG is a voltage measurement between a reference electrode and one or more measuring electrodes attached to scalp. Usually a standardised electrode setup is used.

In majority of EEG measurements there is no direct contact of scalp electrode and tissue. Instead an electrolytic gel or paste is applied to connect the electrode indirectly to the tissue [8]. Electrochemical properties in the electrode/electrolyte junction and electrolyte/tissue junction help to keep a steady potential and impedance between the tissue and measurement device.

All electropotential measurements of living organisms are prone to signal artifacts. Due to small amplitude of EEG signal artifacts can be especially disruptive. Artifacts can be either biological or technical in nature. Biological artifact includes thermodynamic noise and other bioelectric signals such as electrocardiogram and electro-oculogram or artifact. Technical artifacts arise of the properties of electrodes, cables, amplifiers and analogue/digital conversion. Possible artifacts include DC offset voltage, electrode movement artifacts, capacitive and inductive coupling and various noise sources. A quintessential technical artifact is mains interference that arises from capacitive and inductive coupling. Mains interference causes sinusoid artifact in mains frequency, either 50 Hz or 60 Hz depending on country, and the harmonic frequencies. A typical biological artifact in EEG measurement are EMG signals from facial or neck muscles.

2.2.2 Electromyography

Electromyography (EMG) is a biopotential measurement in many ways similar to EEG. EMG measures electric potential changes directly from a muscle or from a site close to the desired muscle or muscle group.

EMG signal can be measured either with surface electrodes attached to skin or with subcutaneously or intramuscularly placed needle electrodes. With needle electrodes

it is possible to measure activity of a single motor unit or even single muscle fiber. Then again, surface electrodes give a more general picture of total activity of muscle or muscles it is placed over [11]. The strength of surface potential is proportional to the number and size of activated motor units and on frequency of MUAPTs [12] and the integral of full-wave rectified surface electromyographic record is proportional to the force generated by a muscle [11]. Hence power of the surface potentials follows the muscle activity near the measurement site.

Like all skin potential measurements surface EMG is prone to signal artifacts. If surface electrodes are used artifacts are similar to EEG artifacts. Technological artifacts are the same as in EEG. Biological artifacts are caused by other bioelectric processes than the measured muscle activity. In frontal electromyography the main biological artifact sources are EEG, EOG, ECG and other nearby EMG signal sources.

In this study potentials were measured with relatively large surface electrodes placed on forehead of the patient. The measured potential contains both EEG and EMG signal. However EEG signal bandwidth is limited and quite well known. EEG amplitude in frequencies over 50 Hz or so is very low compared to EMG amplitude measured on forehead. Therefore power in higher frequencies can be quite safely be assumed to be mainly from EMG signal.

2.3 Sedation in intensive care units

Patients in intensive care unit (ICU) typically undergo significant amounts of anxiety, stress, pain and discomfort during their stay in ICU. Also lack or poor quality of sleep is often reported. These problems can be alleviated with administration of analgesic and sedative drugs. Administration of analgesic drugs or analgesia is used to remove or to blunt the pain. Indications for administrating sedatives or sedation are not as clear. Usually sedatives are needed in situations where patient's behaviour risks safety of the patient or nursing staff. Potentially harmful behaviour includes aggressiveness, agitation and removal of ventilator tubes or intravenous needles. Problems in ventilator compliance occur often with unsedated patients and therefore mechanical ventilation almost always requires at least some sedation. It is clear that in ICU treatment both analgesia and sedation are usually required for patient safety and comfort. On the other hand, overdosage may lead to longer stay in the ICU, increased duration of mechanical ventilation and morbidity [13] Both overdosage as well as underdosage of analgesic and sedative drugs have severe risks so caution is needed in adjusting doses.

Use of continuous intravenous sedation correlates with length of mechanical ventilation and general length of the stay in ICU and hospital [14, 13]. Length of the ventilation in turn is associated with several adverse conditions such as ventilator associated pneumonia [15], respiratory depression, hypotension, immunosupression and renal dysfunction.

Protocols and drugs used in sedation vary greatly across the intensive care units [16,

17]. Surprisingly only relatively small portion of sedative agents reportedly used in intensive care have been rigorously tested in several studies [16]. Drug combinations can also change during the the stay in ICU [17]. Especially the drugs used in long term sedation tend to differ from the drugs used in short term one. The most widely used drugs in British practice are propofol, midazolam, fentanyl, alfentanil and morphine [17]. Propofol and midazolam are used in sedation while fentanyl, alfentanil and morphine are all analgesics. Neuromuscular blocking agents are used only in a median of 10 % of patients [17].

Studies show that following a strict protocol on adjusting sedative levels can reduce risk of ventilator associated pneumonia, duration of mechanical ventilation, and length of stay in intensive care unit and hospital compared to adjusting sedative levels on physician's decision alone [13, 18, 19]. Similar effects can be achieved by daily interruption of sedative infusions [20]. Also daily spontaneous awakening trial followed by spontaneous breathing trial gives similar results. It leads to shorter coma and mechanical ventilation, less time spent in hospital and ICU and it improves one year survival compared to contemporary methods [21]. However it seems that neither strict protocols nor daily interruption of sedative administration are commonly used [21].

2.4 Subjective methods for assessing level of sedation

Current clinical practice for assessing level of sedation is based on subjective assessment of patient's reactions to a given stimulus protocol. There are several different assessment scales. Practically all of the sedation scoring systems are based on nurse observations of patient behaviour and reactions to certain stimuli. Used stimuli vary between the scoring systems used.

Review done by Murdoch & Cohen [17] reveals that in the year 2000 in United Kingdoms 67 % of adult general intensive care units (ICU) used a systematic sedation scoring system. In cardiac ICUs the percentage was 56. Out of the reported systems in the study a little over a half of the general ICUs used Ramsay sedation score. Other scoring systems reportedly used in more than one unit were Addenbrookes, Cook and Sheffield scores. Another review by Ostermann et al. [16] shows that almost all clinical studies done between 1980 and 1998 used some kind of scoring system and that Ramsay sedation was also the most used sedation scoring system. Other systems used were Cook and Palma modified Glasgow Coma Scales, modified Ramsay score and several ad hoc scores created for a specific study.

Due to their prevalence in contemporary ICU practice Ramsay sedation score and more recent Richmond agitation sedation scale [22] will be further discussed.

2.4.1 Ramsay sedation score

One of the earliest documented methodical scoring systems for assessing depth of sedation is Ramsay sedation score [23]. The scoring system was originally developed

for assessing the level of sedation in a study on use of Alphadolone-Alphaxalone as sedative. Since then Ramsay sedation score has become widely used in clinical practice and is now one of the most widely adapted measurements of depth of sedation. Although Ramsay sedation score is very widely adapted it has never been rigorously tested [24].

Ramsay sedation score is measured by stimulating patient with gradually strengthening stimuli and observing patient reactions to given stimuli. Ramsay sedation score values and their descriptions can be seen in Table 2.1. Although Ramsay defines an acceptable sedation level to be anything from Ramsay score 2 to 5 more recent studies recommend target level of Ramsay score of 3 [13] or from 3 to 4 [20] in general. Other guidelines however recommend that need of sedation should be assessed according to patient and situation [19].

Table 2.1: *Ramsay sedation scale*

Score	Description
1	Patient awake, anxious, agitated or restless
2	Patient awake, cooperative, orientated and tranquil
3	Patient drowsy with response to commands
4	Patient asleep, brisk response to glabella tap or loud auditory stimulus
5	Patient asleep, sluggish response to stimulus
6	No response to firm nail-bed pressure or other noxious stimuli

2.4.2 Richmond agitation sedation scale

Richmond agitation sedation scale (RASS) is an observation based sedation scoring system in many ways similar to Ramsay sedation score. RASS was developed by a multidisciplinary group consisting of physicians, nurses and pharmacists [22].

RASS is measured first by observing whether the patient is awake or not. If patient is awake the observer assesses the level of agitation by observing patient behaviour and gives score of 0 to +4 according to the behaviour. If patient is not visibly awake patient is spoken to and strength of reaction is used to assess level of sedation. If patient is unresponsive to voice a physical stimulus is given to separate deep sedation from completely unarousable patients. Complete listing of RASS scores and respective terms and descriptions is shown in Table 2.2.

RASS is very similar to Ramsay score. Most notable difference between the scoring systems is that RASS has four scores for measuring agitation or anxiety level whereas Ramsay score has only one. Another differences is that in RASS quality of response to voice is used for scoring and noxious stimulus is not used at all.

Table 2.2: *Richmond agitation–sedation scale . Table reprinted from [22].*

Score	Term	Description
+4	Combative	Overtly combative or violent; immediate danger to staff
+3	Very agitated	Pulls on or removes tube(s) or catheter(s) or has aggressive behaviour toward staff
+2	Agitated	Frequent non-purposeful movement or patient-ventilator dyssynchrony
+1	Restless	Anxious or apprehensive but movements not aggressive or vigorous
0	Alert and calm	
–1	Drowsy	Not fully alert, but has sustained (more than 10 seconds) awakening, with eye contact, to voice
–2	Light sedation	Briefly (less than 10 seconds) awakens with eye contact to voice
–3	Moderate sedation	Any movement (but no eye contact) to voice
–4	Deep sedation	No response to voice, but any movement to physical stimulation
–5	Unarousable	No response to voice or physical stimulation

2.5 Objective methods for assessing level of sedation

At present there is no gold standard for objective measure of sedation [1]. Although there has been some effort to develop objective methods to assess the level of sedation, as of now no proven method is available for clinical use. This chapter describes some suggested methods for objective assessment of level of sedation and some of the problems these methods face.

2.5.1 Bispectral index

Bispectral index (BIS) is an EEG analysis method which was originally developed and validated as measure of depth of anaesthesia. Since drugs used in sedation of ICU patients are in large extent the same as drugs used in general anaesthesia it seems reasonable to test validity of the algorithm in ICU setting.

Following the development of BIS there have been several studies concerning validity of BIS in intensive care patients. Some of the initial studies on using BIS as measure of sedation were promising [25]. However more thorough studies have shown significant inter-individual differences [26, 27]. With some patients BIS correlates moderately with clinical sedation scores. With others there is no correlation at all. Moreover there seems to be no differentiating factor between these groups [27].

One of the main confounding factors with BIS in intensive care unit is facial muscle

activity. Since BIS electrodes are attached to patient's forehead, measured signal can be seen as compound signal of cortical electric activity and electric activity of facial muscles mainly *frontalis* muscle [26]. In general anaesthesia this is usually not a problem, since generally neuromuscular blocking agents are used to ensure no patient movement during the operation, which lowers the EMG artifact in BIS measurement. However in general ICU practice neuromuscular blockade is used relatively seldom. In the year 2000 in United Kingdoms median of 10 % of the general adult intensive care patients received therapeutical paralysis [17]. EMG activity therefore remains a significant factor in frontal EEG based indicators.

2.5.2 Entropy

Entropy is a parameter for measuring the depth of anaesthesia similar to BIS. Entropy is is based on spectral entropy measure [28]. Entropy monitor gives two parameters state entropy (SE) and response entropy (RE). SE is computed over frequency range from 0.8 Hz to 32 Hz and RE from 0.8 Hz to 47 Hz [28]. The entropy values are normalised so they are equal when entropy from 32 Hz to 47 Hz is zero. Hence difference of these two values can be considered to indicate amount of EMG activity. Difference of the parameters can easily be seen visually when the parameters are plotted on a same display.

Normalised spectral entropy is in range between 0 and 1. Entropy as presented in patient monitors is scaled with a spline transformation to a scale from 0 to 100 to emphasise the interesting range.

Entropy measurement has been tried out in intensive care settings. It has faced the same difficulties as BIS in measuring depth of sedation. Similar to BIS, entropy is confounded by frontal EMG activity and therefore is not a reliable measure of depth of sedation in intensive care unit [29].

2.5.3 Auditory evoked potential

Auditory evoked potential (AEP) is a stimulus related EEG based measure of depth of anesthesia and sedation. So far AEP has mostly been used in research. AEP is calculated averaging EEG signal over several samples aligned to a given standard auditory stimulus. Increase in latency and decrease of amplitude of mid-latency auditory evoked waves seem to indicate deepening of anesthesia or sedation [30].

Auditory evoked potential seems to work quite well in clinical setting. It performs rather well at deeper levels of sedation but poorly in lighter levels of sedation [30, 1]. Clinical viability of AEP system has yet to be proven [1]. EMG is not that big of an issue with AEP as it is with BIS and Entropy due to electrode setup and possibility of rejecting all EMG confounded samples before averaging. However the electrode setting is a bit more complicated than with either BIS or Entropy. For example in Sculte-Tamburen et al. [30] electrodes were placed at standardised locations at back of the head and both auricles and grounded to forehead.

Overall AEP might be a good tool for monitoring deeply sedated patients, but as it is now, it is not validated for overall measurement of depth of sedation in ICU.

2.5.4 EMG responsiveness

Like electroencephalography, electromyography can be used as objective measure of level of anesthesia [31]. Recently a novel method for measuring depth of sedation based on *frontalis* EMG signal has been developed [2].

First published version of responsiveness measurement is a two fold process. The measurement is based on forehead EMG frequency band power values. First, arousals are estimated by reducing a time interval minimum baseline power from the current power value (2.1). Then their responsiveness value is calculated as a long time mean logarithmic arousal value (2.2). Responsiveness value was shown to correlate well with the observed modified Ramsay scores [2]. Especially well responsiveness worked in classifying patients to lightly sedated (Ramsay 3 or less) and deeply sedated ones (Ramsay 4 or above).

$$P_{arousal}(t_i) = P_{EEG/EMG}(t_i) - \min_{[t_i - \delta t_m \leq t_j \leq t_i]} [P_{EEG/EMG}(t_j)] \tag{2.1}$$

$$R(t_i) = mean[\log_{10} P_{arousal}(t_j)], t_i - \Delta t_a \leq t_j \leq t_i \tag{2.2}$$

Responsiveness value was extensively analysed and it was proven to work well in all patients except ones with some form of brain problem, mainly encephalopathia. Frequency band from which the power was calculated did not matter a lot. Responsiveness measure worked best at the 55 Hz to 145 Hz frequency band which has only minimal EEG power left and is therefore composed mainly from EMG signal. The best correlation was achieved with relatively small δt_m (one minute interval or so) and quite large Δt_a (half an hour or so).

Although an extensive clinical tests on clinical relevance of responsiveness measurement remain do be done the responsiveness value seems promising. It outperformed both BIS and entropy measurements measured from the same patients and even multi-parameter neural network classification was not significantly better in performance [2]. Even though there is a problematic patient group with which the responsiveness is not as reliable, the patient group is clearly identifiable.

Although responsiveness measure is still under development, due to good performance of the measure it will be used as a standard against which the results of this study will be compared. There has been further refinement of responsiveness parameter since Lapinlampi's Master's thesis but since the thesis is still the only published version of the algorithm that version will be used in this analysis.

3 Descriptive analysis of EMG power changes after stimulus

The objective of the first part of the study is to find typical stimulus related changes in EMG power. It is clear that at least to some extent it is possible to use EMG response to stimuli as a measure of reaction. When patient hears a command to open their eyes and obeys this most probably yields in both visible response and a measurable potential change. However, it is not known to what extent correlating EMG measures with detectable reactions is possible. Also typical form and properties of frontal EMG responses has not been extensively studied. This first part of the study therefore concentrates on identifying typical EMG amplitude changes related to stimuli and finding out phenomena that could be used to develop an automatic measure of reactions to the stimuli.

3.1 Materials and Methods

3.1.1 Clinical Trials

This study uses patient data collected by research nurses at the Surgical Hospital of Hospital District of Helsinki and Uusimaa (HUS). The acceptance criteria for the patients were:

- Patient admitted for unit 5 or ICU after hepatic failure
- The stage of hepatic encephalopathy is two or higher
- Patient has not received any sedatives in previous 12 hours period (in case of Remifentanil period is one hour)
- Assent for the study from the patients or relatives

Nineteen patient (11 females, 8 males) with clinically diagnosed hepatic encephalopathy were admitted to the study. Mean (SD) age of patients was 46.5 (17.1).

Each subject was monitored for a minimum of two days and a maximum of one week. Recording was stopped when patient was released from test site, maximum recording period was reached or research nurses were needed in other actions of test site.

Physiological signals were recorded with Datex-Ohmeda S/5 Critical Care Monitor, S/5 Anesthesia Monitor and S/5 Compact Monitor. Data from S/5 Critical Care Monitor was streamed via serial interface to a laptop computer with S/5 Collect program. Data from other two monitors was streamed to the same laptop computer and to S/5 Collect program via Ethernet. All data was saved to hard drive of the laptop. Collected physiological signals were 400 Hz EEG/EMG waveform, 300

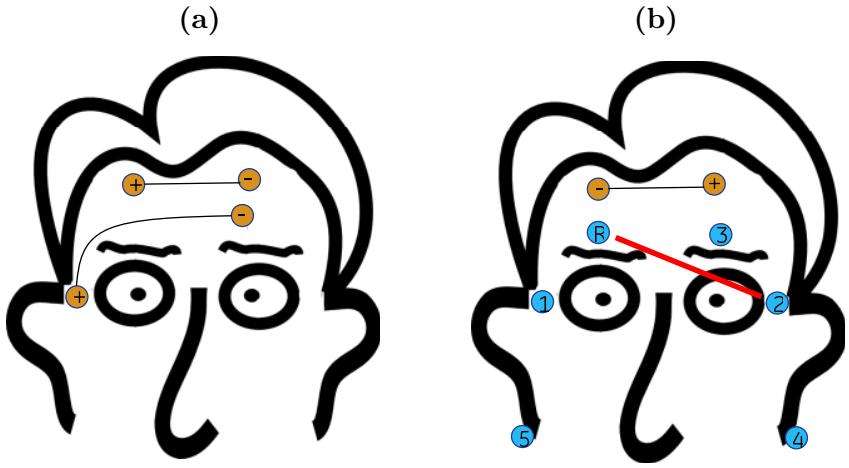

Figure 3.1: *EEG/EMG electrode placements in the first 6 recordings (a) and in the remaining 13 recordings (b).*

Hz ECG waveform, 100 Hz photoplethysmographic waveform and data from an acceleration sensor attached on top of an EEG/EMG electrode.

Also another EEG/EMG waveform was recorded from all patients. In first six patient recordings the waveform was another one channel 400 Hz waveform. Electrodes were placed as shown in Figure 3.1 (a). In remaining thirteen recordings the waveform was 500 Hz five channel measurement. EEG electrode placements are shown in Figure 3.1 (b). Although channel 2 in these measurements is a mirror image of the inferior channel in the previous setting, amount of EMG activity in the channels can be assumed to be close enough to treat these channels as equivalent measurements. From now on the near hairline electrode placement is called the top channel and inferior placement in Figure 3.1 (a) and channel two in Figure 3.1 (b) is called the bottom channel.

Recording was divided to passive and active periods. During passive periods research nurses were not present but data was collected automatically. During active periods one of the research nurses annotated treatment of the patient and functioning of the monitors. Active periods lasted 8 hours a day in average. During the active period research nurses assessed state of the patient every three hours or when clinical status was changed. Assessment consisted of following steps:

1. Standardised stimulation sequence
2. Glasgow Coma Score
3. Pupil reflex
4. Babinski's reflex

During a standardised stimulation sequence up to four stimuli are given to the patient as described in the flowchart (Figure 3.2). All stimulation times were recorded

together with a research nurse's annotation whether the patient responded to stimulus or not. According to the protocol, reaction to stimulus was defined as any gross purposeful movement of face or extremities. However, interviews with a research nurse showed that also large change in hemodynamic measures was sometimes annotated as a reaction to the stimulus. Although in most of the assessments the protocol was followed accurately, there were some exceptions. With first three patients TOF stimulus was not used at all. In some cases when patient was already awake at the time of assessment vocal stimulus was given by nurse and not with headphones. Also in some cases when for some reason headphones were not usable vocal stimulus was given by nurse.

The stimuli were given in assumed order from lightest to strongest. Time stamp was recorded for each stimulus. First a vocal stimulus was given to patient. Vocal stimulus was played with headphones at 85 decibel (dB) intensity level. Given stimulus was spoken command "Avaa silmäsi", which is "Open your eyes" in Finnish. If the patient did not respond to vocal stimulus a loud noise stimulus was given. Stimulus consisted of white noise which was played with headphones at 105 dB intensity level. If the patient did not respond to noise stimulus a Train-Of-Four (TOF) stimulus was given to patient's hand. Stimulus consisted of four consecutive short electrical shocks alternating current at rapid intervals. Active electrode was a skin electrode which was attached at patients wrist over the ulnar nerve. Ground electrode was also attached to the wrist. If the patient did not respond to TOF stimulus a tetanic stimulus was given. The tetanic stimulus is otherwise identical to TOF stimulus except for only one 30 s long electric shock was given.

3.1.2 Data processing

Epochs were extracted form EEG/EMG waveform, each epoch starting two minutes prior to a given stimulus and ending two minutes after stimulus. Epochs were grouped based on given stimulus and nurse annotations of patient response. A visual inspection was performed to remove epochs that contained artifacts. Most of these artifacts included poor electrode contacts and impedance tests done by other devices than the measuring one. There are also a number of reasons why a part of the data can be missing. For example if electrode comes loose or measuring module performs an electrode impedance test no recordings are made in such period. Also some of the measurements were performed in the beginning or in the end of a recording session and could be shorter than required epoch length of four minutes. All epochs containing artifacts or more than 20 seconds of missing data were removed prior to the analysis. The number of valid epochs used in the analysis is shown in table 3.1. Due to small number of positive responses to TOF stimulus only some of the following analysis was done to TOF responses.

An estimate of EMG power signal was calculated by dividing each epoch to 0.5 second windows. RMS power of 65 - 95 Hz frequency band was calculated using fast Fourier transform. Frequency band was chosen based on Lapinlampi's results [2] and to avoid mains frequency (50 Hz in Finland) and harmonics. Hence acquired

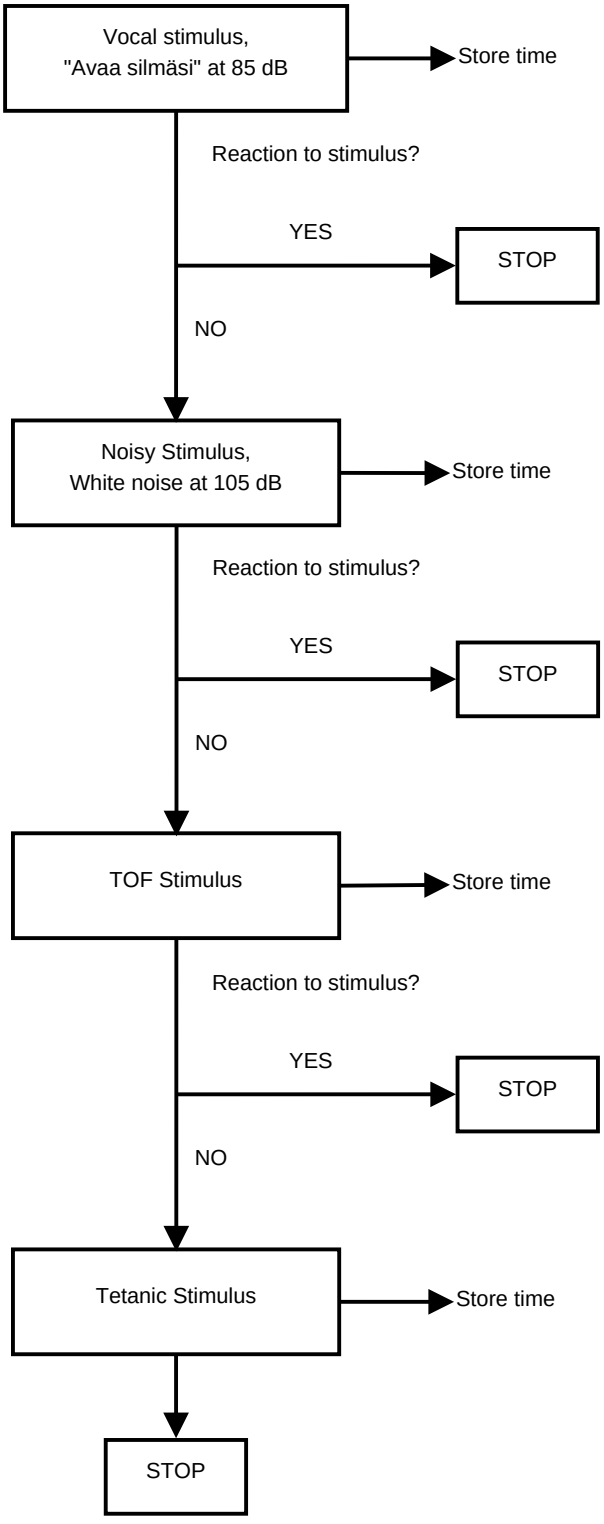

Figure 3.2: *Flowchart describing standardised stimulation sequence.*

Table 3.1: *Number of assessments used in the study.*

	Response	No response
Vocal stimulus	244	171
Noisy stimulus	30	140
TOF	3	107
Tetanic stimulus	62	77

EMG power signal has sampling rate of 2 Hz.

A stimulus related median and lower and upper quartile powers were calculated for vocal, noisy and tetanic stimuli. For each stimulus data was grouped according to given stimulus and whether the patient responded to the stimulus or not.

Median and lower quartile and upper quartile differences of 30 second median power after the stimulus and 30 second median power before the stimulus were calculated. Wilcoxon signed rank test [32] was performed to 30 second median power before and after stimulus to analyse significance of differences in the power levels.

The delay from annotated stimulus time to onset of EMG response can vary for two different reasons. There may be a time difference between annotation time and actual moment of stimulus in some cases where time labeling was not automatic. Physiological delay between stimulus and reaction is also possible. Delay between stimulus and EMG onset was studied in cases where patient responded to stimulus. Onset time was defined as the first time after stimulus when EMG power is greater than 20 second mean power before stimulus plus three times 20 second standard deviation before stimulus. Only the cases where onset happened in 30 seconds or less after the stimulus were used (Table 3.2). A feature aligned median and 25 percentile and 75 percentile powers were calculated to better describe the form of actual EMG response.

Amplitude spectrum of EMG power waveform was calculated to estimate whether there is difference between slow and fast activity.

Data processing was done with Matlab version 6.5.

Table 3.2: *Number of assessments used in feature aligned analysis.*

Vocal stimulus	191
Noisy stimulus	17
Tetanic stimulus	25

3.2 Results

In the figures of EMG power sequences the following form is used unless otherwise noted. Stimulus is given at $t = 0$ and is illustrated by a vertical dashed line at

the time of the stimulus. Example of stimulus related EEG/EMG signal and corresponding EMG powers can be seen in Figure 3.3. The example shows a rise and some spikes in both top and the bottom channel. In this particular situation there is a rise in the level that is more pronounced in the top channel and some spikes can seen in the bottom channel. Also in bottom channel raw signal there is a lot of high amplitude but low frequency activity, possibly due to eye movements (Figure 3.3 (c)). However this has only a marginal effect to the EMG power (Figure 3.3 (d)).

Median stimulus related EMG powers to different stimuli and corresponding upper and lower quartile powers at top channel are shown in Figure 3.4. Median, upper quartile and lower quartile stimulus related EMG powers at bottom channel are shown in Figure 3.5. Figures for the TOF stimulus are excluded due to the small number of positive responses to the stimulus and since median and percentile graphs for TOF stimulus when no response is detected are essentially similar to the cases where there is no response to noisy stimulus except for approximately one minute time shift. P values for Wilcoxon signed rank tests for top channel powers are shown in Table 3.3 and for bottom channel powers are shown in Table 3.4.

Regardless of which channel was used in all responding cases median power rises sharply shortly after stimulus is given and power stays elevated for rest two minutes of the epoch length. Initial level of power is higher in responding cases than it is in non-responsive ones. All responsive cases show statistically significant ($P < 0.05$) difference between 30 second median EMG power after stimulus and 30 second median power before stimulus.

Median power from cases where patient responds to noisy stimulus (Figure 3.4 (e)) shows an obvious raise in EMG power approximately 60 seconds before stimulus. This is around the time where vocal stimulus is given. Even though the nurse did not detect a visible response it is likely that patient awoke and reacted to stimulus in some cases. A response similar to reacting cases was sometimes seen when visually examining EMG power of the patients annotated non-responsive to vocal stimulus but reacted to noisy stimulus. It is also plausible that in these cases patient was more likely to react visibly when the noisy stimulus was given.

In cases where patient showed no response reaction in median power depends on given stimulus. When vocal stimulus is given (Figure 3.4 (b)) there is little change to median EMG power. After noisy stimulus reaction is also negligible until about a hundred seconds after the stimulus which is approximately the time when tetanic stimulus is given (Figure 3.4 (d)). There was neither statistically significant change in the 30 second median power after stimulus compared to 30 second median power before stimulus. Tetanic stimulus however causes similar reactions in EMG power as with responding cases (Figure 3.4 (e)). Also statistically significant changes can be seen after tetanic stimulus even when patients are not responding to the stimulus.

Median feature aligned EMG power response is shown in Figure 3.6. In this figure the moment $t = 0$ is the moment of first outliers after the stimulus. Figure shows a very sharp rise followed by an elevated EMG power level, which is a typical change when a change occurs after stimulus. Although a typical response resembles the

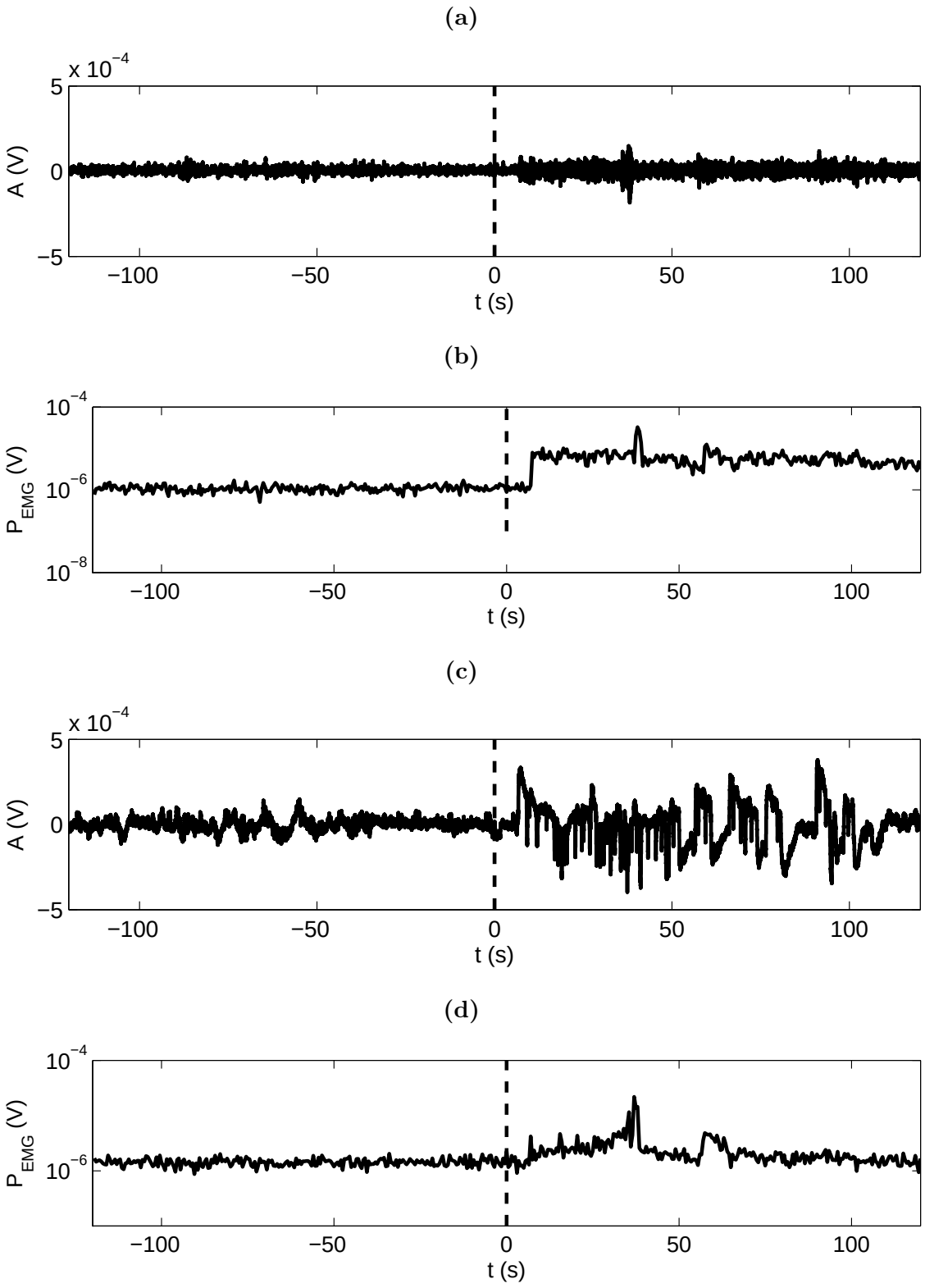

Figure 3.3: *An example of EMG responses to a vocal stimulus in top channel (a and b) and in bottom channel (c and d). Graphs a and c show the raw recorded signal while graphs b and d show the derived EMG band power.*

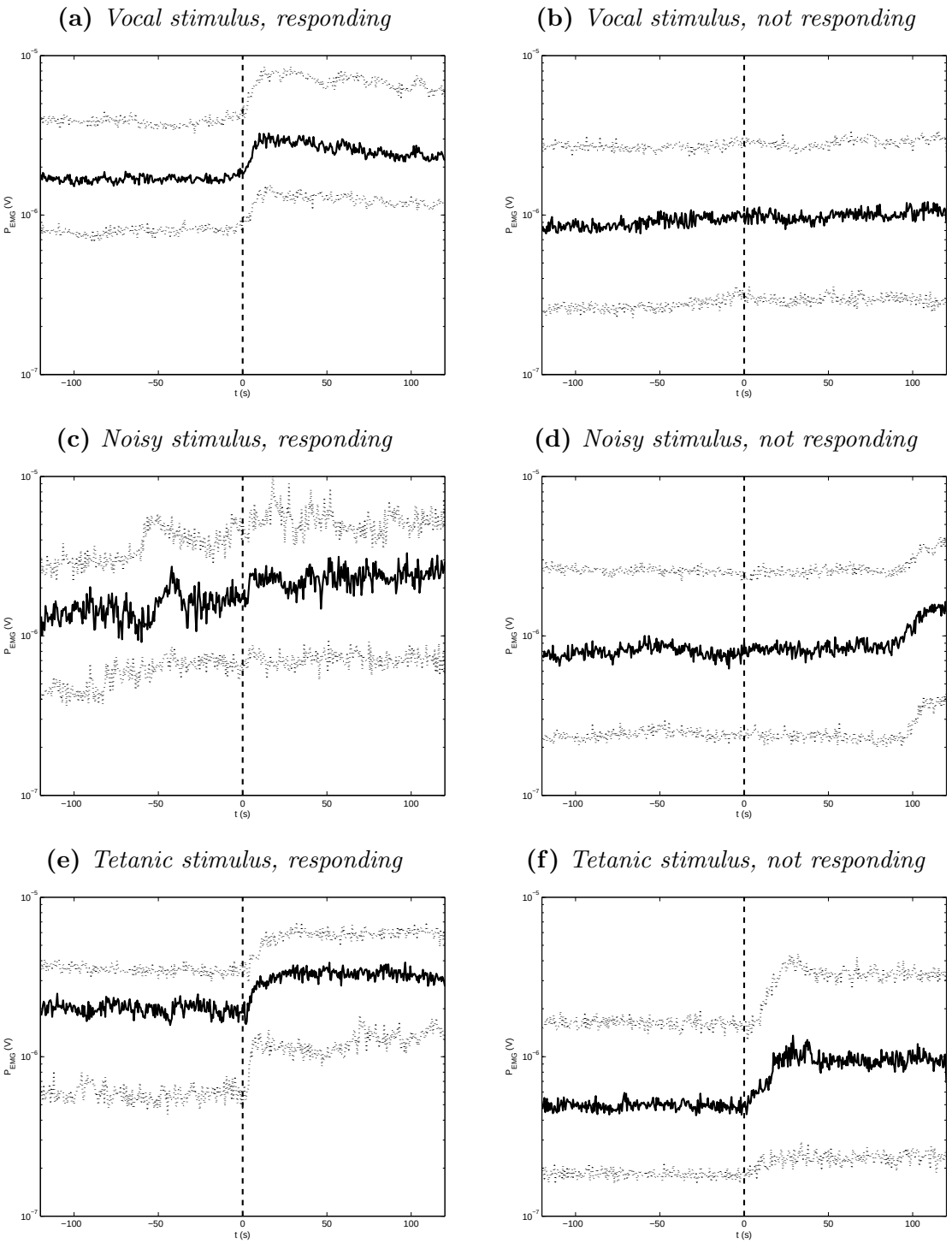

Figure 3.4: *Median, upper quartile and lower quartile stimulus related powers at top channel for different stimuli.*

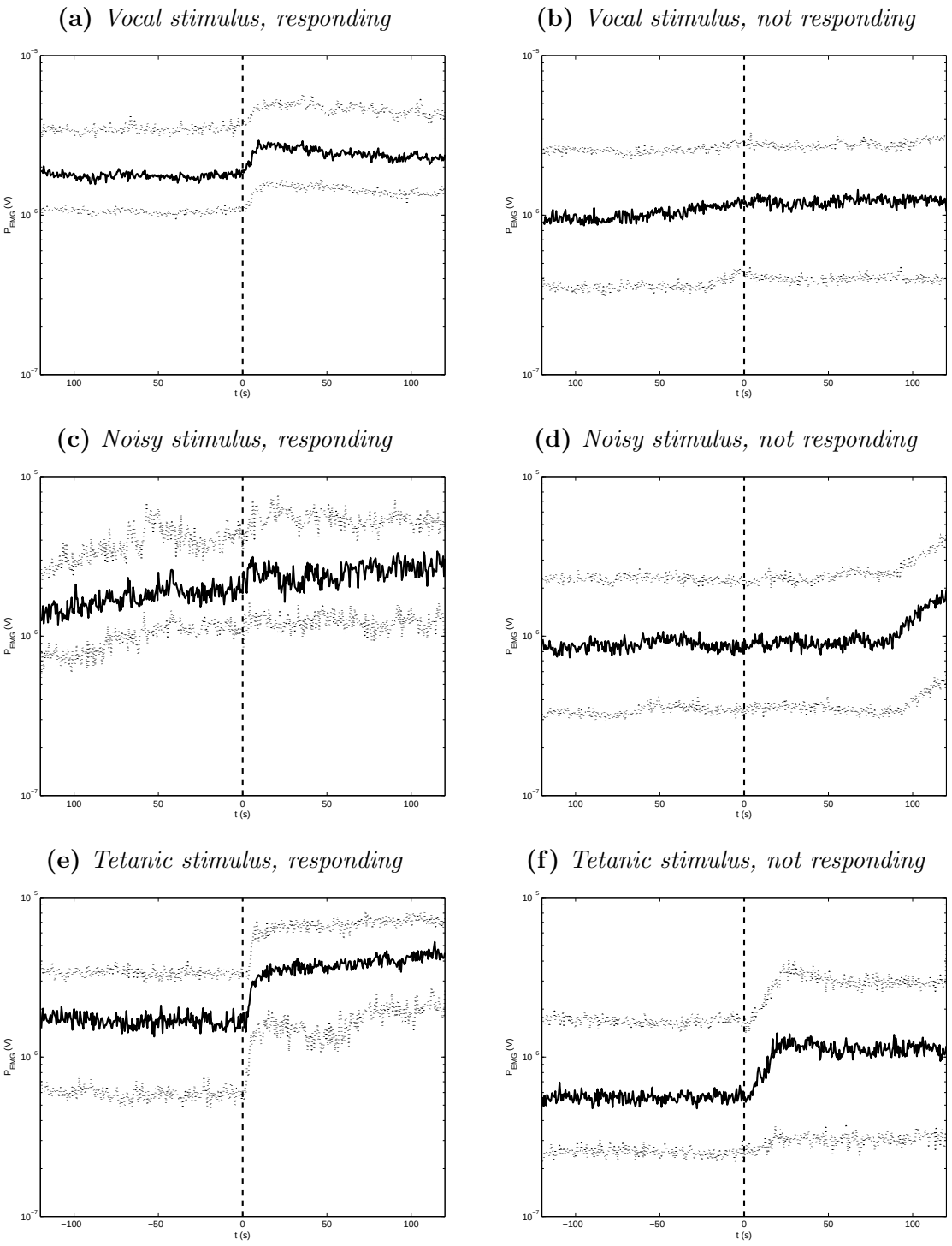

Figure 3.5: *Median, upper quartile and lower quartile stimulus related powers at bottom channel for different stimuli.*

Table 3.3: *Median (25 percentile, 75 percentile) difference between 30 seconds median power before and 30 seconds median power after stimulus at top channel in microvolts.*

	Response to stimulus	No response to stimulus
Vocal stimulus	$0.33(-0.00, 1.38), P = 0.000$	$0.00(-0.04, 0.05), P = 0.727$
Noisy stimulus	$0.06(0.00, 0.72), P = 0.010$	$0.00(-0.03, 0.05), P = 0.13$
TOF stimulus	-	$0.00(-0.02, 0.02), P = 0.248$
Tetanic stimulus	$0.66(0.04, 1.39), P = 0.000$	$0.07(0.00, 0.46), P = 0.000$

Table 3.4: *Median (25 percentile, 75 percentile) difference between 30 seconds median power before and 30 seconds median after stimulus at bottom channel.*

	Response to stimulus	No response to stimulus
Vocal stimulus	$0.23(-0.07, 0.86), P = 0.000$	$0.00(-0.05, 0.05), P = 0.964$
Noisy stimulus	$0.10(0.00, 0.71), P = 0.006$	$0.01(-0.02, 0.04), P = 0.092$
TOF stimulus	-	$0.01(-0.03, 0.02), P = 0.456$
Tetanic stimulus	$0.75(0.19, 2.40), P = 0.000$	$0.09(0.01, 0.47), P = 0.000$

feature aligned power, a visual examination of individual EMG responses reveals that there are several different kinds of responses to all stimuli.

Examples of individual EMG responses are shown in Appendix A. From these examples it is evident that in addition to form presented in median responses there are some fast changes and spiky activity in the EMG power. The fast action may be one cause of the sharp spike in time $t = 0$ of feature aligned response. There are also several other interesting phenomena that can be seen from individual cases but not from the median curves. There are several cases where EMG level rises before the stimulus. This may be due to activity that happens in the room before the stimulation. In case of the vocal stimulus one reason for pre-stimulus power onset could be nurse setting the headphones on the patient. At the other stimuli the activity before the stimulus could be due to preceding stimulus. There is also evidence on cases where there is a clear activity on one channel but not on the other.

Latencies to first outlier after the stimulus (Figure 3.7) are concentrated within first 15 seconds after stimulus. It seems that if there is a response, it is likely to occur relatively fast after the stimulus is given.

Median amplitude spectra of EMG power waveform at top channel is shown in Figure 3.8. Median amplitude spectra at bottom channel are essentially identical with top channel spectra. In all cases amplitude of the EMG power signal decreases almost linearly in a log-log plot as a function of frequency. In responsive patients, amplitude is at all frequencies clearly higher than when no response was detected by nurse. When given vocal or noisy stimulus (Figure 3.8 (a)) difference between responding and not responding cases is bigger at lower frequencies. At assessments where tetanic stimulus was given difference between responding and not responding

cases does not depend on frequency.

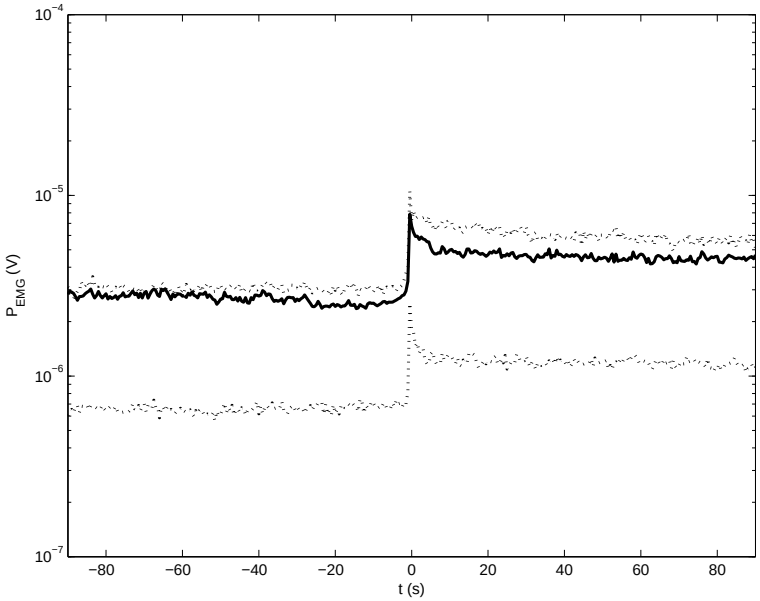

Figure 3.6: *Median, lower quartile and upper quartile feature aligned EMG power.*

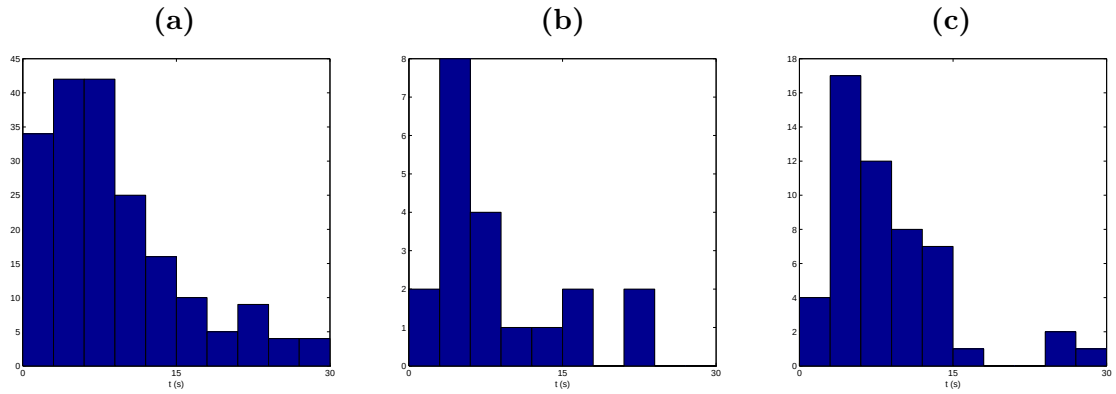

Figure 3.7: *Histograms of latency to first outlier (value after stimulus that is greater than 20 s average power plus three times the 20 s standard deviation before the stimulus). Latencies are measured from assessments where patients responded to vocal stimulus (a), noisy stimulus (b) or tetanic stimulus (c) and where first outlier was detected within 30 seconds after stimulus.*

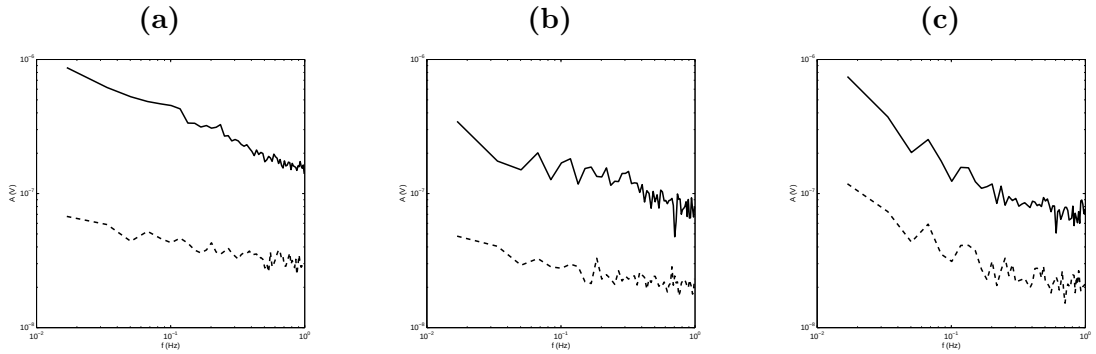

Figure 3.8: *Median amplitude spectra of EMG power waveforms for patients responding (solid line) and not responding (dashed line) to vocal stimulus (a), noisy stimulus (b) and tetanic stimulus (c).*

3.3 Conclusions

There are two distinctive differences in EMG waveforms between assessment groups where nurse has annotated a response to stimulus and where no visible response occurs. Patients responding to auditory or noxious stimulus typically show an obvious rise in level of EMG power. After vocal or tetanic stimulus the rise in EMG power is usually abrupt. However, after noisy stimulus the rise does not seem to be as clear. In part this this can be because of relatively small number samples (N = 30). Another possible explanation is that preceding vocal stimulus can change status of the patient in a way that can affect EMG power but does not cause a visible response. When patients do not respond to vocal or noisy stimulus there seldom is a rise in EMG power. Yet, when tetanic stimulus was given an abrupt rise in EMG power level occurred regardless whether the patient responded to the stimulus or not.

Another difference in EMG waveforms was that overall power level even before the stimulus was somewhat higher in responding patients than in not responding patients. This relation was similar in vocal, noisy and tetanic stimuli.

Train-of-four (TOF) stimulus did cause only three responses noted by nurse out of total 110 times the stimulus was given. A likely reason for this is that due to the study protocol TOF stimulus is preceded by vocal stimulus and by 105 dB noise stimulus. It is possible that all the patients who were liable to response to TOF stimulus already responded to either auditory stimuli.

Latency to the first outlier is most of the time less than fifteen seconds and mainly very small. Due to stimulus protocol and due to the fact that the time between the stimuli sometimes was not even a full minute it unfortunately is impossible to conclusively assess exactly how much there are late responses. Fortunately it is evident that if there is a response to the stimulus it is likely to be quite fast after the stimulus. A 30 second window should therefore be large enough to assess whether there is a response to stimulus and small enough to discount any activity due to

other stimuli. This window will be used in the second part of the study.

Amplitude spectra of data shows that both fast and slow components of EMG power wave differ between cases where patient responds to a given stimulus and where patient does not respond to stimulus. Difference between median spectra is greater in lower frequencies. This implies that long term activity is more pronounced differentiating factor.

Individual visual examination of EMG powers also shows that there are responsive patients whose EMG power level does not rise for a long period of time but instead there are only one or more transient power spikes after the stimulus. These spikes cannot obviously be seen clearly in median powers since the likelihood of transient spikes occurring at the same time after the stimulus is quite low. However the transient spikes may be important indication of patient awakening and reacting to given stimulus.

Research nurses were to annotate patients responsive to the stimuli if there was any gross purposeful movement to the stimulus. An interview of a research nurse revealed though that some subtler signals, such as small movements or changes in hemodynamic measures, were also sometimes marked as responses. This also implies that even the short time changes in the power may be significant.

4 Quantitative analysis of EMG changes after stimulus

The objective of this chapter is to derive computational methods for assessing patient responsiveness to given stimuli from descriptive properties of EMG power responses and to assess the predictive value of derived methods in assessing the depth of sedation in intensive care patients.

Descriptive analysis shows two distinct features in EMG power that are good candidates for automatic assessment of depth of sedation in intensive care patients. Typically after any given stimulus sharp rise in EMG power occurs followed by subsequent elevated power level if patient is responsive to the stimulus. However, similar reaction also occurs after tetanic stimulus even when patients do not visibly respond to the stimulus. Therefore parameters measuring mainly the change in power level are likely to have lower performance at these cases

Another feature that can be used to predict depth of sedation is general level of EMG power. Even power baseline before stimulus seems to be significantly higher in responsive patients than it is in non-responsive ones.

4.1 Materials and methods

Clinical data used in the analysis is the same that was used in descriptive analysis. Data preprocessing and artifact removal was also done in the same way.

4.1.1 Statistical methods

Two related performance measures were used to analyze efficiency of derived parameters as depth of sedation indicator. Prediction probability P_K [33] is a measure commonly used assessing performance of aesthetic depth indicators. P_K is a scaled variant of Kim's $d_{x \cdot y}$ [34] measure of association. Ideal anaesthetic depth indicator would be a monotonical function of underlying anaesthetic depth. However there is no direct way of measuring underlying anaesthetic depth but only indirect quantal observations can be used as reference value. Therefore an indicator can be seen being ideal if it perfectly predicts the observed anaesthetic depth. An illustration of this relationship can be seen in Figure 4.1.

Smith defines P_K as follows: "Let P_c, P_d, and P_{tx} be the respective probabilities that two data points drawn at random, independently and with replacement, from the population are a concordance, a discordance, or an x-only tie. The only other possibility is that the two data points are tied in observed depth y; therefore, the sum of P_c, P_d, and P_{tx} is the probability that the two data points have distinct values of observed anaesthetic depth, that is, that they are not tied in y." [33] After inserting these to definition of $d_{x \cdot y}$ and rescaling P_K can be calculated with Equation 4.1.

$$P_K = \frac{P_c + \frac{1}{2}P_{tx}}{P_c + P_d + P_{tx}} \tag{4.1}$$

In practice prediction probability P_K is calculated by replacing the probabilities in 4.1 with sample estimates. Estimates of P_K and its SE can be derived with several methods. In this study a Matlab program that uses jackknife method in the estimation. Jackknife method allows testing hypotheses with Student's t-test. For testing whether P_K differs statistically from 0.5, t-statistics $(P_{Kjack} - 0.5)/\sigma_{PKjack}$ can be used [33]. Comparing two indicators is also possible using the jackknife data.

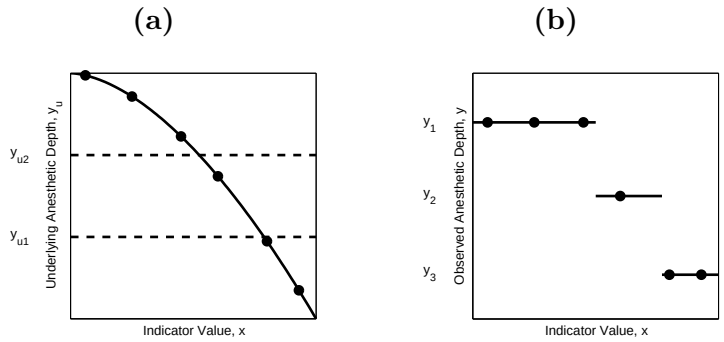

Figure 4.1: *Ideal relationships between indicator value and underlying anaesthetic depth (a) and quantal observed anaesthetic depth (b). Figure redrawn from [33].*

Similar to P_K receiver operating characteristic (ROC) measures such as specifity and sensitivity can be used to assess performance. Sensitivity is defined as number of true positives divided by sum of true positives and false negatives. Similarly specifity is number of true negatives divided by sum of true negatives and false positives. One drawback of ROC measures is that they can only be used with dichotomous reference. In this study used reference data is nurse annotations whether the patient responded to given stimulus or not. Therefore ROC measures can be used as well as P_K.

Specifity and sensitivity are usually plotted in a ROC curve where sensitivity is plotted against 1 - specifity. Percentages are also used regularly as scale instead of the fraction. One measure of prediction accuracy is so called area under curve (AUC), which is the fraction of unit square under ROC curve. This measure is closely related to P_K. For dichotomous data parametrical or trapezoidal AUC is numerically equivalent to P_K [35]. In this thesis P_K is used to assess the performance of all derived parameters and ROC curves are plotted for the best performing parameter for vocal, noisy and tetanic stimuli.

4.1.2 Data processing

Several different parameters were derived from the features identified in chapter three. Performance of the parameters was assessed with Smith's P_K value. No performance analysis was done with TOF stimulus due to a small number of patients responding to the stimulus.

A 60 second median power before the stimulus was chosen as a baseline parameter. This parameter is independent of the stimulus. Length of the baseline was chosen to be relatively long so that the parameter would give a better estimate of overall EMG power level preceding the stimulus.

Another basic parameter was 30 second median power after stimulus. Length of the window after stimulus was chosen based on the results of descriptive analysis. 30 seconds is long enough to catch most of the relevant changes but short enough to avoid changes actually resulting from later stimulation. Median power following the stimulus can also be seen as a combination of EMG baseline and chance following the stimulus.

As a measure of absolute change in EMG power level 60 second median before the stimulus was reduced from the 30 second median power after the stimulus. Respectively as measure of relative change of power level the 30 second median after the stimulus was divided by the 60 second median after the stimulus.

Median is a good parameter for testing changes in overall level of the EMG power. However it does not detect fast changes regardless however big they are. Maximum power after stimulus is therefore analysed as well. Similarly the changes to the baseline level are analysed. The same baseline of 60 median before the stimulus is reduced to measure absolute change to the maximum power. Relative change is analysed similarly as well.

EMG power in the two time windows are defined as $P_0 = P(t), -60s \leq t < 0$ (baseline power) and $P_x = P(t), 0 \leq t < 30$ (response power). A summary of tested parameters is shown in Table 4.1.

Performance of responsiveness value (Equations 2.1 and 2.2) prior to assessments was compared to best performing parameter. One minute minimum filtering was used calculating $P_{arousal}$ (Equation 2.1) and 30 minutes moving average filter was used calculating responsiveness value (Equation 2.2). Because 30 minutes of valid power data prior to assessment was needed, responsiveness value could not be calculated from all assessments. Number of the assessments where responsiveness value was available is shown in Table 4.2. Prediction probability P_K of responsiveness is compared to corresponding P_K of the best performing stimulus related response measurement.

Table 4.1: *Parameters used as measure of depth of sedation.*

Parameter	Description
$\mu_{1/2}(P_0)$	60 second median before stimulus.
$\mu_{1/2}(P_x)$	30 second median after stimulus.
$\mu_{1/2}(P_x) - \mu_{1/2}(P_0)$	30 second median after stimulus minus 60 second median before stimulus.
$\mu_{1/2}(P_x)/\mu_{1/2}(P_0)$	30 second median after stimulus divided by 60 second median before stimulus.
$\max(P_x)$	30 second maximum after stimulus.
$\max(P_x) - \mu_{1/2}(P_0)$	30 second maximum after stimulus minus 60 second median before stimulus.
$\max(P_x)/\mu_{1/2}(P_0)$	30 second maximum after stimulus divided by 60 second median before stimulus.

Table 4.2: *Number of assessments where responsiveness value was available.*

	Response	No response
Vocal stimulus	214	157
Noisy stimulus	25	118
Tetanic stimulus	56	65

4.2 Results

Prediction probabilities for all derived parameters can be seen in Table 4.3. All P_K values differ statistically significantly from 0.5 ($P \leq 0.001$). All of the parameters have at least some predictive value.

For vocal stimulus 60 seconds median power before stimulus performs relatively poorly but even this parameter still has some predictive value. Median level after the stimulus and changes in the median level have prediction probabilities around 0.7 showing clear correlation. Maximum after stimulus and maximum compared to baseline level are clearly the best predictors for the vocal stimulus having P_K of around 0.9. It seems clear that for differentiating lightly and deeply sedated patients maximum EMG power after the stimulus is a very good parameter.

For most of the parameters the top channel seems to perform a bit better than the bottom channel as a indicator of response to vocal stimulus. However the difference is statistically significant (P ¡ 0.05) only with the parameter $\mu_{1/2}(P_x)/\mu_{1/2}(P_0)$. Better performance of top channel and the insignificance of the difference implies that activity of *Musculus Frontalis* in response to the vocal stimulus is most likely the dominant source of EMG around the electrodes.

Prediction probabilities for noisy stimulus are in general lower than for vocal stimulus. Maximum and difference of maximum and median baseline have the highest

Table 4.3: *$P_K(SE)$ values for all tested parameters at top and bottom channels*

	Top channel	Bottom channel
Vocal stimulus		
$\mu_{1/2}(P_0)$	0.64(0.03)	0.65(0.03)
$\mu_{1/2}(P_x)$	0.73(0.03)	0.71(0.03)
$\mu_{1/2}(P_x) - \mu_{1/2}(P_0)$	0.70(0.03)	0.66(0.03)
$\mu_{1/2}(P_x)/\mu_{1/2}(P_0)$	0.72(0.03)	0.67(0.03)
$\max(P_x)$	0.88(0.02)	0.87(0.02)
$\max(P_x) - \mu_{1/2}(P_0)$	0.92(0.01)	0.90(0.02)
$\max(P_x)/\mu_{1/2}(P_0)$	0.85(0.02)	0.86(0.02)
Noisy stimulus		
$\mu_{1/2}(P_0)$	0.66(0.05)	0.69(0.05)
$\mu_{1/2}(P_x)$	0.70(0.05)	0.70(0.05)
$\mu_{1/2}(P_x) - \mu_{1/2}(P_0)$	0.64(0.07)	0.72(0.06)
$\mu_{1/2}(P_x)/\mu_{1/2}(P_0)$	0.65(0.07)	0.71(0.06)
$\max(P_x)$	0.77(0.05)	0.81(0.04)
$\max(P_x) - \mu_{1/2}(P_0)$	0.77(0.06)	0.85(0.04)
$\max(P_x)/\mu_{1/2}(P_0)$	0.73(0.06)	0.79(0.05)
Tetanic stimulus		
$\mu_{1/2}(P_0)$	0.71(0.04)	0.70(0.04)
$\mu_{1/2}(P_x)$	0.74(0.04)	0.77(0.04)
$\mu_{1/2}(P_x) - \mu_{1/2}(P_0)$	0.67(0.05)	0.70(0.05)
$\mu_{1/2}(P_x)/\mu_{1/2}(P_0)$	0.63(0.05)	0.66(0.05)
$\max(P_x)$	0.75(0.04)	0.79(0.04)
$\max(P_x) - \mu_{1/2}(P_0)$	0.75(0.04)	0.79(0.04)
$\max(P_x)/\mu_{1/2}(P_0)$	0.65(0.05)	0.71(0.04)

P_K values as well. Interestingly median before the stimulus seems to have similar predictive power to changes in median and median after the stimulus. For every parameter bottom channel gives greater or as good P_K values than the top channel but the difference is statistically significant with the parameter $\max(P_x) - \mu_{1/2}(P_0)$. This indicates that *Musculus Corrugator Supercilii* activity might have some role in these cases.

Prediction probabilities in cases where tetanic stimulus was given were lower than for either auditory stimulus. This is in line with the finding in descriptive analysis that the median EMG powers for responsive and non-responsive patients are similar. However the P_K values still differed from 0.5 (P ¡ 0.05). The best performing parameters were again $\max(P_x)$ and $\max(P_x) - \mu_{1/2}(P_0)$ with $P_K = 0.79$. With all parameters except for 60 s median before the stimulus the bottom channel performs better than the top channel. The difference is statistically significant (P ¡ 0.05) for parameters $\mu_{1/2}(P_x)$, $\max(P_x)$, $\max(P_x) - \mu_{1/2}(P_0)$ and $\max(P_x)/\mu_{1/2}(P_0)$. Although the difference is not statistically significant for all the parameters this im-

plies significant corrugator activity after the stimulus. This is in line with the role of corrugator in expressing pain and discomfort.

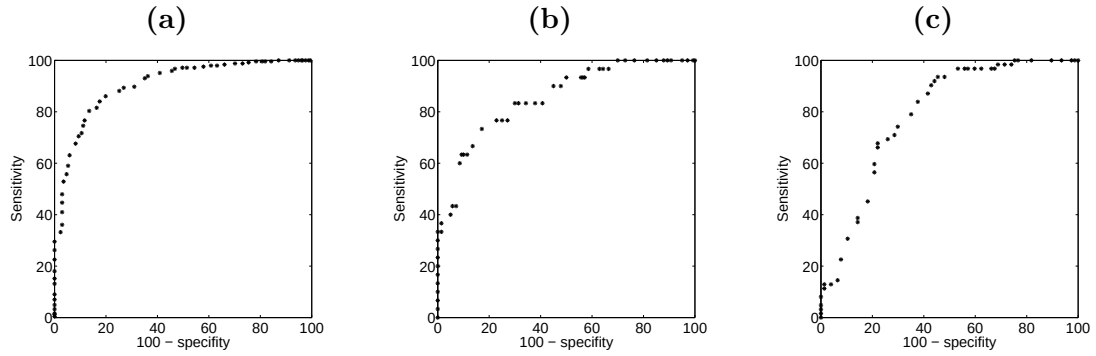

Figure 4.2: *ROC curves for 30 second maximum after the stimulus EMG minus 60 second median before the stimulus when vocal (a), noisy (b) and tetanic (c) stimulus are given.*

The best performing parameter is 30 second maximum after the stimulus minus 60 second median before the stimulus. Different linear combinations of 30 second maximum after the stimulus and 60 second median before the stimulus, were also tested. P_K values were calculated from all linear combinations $A \cdot \max P_x + B \cdot \mu_{1/2} P_0$, where A and B were real values at 0.1 intervals between -2 and 2. P_K values were the best when $B \approx -A, A \neq 0$ and $B \neq 0$. Differences between the P_K values when A/B was close to -1 were negligible.

To further illustrate properties of difference of 30 second maximum after the stimulus and 60 second median before the stimulus as response detection parameter ROC curves were drawn. The curves were drawn only based on the bottom channel data. Sensitivity was defined as percentage of correct assessments to responsive group with a given threshold. Corresponding specifity value was defined as percentage of correct assessments to non-responsive group with a given threshold. Sensitivity was plotted as function of 100 - specifity. Resulting ROC curves are shown in Figure 4.2. ROC curves show that both good sensitivity and specifity can be achieved in detecting response to vocal stimulus. Relatively good specifity and sensitivity related to response to a noisy stimulus are also achievable. However ROC curve related to tetanic stimulus is quite asymmetric. While a good sensitivity can be achieved with 50 % or so specifity, increases in specifity after 50 % cause a sharp drop in sensitivity. This finding is in line with qualitative results that there is a good amount of EMG power changes in unresponsive patients similar to the changes in responsive ones.

The best performing parameter performs statistically significantly better than corresponding responsiveness value before the stimulus with both auditory stimuli (Table 4.4). Although the best performing parameter has higher P_K than responsiveness has in cases where tetanic stimulus was given the difference was not statistically significant. Calculated responsiveness values are in line with those reported by

Lapinlampi [2] for patients with suspected brain dysfunction.

Table 4.4: *$P_K(SE)$ comparison of responsiveness value and the best performing parameter*

	$max(P_x) - \mu_{1/2}(P_0)$	Responsiveness value	P
Vocal stimulus	0.91(0.01)	0.70(0.03)	0.00
Noisy stimulus	0.82(0.05)	0.64(0.05)	0.00
Tetanic stimulus	0.78(0.04)	0.73(0.05)	0.23

Difference between 30 second maximum EMG power after the vocal stimulus and 60 second median power before the stimulus as predictor of patient reaction to noisy stimulus gives $P_K(SE)$ values of 0.72(0.05) at top channel and 0.74(0.05) at bottom channel. P_K values of both channels differ statistically significantly from 0.5 P value for both channels being lower than 0.001. Although corresponding $P_K(SE)$ values for 60 second medium before vocal stimulus are 0.65(0.05) at top channel and 0.68(0.05) at bottom channel, the difference between predictive properties is not statistically significant. P values of pairwise comparison are 0.16 at top channel and 0.09 at bottom channel.

4.3 Conclusions

The results prove that it is possible to design an algorithm capable of reliably detecting patient reactions to given stimuli. Along with a stimulation protocol this algorithm could be used as an depth of sedation measure similar to Ramsay or RASS scoring systems. From the analysis it is evident that a parameter that takes in account both fast changes and the baseline change will perform likely most similarly to nurse's assessment of whether the patient responded or not. Maximum EMG power after the stimulus is one example of that kind of parameter, but it may be possible to find even better parameters.

The differences between the channels were in most of the cases statistically significant. However the differences were usually consistent. Top channel was a bit better with vocal stimulus and the bottom channel was better with the others. This indicates that the *frontalis* activity is related to reactions to vocal stimulus. This may be due to startlement or due to the fact that the stimulus tells the patient to open their eyes. The results also confirm Lapinlampi's results that EMG activity of *Musculus Frontalis* is a good classifier of lightly and deeply sedated patients.

On the other hand the bottom channel was better with both tetanic and noisy stimulus. In both cases the differences are most likely due to *corrugator* activity. *Corrugator* is associated with expressions of pain, which explains the differences with tetanic stimulus. The differences in after noisy stimulus can also be due to pain or discomfort. The stimulus was white noise at 105 dB that can be almost painfully loud. In any case it is highly uncomfortable to the most people.

Most of the methods work best with vocal stimulus which corresponds to separating

between Ramsay levels of 1 to 3 and 4 to 6. In contemporary clinical practice this is the most significant classification since patients with Ramsay score of four or higher are usually considered oversedated. The good performance with vocal stimulus implies that the parameters are clinically relevant.

As far as P_K values go none of the parameters were particularly good with tetanic stimulus. Descriptive analysis shows that there were significant number of cases where there was clear EMG reaction to stimulus although nurse did not detect any reaction. This can be confirmed to be the reason of poor P_K values by ROC analysis. ROC curve of tetanic stimulus shows that high sensitivity is easily attainable with reasonable specifity, but high specifity can be achieved only by almost completely disregarding the sensitivity.

5 Discussion and conclusions

Measuring depth of sedation in intensive care patients is a difficult task for number of reasons. There is still some uncertainty of how regulation of consciousness happens and what kinds of physiological changes loss of consciousness causes. Intensive care patients are severely ill on arrival and tend to be in ICU for long periods of time. This likely causes several differences in biomedical measurements compared to healthy patients or even ones under general anesthesia. This has caused a lot of difficulty in developing reliable measure of depth of sedation. The measures applicable in general practice or anesthesia may not be suitable for intensive care patients at all.

This study has focused on understanding the connection between subjective measurements of depth of sedation and electromyography (EMG) of facial muscles. The results conclusively prove that a *frontalis* EMG power rise correlates highly with nurse observations when vocal stimulus is given. Correlation is also clear when a noisy stimulus is given. When tetanic stimulus is given there seems to be a lot of patients whose EMG power rises clearly but who the nurses have classified as unresponsive.

Several measures were developed based on characteristics of EMG power response to stimulus. Out of the measures the most reliable one was 30 second maximum EMG power after the stimulus minus 60 second median power baseline before the stimulus. This measure reliably classifies patients to responsive and non-responsive ones when a vocal stimulus is given. The measure works still well with noisy stimulus but it is not as reliable as with vocal stimulus. Part of the unreliability may be due some of cases where EMG power rose after vocal stimulus and was therefore already elevated when noisy stimulus was given. Tetanic stimulus was the most difficult to classify reliably. This is most likely due to the fact that there was a clear EMG response in many cases where nurse did not notice any response.

5.1 Critical analysis of the study

There are several limitations to this study. First, the studied patient group was limited to a specific illness and does not represent well whole intensive care patient population. Moreover all patients suffered from hepatic encephalopathy, which may cause abnormalities in neurological activity, which in turn could change behaviour of facial muscles and measured EMG. Although further studies will be needed to test whether these results are applicable to different patient groups, the results are encouraging. In previous studies encephalopathic patients have proven to be especially difficult patient group to monitor automatically. Successful classification with this patient group suggests that the measure is robust and could well be suitable for other patients as well.

The second limitation of the study is that only one stimulation protocol was used. Hence no comprehensive analysis of sensitivity to different stimuli was done. In

particular the train-of-four stimulus did not cause any responses and could not be analysed in same length that other stimuli. This suggests that a train-of-four probably is not a stronger stimulus than a loud noise, as was believed prior to the study. However the strength of stimuli was not compared in the study and therefore no conclusive findings were made.

One limitation to the study is that the proposed method was not compared with all relevant technologies directly. However, a comparison between responsiveness measure was carried out. Responsiveness itself has been compared with BIS and entropy measurements and performed significantly better than either. Since EMG stimulus response was more reliable than responsiveness parameter with the patients involved in the study this suggests that it would probably be also more reliable than BIS or entropy.

5.2 An alternative method for measuring level of sedation

This study proves conclusively that upper frontal EMG power responses correlate with visible responses that are key in subjective sedation measurements. Although no clinical study on effectiveness of automated sedation measurement was done clinical viability of the method can be assessed.

Especially interesting classification is separating patients from RASS 3 or higher from those with RASS 4 or lower. This distinction is equal with classifying patients to those responsive to vocal stimulus and those unresponsive to vocal stimulus. The study shows that stimulus related EMG power response is a good measure for this purpose.

The best performing parameter, 30 second maximum after the stimulus minus 60 second minimum before the stimulus, is closely related to $P_{arousal}$ 2.1 defined by Lapinlampi [2]. Maximum $P_{arousal}$ 30 seconds after the stimulus with 60 seconds minimum window will likely give almost equal P_K values.

Although the measure seems relatively similar to auditory evoked potential, there is a significant difference in the approach. Auditory evoked potential is EEG measurement which requires several samples and high signal quality to work properly. Stimulus related EMG response on the other hand requires only one run of the protocol and does not require high quality control on the signal. From the comparison of top and bottom channels it seems that EMG power is relatively insensitive to some artifacts such as eye movements.

The study shows that it is possible to determine the depth of sedation from EMG amplitude. Moreover a reliable mathematical method for finding stimulus response was found. This study also shows that stimulus related EMG responses work well as measure of depth of sedation in encephalopathic patients, a patient group which has proven difficult for other methods. In future stimulus related EMG response could be used in conjunction of a continuous sedation measurement e.g. responsiveness. This approach would give continuous but probably lagging estimate of patient's sedation

level and on demand discrete reaction detection that would give an instantaneous value of patient response.

5.3 Painful stimulus in measuring depth of sedation

One interesting finding of this study is the relatively poor correlation between nurse annotations and frontal EMG power when patient was given a tetanic stimulus. More careful analysis shows that there are similar rises in EMG power in unresponsive patients as in responsive ones. There are several possible explanations for this phenomenon.

One possible explanation is that all patients in this study have at least some level of encephalopathy. In severe cases encephalopathy may lead to a coma which may also cause patients to be unresponsiveness. Since the patients who the painful stimulus was given are clearly quite unconscious it is extremely hard to assess whether this is due to sedation or due to underlying illness. It is possible that comatic patients' pain reflexes differ from deeply sedated ones.

Another possibility is that the long time in intensive care unit may cause some of changes to muscle responses. Specifically polyneuropathies and polymyopathies are known to cause problems to intensive care patients [36]. It is possible that these or other conditions developing in ICU might cause degradation of muscular responses in such manner that patient shows no visible response while the reaction to the stimulus is evident in EMG power.

The results of this study indicate that EMG after painful stimulus differs significantly from other stimuli especially in the cases where nurses cannot see any response from the patient. It is evident that more clinical studies in relevance of painful stimuli in level of sedation measures are needed.

5.4 Future work

Although the study shows that automated stimulus protocol coupled with upper frontal EMG measurement can be used as a reliable measure of level of sedation, no work was done to optimise the performance of the signal processing algorithm. This is an obvious line for future research. In this study it was shown that a very reliable classification of patients to lightly sedated and deeply sedated is possible. The parameters developed can still be improved. The study to improve the parameter should also use larger corpus of data including intensive care patients with several different conditions.

More importantly the study did not compare different stimulus protocols. The results indicate that train-of-four stimulus is not a stronger stimulus than loud auditory noise. Other than that the protocol seemed to be appropriate, but there is still room for improvement. It is also not known what should be the time interval between the stimuli. Future research should address the issues with the protocols.

There may also be other stimuli that could be used as well. A comprehensive study on responses to different stimuli protocols would likely be insightful.

Before the results of this study are clinically applicable comparative study between present day sedation practice and automated sedation measurement is required. The study shows promising correlation with nurse assessed responses to stimuli, which is the basis of the current depth of sedation measures. This suggests that EMG based stimulus responses might have much clinical potential. This has to be conclusively proven in future research.

Results also give insights for developing the responsiveness parameter further.

References

[1] Genis Carrasco. Instruments for monitoring intensive care unit sedation. *Critical Care*, 4:217–225, 2000.

[2] Petteri Lapinlampi. Methods for assessing the depth of natural and drug-induced sleep based on electroencephalographic signal. Master's thesis, Helsinki University of Technology, Espoo, 2005.

[3] Gerard J. Tortora and Sandra Reynolds Grabowski. *Principles of Anatomy & Physiology.* John Wiley & Sons, Inc, Hoboken, 10th edition, 2003.

[4] Henry Gray. *Gray's Anatomy.* John W. Parker and son., 1858.

[5] Eric R Kandel and James H Schwartz and Thomas M. Jessell, editor. *Principles of neural science.* McGraw-Hill, New York, 4th edition, 2000.

[6] P. Brown, J. C. Rothwell, P. D. Thompson, T. C. Britton, B. L. Day, and C. D. Marsden. New Observations on the Normal Auditory Startle Reflex in Man. *Brain*, 114(4):1891–1902, 1991.

[7] Jörg B. Schulz and Allen I. Arieff. Metabolic and Toxic Encephalopathies. In Thomas Brandt and Louis R. Caplan and Johannes Dichgans and H. Christoph Diener and Christopher Kennard, editor, *Neurological Disorders*, pages 991–1009. Elsevier Inc., 2nd edition, 2003.

[8] Ernst Niedermeyer and Fernando Lopes da Silva. *Electroencephalography : Basic Principles, Clinical Applications, and Related Fields.* Lippincot Williams & Wilkins, Philadelphia, 5th edition, 2005.

[9] Reginald G. Bickford and Hugh R. Butt. Hepatic Coma: The Electroencephalographic Pattern. *The Journal of Clinical Investigation*, 34(6):790–799, 1955.

[10] Stephen M. Riordan and Roger Williams. Treatment of Hepatic Encephalopathy. *The New England Journal of Medicine*, 337(7):473–479, 1997.

[11] William F Brown. *The Physiological and Technical Basis of Electromyography.* Butterworth Publishers, Stoneham, 1984.

[12] John T. Cacioppo and Louis T. Tassinary and Alan J. Fridlund. The skeletomotor system. In John T. Cacioppo and Louis T. Tassinary, editors, *Principles of psychophysiology: Physical, social, and inferential elements* , pages 325–384. Cambridge University Press, New York, 1990.

[13] Alan D. Brook and Thomas S. Ahrens and Robyn Schaiff and Donna Prentice and Glenda RN Sherman and William Shannon and Marin H Kollef. Effect of a nursing-implemented sedation protocol on the duration of mechanical ventilation. *Critical Care Medicine*, 27(12):2609–2615, 1999.

[14] Marin H. Kollef and Nat T. Levy and Thomas S. Ahrens and Robyn Schaiff and Donna Prentice and Glenda Sherman. The use of continuous IV sedation is associated with prolongation of mechanical ventilation. *Chest*, 114:541–548, 1998.

[15] Deborah J. Cook, Stephen D. Walter, Richard J. Cook, Lauren E. Griffith, Gordon H. Guyatt, David Leasa, Roman Z. Jaeschke, and Christian Brun-Buisson. Incidence of and Risk Factors for Ventilator-Associated Pneumonia in Critically Ill Patients. *Annals of Internal Medicine*, 129(6):433–440, 1998.

[16] Marlies E. Ostermann and Sean P. Keenan and Roxanne A. Seiferling and William J. Sibbald. Sedation in the intensive care unit: a systematic review. *Journal of American Medical Association*, 283(11):1451–1459, 2000.

[17] S. Murdoch and A. Cohen. Intensive care sedation: a review of current British practice. *Intensive Care Medicine*, 26:922–928, 2000.

[18] Bernard de Jonghe and Sylvie Bastuji-Garin and Pascal Fangio and Jean-Claude Lacherade and Julien Jabot and Corinne Appéré-De-Vecchi and Nathalie Rocha and Hervé Outin. Sedation algorithm in critically ill patients without acute brain injury. *Critical Care Medicine*, 33(1):120–127, 2005.

[19] Judith Jacobi and Gilles L. Fraser and Douglas B. Coursin and Richard R. Riker and Dorrie Fontaine and Eric T. Wittbrodt and Donald B. Chalfin and Michael F. Masica and H. Scott Bjerke and William M. Coplin and David W. Crippen and Barry D. Fuchs and Ruth M. Kelleher and Paul E. Marik and Stanley A. Nasraway, Jr and Michael J. Murray and William T. Peruzzi and Philip D. Lumb. Clinical practice guidelines for the sustained use of sedatives and analgesics in the critically ill adult. *Critical Care Medicine*, 30(1):119–141, 2002.

[20] John P. Kress and Anne S. Pohlman and Michael F. O'Connor and Jesse B. Hall. Daily interruption of sedative infusions in critically ill patients undergoing mechanical ventilation. *The New England Journal of Medicine*, 342(20):1471–1477, 2000.

[21] Timothy D Girard, John P Kress, Barry D Fuchs, Jason W W Thomason, William D Schweickert, Brenda T Pun, Darren B Taichman, Jan G Dunn, Anne S Pohlman, Paul A Kinniry, James C Jackson, Angelo E Canonico, Richard W Light, Ayumi K Shintani, Jennifer L Thompson, Sharon M Gordon, Jesse B Hall, Robert S Dittus, Gordon R Bernard, and E Wesley Ely. Efficacy and safety of a paired sedation and ventilator weaning protocol for mechanically ventilated patients in intensive care (Awakening and Breathing Controlled trial): a randomised controlled trial. *Lancet*, 371:126–134, 2008.

[22] Curtis N. Sessler and Mark S. Gosnell and Mary Jo Grap and Gretchen M. Brophy and Pam V. O'Neal and Kimberly A. Keane and Eljim P. Tesoro and R. K. Elswick. The Richmond Agitation-Sedation Scale: Validity and Reliability

in Adult Intensive Care Unit Patients. *American Journal of Respiratory and Critical Care Medicine*, 166:1338–1344, 2002.

[23] M A E Ramsay and T M Savege and B R J Simpson and R Goodwin. Controlled Sedation with Alphaxalone-Alphadolone. *British Medical Journal*, 2(5920):656–659, 1974.

[24] B. de Jonghe and D. Cook and C. Appere-De-Vecchi and G. Guyatt and Maureen Meade and H. Outin. Using and understanding sedation scoring systems: a systematic review. *Intensive Care Medicine*, 26(3):275–285, 2000.

[25] C de Deyne and M Struys and J Decruyenaere and J Creupelandt and E Hoste and F Colardyn. Use of continuous bispectral EEG monitoring to assess depth of sedation in ICU patients. *Intensive Care Medicine*, 24:1294–1298, 1998.

[26] M L Riess and U A Graefe and C Goeters and H van Aken and H G Bone. Sedation assessment in critically ill patients with bispectral index. *European Journal of Anaesthesiology*, 19:18–22, 2002.

[27] Dirk Frenzel and Clemens-A. Greim and Christian Sommer and Kerstin Bauerle and Norbert Roewer. Is the bispectral index appropriate for monitoring the sedation level of mechanically ventilated surgical ICU patients? *Intensive Care Medicine*, 28:178–183, 2002.

[28] H. Viertiö-Oja, V. Maja, M. Särkelä, P. Talja, N. Tenkanen, H. Tolvanen-Laakso, M. Paloheimo, A. Vakkuri, A. Yli-Hankala, and P. Meriläinen. Description of the Entropy algorithm as applied in the Datex-Ohmeda S/5 Entropy module. *Acta Anaesthesiologica Scandinavica*, 48(2):154 – 161, 2004.

[29] Timothy S Walsh and Pamela Ramsay and T Petteri Lapinlampi and Mika O K Särkelä and Hanna E Viertiö-Oja and Pekka T Meriläinen. An assessment of the validity of spectral entropy as a measure of sedation state in mechanically ventilated critically ill patients. *Intensive Care Medicine*, 34(2):308–315, 2008.

[30] A.M. Schulte-Tamburen, J. Scheier, J. Briegel, D. Schwender, and K. Peter. Comparison of five sedation scoring systems by means of auditory evoked potentials. *Intensive Care Medicine*, 25:377–382, 1998.

[31] Markku Paloheimo. Quantitative surface electromyography (qEMG): applications in anaesthesiology and critical care. *Acta Anaesthesiologica Scandinavica*, 34:suppl 93, 1990.

[32] Frank Wilcoxon. Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6):80–83, 1945.

[33] Warren D Smith and Robert C Dutton and Ty N Smith. Measuring the Performance of Anesthetic Depth Indicators. *Anesthesiology*, 84(1):38–51, 1996.

[34] Jae-On Kim. Predictive Measures of Ordinal Association. *The American Journal of Sociology*, 76(5):891–907, 1971.

[35] Warren D. Smith, Robert C. Dutton, and N. Ty Smith. A measure of association for assessing prediction accuracy that is a generalization of non-parametric ROC area. *Statistics in Medicine*, 15:1199–1215, 1996.

[36] K. Berek, J. Margreiter, J. Willeit, A. Berek, E. Schmutzhard, and N.J. Mutz. Polyneuropathies in critically ill patients: a prospective evaluation. *Intensive Care Medicine*, 22(9):849–855, 1996.

Appendix A Examples of different stimulus related EMG powers

This appendix shows several examples of stimulus related EMG power responses. The graphs shown were picked to illustrate some phenomena evident in the data that cannot be seen from the median or quartile powers as described in chapter 3. The figures include all three cases where patient was responsive to train-of-four (TOF) stimulus. The graphs are drawn from the signal at the bottom channel except for one figure demonstrating distinct difference between the channels.

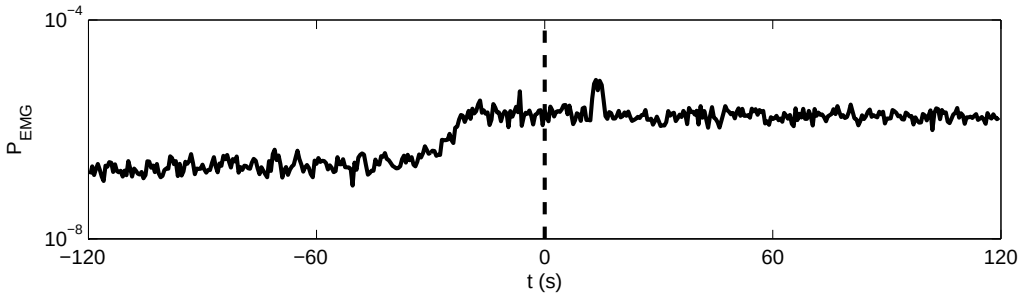

Figure A1: *A response to vocal stimulus where EMG onset is well before the stimulus is given. This graph still shows a short spike after the stimulus.*

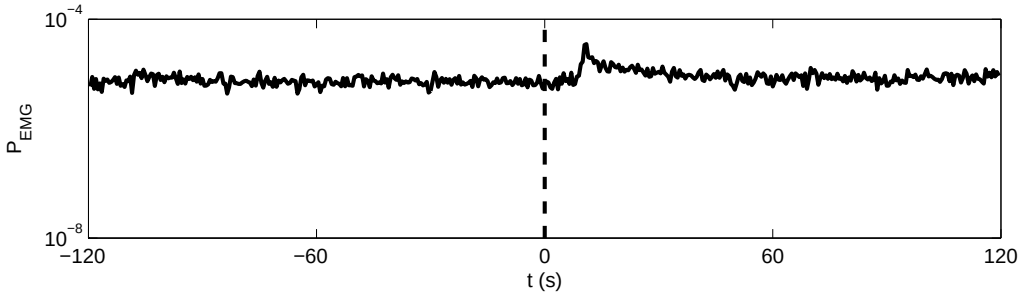

Figure A2: *A response to vocal stimulus with very high baseline power.*

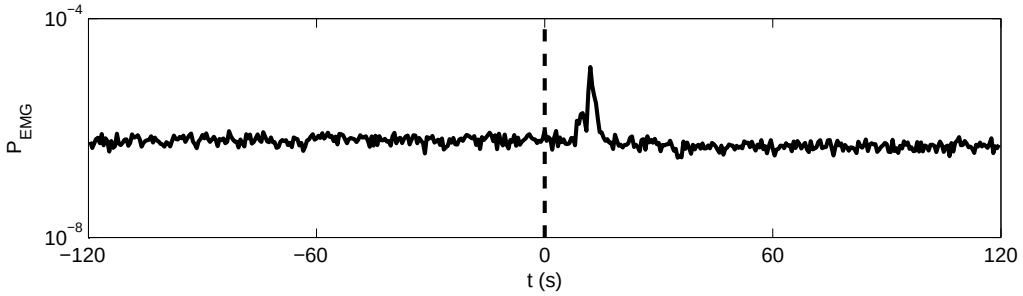

Figure A3: *A response to vocal stimulus, which shows only a few second long spike after the stimulus.*

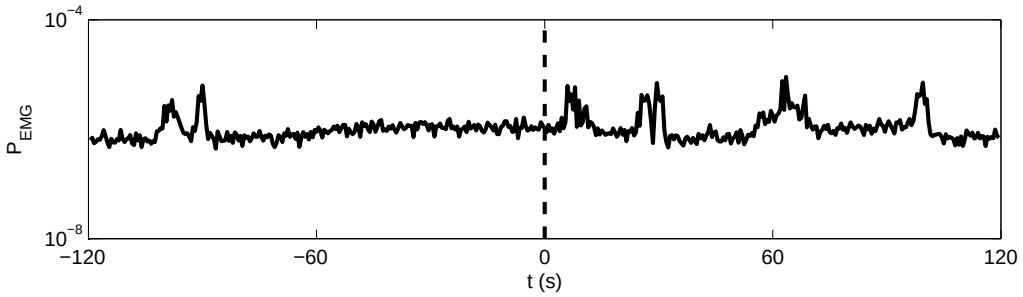

Figure A4: *A response to vocal stimulus, which has several quick spikes after the stimulus.*

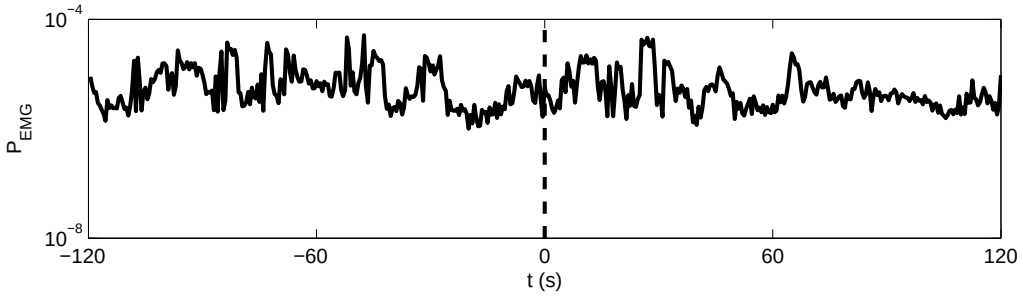

Figure A5: *A response to vocal stimulus showing very varying activity both before and after the stimulus.*

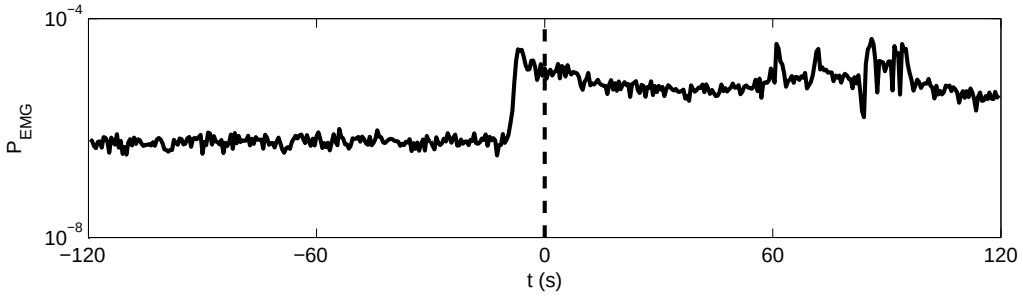

Figure A6: *Vocal stimulus, no response to stimulus. A clear rise of EMG power occurs just before the stimulus.*

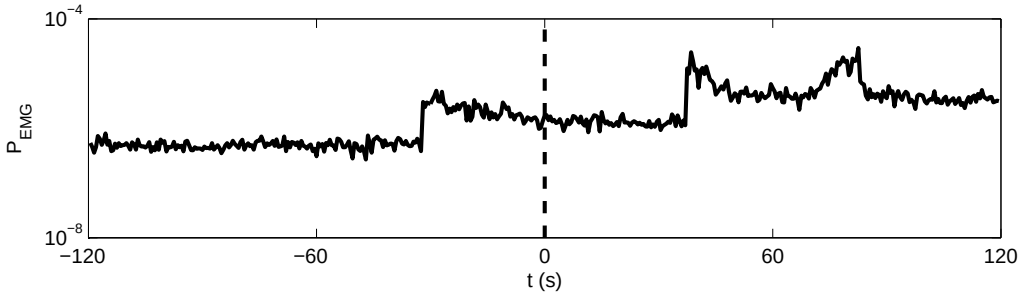

Figure A7: *Vocal stimulus, no response to stimulus. A small rise of EMG power occurs before the stimulus and a bigger increase about 30 s after the stimulus.*

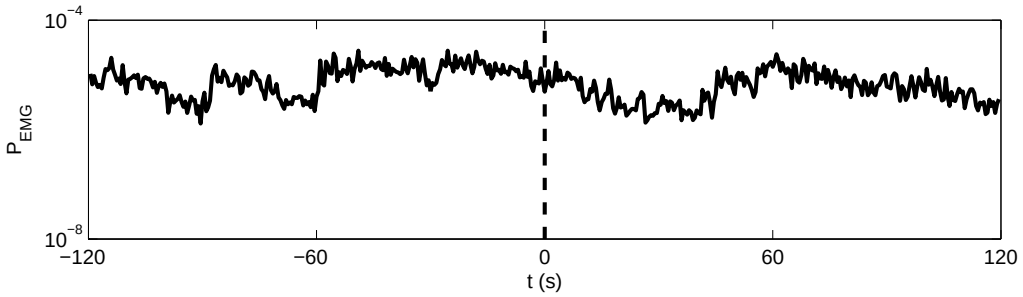

Figure A8: *Vocal stimulus, no response to stimulus. High and slowly changing EMG power.*

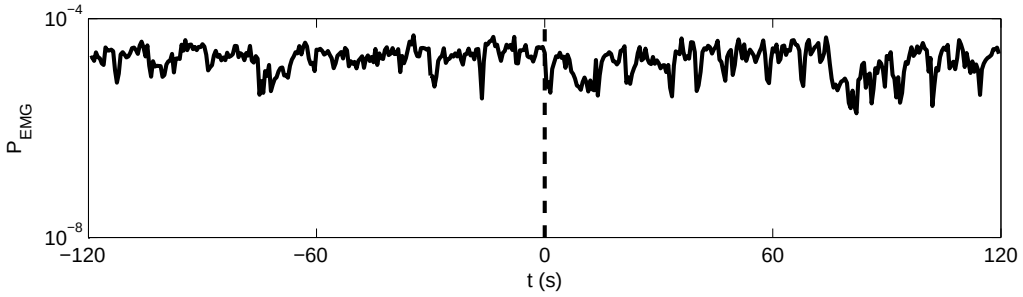

Figure A9: *Noisy stimulus, response to stimulus. High level of EMG power with a lot of variance.*

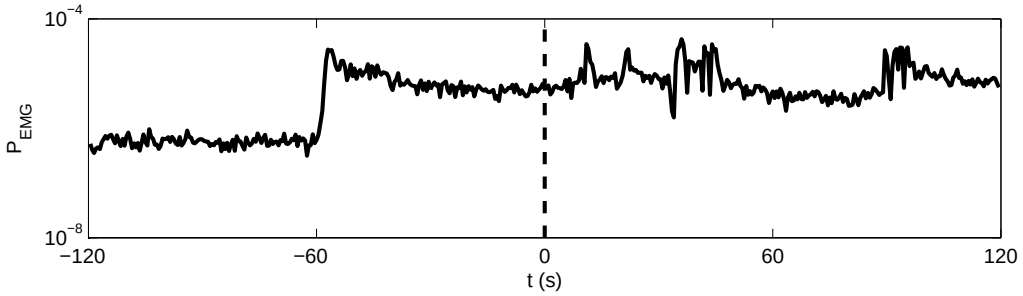

Figure A10: *Noisy stimulus, response to stimulus. EMG onset happens 60 seconds before the stimulus, which is roughly when vocal stimulus is given.*

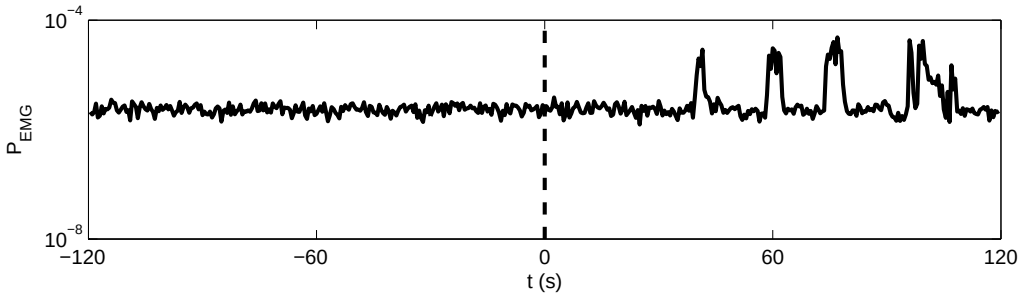

Figure A11: *Noisy stimulus, response to stimulus. EMG power shows several short spikes starting 30 s after the stimulus.*

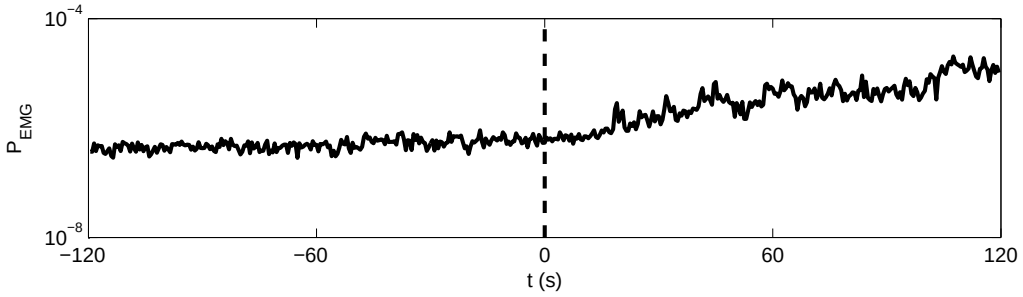

Figure A12: *Noisy stimulus, no response to the stimulus. Graph shows a clear but slow rise of EMG power after the stimulus. This patient later responded to the tetanic stimulus as is shown in Figure A13*

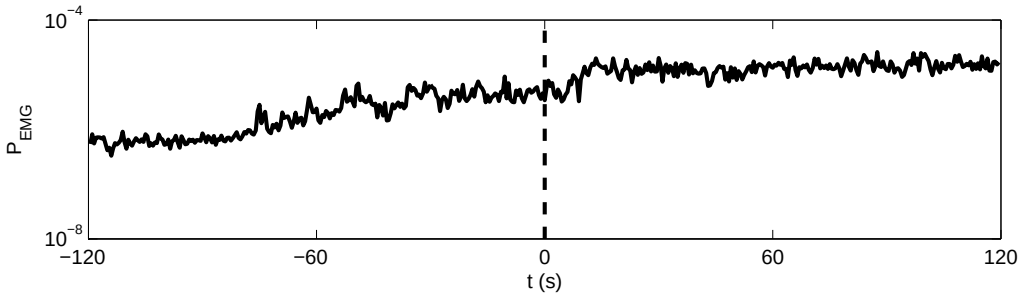

Figure A13: *Tetanic stimulus, response to the stimulus. The EMG shows rise after noisy stimulus as shown in Figure A12. A small rise after tetanic stimulus also occurs.*

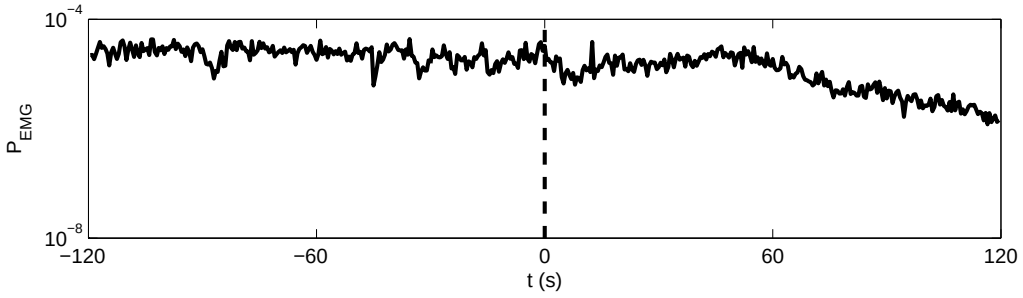

Figure A14: *Tetanic stimulus, response to the stimulus. This graph shows slow but clear drop of EMG power after the stimulus. No explanation was found for this phenomenon.*

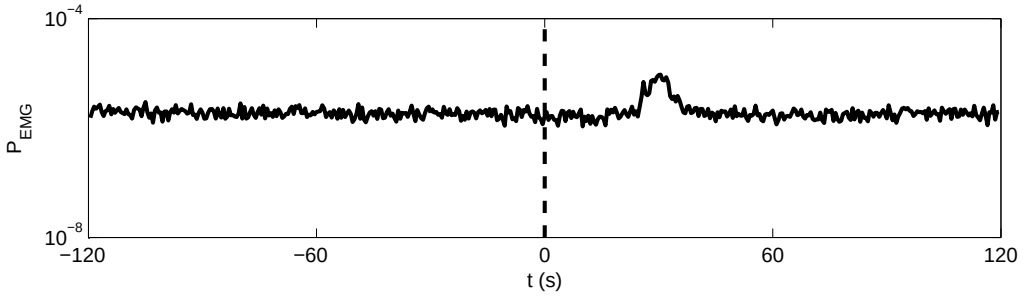

Figure A15: *Tetanic stimulus, no response to the stimulus. A small and short rise of EMG power occurs after the stimulus.*

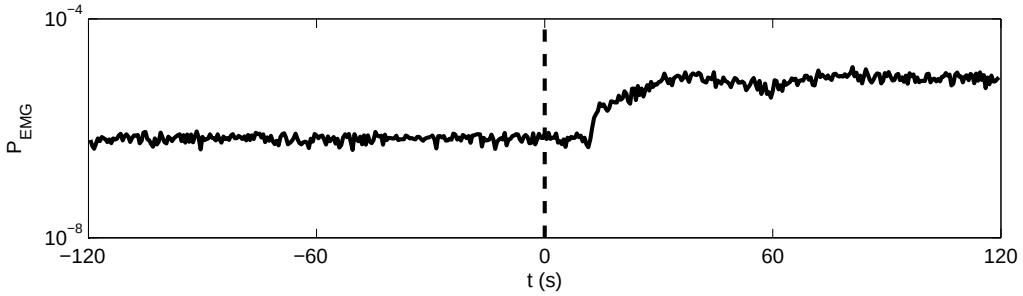

Figure A16: *Tetanic stimulus, no response to the stimulus. Clear rise of EMG power shortly after the stimulus. EMG power stays elevated for the whole time window.*

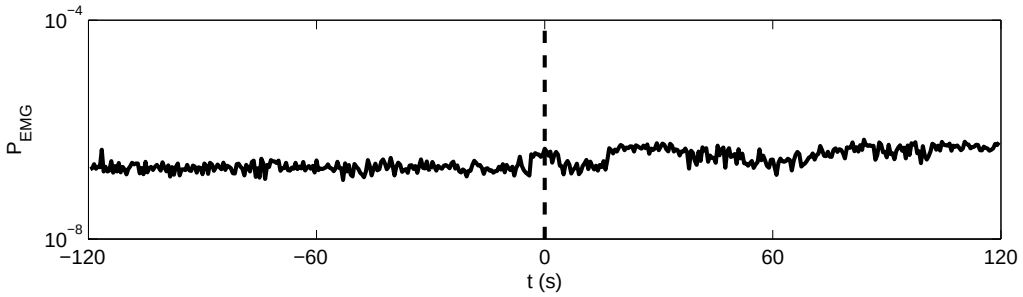

Figure A17: *Tetanic stimulus, no response to the stimulus. A small but long rise in EMG power after the stimulus.*

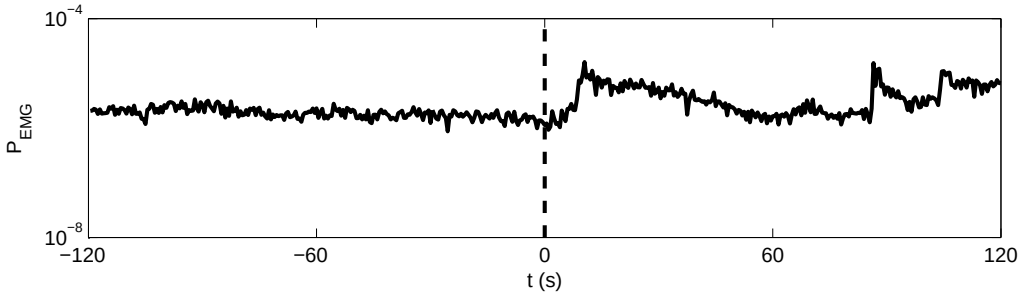

Figure A18: *TOF stimulus, response to stimulus. A clear response to the stimulus similar to typical responses to other stimuli.*

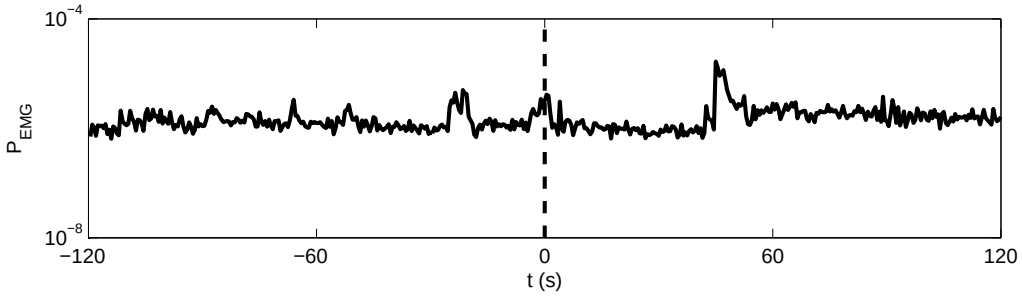

Figure A19: *TOF stimulus, response to stimulus. A short spike occurs well after the stimulus and is followed by slightly elevated EMG power level.*

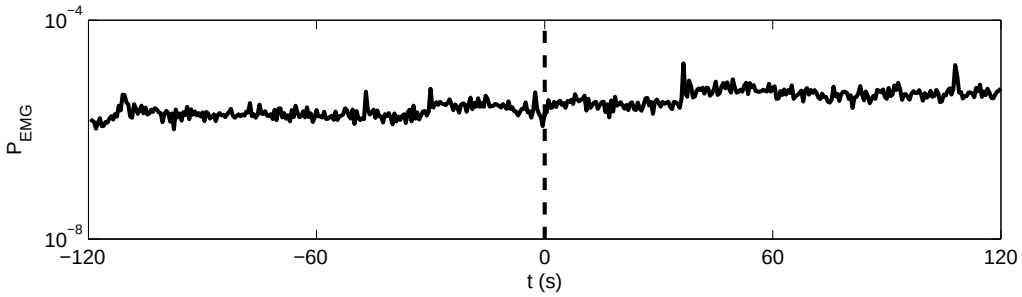

Figure A20: *TOF stimulus, response to stimulus. A small rise to the EMG power after the stimulus.*

Figure A21

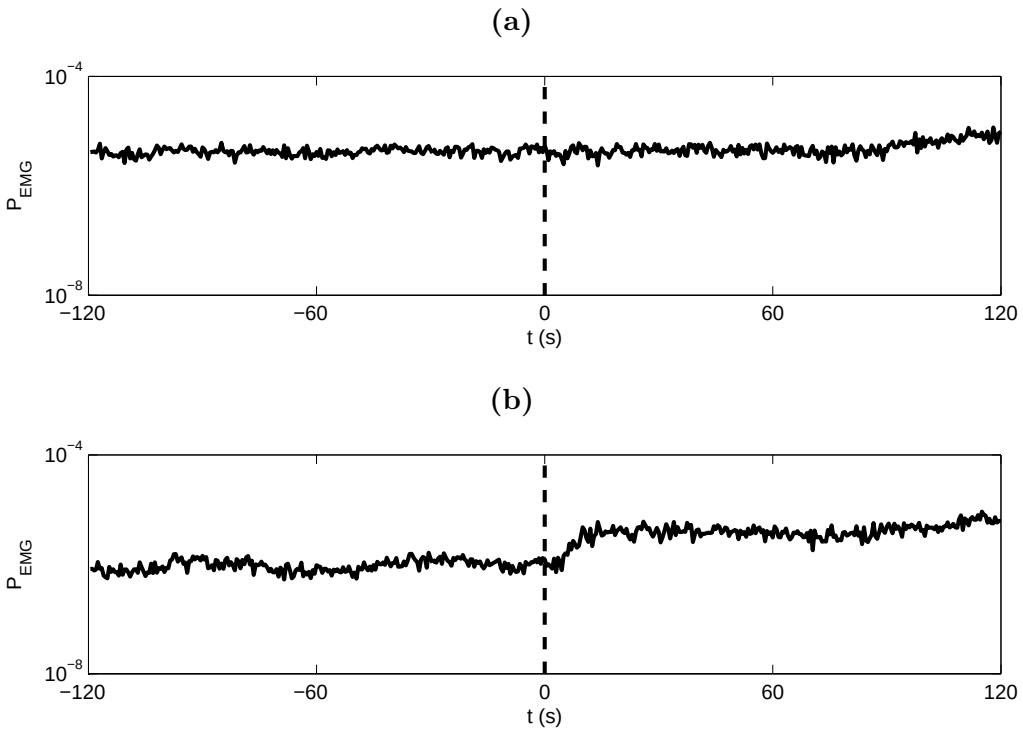

Figure A22: *Tetanic stimulus, response to the stimulus. A clear EMG response is show at the bottom channel (b) but no response occurs at the top channel (a)*