UNIVERSIDAD AUTÓNOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR

PROYECTO FIN DE CARRERA

QUANTIFICATION OF ELECTROENCEPHALOGRAPHIC CHANGES DURING HYPOGLYCAEMIA

María Riesco García

March 2009

QUANTIFICATION OF ELECTROENCEPHALOGRAPHIC CHANGES DURING HYPOGLYCAEMIA

AUTOR: María Riesco García
SUPERVISORES: Helge B.D. Sørensen
Carsten E. Thomsen
PONENTE: Pablo Varona

Biomedical Engineering Group
Department of Electrical Engineering
Technical University of Denmark

In collaboration with:

Faculty of Health Sciences
Copenhagen University

March 2009

This Master thesis is part of de requirements to complete Master studies at the Technical University of Denmark. It represents 35 ECTS.

Palabras clave

Aprendizaje automático, Clasificación, Clustering, Diabetes, EEG, Electroencefalograma, Extracción de artefactos, Insulina, Método no supervisado, Reconocimiento de patrones, Tratamiento digital de señales.

Key words

Artefact extraction, Classification, Clustering, Diabetes, Digital signal processing, EEG, Electroencephalogram, Insulin, Machine learning, Pattern recognition, Unsupervised method.

Resumen

Diabetes mellitus es una enfermedad en la cuál el cuerpo no produce suficiente o no usa apropiadamente insulina, una hormona producida en el páncreas. La insulina es necesaria para convertir el azúcar y otra comida en energía. La diabetes es normalmente una enfermedad crónica sin cura. El paciente juega un papel excepcionalmente importante en cuanto a educación, apoyo dietético, ejercicio físico y auto monitorización del azúcar en sangre, con el objetivo de mantener dicho nivel de azúcar dentro de unos límites aceptables.
Cuando la concentración de azúcar en la sangre es más baja que el nivel normal se conoce como hipoglucemia. La hipoglucemia más común surge como una complicación del tratamiento de la diabetes mellitus con insulina. Sin embargo, en personas no diabéticas, la hipoglucemia puede surgir por varias causas a cualquier edad.
En el año 2000, según la Organización Mundial de la Salud, al menos 171 millones de personas en el mundo sufren diabetes, un 2.8% de la población. Su incidencia está aumentando rápidamente, y se estima que para el año 2030, el número se habrá casi duplicado [20].
Los síntomas y manifestaciones de la hipoglucemia se dividen en aquellos producidos por las hormonas afectadas por bajos niveles de azúcar, y los efectos neuronales producidos por la falta de azúcar en el cerebro.
En este estudio, el objetivo principal es analizar y cuantificar cómo afecta la falta de azúcar en la sangre al cerebro, a través del análisis de señales encefalográficas (EEG).
Para investigar cambios en los patrones encefalográficos, un grupo de pacientes (diabéticos y no diabéticos) ha sido examinado durante normoglucemia, hipoglucemia y durante estados intermedios mientras estaban desarrollando diferentes ejercicios. Se grabaron señales encefalográficas (EEG) para cada uno de ellos.
Se eligieron tres secuencias diferentes para análisis, correspondiendo con dos ciclos en estado normal y el máximo evento hipoglucémico sufrida por cada persona.
Tras la eliminación de artefactos (interferencias de la señal), para cada segmento de 2 segundos de la señal EEG pre-amplificada, un conjunto de muestras, representando la distribución frecuencial y la amplitud, fue extraído. Se llevó a cabo un clustering jerárquico de estas muestras, para cada paciente y ejercicios, y combinando pacientes y ejercicios de distintas maneras. Se clasificaron las señales EEG a través de un clasificador probabilístico, y se asoció la ocurrencia de ciertos patrones a cambios en el nivel de azúcar sanguíneo usando unos correspondientes colores distintos, para diferentes grupos de pacientes. Finalmente, se presentan los resultados y conclusiones.
Para la mayoría de las personas, se encontraron patrones característicos individuales que ocurren principalmente en la secuencia hipoglucémica. Para estos pacientes, se ha observado un contenido más alto de actividad frecuencial desde 8 a 12 Hz (ritmo alfa) durante hipoglucemia, pero la mayor actividad frecuencial parece ocurrir de 6 a 10-12 Hz (en lugar de 8 a 12 Hz).
Las personas diabéticas desarrollan más patrones hipoglucémicos que las personas no diabéticas. La razón podría ser la cantidad de eventos severos hipoglucémicos sufridos en el pasado. Cuando están desarrollando una tarea que demanda concentración y actividad cerebral durante hipoglucemia, las personas diabéticas parecen estar más afectadas.
Para un grupo de pacientes, se observa menos contenido de altas frecuencias (>7Hz) para los patrones hipoglucémicos evidentes y una actividad más alta en el rango de 4 a 7 Hz.

El resultado principal y más importante obtenido es la variabilidad entre pacientes. Los patrones frecuenciales de algunos pacientes durante hipoglucemia corresponden a patrones de otros pacientes durante normoglucemia. Esto significa que hay más diferencias en los patrones frecuenciales de una persona a otra, que de normo a hipoglucemia para la misma persona.
Se han encontrado patrones hipoglucémicos comunes, pero sólo aparecen para unos pocos pacientes.

Abstract

Diabetes mellitus is a disease in which the body does not produce enough, or properly respond to, insulin, a hormone produced in the pancreas. Insulin is needed to turn sugar and other food into energy. Diabetes is currently a chronic disease with no cure. There is an exceptionally important role for patient education, dietetic support, sensible exercise, self monitoring of blood glucose, with the goal of keeping blood glucose levels within acceptable bounds.

When the glucose concentration in the blood is lower than normal level it is called hypoglycaemia. The most common forms of hypoglycemia occur as a complication of treatment of diabetes mellitus with insulin. However, in non-diabetic persons, hypoglycemia can arise from many causes at any age.

In 2000, according to the World Health Organization, at least 171 million people worldwide suffer from diabetes, or 2.8% of the population. Its incidence is increasing rapidly, and it is estimated that by the year 2030, this number will almost double [20].

Hypoglycemic symptoms and manifestations can be divided into those produced by the hormones affected by the falling glucose, and neural effects produced by the reduced brain sugar.

In this research, the main goal is to analyze and quantify how the lack of sugar affects to brain, by means of electroencephalographic (EEG) signal analysis.

To investigate changes in the EEG-pattern, a group of patients (diabetic and non diabetic) was examined during normoglycaemia, hypoglycaemia and intermediate states while they were performing different exercises. EEG signals were recorded from each of them.

Three different sequences were chosen to be analyzed, representing two baselines (normoglycaemic states) and the maximum hypoglycaemic event for each person.

After removing artefacts (interferences of the signal), from each 2 seconds segment of the pre-whitened EEG signal, a set of features, representing both frequency distribution and amplitude was extracted. Hierarchical clustering of these features was carried out, for each patient and exercise, and also combining them in different ways. Classification of the EEG was performed by a probabilistic classifier, and the occurrence of patterns was correlated to changes in the blood glucose using different colours accordingly, for different groups of patients. Finally, results and conclusions were presented.

For most of the persons, individual characteristic patterns occurring principally during the hypoglycaemic sequence were found out. For these patients, it has been observed higher content of 8 to 12 Hz activity (alpha rhythm) during hypoglycaemia, but the highest frequency activity seems to occur from 6 to 10-12 Hz (instead of from 8 to 12 Hz).

Diabetic patients develop more hypoglycaemic patterns than non diabetic patients. The reason may be their many several past hypoglycaemic events. When performing a demanding mental task during hypoglycaemia, diabetic seems to be more affected.

For a group of patients, it was found less high frequency content (>7Hz) in the clear hypoglycaemic patterns and a correspondent higher content of 4 to 7 Hz activity.

The main and most important result obtained is the interpatient variability. The frequency patterns of some patients during hypoglycaemia correspond to other patients' patterns during normoglycaemia. This means that the interindividual differences differ more from person to person, than from normo- to hypoglycaemia for the same person.

Common hypoglycaemic patterns have been found, but they do not seem to appear for more than a few patients.

Acknowledgements

The work presented in this master thesis has been carried out thanks to the great cooperation, advice and support of my supervisors in Denmark, Helge B.D. Sørensen and Carsten E. Thomsen. I would like to express my deepest gratitude to them, for letting me be part of this amazing research. I really appreciate the effort they have made to guide me and help me out, being available any time I needed. Thanks also to Martin and Melissa for their collaboration and kindness.

Special thanks to Pablo Varona, my home university supervisor, for all his help with the Spanish procedures and his support and valuable view points.

In the academic field I also want to thank to all my teachers from my home university (Universidad Autónoma de Madrid), because they have shared much time and knowledge with me, and in these five years, all of them have contributed somehow to form the person I am now. Thanks to the international office from the EPS for giving me the chance of spend one of my most amazing years in Denmark.

Speaking of one of the best years in my life, I want to say thank you to all the people that made it possible. Thanks Mai and my international group (especially Virgile, Nick, Pavel and Carlos), I spent really great moments with you. Thanks to Javi, Jose, Miguel and Thomas for being my family in Denmark. And thanks to the Spanish crowd in Denmark: Tania, Lili, Sara, Cris, Itzi, Laura, Lucía, Miki... Thank you for the trips, your smiles, your laughs and all the good moments we spent together.

I do not forget my best friends in Spain, María, Esther and Rosanna, for sharing their lives, joys, hopes and dreams with me.

In an especial way, I want to show my gratitude to Héctor. Thanks for being always there for me. I will never thank you enough for letting me to live my life, but at the same time sharing it with you. Just thank you for being who you are.

And last, but not least, never enough thanks to my family: my brother and my parents. Let me thank them in Spanish. Mi agradecimiento más importante, y nunca suficiente, es para mi familia: padres y hermano. Con este proyecto y tras cinco años me convierto en ingeniera, pero la mejor educación la he recibido siempre en casa. Soy la persona que soy gracias a vosotros. Gracias por vuestro constante apoyo, y porque vosotros habéis hecho que siempre haya sido tan feliz. Os quiero muchísimo.

Abbreviations

AC:	Alternating current.
ACC:	Autocorrelation coefficient.
AEP:	Auditory Evoked Potentials.
AQT:	Alzheimer Quick Test.
AR:	Auto-regressive.
b1:	Baseline 1.
b2:	Baseline 2.
bsl1c4:	Blood sugar level 1, cycle 4.
bslc1:	Blood sugar level, cycle 1.
C:	Central region.
C1:	Baseline 1, CalCAP.
C2:	Baseline 2, CalCAP.
CalCAP:	California Cognitive Assessment Package.
DFT:	Discrete Fourier Transform.
ECG:	Electrocardiogram.
EGG:	Electroencephalographic.
F:	Frontal region.
Fig:	Figure.
h:	Hypoglycaemic sequence.
HC:	Hypoglycaemic sequence, CalCAP.
HM:	Hypoglycaemic sequence, Monitoring.
M1:	Baseline 1, Monitoring.
M2:	Baseline 2, Monitoring.
O:	Occipital region.
P:	Parietal region.
PC:	Prediction coefficients.
RMS:	Root Mean Squared.
T:	Temporal region.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1A. Introducción

1.1 Descripción del problema

La diabetes es una enfermedad en la cual el cuerpo no produce, o usa apropiadamente, insulina, dando como resultado niveles de azúcar anormalmente altos en la sangre. Insulina es la hormona necesaria para convertir azúcar, almidones y otros componentes de la comida en la energía necesaria para el funcionamiento celular normal.
Es importante para los pacientes saber si el nivel de glucosa en la sangre está fuera del rango debido, el cual podría ser diferente para cada persona. Cuando el nivel de azúcar en la sangre está por debajo de dicho rango, se conoce como hipoglucemia.

La hipoglucemia puede ocurrir debido a diferentes factores: demasiada insulina, no suficiente comida, demasiado ejercicio, comer tarde, o comer demasiados pocos carbohidratos. Es decir, sucede cuando la insulina y el azúcar en la sangre están fuera de balance.
La gente sin diabetes normalmente no sufre hipoglucemia. Su cuerpo detecta cuándo tiene suficiente insulina y deja de producirla automáticamente. Pero la gente diabética tiene que descubrir cuánta insulina necesitará su cuerpo. Una vez que la insulina ha sido inyectada, sigue funcionando hasta que ha sido consumida, incluso si el nivel de glucosa en la sangre desciende demasiado.
Cuando el azúcar en la sangre baja hasta cierto nivel, normalmente 3.6mm/L, se produce una falta de fuel disponible en el cerebro y ocurrirán síntomas de inanición en el cerebro. Es esencial tratar hipoglucemia al primer signo detectado para prevenir manifestaciones más serias como confusión, pérdida de consciencia y ataques. Los primeros síntomas incluyen inestabilidad, vértigos, hambre, dolor de cabeza, mareos, mal humor, palidez y confusión. Si los niveles de glucosa descienden mucho más, puede tener lugar pérdida de consciencia y ataques. Para tratar la hipoglucemia, las personas diabéticas deberían tomar rápidamente azúcar absorbible oralmente en forma de zumo de frutas, refrescos regulares (no dietéticos), tabletas de chocolate o gel de azúcar.

La grabación de la actividad eléctrica del cerebro, el electroencefalograma (EEG), representa una técnica no invasiva bien conocida, usada para diagnósticos e investigación.
La actividad electroencefalográfica sobre el cuero cabelludo muestra oscilaciones en frecuencias variadas. Varias de estas oscilaciones tienen rangos de frecuencia características, distribuciones espaciales y están asociadas con diferentes estados del funcionamiento del cerebro (por ejemplo despierto y varios estados mientras la persona está dormida). Estas oscilaciones representan más o menos actividad sincronizada sobre una red de neuronas.
El análisis encefalográfico tiene muchos usos pero su utilización durante hipoglucemia todavía se encuentra lejos de una exploración profunda. La información “escondida” en el electroencefalograma tiene el potencial de ser un importante punto de partida para desarrollar in método de detección de hipoglucemia, y se trataría de un método no invasivo. Además, con el uso del electroencefalograma sería posible detectar la reacción del cerebro

y del sistema nervioso central a niveles bajos de glucosa en la sangre, individualmente para cada paciente.
El patrón encefalográfico seguramente va a ser diferente parea cada persona y para cada estado del paciente. Depende del metabolismo y de las condiciones bajo las que han sido grabadas las señales encefalográficas. La reacción del cerebro a bajos niveles de azúcar en la sangre también podría ser diferente para cada persona. Por lo tanto, es necesario estudiar primero cada paciente por separado, y determinar qué niveles de hipoglucemia afectan al cerebro de cada persona. En segundo lugar, sería posible desarrollar características electroencefalográficas comunes, útiles para detectar signos tempranos de hipoglucemia para un grupo de pacientes mayor.

1.2 Objetivos de la investigación

El objetivo principal de este estudio es investigar sobre el comportamiento del cerebro, usando análisis de señales encefalográficas, durante normo e hipoglucemia para descubrir cómo afecta al cerebro la hipoglucemia. El patrón encefalográfico está cercanamente relacionado con el metabolismo celular del cerebro y la forma de las ondas varía considerablemente según el estado del paciente [2]. En general quiere decir que un paciente que está profundamente inconsciente muestra grandes diferencias en su patrón encefalográfico que uno que está despierto.

Debido a la variabilidad entre pacientes, cada paciente individual será estudiado por separado y se intentarán desarrollar características comunes y encontrar patrones encefalográficos típicos (o clases) a partir de diferentes reacciones cerebrales y del sistema nervioso debido a diferentes niveles de azúcar para clasificarlos. Las clases serán agrupadas y se les asignarán colores de acuerdo con cada nivel de azúcar que se encuentre bajo un esperado umbral hipoglucémico, o bajo el estado normal del cerebro durante periodos con normoglucemia. La mayoría de las distintas clases se espera que describan un estado normal, y nuestro objetivo será encontrar síntomas característicos de la hipoglucemia en unas pocas de ellas.

Hay que resaltar que el objetivo principal no es detectar hipoglucemia, sino dar un temprano aviso de disfunción cerebral en esos estados en los que la hipoglucemia se encuentra bajo un umbral que no es aconsejable para el paciente. Este umbral puede ser diferente para cada paciente dependiendo de sus características personales (metabolismo) y condiciones (vida normal, inconsciente…) y por lo tanto será estudiado dependiendo de características individuales.

La presentación de los resultados debe ser comprensible y al mismo tiempo, adecuada para mostrar los resultados del análisis y clasificación de señales electroencefalográficas.

Si se encontraran patrones hipoglucémicos se daría al personal clínico valiosa información sobre las condiciones del paciente y sería posible mejorar el tratamiento individual del paciente. Una limitación importante es que muchos de los pacientes que desarrollan hipoglucemia espontánea no son conscientes de su ocurrencia, y por lo tanto no son capaces de tratarse apropiadamente a sí mismos.

Tabla 1: Conjunto de objetivos de la investigación.

Conjunto de objetivos
1. Descubrir las diferencias entre los estados de normoglucemia e hipoglucemia para cada paciente durante el ejercicio de monitorización.
2. Si estas diferencias existen, analizar cómo afecta al funcionamiento del cerebro la hipoglucemia para cada paciente y extraer resultados y conclusiones.
3. Descubrir cómo afecta al funcionamiento del cerebro la hipoglucemia para el ejercicio CalCAP para cada paciente.
4. Averiguar si hay alguna relación entre patrones hipoglucémicos de diferentes ejercicios (CalCAP y monitoring).
5. Analizar si nuestras conclusiones son válidas para un grupo de pacientes y si es posible hacer alguna generalización.

1.3 Organización de la memoria

Para llevar a cabo estos objetivos seguiremos los siguientes pasos (Fig 1):

- Extracción de datos de la base de datos, empezando solamente con unos pocos canales (4 + 1 de referencia).
- Etiquetar los datos y extraer los artefactos: Las señales electrofalográficas grabadas durante la vida diaria, incluyen una cantidad sustancial de datos perturbados por varios artefactos. Dichos datos no son útiles para la detección de hipoglucemia.
  Por lo tanto, los segmentos de señal serán preprocesados para separar y extraer la cantidad de datos útiles para nuestro estudio (detección y eliminación de artefactos).
  En segundo lugar, etiquetar los datos especificando qué ocurre en qué tiempo exacto.
- Extracción de patrones: Extraer los patrones útiles de las señales EEG.
- *Clustering* o agrupamiento: No existe un patrón de referencia para detectar la disfunción cerebral durante hipoglucemia, así que nos centraremos en la utilización de métodos no supervisados.
- Procesamiento de datos: Analizaremos primero una pequeña cantidad de datos para conseguir los primeros resultados y conclusiones, y después utilizar el método desarrollado para la cantidad de datos restante.
- Clasificador: Consiste en clasificar patrones de acuerdo con el estado del azúcar en la sangre del paciente. De ser exitoso, intentaremos detectar un conjunto de patrones comunes para varios pacientes y averiguar cómo es la variabilidad entre pacientes, y por lo tanto, cómo bajos niveles de azúcar en la sangre afectan al cerebro.

4 Quantification of electroencephalographic changes during hypoglycaemia

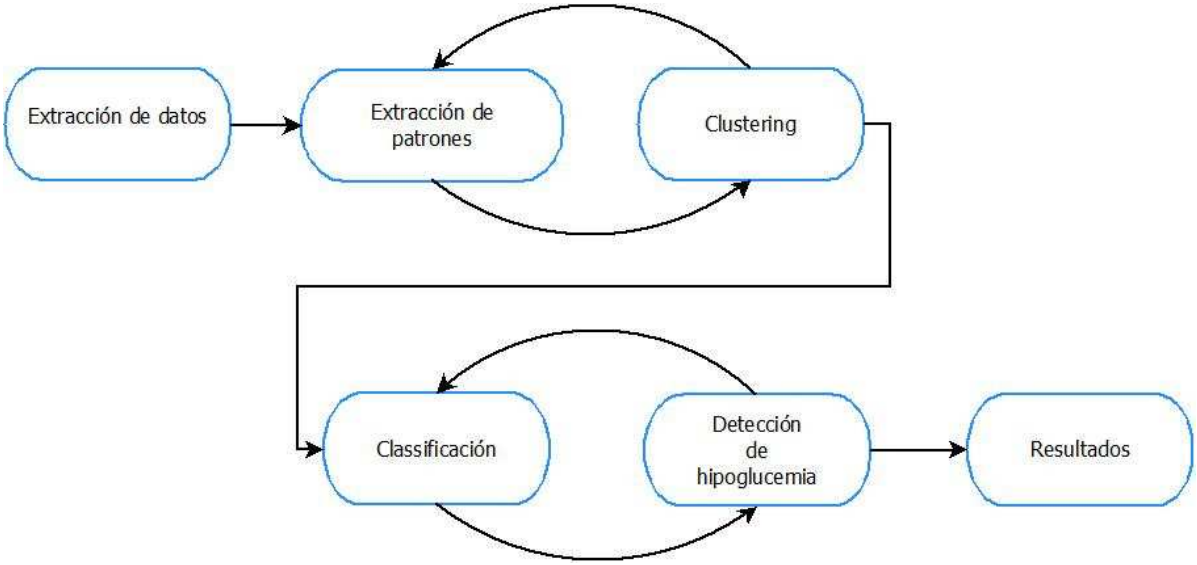

Figura 1: Organización de la memoria

1B. Introduction

1.1 Problem description

Diabetes is a disease in which the body does not produce or properly use insulin, resulting in abnormally high blood sugar levels (hyperglycaemia). Insulin is a hormone that is needed to convert sugar, starches and other food components into energy needed for normal cell function.
It is important for the patients to know if the blood glucose level is outside the target range, which might be different for each person. When blood glucose is below the target range, it is called hypoglycaemia.

Hypoglycaemia can be caused by a number of factors: too much insulin, not enough food, too much exercise, eating late, or eating too little carbohydrates. In short, it can happen when insulin and blood glucose are out of balance.
People without diabetes usually do not get hypoglycaemia. Their body can tell when it has enough insulin and stops releasing it automatically. But people with diabetes have to figure out how much insulin their body will need. Once the insulin is injected, it keeps working until it has gone, even if the blood glucose level goes too low.
When the blood glucose falls below a certain level, usually 3.6mm/L, there is a lack of available fuel to the brain and symptoms of brain starvation will occur. Treating hypoglycaemia at the first warning sign is essential in preventing more serious manifestations, such as confusion, loss of consciousness, and seizures. Early symptoms include shakiness, dizziness, hunger, headache, light-headedness, moodiness, pallor, and confusion. As glucose levels fall further, loss of consciousness and seizures may result. To treat hypoglycaemia, people with diabetes should take rapidly absorbable sugar orally in the form of fruit juice, regular (not diet) soda, glucose tablets, or glucose gel.

Recording of the electrical activity from the brain, the electroencephalogram (EEG), represents a well-known non invasive technique, used for diagnostic and research purposes. Scalp EEG activity shows oscillations at a variety of frequencies. Several of these oscillations have characteristic frequency ranges, spatial distributions and are associated with different states of brain functioning (e.g., awake and the various sleep stages). These oscillations represent more or less synchronized activity over a network of neurons.
EEG analysis has many uses but its use during hypoglycaemia is still far from fully explored. The information ‘hidden’ in the EEG has the potential to be an important input to develop a hypoglycaemia detection method, and it would be a non invasive method. Furthermore, with the use of EEG it might be possible to detect the reaction of the brain and the central nervous system to low blood glucose levels individually for each patient.
The EEG pattern is likely to be different for each patient and for each state of the patient. It depends on the metabolism and on the conditions under which the EEG signals are recorded, and also the brain reaction to low blood glucose levels might be different for each patient. Therefore it is necessary firstly to study each patient separately, and to determine what levels of hypoglycaemia affect the brain for each person. Secondly it might be possible to develop common EEG characteristics, useful for detecting early signs of hypoglycaemia for a larger group of patients.

1.2 Goals of the research

The primary aim of this study is to do research into brain functioning by using EEG signal analysis during normo and hypoglycaemia in order to find out how hypoglycaemia affects to the brain. The EEG pattern is closely related to the cell metabolism in the brain and the waveforms vary considerably with the state of the patient [2]. In general this means that a patient who is deeply unconscious shows highly different EEG patterns to one who is awake.

Due to inter-patient variability, each single patient will be studied separately and it will be tried to develop common characteristics and find typical EEG patterns (or classes) from different brain and neural system reactions due to different glucose levels for classification. The classes will be grouped and assigned colours according to either glucose levels below the expected hypoglycaemic threshold or the general state of the brain for periods with normo-glycaemia. Most of the different classes are expected to describe a normal state, and our goal will be to find characteristic hypoglycaemia "symptoms" in a few of them.

Recall that the main goal is not to detect hypoglycaemia in itself, but to give early warning for cerebral dysfunction in those states in which hypoglycaemia is below a threshold which is not advisable for the patient. This threshold can be different for each patient depending on their personal characteristics (metabolism) and conditions (normal life, unaware…) so it will be studied depending on individual characters.

The display of results should be comprehensible but also adequate for showing the results of the analysis and classification of the EEG signals.

If hypoglycaemic patterns could be found it would give the clinicians valuable information about patient's condition and it would make it possible to improve the individual treatment of the patient. An important limitation is that many of the patients developing spontaneous hypoglycaemia are unaware of the hypoglycaemic event, and therefore not able to treat themselves properly.

Table 2: Set of goals of the research.

Set of goals
6. Find out the differences between normoglycaemia and hypoglycaemia state for each patient during monitoring exercise.
7. If the differences exist, analyze how hypoglycaemia affects to brain functioning for each different patient and extract conclusions and results.
8. Find out how hypoglycaemia affects to brain functioning for the CalCAP exercise for each patient.
9. Find out if there is some relation between hypoglycaemic patterns for different exercises (CalCAP and monitoring).
10. Analyze if our conclusions are valid for a group of patients and if it is possible to do some generalization.

1.3 Outline of the report

For these goals we will follow some steps (Fig 1):

- Data extraction from the database, starting with just a few channels (4 + 1 of reference).
- Database labelling and Artefact extraction: The EEG signals recorded during daily life potentially include a substantial amount of data disturbed by various artefacts, and data from these periods might be useless to detect hypoglycaemia. Thus, the signal segments will be pre-processed in order to separate and extract the useful amount of data for our study (artefact detection and removal). Secondly, to label them specifying what it might occur at what exactly time
- Feature extraction: Extract useful features from EEG signals.
- Clustering: There exists no reference EEG pattern for detecting cerebral dysfunction during hypoglycaemia, so the focus will be on using unsupervised methods.
- Data processing: Analyzing first a lower amount of data to get the first results and conclusions, and after that using the developed method for the rest of the data amount.
- Classifier: To classify patterns according to patients' blood glucose state. If this is a success we will try to detect set of common patterns for some patients and find out how the variability is between patients, and thus, how the brain is affected by low blood sugar levels.

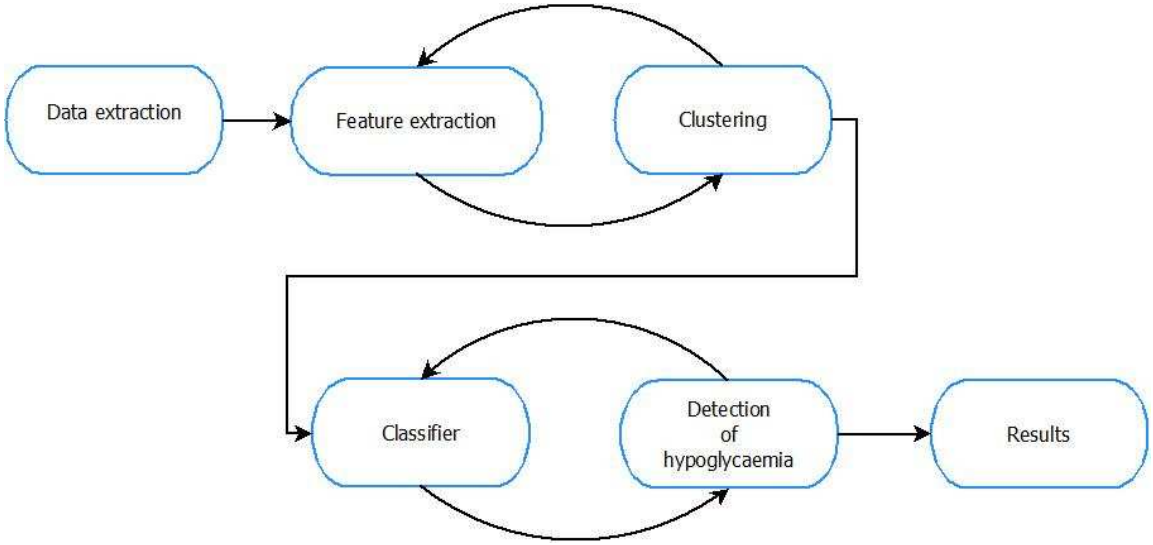

Figure 2: Outline of the report

2. Medical background

2.1 Diabetes

The pancreas is an organ that has two main functions. First, it helps the body to digest food by producing enzymes. And second, the pancreas also manufactures insulin. Insulin is a hormone and is the key that enables the body to use the food that is its fuel.
When food is eaten and digested, the nutrients are turned into a special kind of sugar called glucose. Glucose is the food and energy for all the cells that make up bodies.
Glucose is carried throughout the body by the bloodstream, but it can not enter the cells without the help of insulin.
A healthy pancreas never allows too much or too little glucose to exist in the blood. Just enough insulin is produced so that the glucose in the blood can be used by the cells. Insulin keeps the blood glucose level in balance all the time. It ensures that cells never lack energy.

Mayor types of diabetes:

- Type 1 diabetes:

It results from the body's failure to produce insulin, the hormone that "unlocks" the cells of the body, allowing glucose to enter and fuel them.
Glucose in the blood rises higher and higher without ever entering the cells. A person with type 1 diabetes must get insulin in some way to survive, usually by insulin injections. These shots provide the insulin that the pancreas is no longer producing.
For the rest of their lives, people with diabetes carefully balance the insulin they inject, the food they eat, and the exercise they get so that they can keep their bodies healthy. Once this is achieved, most can live long, active lives.

- Type 2 diabetes:

It results from insulin resistance (a condition in which the body fails to properly use insulin), combined with relative insulin deficiency. The cells are resistant to the insulin, and it is hard for glucose to get inside the cells. It also causes high blood sugar levels.
In response, the pancreas works harder and harder, pumping out more and more insulin and trying to lower glucose levels. As time goes on, beta cells (cells which produce insulin) may burn out completely from this overwork, and therefore less and less insulin can be made. If too many beta cells die, a person with type 2 diabetes may require insulin shots.

Normal blood glucose levels are between 60 and 100 milligrams per decilitre. People with a glucose level above 125 mg/dl are considered to have diabetes [7].

2.2 Hypoglycaemia

Blood glucose checking is one of the best tools for keeping diabetes in control. Frequent testing and good record-keeping give to the patient the most accurate possible picture of diabetes control. One of the purposes of blood glucose testing is checking out how often

blood glucose levels are in the target range. When blood glucose is below the target range, it is called hypoglycaemia.

Hypoglycaemia is a plasma glucose concentration low enough to cause symptoms and/or signs, including impairment of brain function.

In healthy individuals, symptoms of hypoglycaemia develop at an arterialized venous plasma glucose concentration of ~54 mg/dl (~3.0 mmol/l).

However, the critical level for hypoglycaemia might not be the same for each individual patient.

Iatrogenic hypoglycaemia is the limiting factor in the glycaemic management of diabetes. Thus, hypoglycaemia is a distinctly uncommon clinical event, except in people who use drugs that lower the plasma glucose concentration, particularly insulin to treat diabetes. All people with type 1 diabetes must be treated with insulin. Most people with type 2 diabetes ultimately require treatment with insulin.

Asymptomatic episodes of hypoglycaemia will be missed unless they are detected by routine self-plasma glucose monitoring. Because the symptoms of hypoglycaemia are nonspecific, symptomatic episodes may not be recognized as the result of hypoglycaemia. Even if they are recognized, mild-to-moderate self-treated episodes are often not long remembered and therefore may not be reported accurately at periodic clinic visits. Episodes of severe hypoglycaemia are more dramatic events that are much more likely to be recalled and reported. Thus, although they represent only a small fraction of the total hypoglycaemic experience, estimates of the incidence of severe hypoglycaemia are the most reliable [8].

2.3 Electroencephalographic signals

An electroencephalogram is a graphic record of brain waves representing electrical activity in the brain. It is generated by measuring electric signals using a set of electrodes attached to the scalp that act as transducers. Differences of electric potential between different parts of the brain are measured by a portable set of galvanometers and printed or reproduced as multiple simultaneous waveform tracings that have standard configurations in the normal brain.

For reasons of standardization, the locations of the electrodes used in recording EEGs are defined by international agreement as the “10-20 system” of electrode placement (Fig 2)

This system uses measurements of the head referenced to visible anatomic landmarks to minimize the variation in electrode placement among recording technologists and to provide the maximal uniformity in electrode to brain structure correspondence among patients.

A standard nomenclature is used for the electrode locations based on a letter prefix, which indicates the region of the head, and a number suffix, which indicates the exact location within that region. The common letter prefixes are as follows: F for the frontal region, C for the central region, P for the parietal region, T for the temporal region, O for the occipital region, and A for the ears. The most commonly used numbering system is also illustrated in Fig 2. As is evident, an odd number suffix indicates the left side of the head, an even number suffix indicates the right side. The suffix z indicates the sagittal midline, and a suffix that includes p indicates the frontal pole.

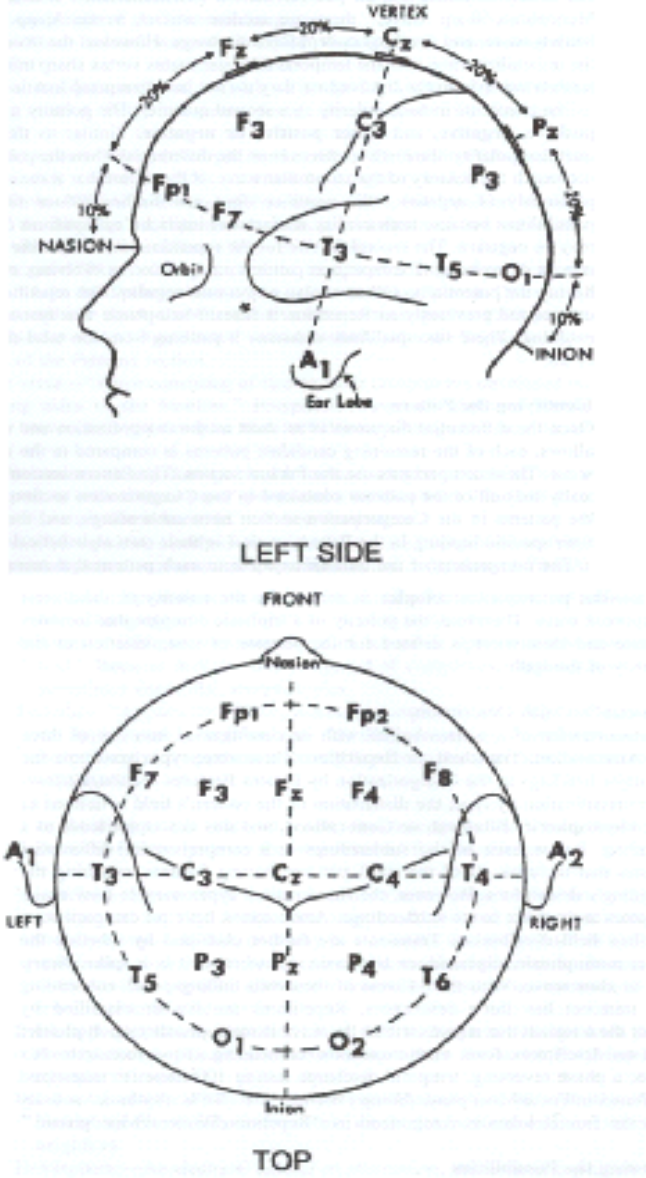

Figure 3: International 10-20 Electrode Placement System. [9] Image extracted from the book "Atlas of EEG Patterns" from John M. Stern.

The EEG record, also called a tracing, truly is a polygraph composed of multiple horizontal lines, each of which is generated by two electrode inputs and is called a channel. In our case, each patient has been recorded from 22 scalp electrodes (including reference and ground), giving 19 configurable EEG channels for each recording.

The specific electrode locations used for the creation of each channel are termed montages and are divided into two general approaches: bipolar and referential (Fig 3 and 4) [9]. We have used bipolar montage in our EEG recording.

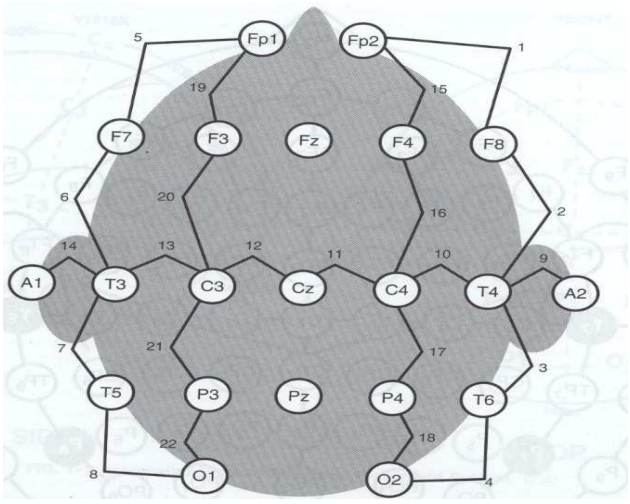

Figure 4: One type of Bipolar Montage: Electrodes are in bold and channels illustrated by lines linking the electrodes. [9] Image extracted from the book "Atlas of EEG Patterns" from John M. Stern.

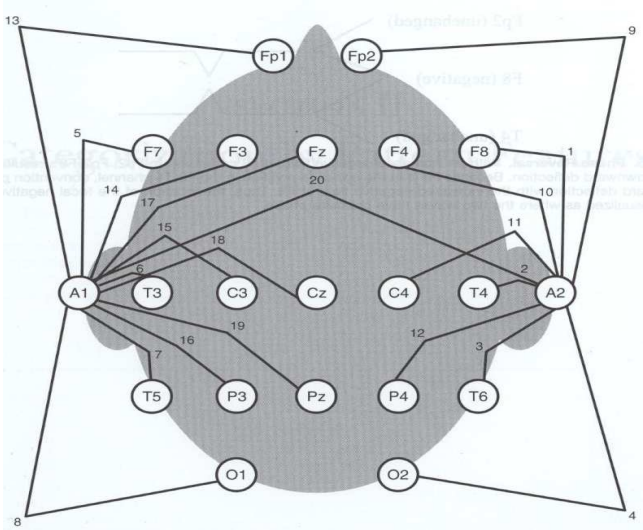

Figure 5: One type of Referential Montage: An "ipsilateral ear" montage with the channels listed as they may appear on an EEG paged. [9] Image extracted from the book "Atlas of EEG Pattern" from John M. Stern.

In EEG signals, the spectrum of frequencies is typically classified in four frequency ranges, as we can see in Table 2:

Table 3: Frequency ranges of the EEG signals.

Activity	Frequency Range
Delta	1.3-3.5 Hz
Theta	3.5-7.5 Hz
Alpha	7.5-13 Hz
Beta	13-30 Hz

3. Database description

3.1 Subjects

The data material for the study consists in two different groups of patients.
Group A: 12 insulin diabetic patients (Type 1).
Group B: 20 control patients.
Each patient has been recorded from 22 scalp electrodes (including reference and ground), giving 19 configurable EEG channels for each recording.

3.2 Experimental protocol

During the week preceding the experiment, subjects were instructed to live and eat as normal as possible and to avoid any rigorous exercise and use of alcohol or psychoactive drugs.

Subjects faced three different exercises in three different days: Hypoxia, placebo and hypoglycaemia.

We will focus in the hypoglycaemia exercise. During the night before the experiment a continuous glucose monitoring device was mounted. If hypoglycaemia took place (blood glucose concentration below 3.5 mmol/l) the experiment was postponed for at least 14 days.
If hypoglycaemia had not been present, the subject was introduced to the experimental setting and scheme, and was equipped as follows:

1) Intravenous cannula in an antecubital vein in both forearms.
2) EEG cap and two precordial ECG leads connected to a digital EEG recorder (Cadwell, Kennewick, Washington, USA).
3) Head phones (in-ear type) connected to a computer delivering auditory stimuli.
4) Ambulatory blood pressure unit.

The experiment was carried out in a cyclic manner with a total of 6 cycles for diabetic patients and 7 cycles for normal patients. These sequences have the following characteristics:

Table 4: Description of cycles: non-diabetic patients.

Hypoglycaemia, non-diabetes	
Cycle	**Type**
1	Baseline
2	Baseline
3	Blood glucose level decreases
4	Hypoglycaemia
5	Hypoglycaemia
6	Recovery
7	Recovery

Table 5: Description of cycles: diabetic patients.

Hypoglycaemia, diabetes	
Cycle	**Type**
1	Baseline
2	Baseline
3	Hypoglycaemia
4	Hypoglycaemia
5	Recovery
6	Recovery

In each cycle, four different exercises were carried out:

Table 6: Description of the exercises carried out during each cycle.

Exercises in each cycle	
Name	**Description**
Monitoring	Patient lied down, awake with close eyes.
Alzheimer Quick Test (AQT)	This test is developed to assess function of the parietal lobe and working memory. The subject must name 40 objects in different colours (for example red square, blue triangle, etc.).
Auditory evoked potentials (AEP)	AEP was obtained by an auditive stimulation paradigm consisting of more than 500 stimuli delivered. In order to maintain attention subjects were asked to count the number of stimuli given to the attended ear.
California Cognitive Assessment Package (CalCAP):	CalCAP test consists of four different reaction time exercises: The first one is a simple reaction time test where subjects must strike a key when a number is presented. The second one is a choice reaction task with reaction to a specific number. The third one is a choice reaction task with reaction to two identical numbers in a sequence. And the fourth and last exercise is a choice reaction test with reaction to two numbers in a sequence (increasing order).

Figure 6 shows the whole hypoglycaemia experiment, divided in cycles and exercises for a non diabetic person.

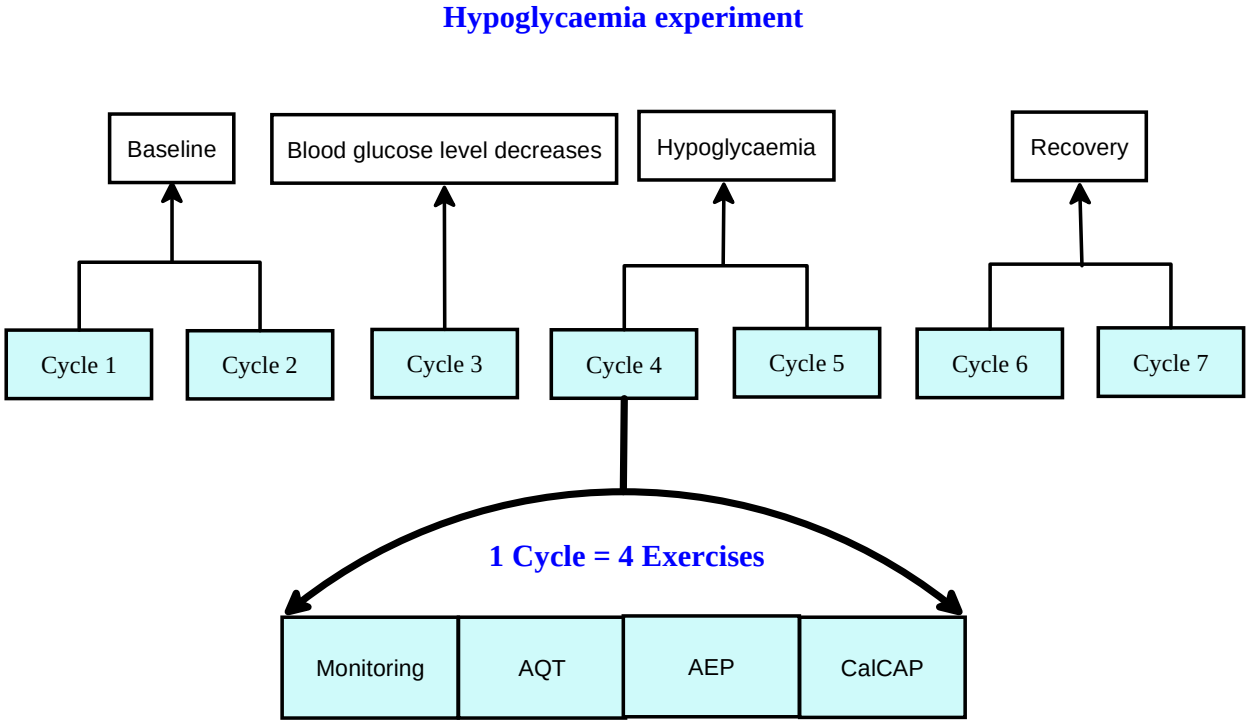

Figure 6: Schematic design of the sequence of cycles during the hypoglycaemia experiment for non diabetic patients. Abbreviations: AQT: Alzheimer Quick Test, AEP: Auditory Evoked Potentials, CalCAP: California Cognitive Assessment Package.

3.3 Pre-processing: Data selection

Section 3.3 and 4.1 was made in collaboration with Martin S. Christiansen and Melissa Larsen, both students of the Bachelor programme in Medical Engineering.

Continuous 19-channel EEG was recorded digitally (200 Hz sampling rate) through the experiment (Plus 1 or 2 of reference). Data were filtered by a first order 0.53 Hz low cut filter and a first order 70 Hz high cut filter. This filter is commonly used in EEG analysis to avoid noise and movement artefacts.

Five derivations were selected for primary power spectral analysis (4 channels + 1 of reference):

Table 7: Description of channels selected for primary analysis.

Channel	Input
1	F3-C3
2	F4-C4
3	C3-P3
4	C4-P4
5	Reference channel

Channels 1 and 2 are situated in the frontal lobe of the brain (left and right side), and channels 3 and 4 are situated in the parietal lobe (left and right side).

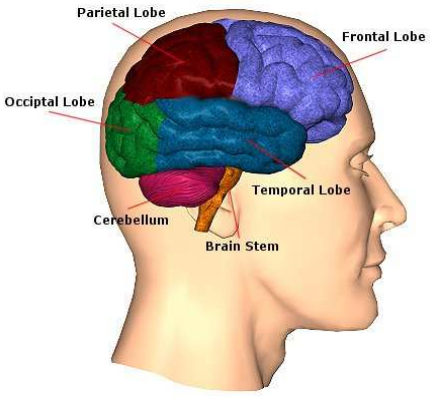

Figure 7: Main brain parts [10].

Within these five channels and for each patient, three cycles containing four exercises were selected to analyze them: Cycle 1 and Cycle 2 as baselines (normoglycaemia), and the Cycle corresponding to the lowest blood sugar level for each patient.

Table 8: Samples of blood sugar level for non diabetic persons. Yellow cells indicate the lowest sugar level of each person.

Non diabetic persons									
File	bslc1	bslc2	bslc3	bsl1c4	bsl2c4	bsl1c5	bsl2c5	bslc6	bslc7
nh17	5	4,9	2,4	2,8	2,3	2,3	2,4	2,9	1,6
nh21	5,4	5,2	4,1	3,5	3,1	3,2	3,2	5,4	4,1
nh3	5	4,8	4,1	2,8	2,6	2,8	2,9	5,7	3,9
nh14	4,8	4,8	3,5	2,5	2,8	3,1	3,2	4,8	3,1
nh11	5	5	3,6	2,3	3,1	3	3	4,4	3,2
nh9	4,7	4,7	3,7	2,4	1,9	2,6	2,2	3,8	3
nh6	5,9	5,8	4,7	1,7	3	3	3	4,3	2,9
nh18	5,5	5,2	3,8	2,8	2,9	3	3,1	5,1	3,2
nh2	5,3	5,4	3,8	2,4	3	2,9	3	4	3,8
nh4	5,2	5,1	4,2	3,1	3,1	3	2,9	6,5	#NULL
nh19	4,9	4,9	3	2,9	2,6	2,8	2,5	5,1	3,2
nh1	4,6	4,6	3,3	2	2,3	2,3	2,1	4,7	3,5
nh15	5	4,8	3,5	2,6	3,2	3	3	3,7	2,5
nh12	4,7	4,5	3,4	2,6	2,8	2,8	2,9	5,1	3,6
nh10	4,3	4,9	4,1	2,8	2,6	2,7	2,7	5,6	4,8
nh5	5,3	5,4	4,8	3,1	2,9	3,4	3,4	5,7	3,9
nh8	4,4	4,5	3,4	2,9	2,7	2,8	2,6	4,9	3,8
nh16	5	4,9	4,3	2,9	3	3,3	3,2	4,9	3,1
nh20	5,4	5,3	3,8	2,6	2,8	3,1	3,1	5,1	2,6
nh7	6,2	6,2	5,7	3,1	2,8	3	3,1	6	5,3
	Baseline 1 and Baseline 2		Sugar level decreases	Hypoglycaemia 1		Hypoglycaemia 2		Recovery 1 and Recovery 2	
bslc1= Blood Sugar Level, Cycle 1 bsl1c4= Blood Sugar Level 1, Cycle 4									

Table 9: Samples of blood sugar level for diabetic patients. Yellow cells indicate the lowest sugar level of each person.

	Diabetic patients							
File	bslc1	bslc2	bsl1c3	bsl2c3	bsl1c4	Bsl2c4	bslc5	bslc6
dh11	13,6	14,3	3,5	2,5	2,2	2,2	6,5	9,2
dh6	19,9	19,9	3,6	2,7	2,3	2,3	7,7	11
dh5	14,4	14,7	3,1	2,85	2,9	3,2	9,2	12,7
dh12	10,5	9,9	2,6	1,7	2,5	2,5	8,7	11,3
dh1	19,2	18,9	2,8	2	2,1	2,5	8,2	12,9
dh8	20	19,5	3,3	3	2,7	3	8,7	9,6
dh2	16,9	16,4	2,9	2,2	2,3	2,4	7,5	11,3
dh4	13,5	14,7	3,1	3	2,7	3	7,1	9
dh10	9,6	10	2,85	2,6	2,6	2,6	6,6	8
dh7	10,7	11,6	2,6	2,6	2,7	2,5	10,1	13,6
dh3	9,9	10,9	2,4	2,6	2,4	2	8,7	13,5
dh9	11,9	12,1	2,2	2	2,5	2,6	8,6	12,1
	Baseline 1 and Baseline 2		Hypoglycaemia 1		Hypoglycaemia 2		Recovery 1 and Recovery 2	
bslc1= Blood Sugar Level, Cycle 1 bsl1c4= Blood Sugar Level 1, Cycle 4								

As each EEG recorded for each patient is around 5 hours long, visual assessment on the EEG was performed on all the patients to determine the exactly starting and ending point for each cycle.

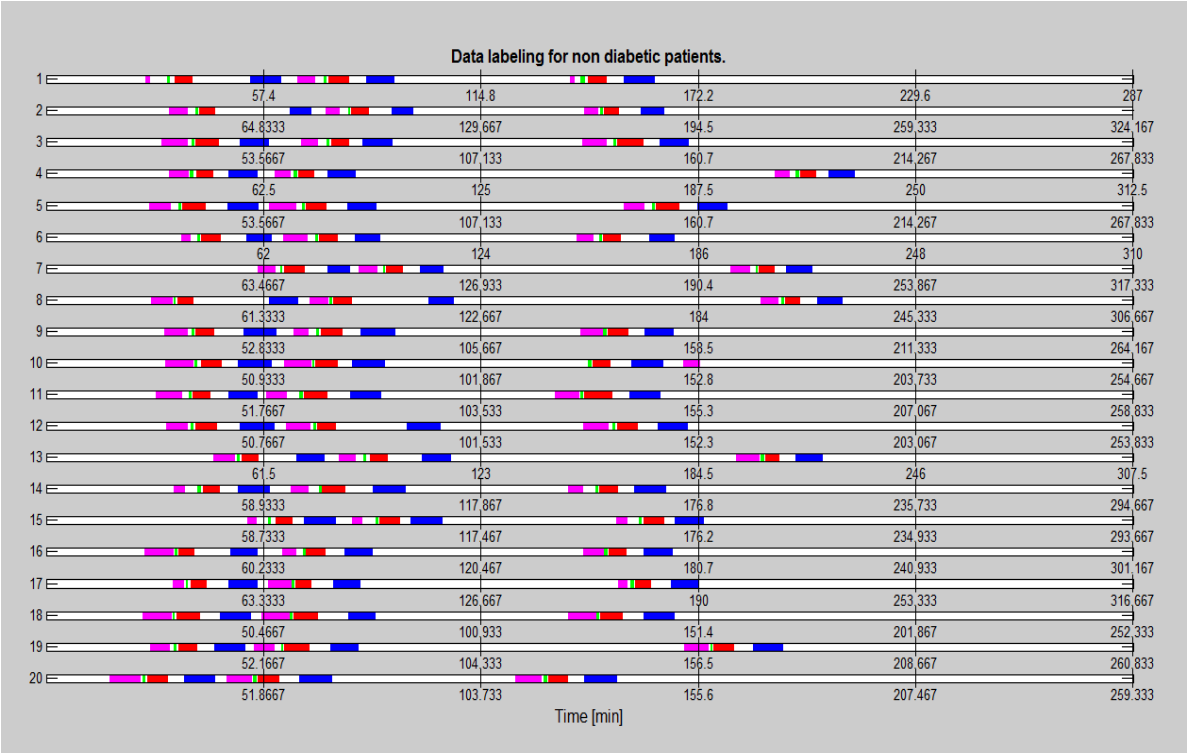

Figure 8: Segments of data extracted for each patient. 3 sequences and 4 exercises for each sequence. Sequences: Baseline 1, baseline 2 and maximum hypoglycaemic period. Exercises: Pink=Monitoring. Green= Alzheimer Quick Test (AQT). Red= Auditory evoked potentials (AEP). Blue=California Cognitive Assessment Package (CalCAP). Recall that for the patient number 10, the monitoring exercise for the maximum hypoglycaemic period was not available so it was taken the next segment with lowest blood sugar level.

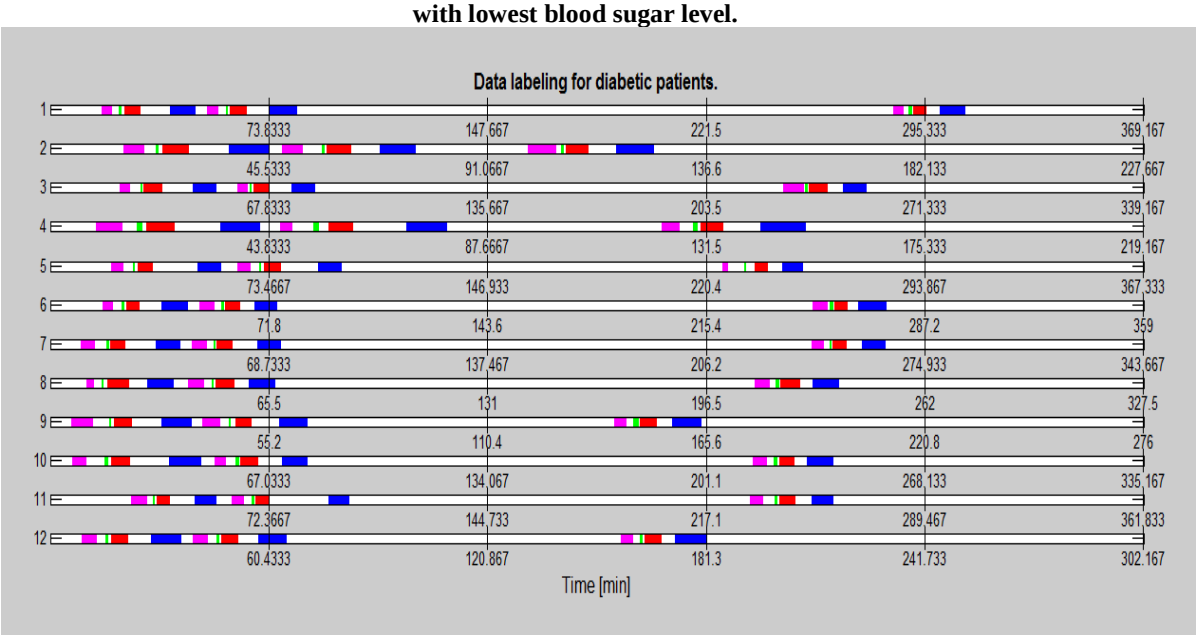

Figure 9: Segments of data extracted for each patient. 3 sequences with their corresponding 4 exercises each one. Sequences: Baseline 1, baseline 2 and maximum hypoglycaemic period. Exercises: Pink=Monitoring. Green= Alzheimer Quick Test (AQT). Red= Auditory evoked potentials (AEP). Blue=California Cognitive Assessment Package (CalCAP).

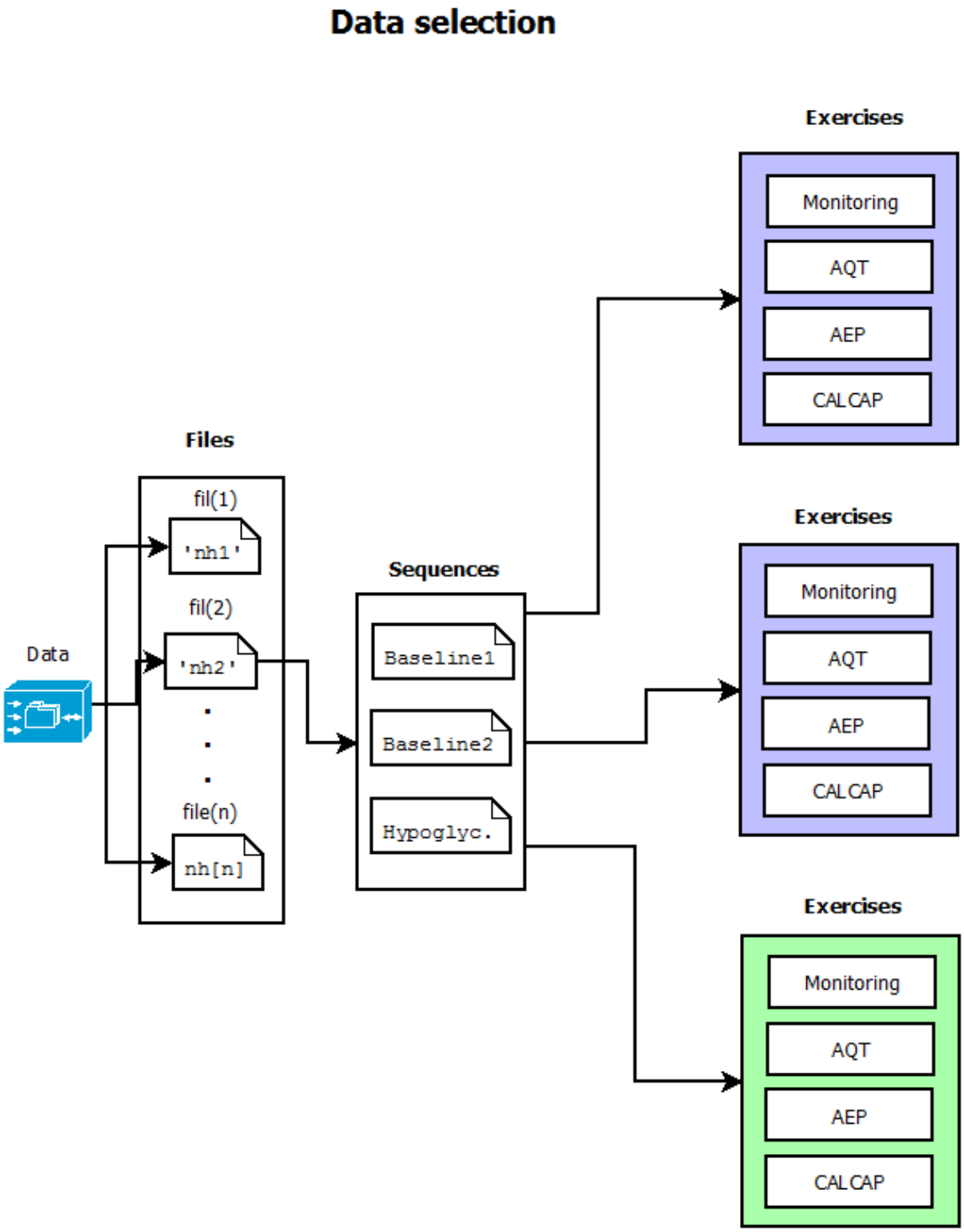

Figure 10: Scheme of data selection.

4. Methodology

4.1 Artefact extraction

EEG signals include a substantial amount of data disturbed by interferences. These interferences are called artefacts. Artefacts can be classified in five types depending on their origin: Cardiac artefacts, Electrode artefacts, External Device artefacts, Muscle artefacts and Ocular artefacts.

4.1.1 Artefact description

The following table is based in Chapter 4 of the book "Atlas of EEG patterns" of John M. Stern [9].

Table 10: Description of artefacts. Table based in Chapter 4 of the book "Atlas of EEG patterns" of John M. Stern [9].

Different types of artefacts		
Name	**Description**	**Characteristics**
Cardiac artefacts	The heart produces two types of EEG artefacts. One consists in an electrocardiogram (ECG) channel due to the differences of potential from the head electrodes to cardiac contractions, and the other arises through the circulatory pulse.	Cardiac artefacts are time-locked to cardiac contractions and are most easily identified by their synchronization with complexes in an electrocardiogram (ECG) channel.
Electrode artefacts	Electrode artefacts are due to spontaneous discharging of an electrical potential that was present between the electrode and the skin, or due to mechanical disturbances to the electrode or its lead.	Electrode artifacts usually manifest as one of two disparate waveforms, brief transients that are limited to one electrode and low-frequency rhythms across the scalp region.
External Device artefacts	Numerous types of external devices produce EEG artefact and may do so through the electrical fields they generate or through mechanical effects on the body. The most common external artefact is due to the alternating current (AC) present in the electrical power supply.	AC: This noise is usually medium to low amplitude and has the monomorphic frequency of the current (60 Hz in North America and 50 Hz in much of the rest of the world). The artefact may be present in all channels or in isolated channels that include electrodes that have poorly matched impedances. Electrical devices: high-amplitude, irregular, polyspike-like, or spike-like artefact. Mechanical devices: Slower components than electrical devices.

Muscle artefacts	Movement during the recording of an EEG may produce artefact through both the electrical fields generated by muscle and through a movement effect on the electrode contacts and their leads. This activity is the most common and significant source of noise in EEG.	High amplitude and frequency. It may appear regular and in the beta frequency band or as repetitive spikes if the high-frequency filter (low-pass filter) is set at 35Hz or less. Without this filtering, it has a more disorganized appearance. The duration of the artefact varies according to the duration of the muscle activity. Muscle artefact most commonly occurs in regions with underlying muscle, specifically the frontal and temporal electrodes.
Ocular artefacts	Most ocular artefacts are due to each eye's inherent 100mV electrical dipole. The dipole is oriented along the corneal-retinal axis and is positive in the direction of the cornea and negative in the direction of the retina. The dipole becomes relevant to the EEG recording when it becomes a moving electrical field, as occurs with changes in gaze and eye opening and closure. Vertical eye movements accompany eye opening and closure with deviation upward on closure. This is called Bell's phenomenon.	The amplitude of the artefact decreases quickly with greater distance from the orbits. The wave is maximum amplitude and surface positive at the frontal poles. Repetitive blinks usually appear as a sequence of the slow wave ocular artefacts and thus resemble rhythmic delta activity. It also appears in lower frequencies.

In order to build a filter for removing these artefacts that are disturbing the useful information, a meeting with the expert Troels W. Kjaer (Department of Neurophysiology. Rigshospitalet, Copenhagen, Denmark) was arranged. Visualization of artefacts and a description of artefact characteristics were performed. Some examples are shown next.

Electrode artefact: Disparate waveform due to patient's movement and thus, electrode movements.

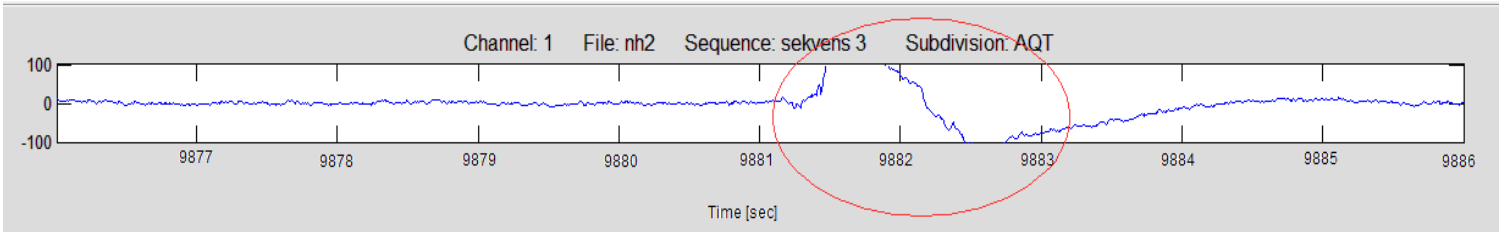

Figure 11: Electrode artefact. X-axis represents time in seconds and Y-axis represents µV. In the figure is represented 10 seconds of EEG recording.

It can be appreciated that Electrode artefacts might present very high amplitudes, and they can present equally low or high frequency components.

Ocular artefact: Eye blinking. High amplitude and low frequency range.

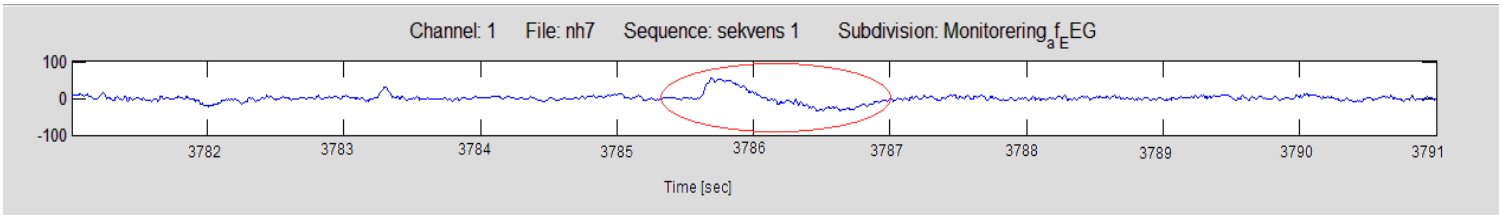

Figure 12: Ocular artefact. X-axis represents time in seconds and Y-axis represents µV. In the figure is represented 10 seconds of EEG recording.

In order to see the differences between a signal disturbed by an artefact and the useful information, one second containing an ocular artefact and one that doesn't contain it were extracted and visualize them in the frequency domain.

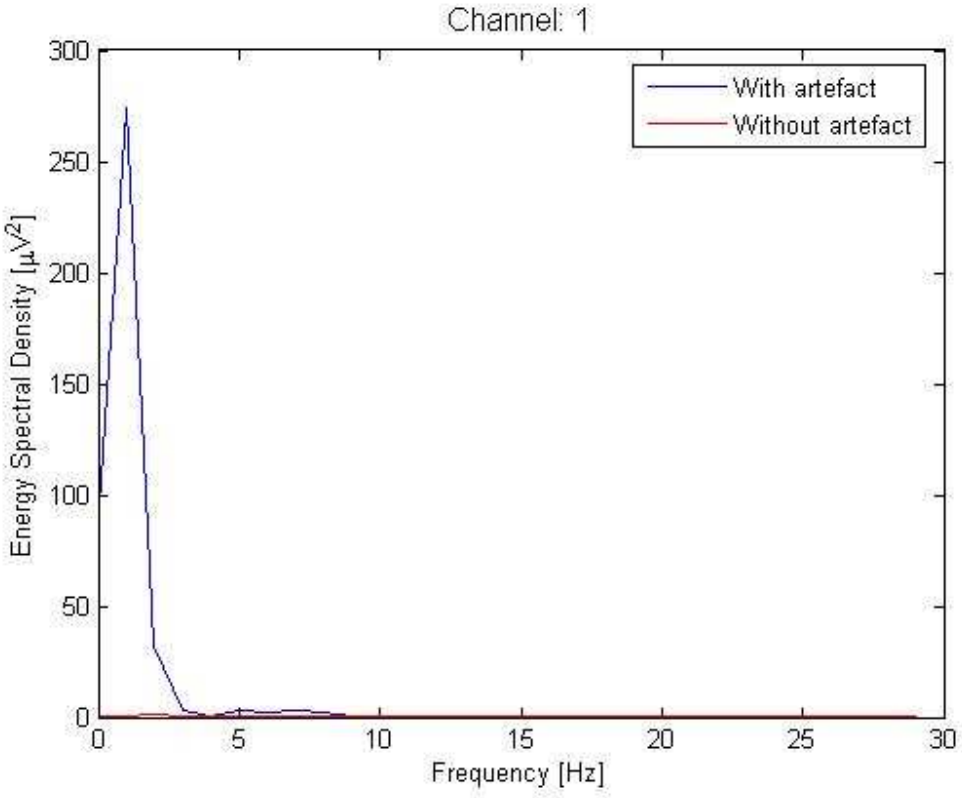

Figure 13: The blue function is the Fourier transform of one second extracted of an EEG signal containing an ocular artefact. The red function is the Fourier transform of one second extracted of an EEG signal containing undisturbed information.

From the last figure is confirmed that ocular artefacts show high energy during low frequency components (1 – 2 Hz).

Muscle artefact: High amplitude: It occurs, for example, when the patient chews.

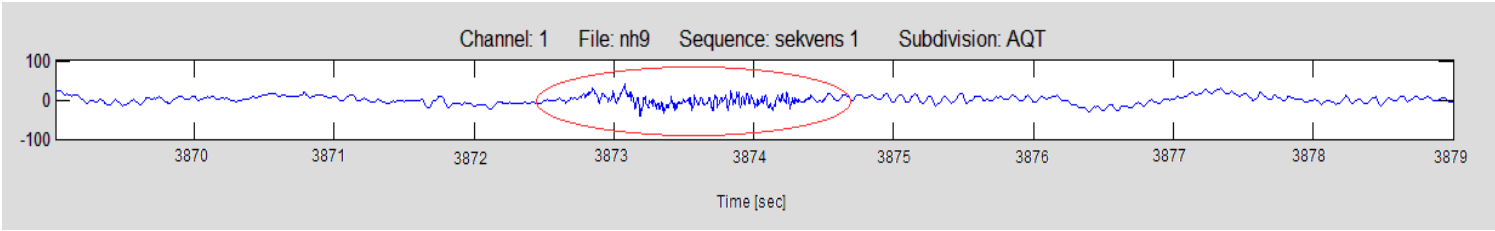

Figure 14: Muscle artefact. . X-axis represents time in seconds and Y-axis represents µV. In the figure is represented 10 seconds of EEG recording.

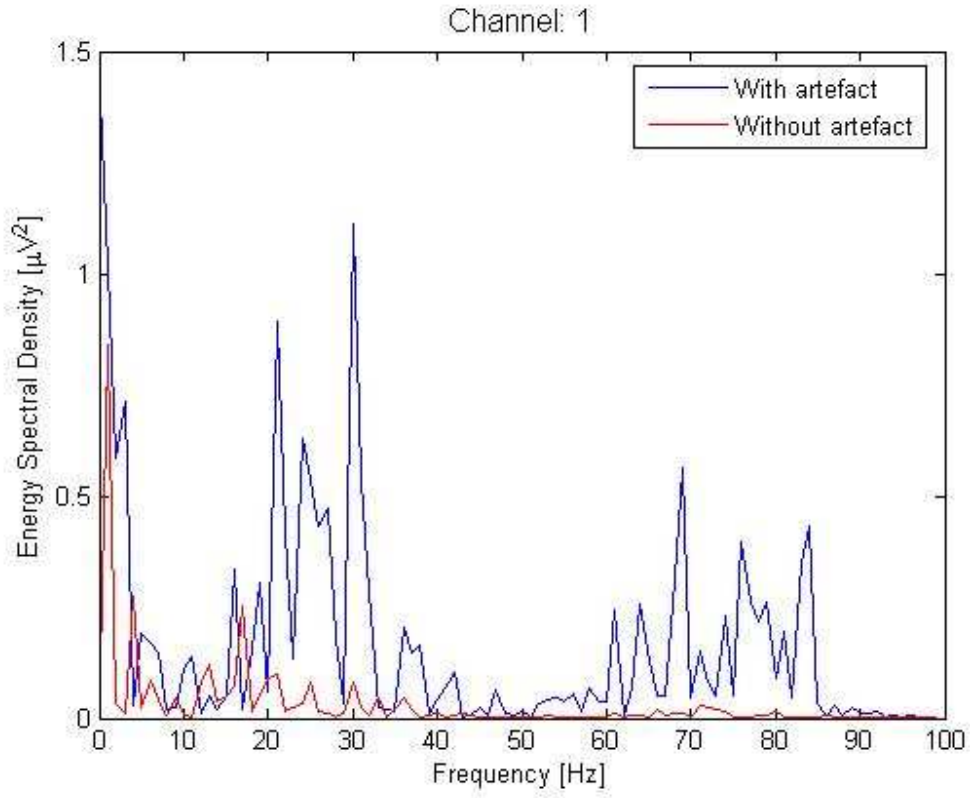

Figure 15: The blue function is the Fourier transform of one second extracted of an EEG signal containing a muscle artefact. The red function is the Fourier transform of one second extracted of an EEG signal containing undisturbed information.

Muscle artefacts might be more difficult to identify, because they look more similar than the undisturbed EEG signal. But, from the last figure, it is seen that they have high energy in higher frequency components than the undisturbed EEG signal.

4.1.2 Artefact filtering

Frequency domain:

To convert our signals from time domain to frequency domain, Welch's method has been used. The signal is split up into overlapping segments, and the overlap is said to be 50%. After that, individual data segments have a window applied to them (in time domain). The windowing of the segments is what makes the Welch method a "modified" periodogram.
A window of 1 second was chosen. Since the sampling rate for the EEG recorded is 200 Hz, the window has a size of 200 samples. This value gives us enough frequency resolution and it also takes a reasonable time to be computed. This is important due to the huge amount of data available.
As it was shown before, ocular artefacts have frequency components of 1 or 2 Hz, so it belongs from 200 to 400 samples, and then a window of 200 samples may be appropriate.
After doing the above, the periodogram is calculated by computing the discrete Fourier transform (DFT), and then computing the squared magnitude of the result. The individual periodograms are then time-averaged, which reduces the variance of the individual power measurements. The end result is an array of power measurements vs. frequency "bin".

Filter No 1: Filter amplitudes above 100 µV

After detailed visual assessment on the EEG signals, it was conclude that EEG signals amplitude are never above ± 100 µV, unless they are electrode artefacts. So a first filter was made in order to remove artefacts which amplitude is above ± 100 µV. If an amplitude signal above ± 100 µV is detected, the filter marks the whole window of 1 second which it belongs to. And when artefacts are removed it is always kept a continuous window of 2 seconds that does not contain any artefacts.
With this filter, movement and big ocular artefacts are removed.

Filter No 2: Filter for ocular artefacts

After detailed examination of sequences containing ocular artefacts and sequences without them, it was concluded that ocular artefacts contain higher energy in 1 and 2 Hz than a sequence without artefacts (as it appears in Figure 13).
Histograms for all the patients were made to see how the energy of 1 and 2 Hz is distributed. Histograms for the different exercises were compared. Monitoring and Auditory Evoked Potential tests the patient is with close eyes, so it is expected to have less ocular artefacts. In the same way, ocular artefacts are abundant during AQT (Alzheimer Quick Test) and CalCAP (California Cognitive Assessment Package**)** tests because patients have opened eyes. All the distributions present a peak around 1 and 2 μV^2, but for the exercises where they have opened eyes it appears much higher energy for 1 and 2 Hz.

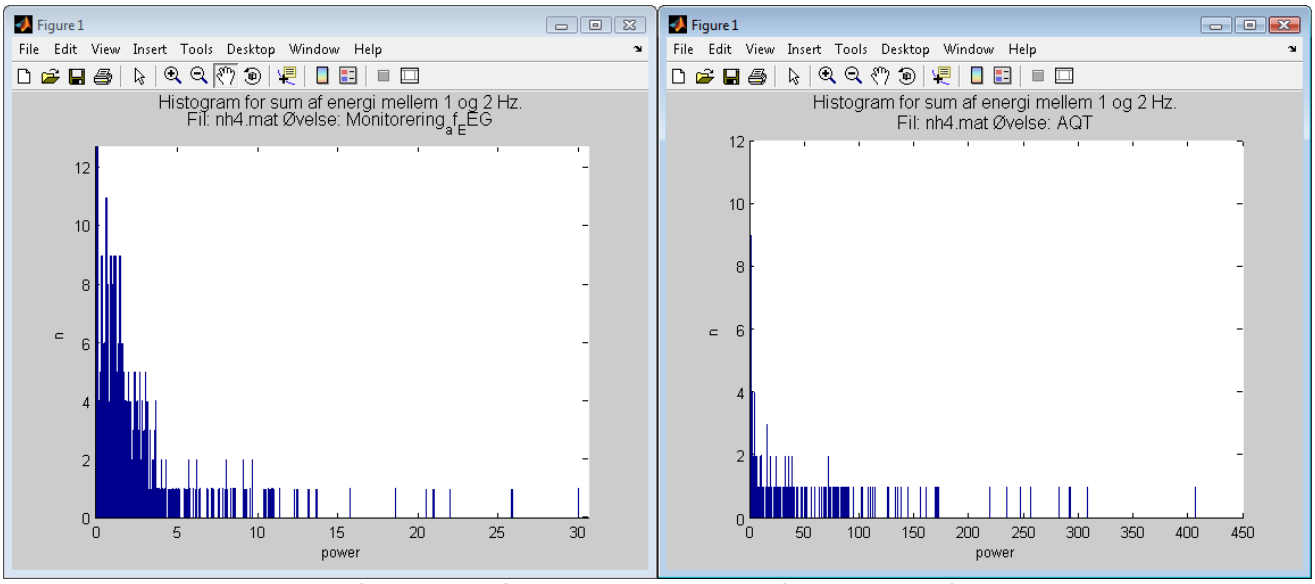

Figure 16: Comparison of histograms for the patient No 4. The figure on the left represents Monitoring exercise (left) and Alzheimer Quick Test (AQT) exercise (right). Comparing the x-axis it is shown that during AQT (opened eyes) the spectrum of our signal contains much higher energy for 1 and 2 Hz than during Monitoring, as it was expected.

A threshold establishing the differences between them in order to remove the artefacts away it is difficult to determine. This threshold should be different for each person, so generalization has been made for these patients very carefully. It has been tested by visual inspection that most of the artefacts have been removed, and there is also a huge amount of data remaining. It has been also compared how much data has been removed in the different sequences (2 baselines and hypoglycaemia) and the amount looks similar. To achieve a better generalization it is purpose to built a filter using an adaptable threshold depending on the energy of each person at each moment as a future work.

Firstly, a mean of the histogram was calculated for every histogram of every patient. A value between 1-8 μV^2 was obtained for the exercises with closed eyes, and a value over it was obtained for opened eyes exercises. But, finally, the threshold was found out by testing how much data it was removing and having two questions into consideration. The first one, is it better to have less amount of data but with higher quality? Or in the other hand, is it better to have as much data as possible even knowing that it still contains some artefacts?
It was preferred to have as much data as possible because hypoglycaemia signs in the brain could appear just in a few samples, if they appear, and that information can not be missed.
For these reason, and testing the data, the more convenient threshold for our data was established as 40.

Thus, the ocular artefacts filter consists in:
Calculate the sum of the power contained in 1 and 2 Hz for each Fourier transform of every 1 second window.
If this value is over the threshold the 1 second window is ruled out, and also the neighbour windows of each side (remember the overlapping), keeping always continuous segments of 2 seconds (2 continued windows of 1 second).

In the next figure, it is shown one minute of EEG signal after the ocular filter (10 seconds each row). Below the signal we can see a colour bar. Blue means it is useful data, and red means it has been removed. It is shown that the threshold is not very strict because there are still some ocular artefacts remaining, but it removes most of them.

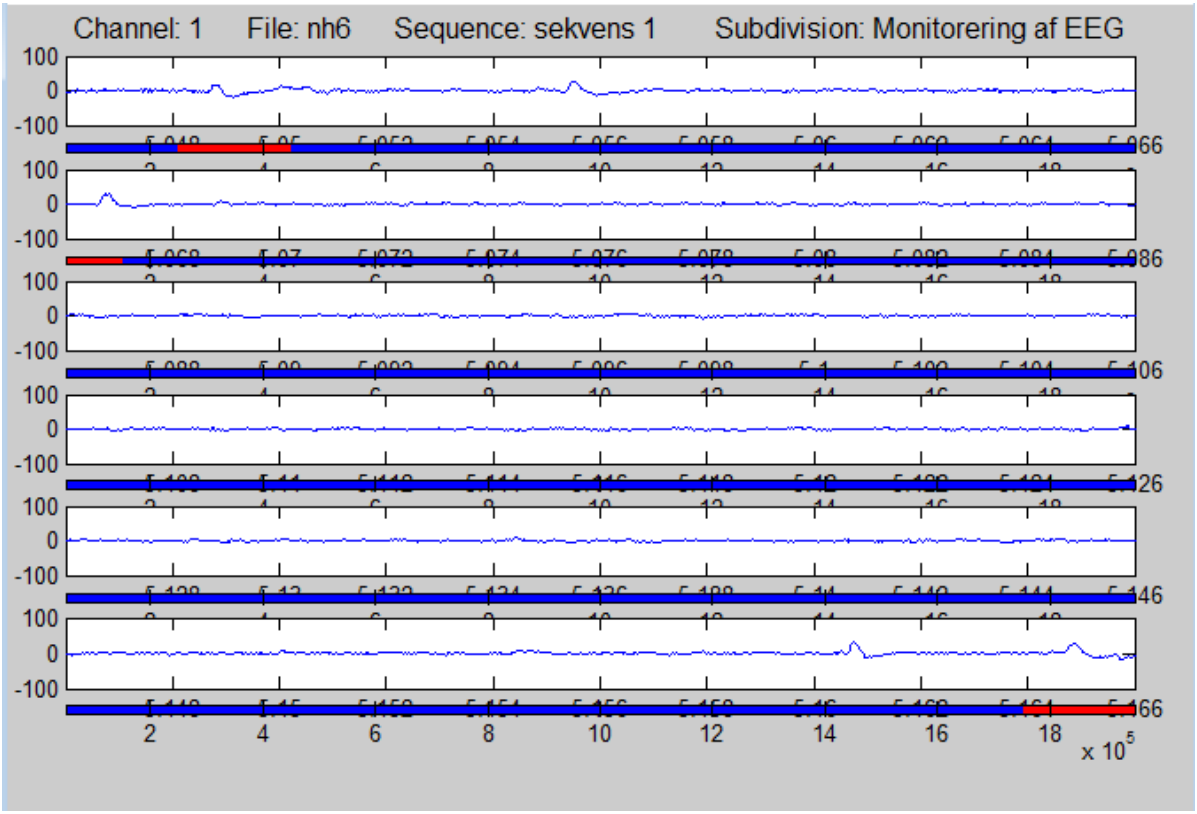

Figure 17: One minute of EEG signal after the ocular filter is represented (10 seconds each row). Below the signal we can see a colour bar. Blue means it is useful data, and red means it has been removed by the ocular filter.

Filter No 3: Filter for muscle artefacts

As it was represented in Figure 15, the segment containing muscle artefacts has much higher energy during approximately 45-90 Hz than a non disturbed segment. For these artefacts, the same reasoning and procedure than for the ocular artefacts were carried out.

Thus, the muscle artefacts filter consists in:
Calculate the sum of the power contained in the interval from 40 and 90 Hz for each Fourier transform of every 1 second window.
If this value is over the threshold the 1 second window is ruled out, and also the neighbour windows of each side (remember the overlapping), keeping continuous segments of 2 windows again.
The threshold was also regulated by testing and visual inspection of the signals, and its value was estimated to be 2.

Figure 18: One minute of EEG signal after the ocular and the muscle filter is represented (10 seconds each row). Below the signal we can see a colour bar. Blue means it is useful data, and red means it has been removed by the ocular filter. The red circle marks a muscle artefact that has been removed.

The data were chosen to be filtered in the following order:

1. 100 µV amplitude signal filter.
2. Ocular filter
3. Muscle filter

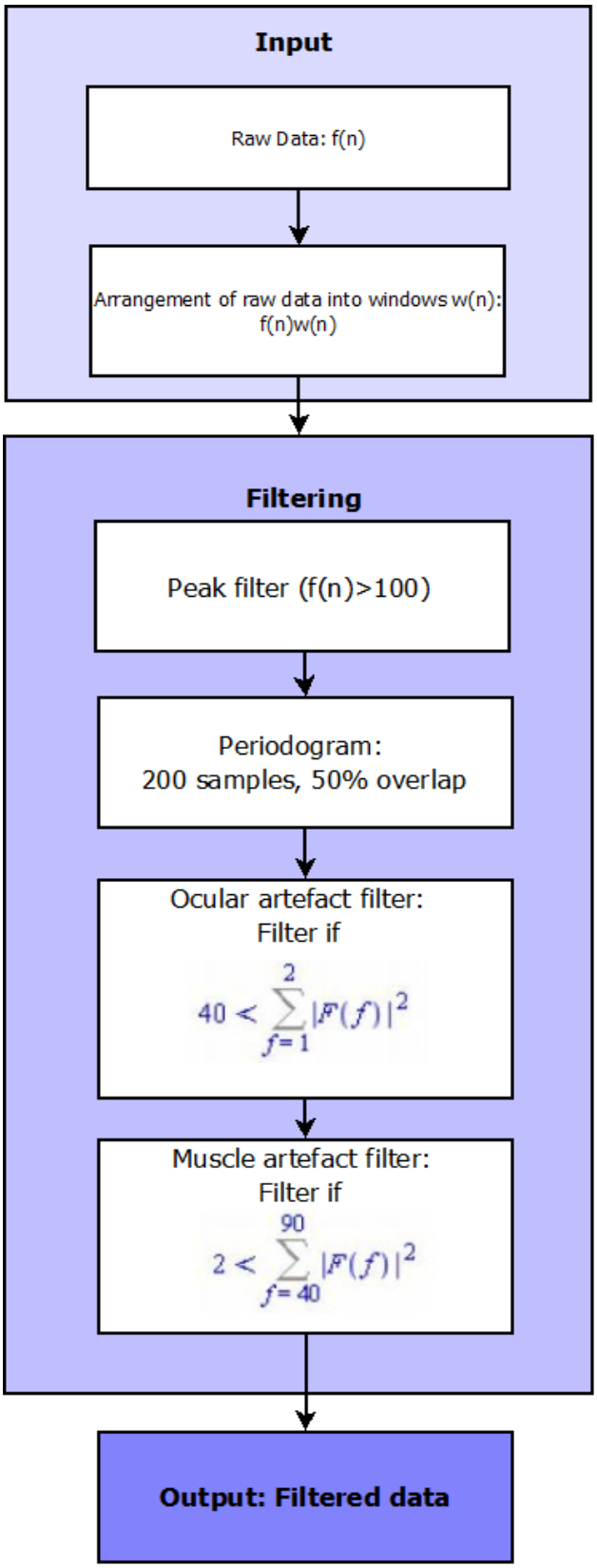

Figure 19: Structure of the artefact filter.

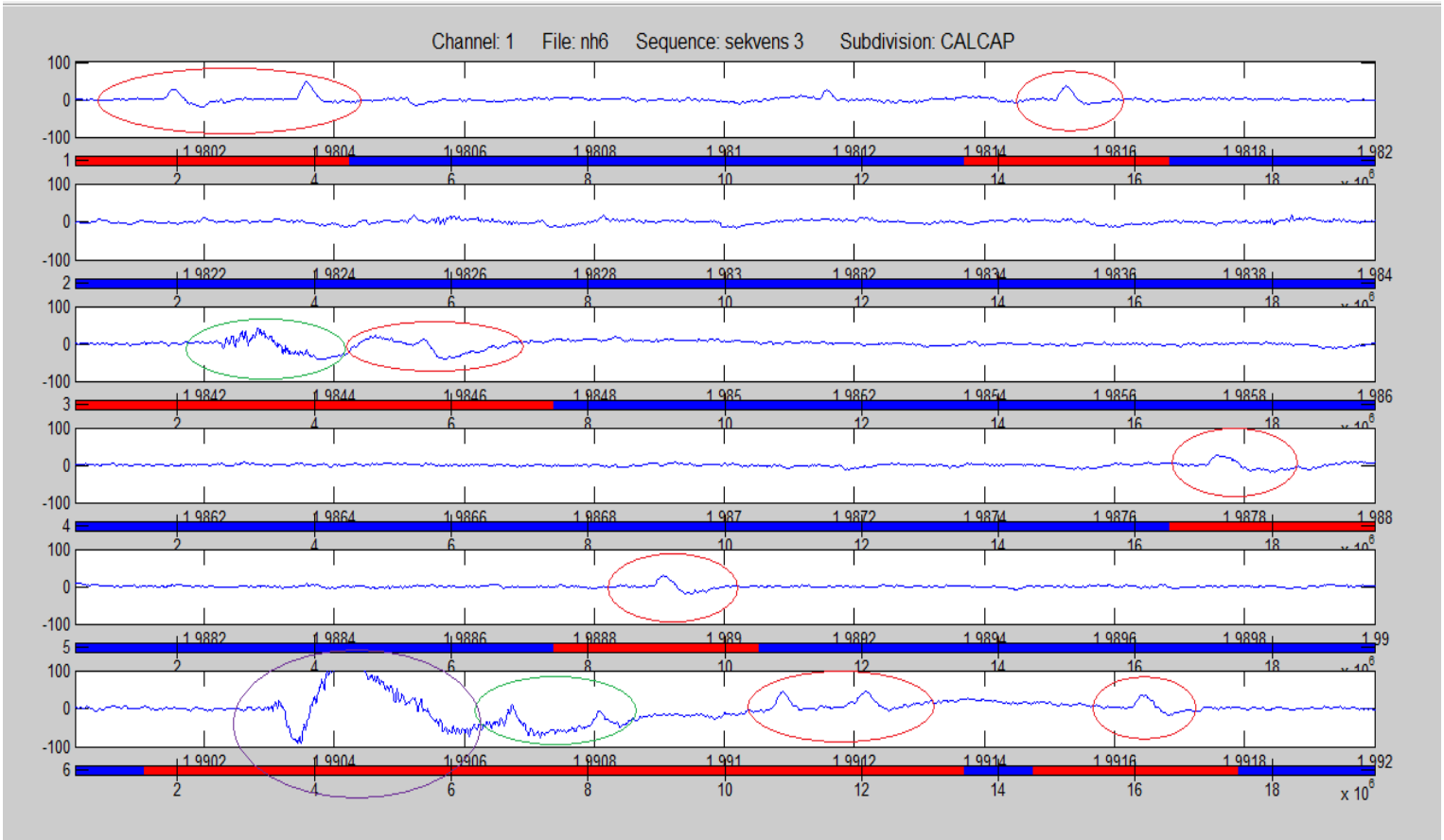

Figure 20: Example of 1 minute sequence filtered by the three different filters. Red circles mark the data removed by the optical filter. Green circle marks the data removed by the muscle filter. Purple circle marks the data removed by the 100 µV amplitude signal filter.

4.2 Noise filter

To avoid aliasing and to reduce the influence of noise and interference from mains supply (50 Hz), the EEG signal without artefacts was filtered with a 45 Hz low-pass filter (squared window) as we can observe in the following figure:

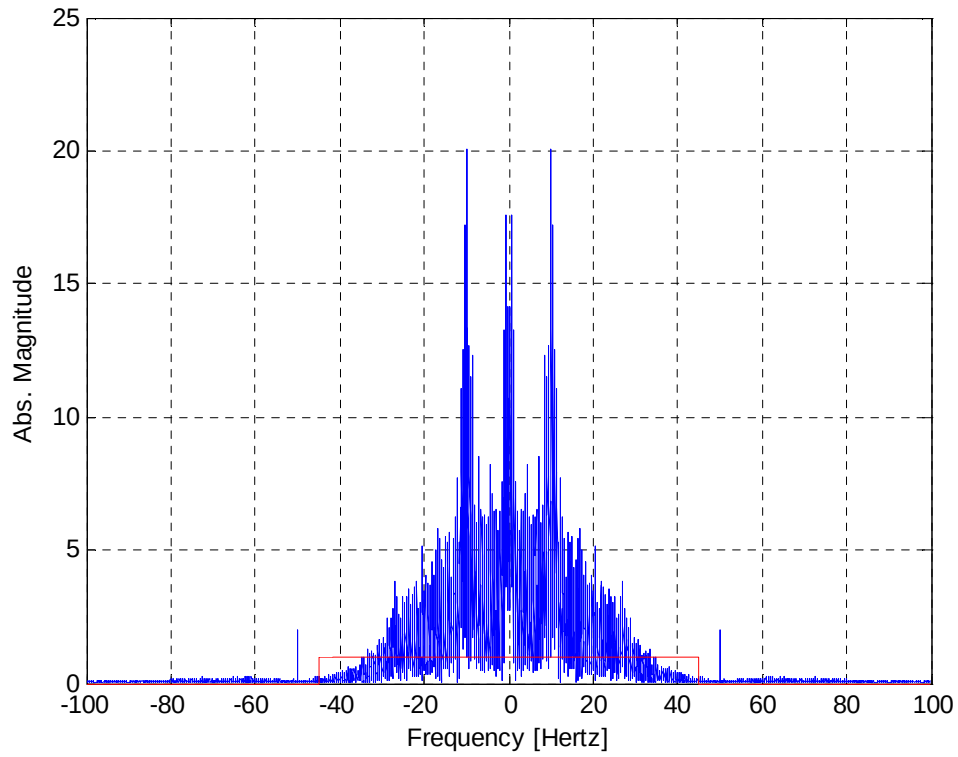

Figure 21: Signal without artefacts in frequency domain (blue color) and filter applied to it in order to avoid aliasing and noise (red line). It can be seen the noise produced by mains supply (50 Hz).

4.3 Pre-emphasising

With signals such as the EEG which have a large range of amplitudes at different frequencies it is good practice to prewhiten the signal before analysis (Blackman and Tukey, 1958). Prewhitening is the process of pre-filtering a signal so that the amplitudes of different frequencies fall within approximately the same range. This has two main advantages. Firstly, it ensures that low amplitude faster frequencies are digitised to the same resolution as the higher amplitude slower frequencies. Secondly, particularly when short epochs and hence wide bandwidth filters are used, it avoids the possibility of a large signal falling within one of the sidelobes, or to one edge of the main lobe, contributing more to the filter output than a low amplitude signal in the middle of the main lobe (passband)[11].

We considered a Butterworth first-order high pass filter with a cut-off frequency at 4.2 Hz. The sampling rate is 200 Hz. The recurrence relation is:

$$y[n] = -1 * x[n-1] + x[n] + 0.8760505959 * y[n-1] \quad [12]$$

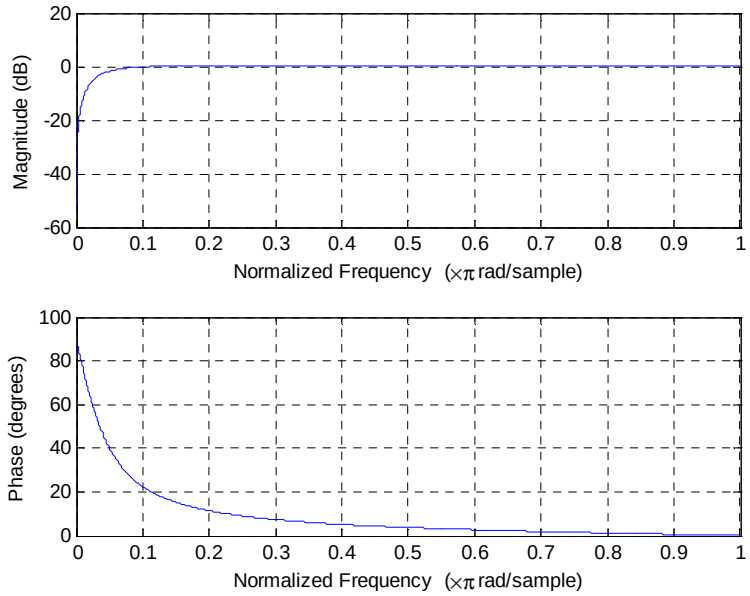

Figure 22: Frequency response of the high-pass filter used to preemphasise the EEG-signal. The filter is a Butterworth first-order high pass filter with a cut-off frequency at 4.2 Hz. The sampling rate is 200 Hz.

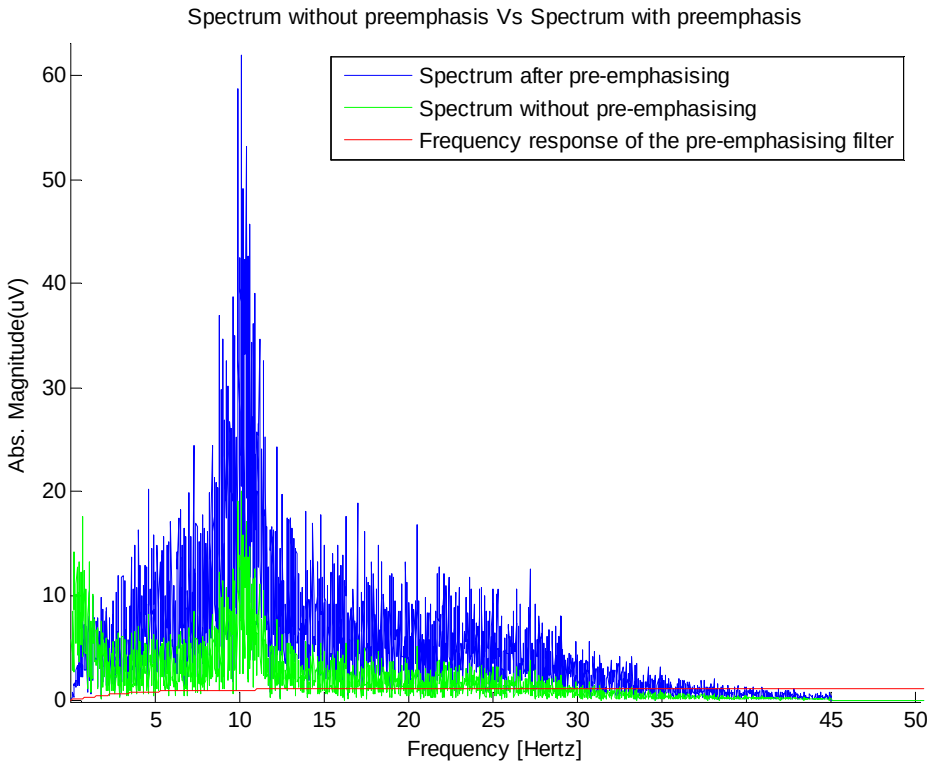

Figure 23: Spectrum of the EEG signal before and after pre-emphasising. The red line represents the frequency response of the pre-emphasising filter (cutoff=4.2Hz).

4.4 Segmentation

Before extraction of descriptive features, the pre-emphasised EEG signal was divided into segments of 2 seconds using a rectangular window function.
The length of the fixed segments must at least be comparable with the wave length of the lowest frequency of interest, and short enough to ensure stationarity within the window. According to Jansen, a window length of 1-5 seconds is suitable, best around 2 seconds. Levy was able to suggest a preference for short (around 2 seconds) segments to indentify changes in EEG signals more rapidly than any longer segments [2].

4.5 Feature extraction

It has been reported in several documents that hypoglycaemia symptoms can be seen as changes in frequency activity in EEG signals [2, 4, 5, 6, 13.]. Instead of define segments of the power spectrum within the conventional frequency bands (see Table 2), it was decided to analyze the entire frequency spectrum (1 - 45 Hz) without assuming changes in any range, to have a more general vision and a more advance statistical analysis.
Discrete wavelet transform was considered as a useful method for feature extraction [14, 15], but it means a more complex analysis and it is expected that it will not be found any important differences compare to a more classical analysis.
Finally, the more simple and efficient analysis consists of extracting from each segment 12 parameters that represent the amplitude and the frequency variability.
Root Mean Squared (RMS) amplitude was chosen as the only parameter that represents the amplitude information, and the auto-correlation function represents the frequency information.
The autocorrelation function is calculated as the expected value of the product of our EEG signal with a time-shifted version of itself.
With a simple calculation and analysis of the autocorrelation function, we can discover a few important characteristics about our signal. These include:
1. How quickly our random signal changes with respect to the time function.
2. Whether our signal has a periodic component and what the expected frequency might be.

Autocorrelation function: $r(i) = \sum_{n=0}^{N-i-1} x(n) \cdot x(n+i)$ *for* $i = 0, ..., p$

Normalised Autocorrelation Coefficients: $R(i) = {r(i)}/{r(0)}$ *for* $i = 1, ..., p$

During this study we calculate 24 Autocorrelation coefficients (lags). But then, we selected 12 out of 24 (lag 0, lag 2, lag 4, up to lag 22). If we have into account all the coefficients, since the sampling rate is 200 Hz, we would be focusing too much in high frequency variations.

To avoid aliasing during downsampling, a previous low pass filter would be needed.

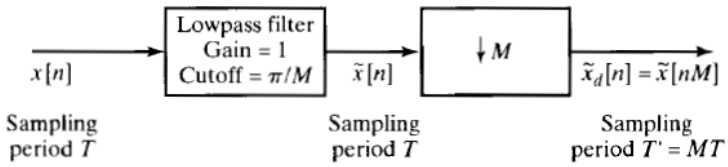

Figure 4.23 General system for sampling rate reduction by M.

Figure 24: Figure extracted from the book "Discrete-Time signal processing", Alan V. Oppenheim [17]. In our case, M=2.

Since we need a cutoff equal to π/2, corresponding to 50 Hz in a continuous signal, the low pass filter applied to avoid noise with cutoff equal to 45 Hz (see Section 4.2) also avoids aliasing during downsampling.

The mean-square value can be found by evaluating the autocorrelation where i=0, that is, lag 0.
The amplitude information, which is present in every coefficient, is then removed through normalisation with lag 0, leaving the auto-correlation coefficients for lag 2-22 to represent only frequency information.

The sets of normalised ACC's were chosen as features because they have proven to describe the feature space with the highest performance seen from a classification point of view in [2], compared to the reflection coefficients and the prediction coefficients (result of auto-regressive modelling of a stochastic process).

The RMS-value has been scaled by 3000 to get the same order of size as the normalised Autocorrelation coefficients.

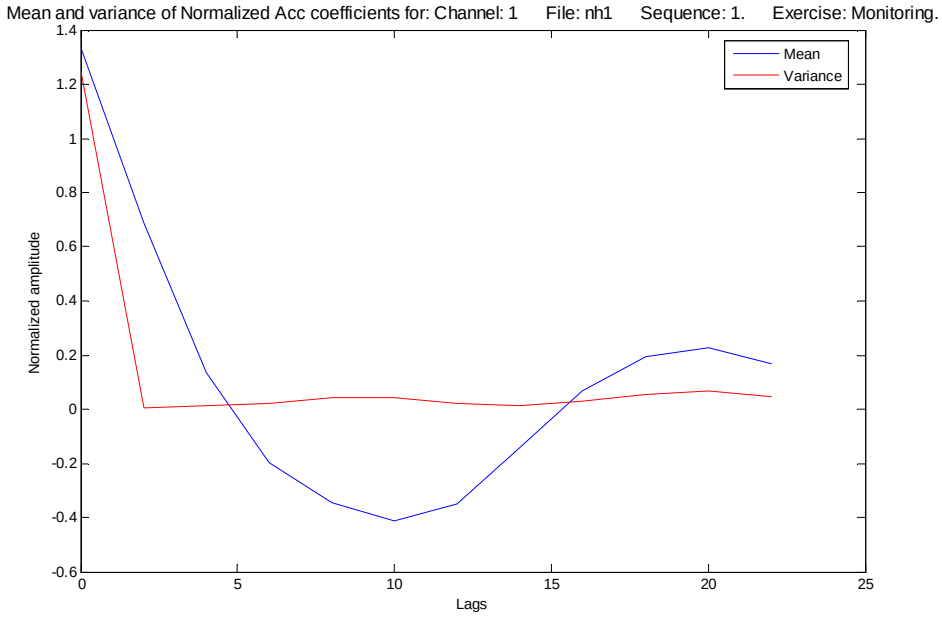

Figure 25: Mean of the Autocorrelation coefficients (Features) extracted for one sequence (baseline1) and its variance.

To show frequency changes the sets of features (ACCs) have been transformed into power spectra using linear prediction, by means of the Levinson Durbin recursion, after reverse the normalization by lag 0.
Linear prediction consists of estimating future values of a discrete-time signal as a linear function of previous samples. The most common way to estimate this value is using the Yule-Walker Auto-Regressive (AR) method, also called the autocorrelation method. It fits an autoregressive (AR) model to the windowed input data by minimizing the forward prediction error in the least squares sense. This formulation leads to the Yule-Walker equations:

$$\sum_{i=1}^{p} a_i R(i-j) = -R(j)$$

for $1 \leq j \leq p$, where R is the autocorrelation matrix, and a_i are the prediction coefficients optimized by minimizing the expected value of the squared error.
The Yule-Walker equations (in matrix formulation: $Ra = -r$) are solved by Levinson-Durbin recursion.

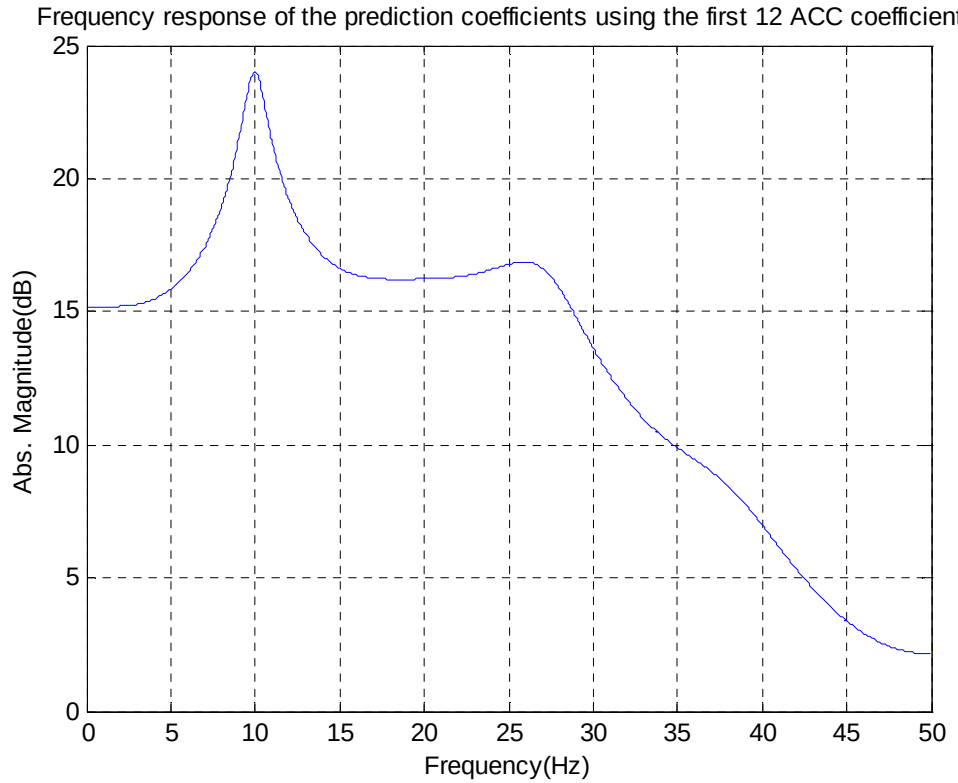

Figure 26: Example of the Frequency response of the prediction coefficients for a set of 12 features (Autocorrelation coefficients).

4.6 Pattern recognition / Unsupervised learning / Cluster analysis

The field of pattern recognition is concerned with the automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions such has classifying the data into different categories [16].
In the hypoglycaemic case, the set of features (called training set) do not have any corresponding target values (it is not known a priori what kind of regularities they may have), so it is called unsupervised learning problems. The goal in such unsupervised learning problems may be to discover groups of similar examples within the data, where it is called clustering.
Unsupervised learning was here performed by hierarchical cluster analysis.

Hierarchical clustering is a method of cluster analysis which seeks to build a hierarchy of clusters. Level one corresponds to n clusters and level n to one. Given any two samples x and x', at some level they will be grouped together in the same cluster, and whenever two samples are grouped in the same cluster, they remain together at all higher levels.

Strategies for hierarchical clustering generally fall into two types:
Agglomerative: This is a "bottom-up" approach: each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy.
Divisive: This is a "top-down" approach: all observations start in one cluster, and splits are performed recursively as one moves down the hierarchy.

In this study it has been used agglomerative procedures.

For every hierarchical clustering there is a corresponding tree, called dendrogram, that shows how the samples are grouped.

In order to decide which clusters should be combined (for agglomerative), a measure of dissimilarity between sets of observations is required. In most methods of hierarchical clustering, this is achieved by use of an appropriate **metric** (a measure of distance between pairs of observations), and a **linkage criteria** which specifies the dissimilarity of sets as a function of the pair-wise distances of observations in the sets.

First of all, each feature vector (12 Autocorrelation coefficients) extracted for each segment (2 seconds) form one class, represented by each class centroid (the mean value vector).

Metric: Having an m-by-n data matrix of Autocorrelation coefficients R, which is treated as m (1-by-n) row vectors x1, x2, ..., x_m, (n=12 features, m= number of segments of 2 seconds) the various distances between the vector x_r and x_s has been chosen as the *Euclidean distance*, according to $d_{rs}^{\,2} = (x_r - x_s)(x_r - x_s)'$.

Linkage criteria: To generate the hierarchical clustering tree it has been used the algorithm known as "Ward linkage". All the linkage algorithms to generate the hierarchical tree are based on different ways of measuring proximity between two groups of objects. If n_r is the number of objects in cluster r and n_s is the number of objects in cluster s, and x_{ri} is the i^{th} object in cluster r, the definitions of these various measurements are as follows:

-*Single linkage*, also called *nearest neighbor*, uses the smallest distance between objects in the two groups.
$d(r,s) = \min(dist(x_{ri}, x_{sj})), i \in (i,...,n_r),\ j \in (1,...,n_s)$

-*Complete linkage*, also called *furthest neighbor*, uses the largest distance between objects in the two groups.
$d(r,s) = \max(dist(x_{ri}, x_{sj})), i \in (i,...,n_r),\ j \in (1,...,n_s)$

The minimum and maximum measures represent two extremes in measuring the distance between clusters. Like all procedures that involve minima or maxima, they tend to be overly sensitive to data placement. This is advantageous when the true clusters are compact and roughly equal in size. However, when this is not the case, the resulting groupings can be meaningless. The use of averaging is an obvious way to ameliorate these problems [19].

-*Average linkage* uses the average distance between all pairs of objects in cluster r and cluster s.

$$d(r,s) = \frac{1}{n_r n_s} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} dist(x_{ri}, x_{sj})$$

-*Centroid linkage* uses the distance between the centroids of the two groups.

$d(r,s) = d(\overline{x_r}, \overline{x_s})$, where $\overline{x_r} = \frac{1}{n_{ri}} \sum_{i=1}^{n} x_{ri}$ and $\overline{x_s}$ is defined similarly.

The centroid method can produce a cluster tree that is not monotonic. This occurs when the distance from the union of two clusters, $r \cup s$, to a third cluster is less than the distance from either r or s to that third cluster. In this case, sections of the dendrogram change direction. This is an indication that you should use another method.

-*Ward linkage* uses the incremental sum of squares; that is, the increase in the total within-group sum of squares as a result of joining groups r and s. In short, this method attempts to minimize the Error Sum of squares (J_e) of any two clusters that can be formed at each step. The within-group sum of squares of a cluster is defined as the sum of the squares of the distance between all objects in the cluster and the centroid of the cluster.
Initially, J_e is 0, since every individual is in a cluster of its own. At each stage the link created is the one that makes the least increase to J_e.
It is given by
$d(r,s) = n_r n_s d_{rs}^2 / (n_r + n_s)$, where d_{rs}^2 is the distance between cluster r and cluster s defined in the Centroid linkage. And n_r and n_s are the number of data units within cluster r and s. This measure can be interpreted as the square of the increase in the error sum of squares if cluster r and s are merged.
The use of this method tends to favour growth by adding singletons or small clusters to large clusters over merging medium-sized clusters. While the final partition may not minimize the error sum of squares (J_e), it usually provides a very good starting point for further iterative optimization [18-19].

Examples for Non diabetic patient number 1, monitoring exercise, sequence 1 (baseline 1):

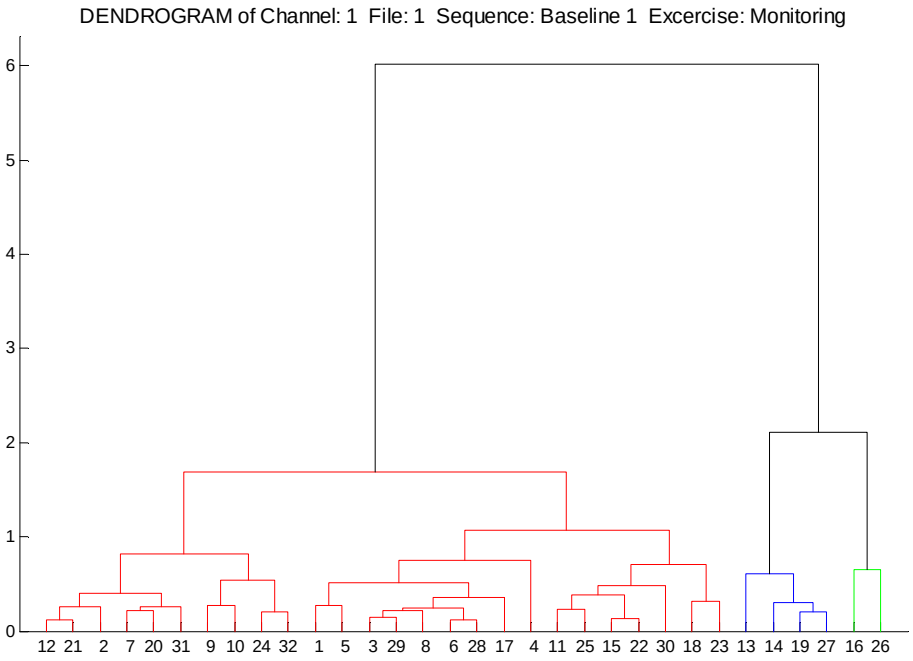

Figure 27: Dendrogram of Channel 1, file 1, baseline 1, Monitoring exercise.

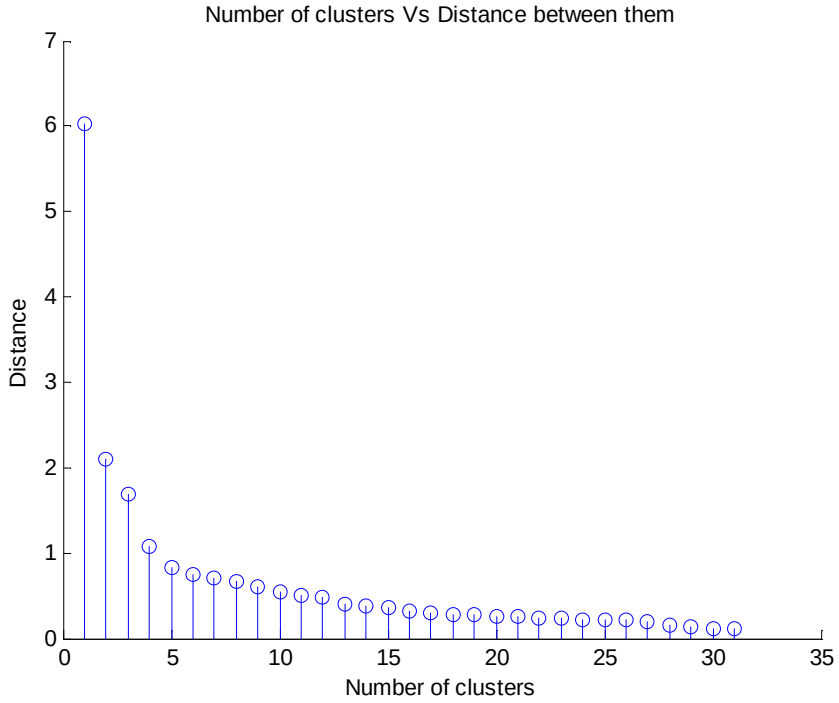

Figure 28: Number of clusters Vs Distance between them.

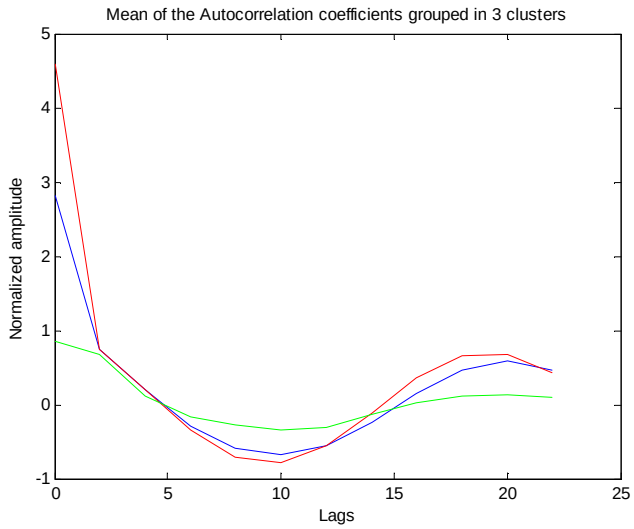

Figure 29: Example of the mean of the Autocorrelation coefficients when the sequence 1 (baseline 1) of Non diabetic patient 1 has been grouped in 3 clusters. The colours correspond to the mergers represented in the dendogram using the same colour.

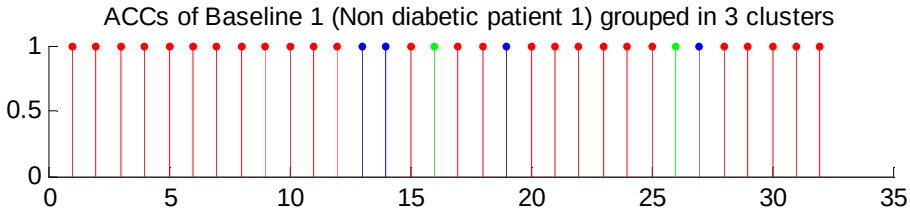

Figure 30: Example of 32 clusters (12 Autocorrelation coefficients each one) grouped in 3 clusters marked with different colours (blue, green or red). The colours correspond to the mergers represented in the dendogram using the same colour.

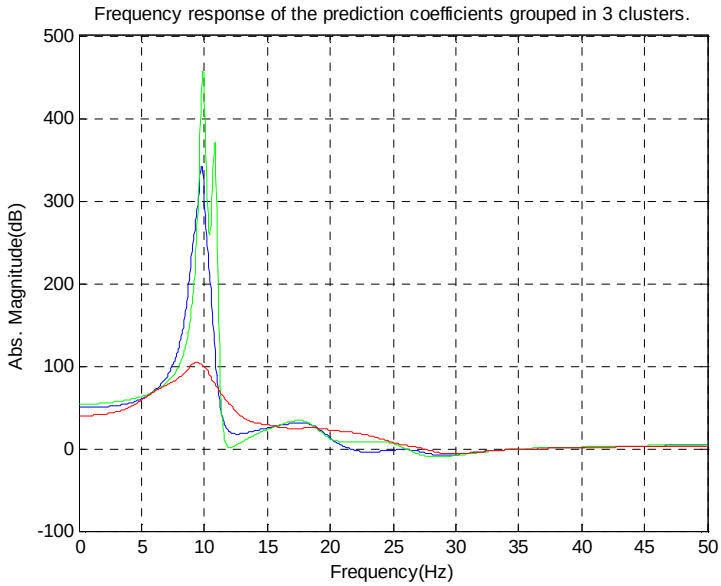

Figure 31: Frequency response of the prediction coefficients for baseline 1 of Non diabetic patient 1, grouped in 3 clusters. The colours correspond to the mergers represented in the dendogram using the same colour.

5. Detection of hypoglycaemia: Results

From the four exercises, two of them (monitoring and AEP) represent rest activity, and the other two (AQT and CalCAP) represent patients submitted to a demanding mental task. Monitoring exercise was selected to represent a patient during rest activity, and CalCAP as the one that shows high brain activity.

From the four pre-processed channels, channel 1 (F3-C3) was selected to perform all the different analysis.

5.1 Data cluster analysis separately for each patient

Taking the data from the 3 different sequences: Baseline 1, baseline 2 and maximum hypoglycaemic period, it has been visually analyzed how clusters are merged for each patient and for different number of clusters (from 4 to 11 clusters). The goal is to visualize that there are one or more group of clusters which occur more often during the hypoglycaemia sequence. This means that those patients show patterns related to hypoglycaemic events.

It will be shown next two examples of individual patients: Non diabetic person number 19 and diabetic patient number 5. Figures of all patients have been added to the additional DVD.

5.1.1 Monitoring exercise

For the Non diabetic person number 19, it is found that when the data are grouped in 5 clusters, the clusters number 3 and 5 show strong patterns that occur principally during the hypoglycaemic period. These are the results for this patient:

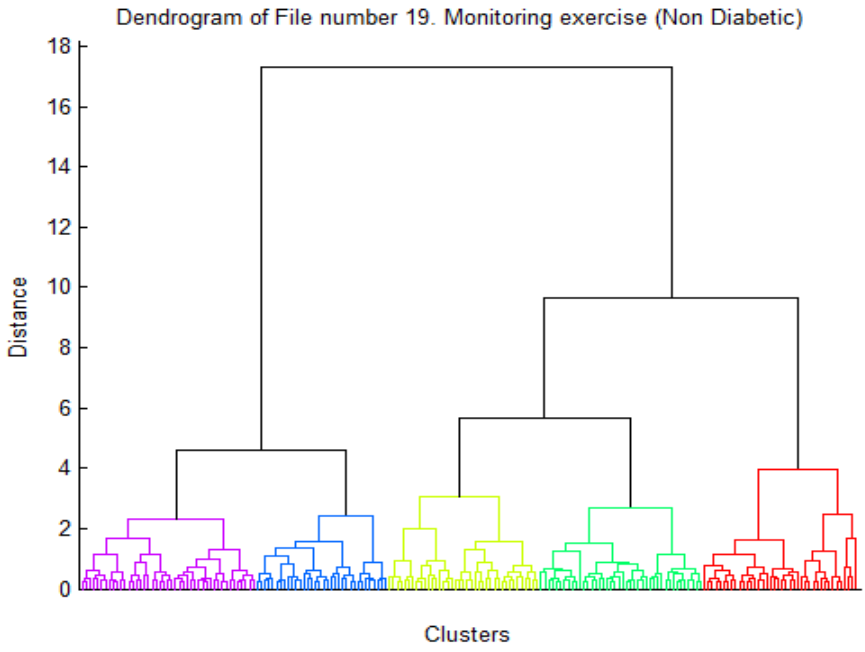

Figure 32: Dendrogram of file number 19 (Non diabetic person).

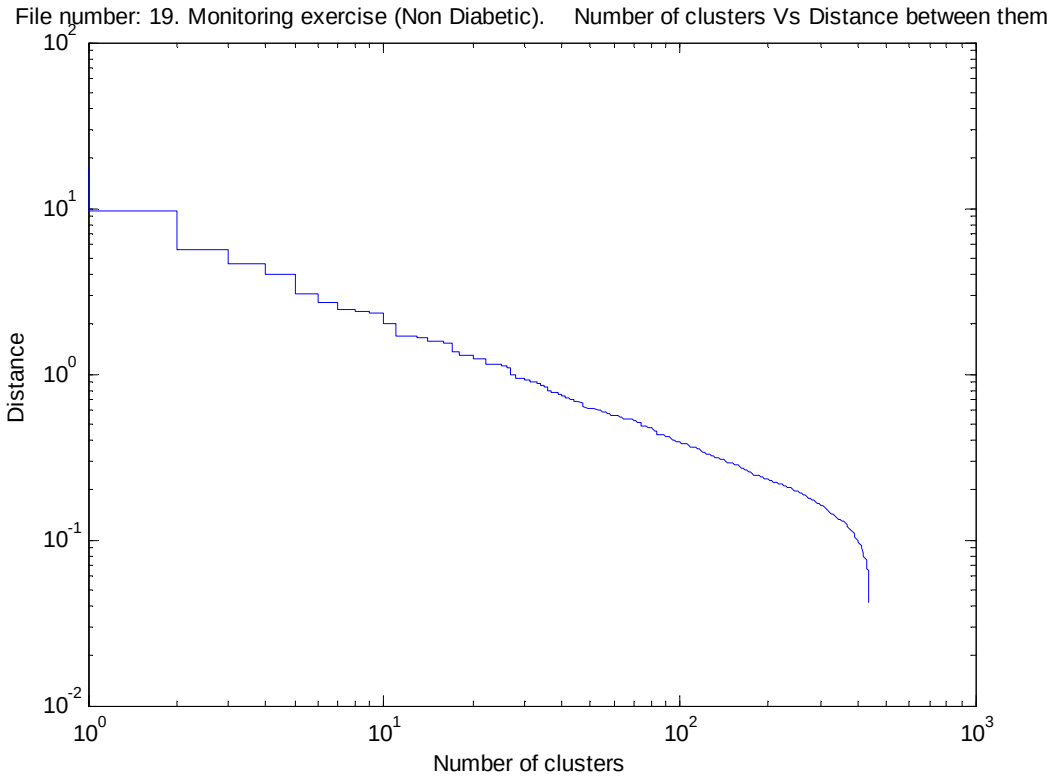

Figure 33: Number of clusters Vs distance between them using a log scale in both axis.

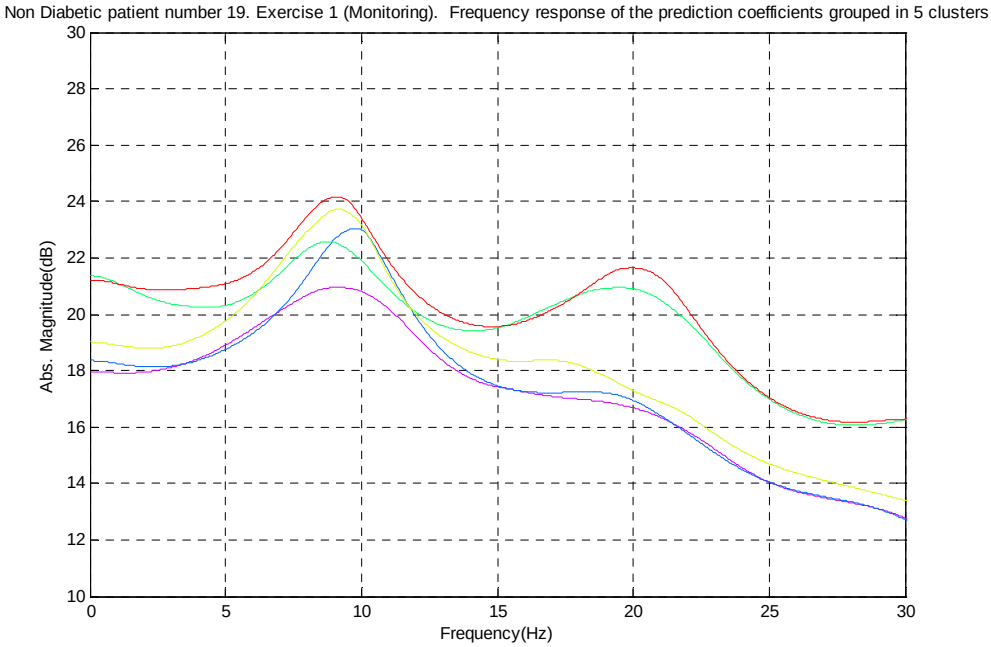

Figure 34: Frequency response of the prediction coefficients grouped in 5 clusters for the Non diabetic patient number 19. The colours correspond with the other figures of this patient, so the red and green lines represent hypoglycaemic patterns.

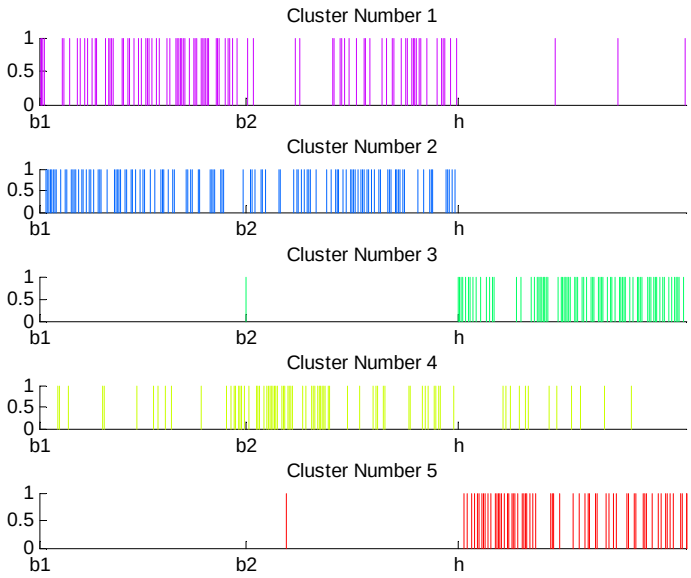

Figure 35: Data from Non diabetic patient number 19 grouped in 5 clusters. X-axis represents the starting point of the baseline 1, baseline 2 and maximum hypoglycaemia period. The five figures represent the five clusters formed (colours according to the dendogram and the other figures of this person). Clusters number 3 and 5 are formed by clusters that only occur during the hypoglycaemia sequence, so it can be said that this patient shows some patterns that differ from normal blood glucose level periods than for the hypoglycaemic period.

For the Diabetic patient number 5 these are the figures obtained:

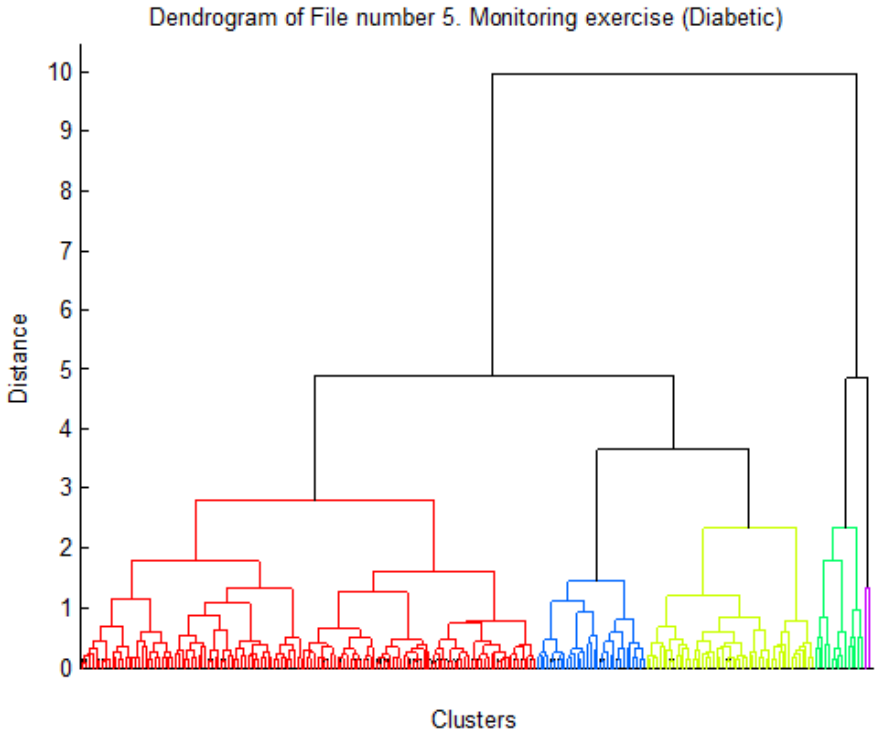

Figure 36: Dendrogram of the Diabetic patient number 5.

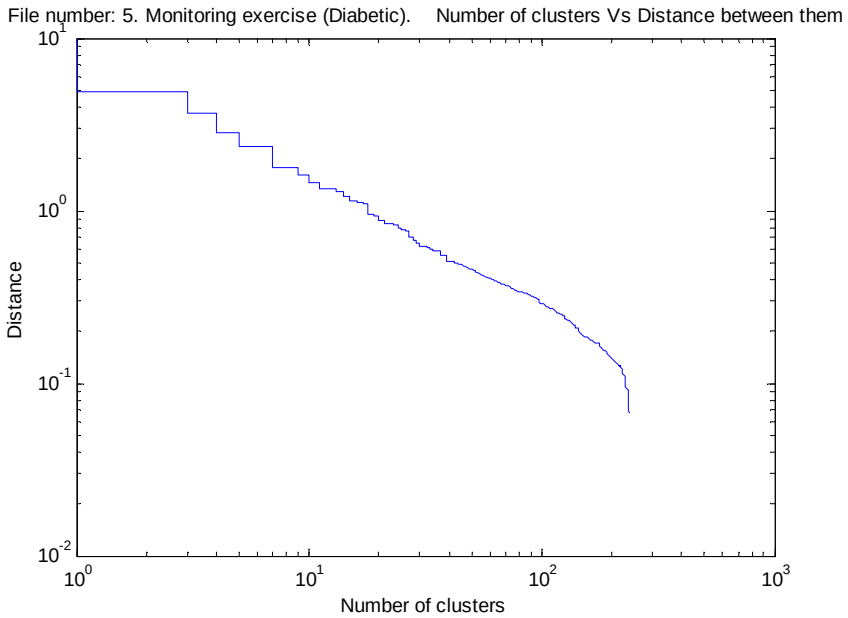

Figure 37: Number of clusters Vs distance between them using a log scale in both axis.

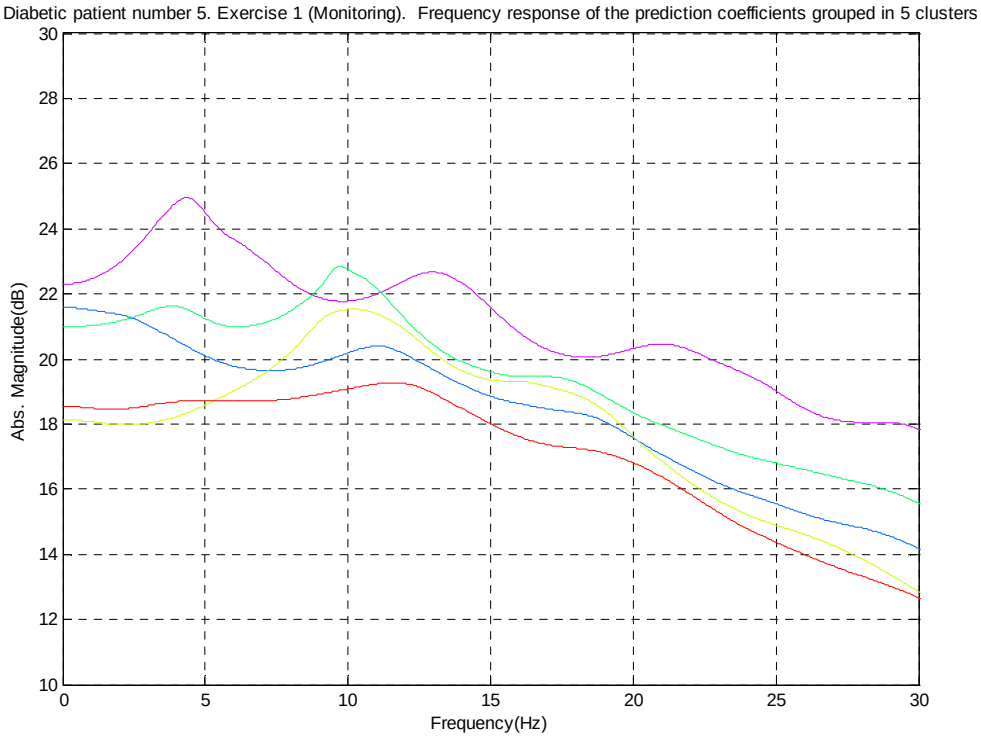

Figure 38: Frequency response of the prediction coefficients grouped in 5 clusters for the Diabetic patient number 5. The colours correspond with the other figures of this patient, so the green and yellow lines represent hypoglycaemic patterns.

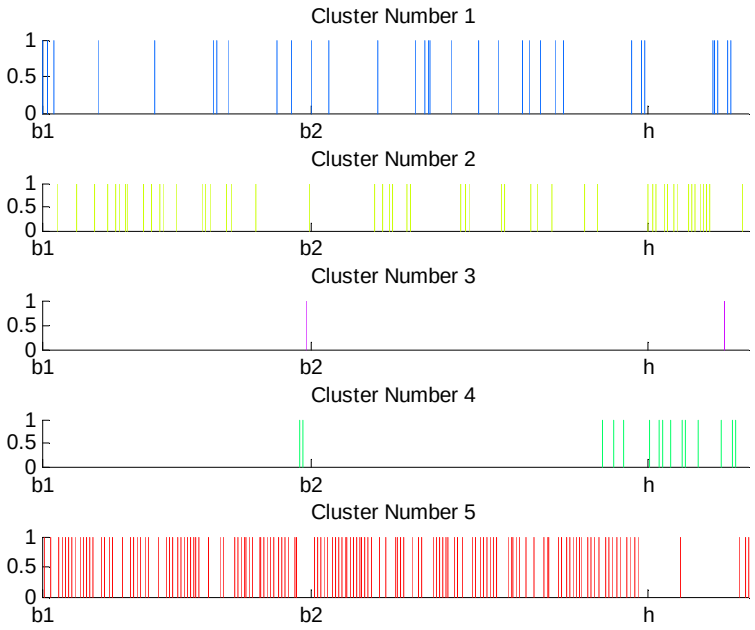

Figure 39: Data from Diabetic patient number 5 grouped in 5 clusters. X-axis represents the starting point of the baseline 1, baseline 2 and maximum hypoglycaemia period. The five figures represent the five clusters formed (colours according to the dendogram and the other figures of this person). Clusters number 2 and 4 are formed by clusters that occur more often during the hypoglycaemia sequence, so it can be said that this patient shows some patterns that differ from normal blood glucose level periods than for the hypoglycaemic period. Cluster number 3 is probably noise or artefacts not removed.

One example of clusters that do not show any hypoglycaemic patterns occur for the diabetic patient number 4:

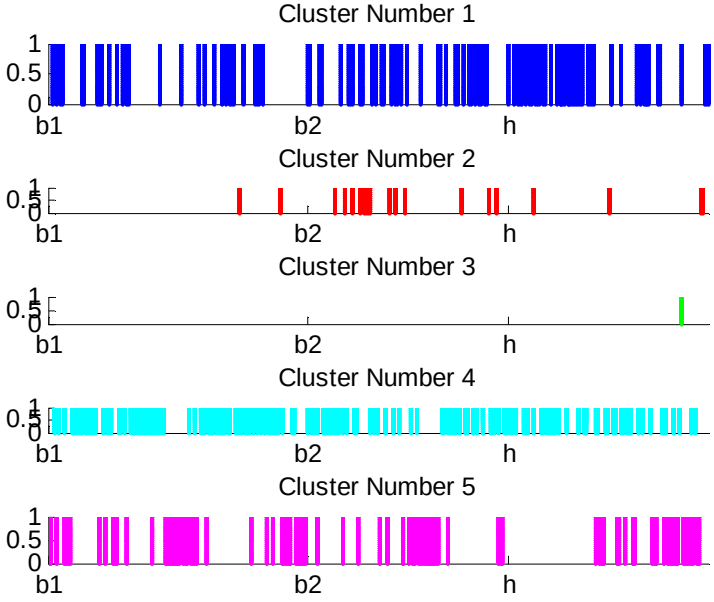

Figure 40: Example of a person (diabetic patient number 4) that does not show hypoglycaemic patterns.

Every patient has been inspected separately for the monitoring exercise and classified in the following table, represented by some stars according to how clearly they show different patterns between normoglycaemia and hypoglycaemia sequences.

☆☆☆☆ = No hypoglycaemic patterns are shown.
★☆☆☆ = Very few hypoglycaemic patterns are shown for 8 or more clusters.
★★☆☆ = Not very clear hypoglycaemic patterns are shown for 6 or more clusters.
★★★☆ = Clear hypoglycaemic patterns appear for 6 or more clusters.
★★★★ = Hypoglycaemic patterns clearly appear for 3, 4 and more clusters.

Table 11: List of patients classified according to how clear they show different patterns between normoglycaemic sequences (baselines 1 and 2) and hypoglycaemic sequence for the monitoring exercise.

Monitoring exercise: Stars represent how clear a person develops patterns which occur more often during the hypoglycaemic sequence.		
Patient Number	**Diabetic patients**	**Non diabetic patients**
1	★★★★	★★★★
2	☆☆☆☆	★★★★
3	★★★★	★☆☆☆
4	☆☆☆☆	☆☆☆☆
5	★★★★	☆☆☆☆
6	★★☆☆	★★★☆
7	★☆☆☆	★☆☆☆
8	★★★★	★★★☆
9	★★★★	☆☆☆☆
10	★☆☆☆	☆☆☆☆
11	★☆☆☆	☆☆☆☆
12	★★★★	★☆☆☆
13		★☆☆☆
14		★★☆☆
15		★☆☆☆
16		☆☆☆☆
17		★☆☆☆
18		★★★☆
19		★★★★
20		☆☆☆☆

5.1.2 CalCAP exercise

Figures for the Non diabetic person number 19:

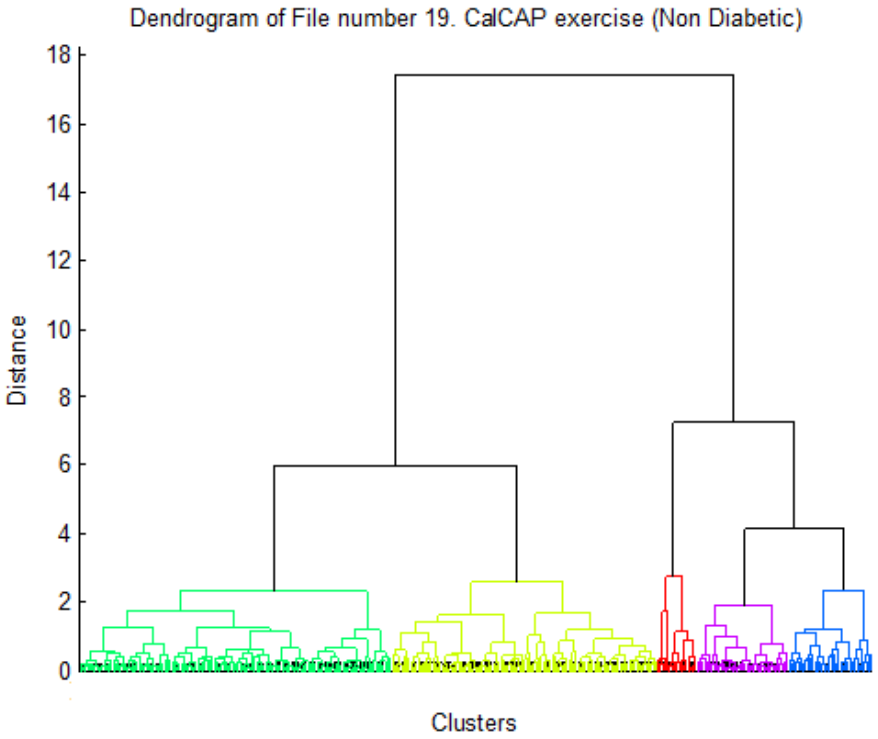

Figure 41: Dendrogram of Non diabetic person number 19 for the CalCAP exercise.

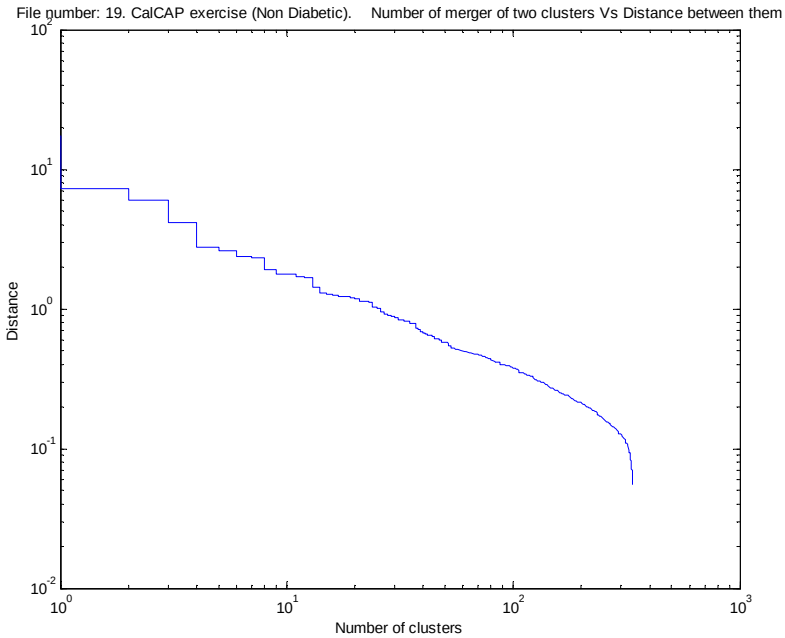

Figure 42: Number of clusters Vs distance between them using a log scale in both axis.

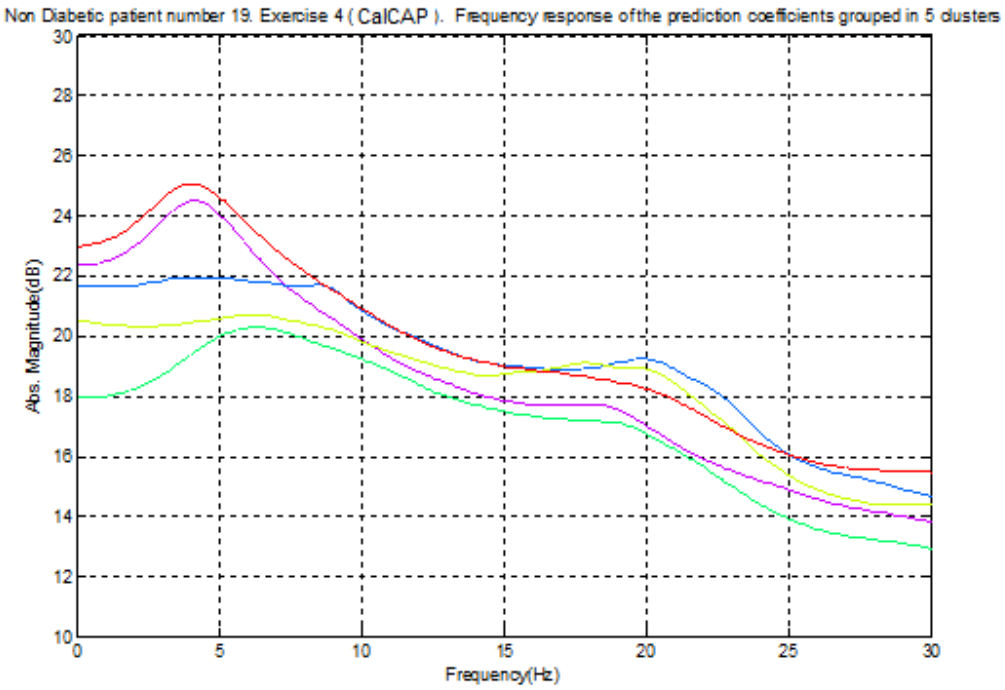

Figure 43: Frequency response of the prediction coefficients grouped in 5 clusters for the Non diabetic patient number 19 for the CalCAP exercise. The colours correspond with the other figures of this patient, so the blue and yellow lines represent hypoglycaemic patterns.

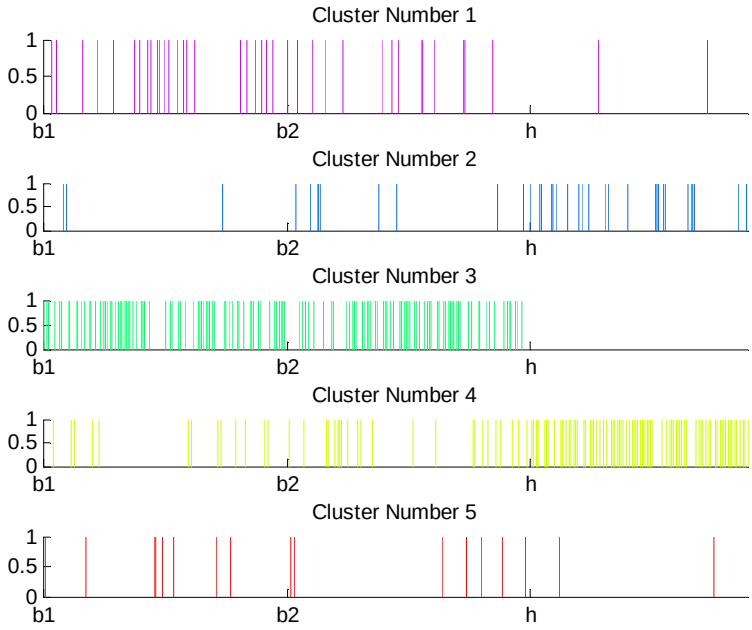

Figure 44: Data from Non diabetic patient number 19 grouped in 5 clusters for the CalCAP exercise. X-axis represents the starting point of the baseline 1, baseline 2 and maximum hypoglycaemia period. The five figures represent the five clusters formed (colours according to the dendogram and the other figures of this person). Clusters number 2 and 4 are formed by clusters that only occur during the hypoglycaemia sequence, so it can be said that this patient shows some patterns that differ from normal blood glucose level periods than for the hypoglycaemic period.

Figures for the Diabetic patient number 5:

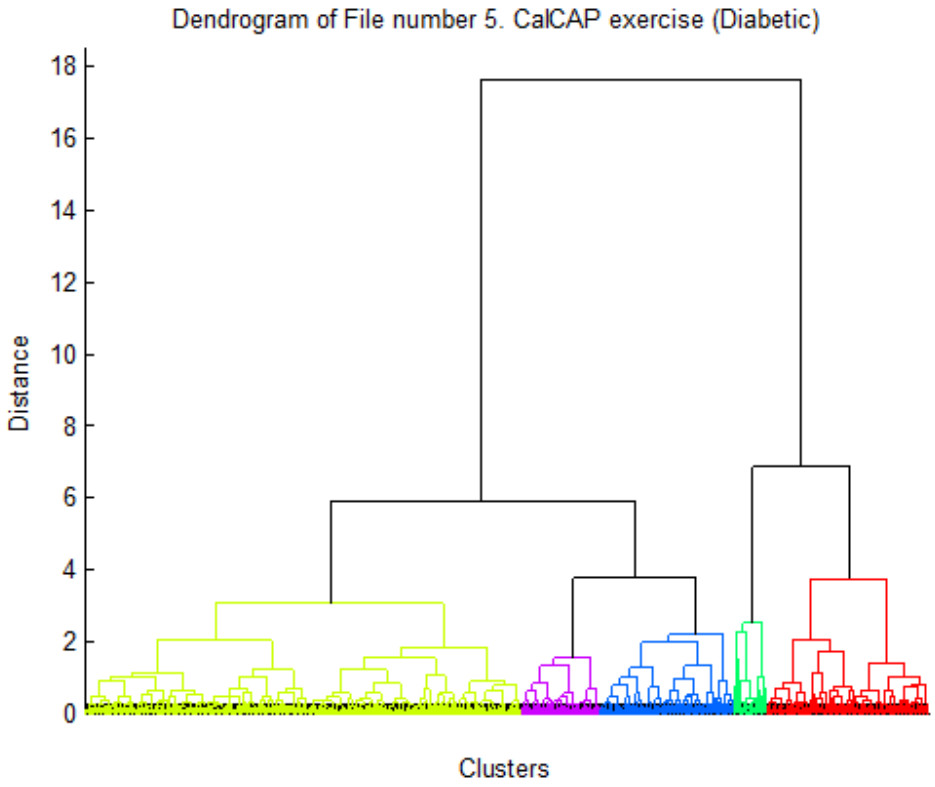

Figure 45: Dendrogram of Diabetic number 5 for the CalCAP exercise.

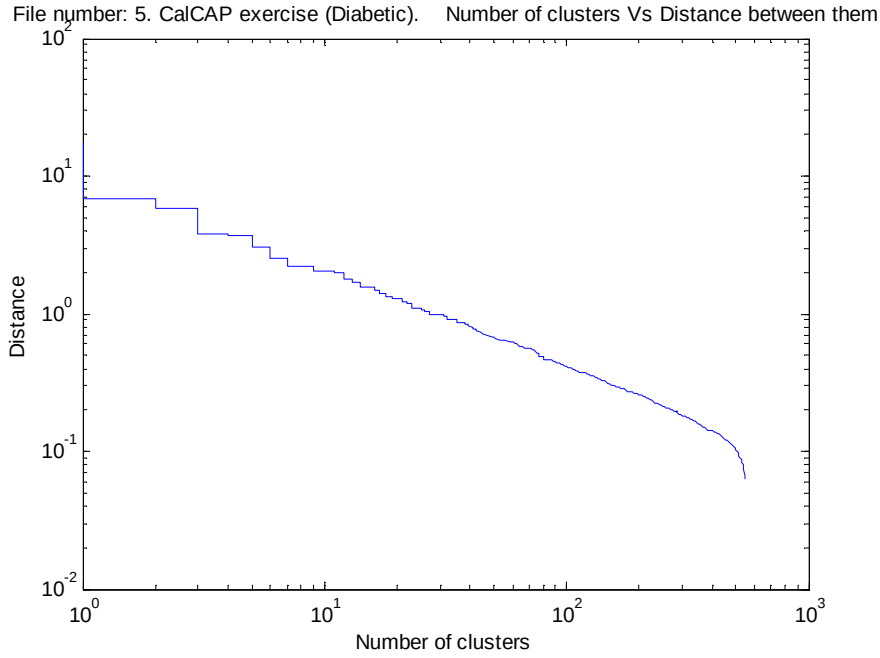

Figure 46: Number of clusters Vs distance between them using a log scale in both axis.

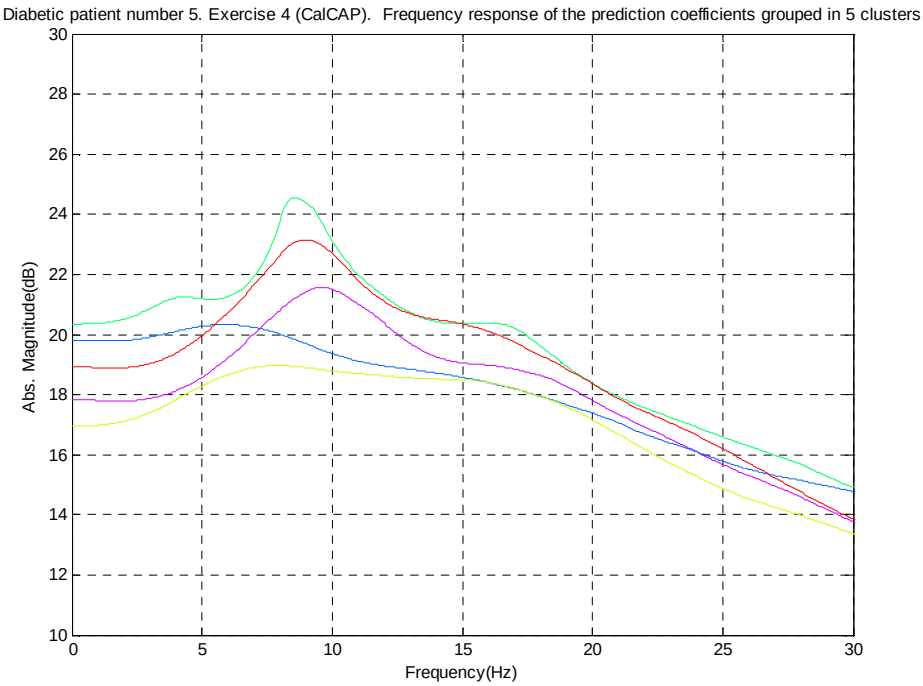

Figure 47: Frequency response of the prediction coefficients grouped in 5 clusters for the Diabetic patient number 5 for CalCAP exercise. The colours correspond with the other figures of this patient, so the green and red lines represent hypoglycaemic patterns.

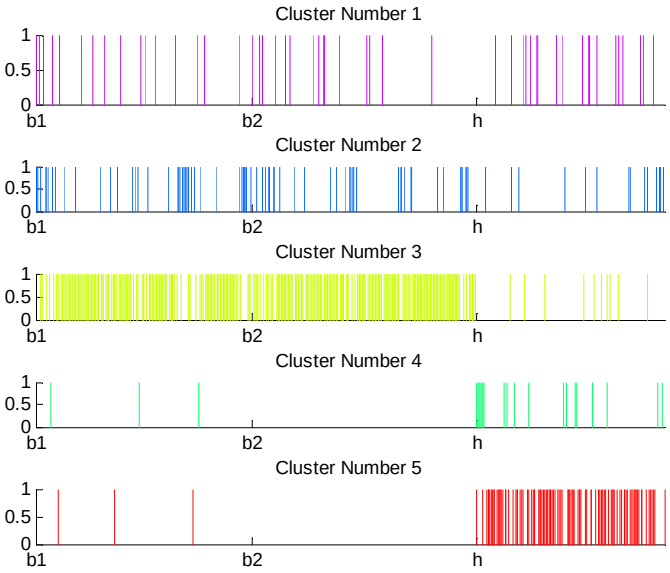

Figure 48: Data from Diabetic patient number 5 grouped in 5 clusters. X-axis represents the starting point of the baseline 1, baseline 2 and maximum hypoglycaemia period. The five figures represent the five clusters formed (colours according to the dendogram and the other figures of this person). Clusters number 4 and 5 are formed by clusters that only occur during the hypoglycaemia sequence, so it can be said that this patient shows some patterns that differ from normal blood glucose level periods than for the hypoglycaemic period.

In the same way than for the monitoring exercise, every patient has been inspected separately for the CalCAP exercise and classified in the following table, represented by

some stars according to how clearly they show different patterns between normoglycaemia and hypoglycaemia sequences.

Table 12: List of patients classified according to how clear they show different patterns between normoglycaemic sequences (baselines 1 and 2) and hypoglycaemic sequence for the CalCAP exercise.

CalCAP exercise: Stars represent how clear a person develops patterns which occur more often during the hypoglycaemic sequence.		
Patient Number	**Diabetic patients**	**Non diabetic patients**
1	★★★★	☆☆☆☆
2	★☆☆☆	★☆☆☆
3	★★★☆	☆☆☆☆
4	★☆☆☆	☆☆☆☆
5	★★★★	☆☆☆☆
6	★★★☆	★★★☆
7	★★★★	☆☆☆☆
8	★★★★	★★★☆
9	★★★★	☆☆☆☆
10	★★★☆	★★☆☆
11	★★★★	☆☆☆☆
12	★★★☆	☆☆☆☆
13		★★☆☆
14		☆☆☆☆
15		☆☆☆☆
16		☆☆☆☆
17		★☆☆☆
18		☆☆☆☆
19		★★★★
20		★☆☆☆

5.1.3 Monitoring & CalCAP exercises together

After visualization of every patient, some of them showed early hypoglycaemic patterns, just classifying the data in 3 or 4 clusters. But for others, the symptoms presented were more limited, and if they appeared, it used to be from 8 or 9 clusters. A threshold of 10 clusters was finally chosen as the best way to visualize hypoglycaemic patterns, whether they appear early or not.

Figures from Non diabetic person number 19:

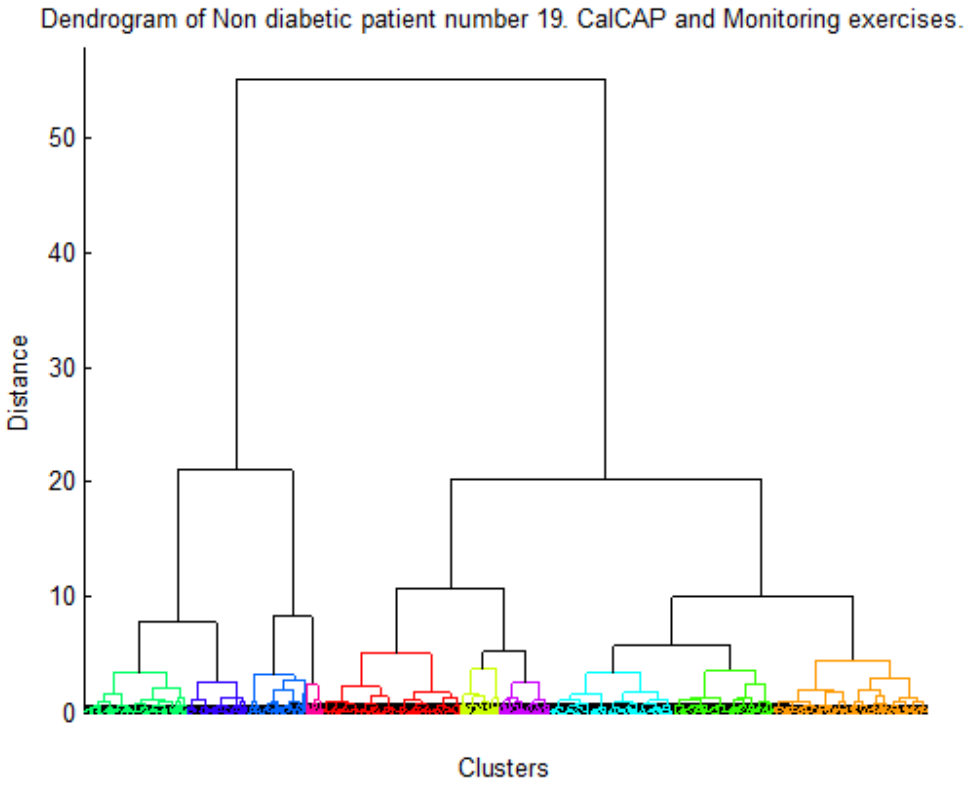

Figure 49: Dendrogram of Non diabetic 19 for CalCAP and Monitoring exercises.

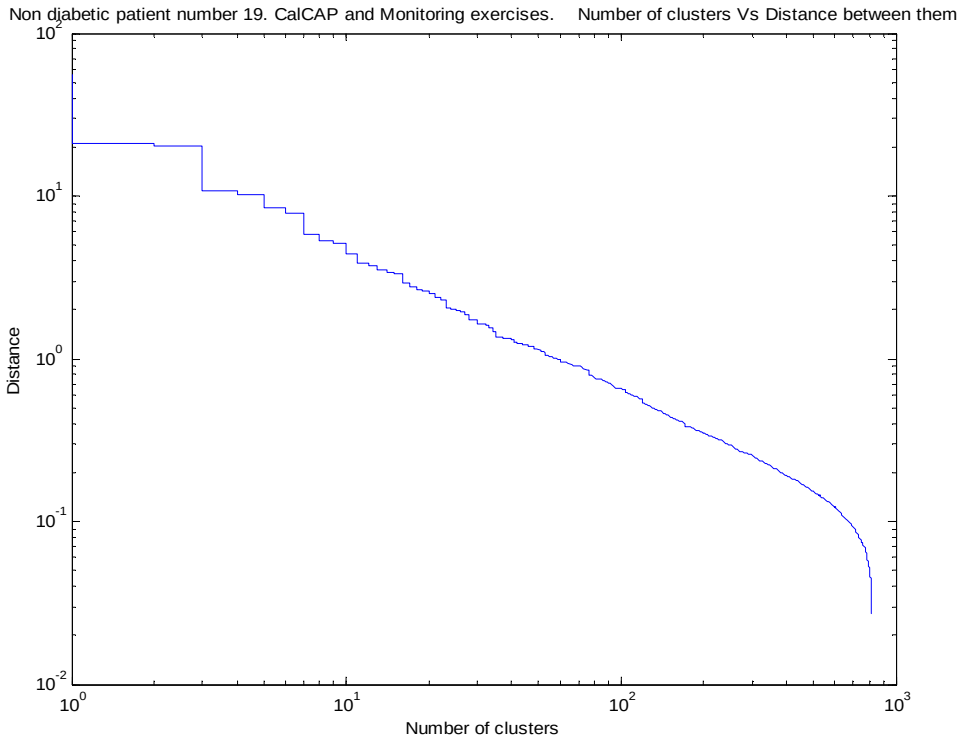

Figure 50: Number of clusters Vs distance between them using a logarithmic scale in both axis.

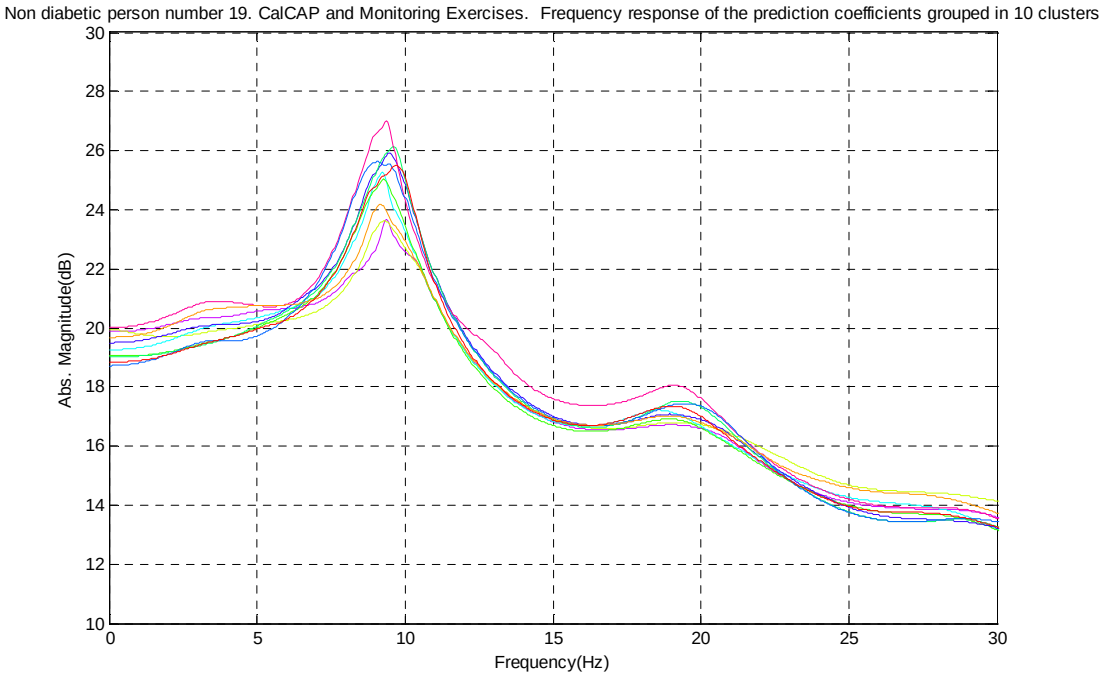

Figure 51: Frequency response of the prediction coefficients for Non diabetic person number 19, for Monitoring and CalCAP exercises together, grouped in 10 clusters.

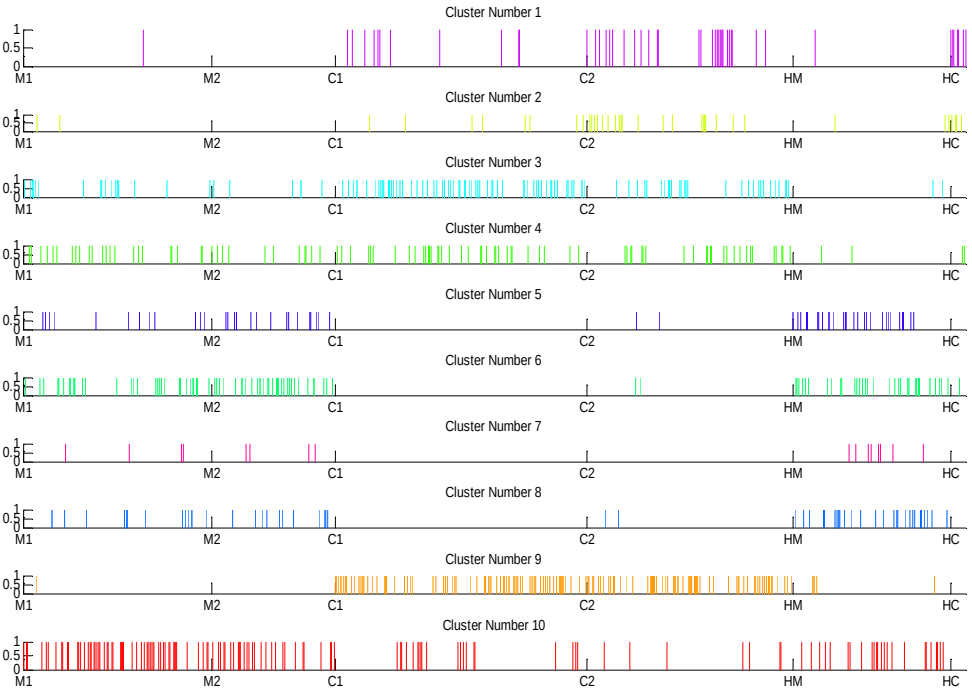

Figure 52: Data from Monitoring and CalCAP exercise for Non diabetic number 19, grouped in 10 clusters. Colours correspond to the ones of the dendogram and the other figures of this patient. In this figure we can appreciate that the differences between exercises are stronger than the differences between hypo and normoglycaemia.

Figures from Diabetic patient number 5:

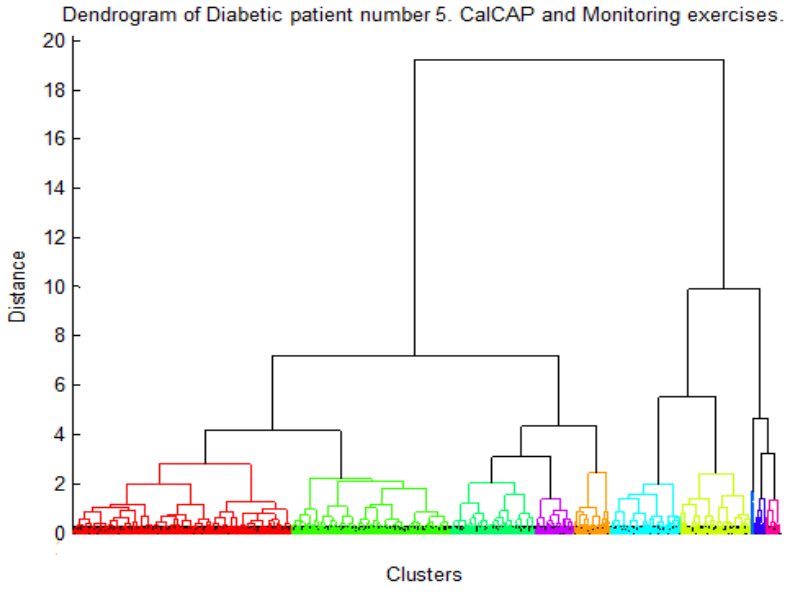

Figure 53: Dendrogram of Diabetic patient 5 for CalCAP and Monitoring exercises.

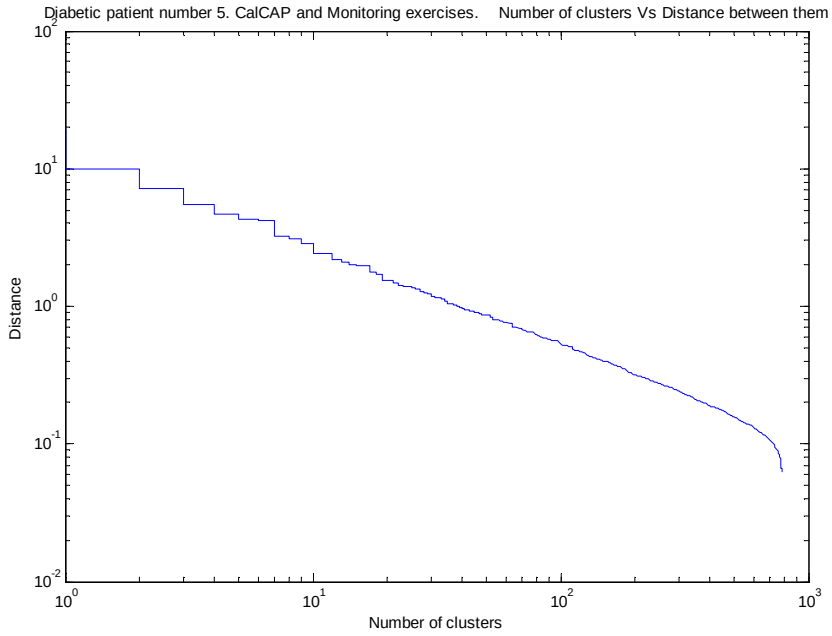

Figure 54: Number of clusters Vs distance between them using a logarithmic scale in both axis.

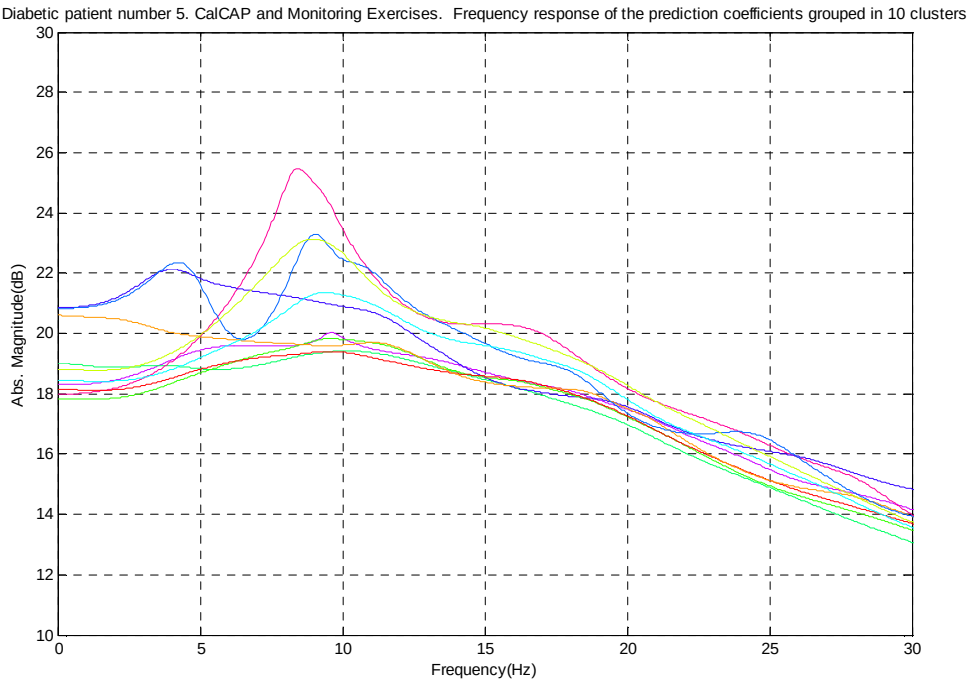

Figure 55: Frequency response of the prediction coefficients for diabetic patient number 5, for Monitoring and CalCAP exercises together, grouped in 10 clusters.

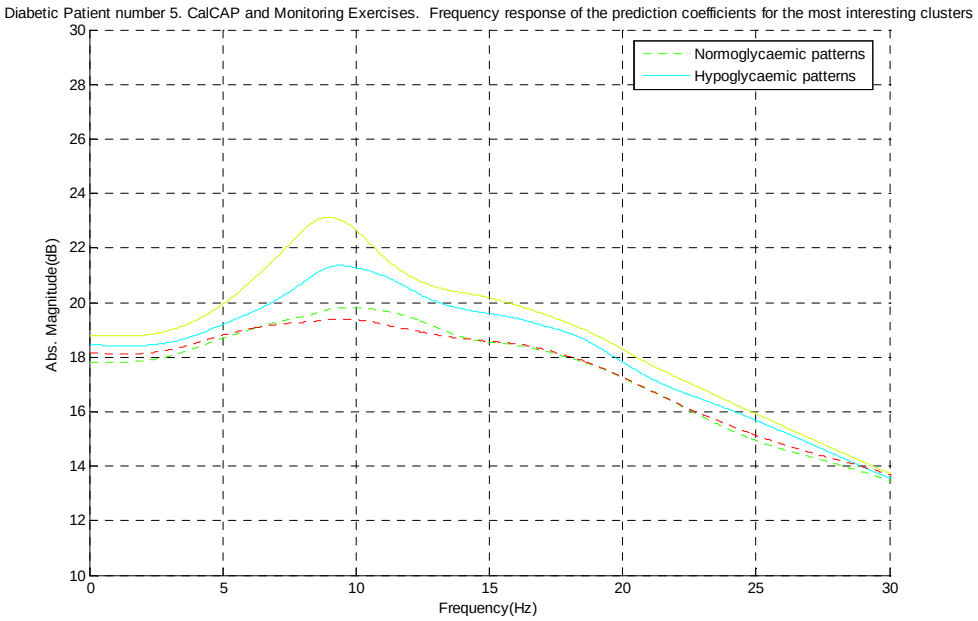

Figure 56: Frequency response of the prediction coefficients for Non diabetic person number 19, for Monitoring and CalCAP exercises together, for the most interesting clusters that represent the differences between hypoglycaemia and normoglycaemia.

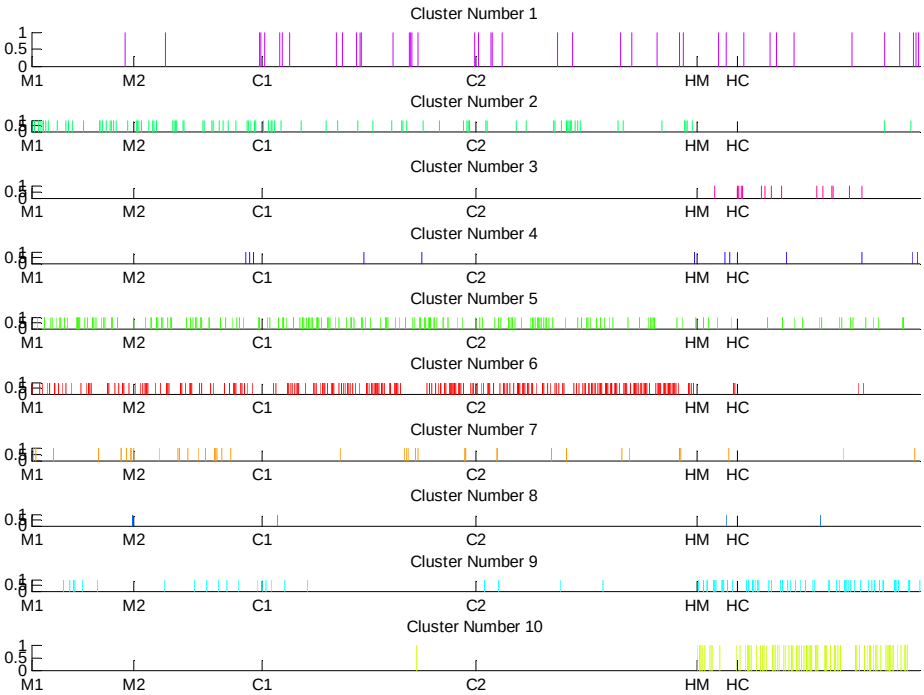

Figure 57: Data from Monitoring and CalCAP exercise for Diabetic number 5, grouped in 10 clusters. Colours correspond to the ones of the dendogram and the other figures of this patient. In this figure we can appreciate differences between hypoglycaemia and normoglycaemia.

5.1.4 Results for individual patients

The conclusions about the graphs of the two previous examples (non diabetic number 15 and diabetic number 9) are next.

Non diabetic number 19:

Monitoring:

It shows higher amplitude from 15 to 25 Hz activity during hypoglycaemia compared to normoglycaemia sequences. It seems that during hypoglycaemia the first peak (8-12Hz) is moved slightly to the left (from 6 to 12 Hz).

CalCAP:

It seems to take place exactly the same than in the monitoring analysis.

Monitoring&CalCAP:

Patterns differ more from exercise to exercise (monitoring to CalCAP) than from normo- to hypoglycaemia for this person.

Diabetic number 5:

Monitoring:

Higher activity from 15 to 20 Hz and from 8 to 12 Hz (alpha rhythm) during hypoglycaemia sequence compared to the two normoglycaemic sequences. It seems that during hypoglycaemia the first peak (alpha rhythm) is moved slightly to the left.

CalCAP:

Higher content from 15 to 25 Hz during hypoglycaemic sequence again.

Monitoring&CalCAP:

In this patient, patterns differ more from normo- to hypoglycaemia than from exercise to exercise (monitoring to CalCAP).

It is easy to appreciate that there is higher amplitude in frequencies from 5 to 20 Hz (mostly in the range of 8 to 12 Hz (alpha rhythm)) and the first peak has been moved again slightly to the left.

General results after analyzing each patient individually:

From table 11 and 12 that show how clearly each patient show hypoglycaemic patterns, it has been created the following table, summarizing how many patients develop clear and early hypoglycaemic pattens, clear but not so early, less clear, a few of them and non existing. Doing this for both exercises analyzed.

Table 13: Summary of how clear patients show hypoglycaemic patterns, for both monitoring and CalCAP exercises.

Monitoring Exercise		CalCAP Exercise	
Number of Diabetic Patients	Number of Non Diabetic Patients	Number of Diabetic Patients	Number of Non Diabetic Patients
6 ★★★★	3 ★★★★	6 ★★★★	1 ★★★★
0 ★★★☆	3 ★★★☆	4 ★★★☆	2 ★★★☆
1 ★★☆☆	1 ★★☆☆	0 ★★☆☆	2 ★★☆☆
3 ★☆☆☆	6 ★☆☆☆	2 ★☆☆☆	3 ★☆☆☆
2 ☆☆☆☆	7 ☆☆☆☆	0 ☆☆☆☆	12 ☆☆☆☆

From this table, some statistics of our group of patients have been calculated.

Table 14: Statistics analyzing how many persons (from our group of subjects) show hypoglycaemic patterns and how clear they do it.

	Monitoring		CalCAP	
	Diabetic	Non diabetic	Diabetic	Non diabetic
% patients who show some patterns (1 or more stars)	10/12 = 83%	13/20 = 65%	12/12 = 100%	8/20 = 40%
% patients who show clearly patterns (3 or 4 stars)	6/12 = 50%	6/20 = 30%	10/12 = 83%	3/20 = 15%

According to these results, it can be concluded that Diabetic patients develop more hypoglycaemic patterns (patterns occurring more often during the hypoglycaemic sequence than during the normoglycaemic sequences) than Non diabetic patients. The reason seems to be their many several past hypoglycaemic events.

During CalCAP (exercise where patients are focus by doing some computer exercises), the differences between Diabetic and Non diabetic are more marked. Diabetic patients show clearer hypoglycaemic patterns and Non diabetic patients almost do not show them. This is an interesting observation that points in the direction that when performing a somehow demanding mental task during hypoglycaemia, diabetic patients (who most likely have had several previous hypoglycaemia attacks) seems to be more affected.

5.2 Data cluster analysis including different groups of patients

Three different programs have been developed:

1. Including monitoring sequence for all the patients that show some hypoglycaemic patterns (one star or more than one in table 11), diabetic and non diabetic patients together.
2. The same than the previous program but for the CalCAP exercise (table 12).
3. Monitoring and CalCAP exercises together including the patients represented by two or more stars (tables 11 and 12).

Patients included:

1. Program 1: Monitoring exercise:
    - Diabetic patients: 1, 3, 5, 6, 7, 8, 9, 10, 11, 12.
    - Control patients: 1, 2, 3, 6, 7, 8, 12, 13, 14, 15, 17, 18, 19.
2. Program 2: CalCAP exercise:
    - Diabetic patients: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12.
    - Control patients: 2, 6, 8, 10, 13, 17, 19, 20.
3. Program 3: Monitoring and CalCAP exercises together:
    - Monitoring exercise:
        - Diabetic patients: 1, 3, 5, 6, 8, 9, 12.
        - Control patients: 1, 2, 6, 7, 8, 14, 18, 19.
    - CalCAP exercise:
        - Diabetic patients: 1, 3, 5, 6, 7, 8, 9, 10, 11, 12.
        - Control patients: 6, 8, 10, 13, 19.

It has been included for each patient the 3 different sequences as before: Baseline 1, baseline 2 and maximum hypoglycaemic period.

To study individual patients a threshold of 10 clusters was finally chosen as the best way to visualize hypoglycaemic patterns, whether they appear early or not. In the same way, when so many patients are included, more clusters are needed to visualize hypoglycaemic patterns. The optimum number of clusters selected for these three programs is 20.

5.2.1 Program 1: Monitoring exercise

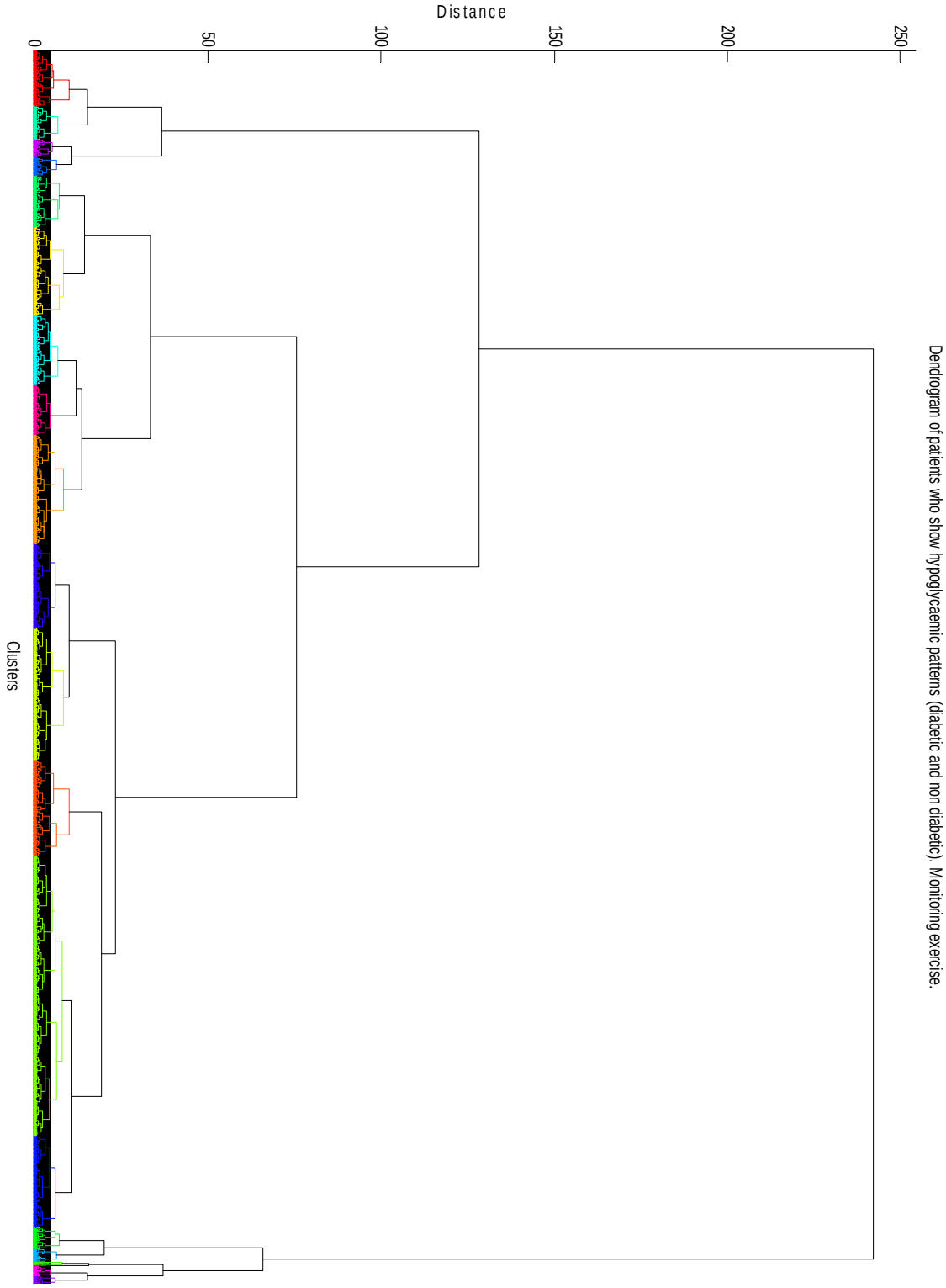

Figure 58: Dendrogram of patients who show some hypoglycaemic patterns. Monitoring exercise

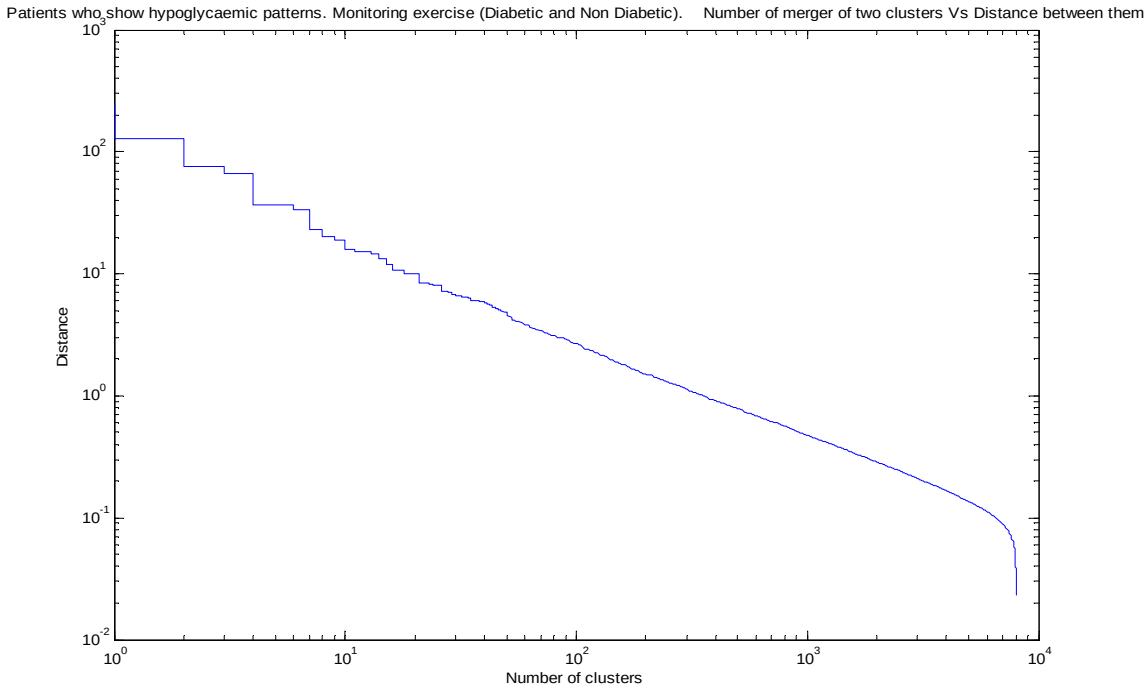

Figure 59: Number of clusters Vs distance between them using logarithmic scale in both axis.

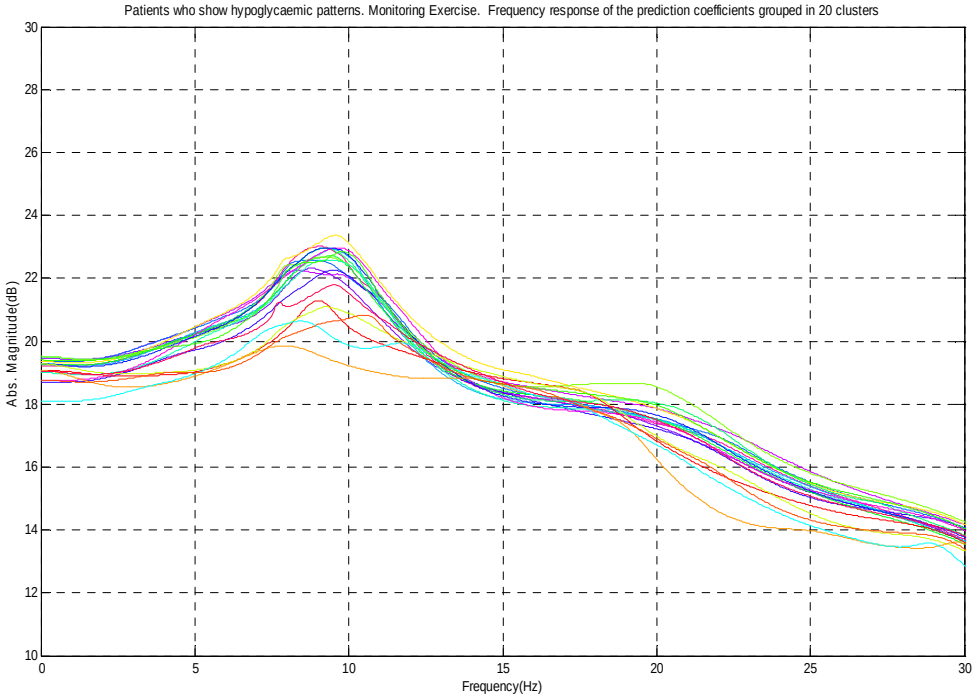

Figure 60: Frequency response of the prediction coefficients for the monitoring exercise grouped in 20 clusters, for patients who show some hypoglycaemic patterns. Colours correspond with the colours used in the other figures of this program.

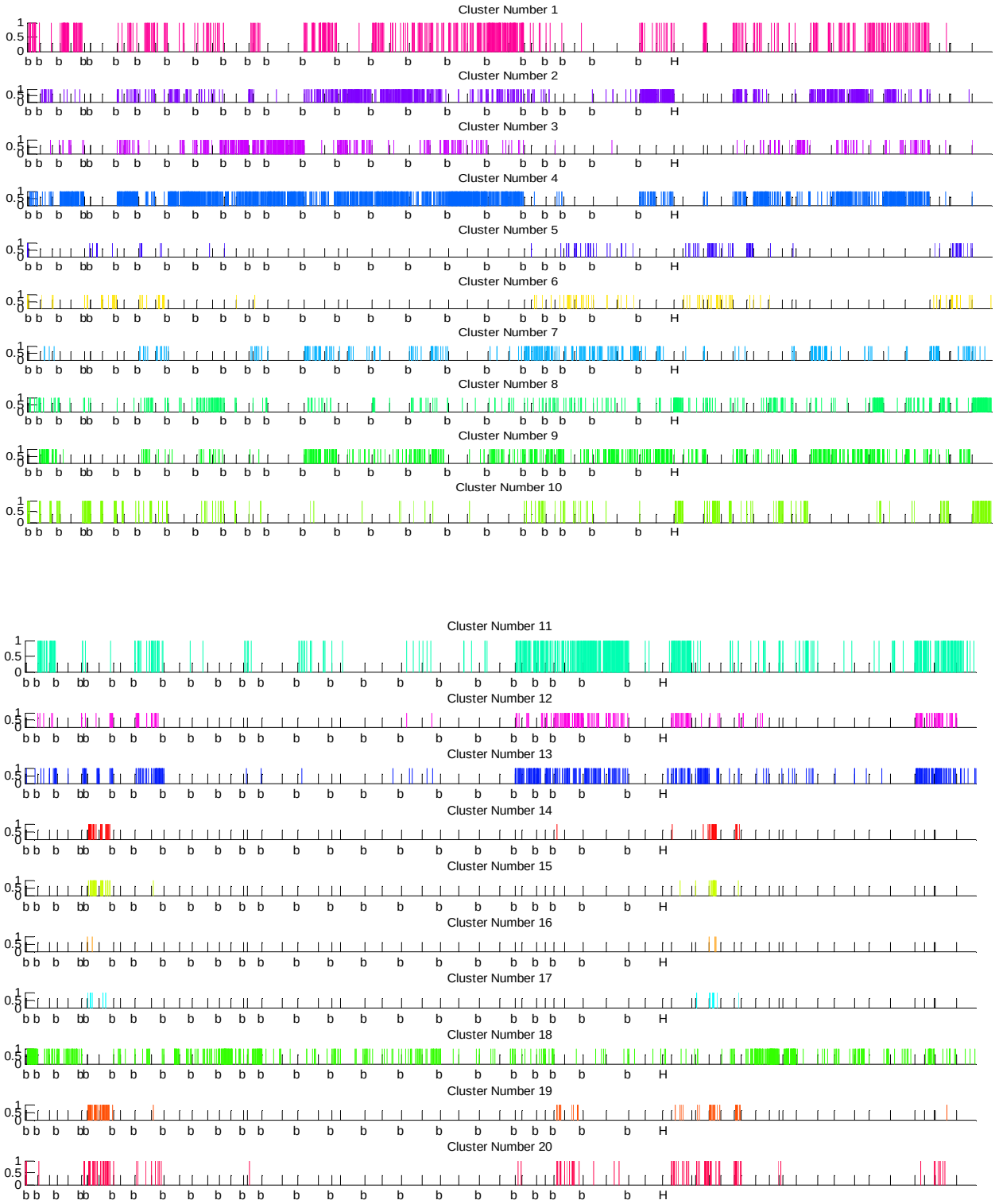

Figure 61: Data from patients who show some hypoglycaemic patterns grouped in 20 clusters, for the monitoring exercise.

5.2.2 Program 2: CalCAP exercise

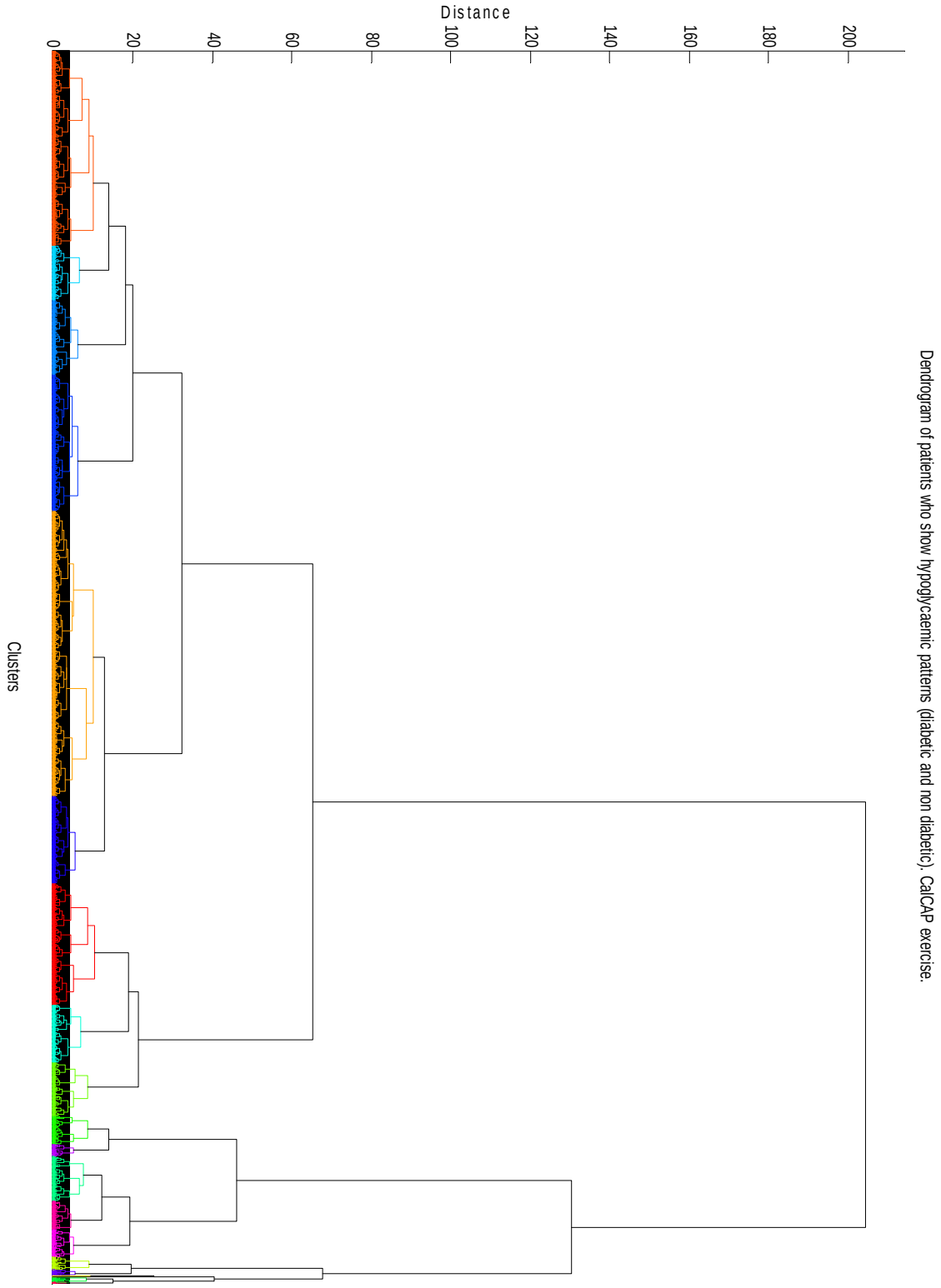

Figure 62: Dendrogram of patients who show some hypoglycaemic patterns for the CalCAP exercise.

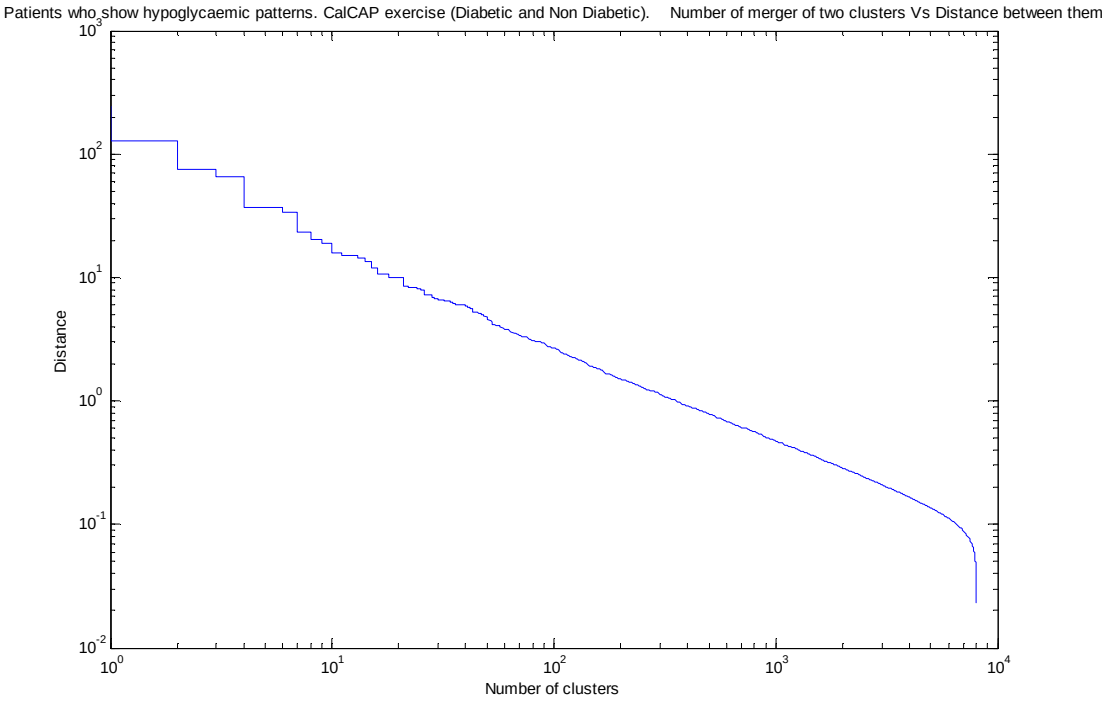

Figure 63: Number of clusters Vs Distance between them using logarithmic scale in both axis.

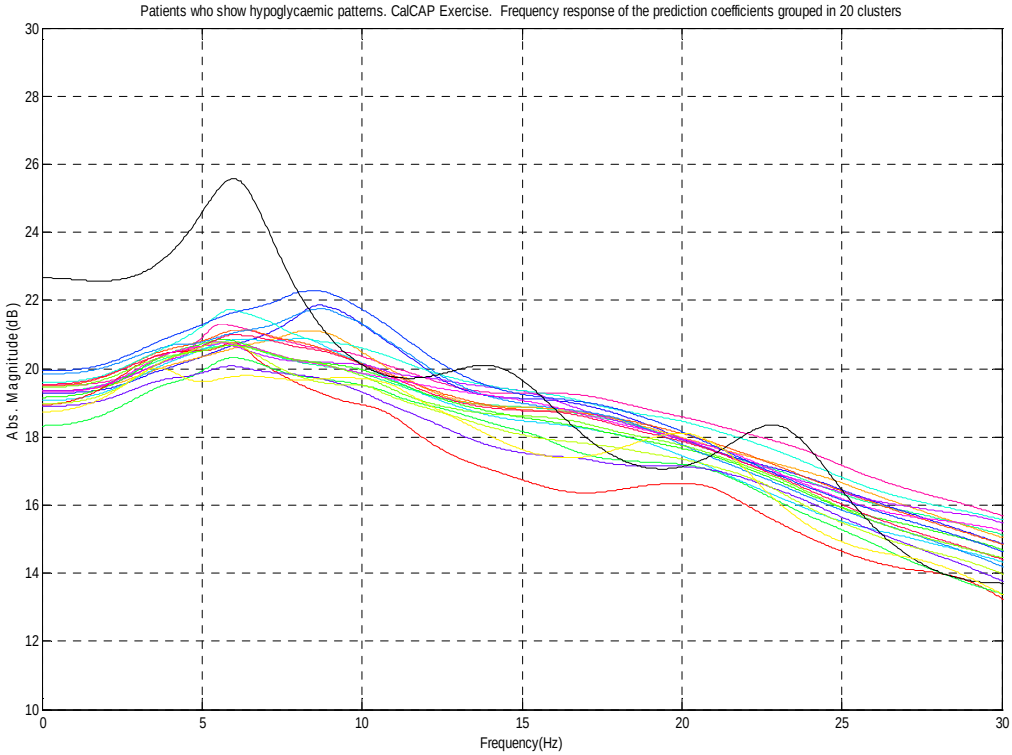

Figure 64: Frequency response of the prediction coefficients grouped in 20 clusters, for the CalCAP exercise.

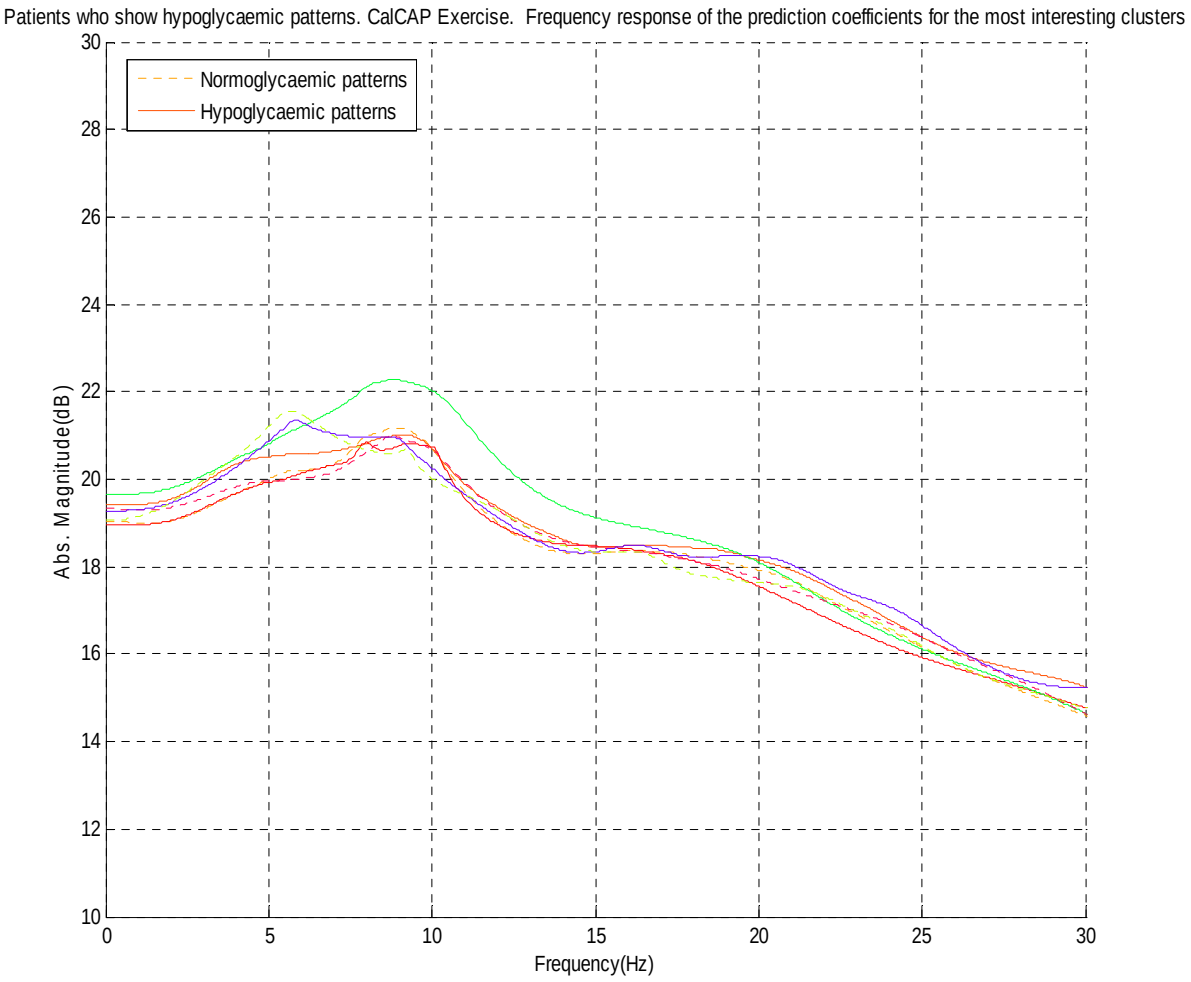

Figure 65: Frequency response of the prediction coefficients of the most interesting clusters for the CalCAP exercise. Solid lines represent hypoglycaemic patterns and dashed lines represent patterns that take place during every sequence (normo- and hypoglycaemic).

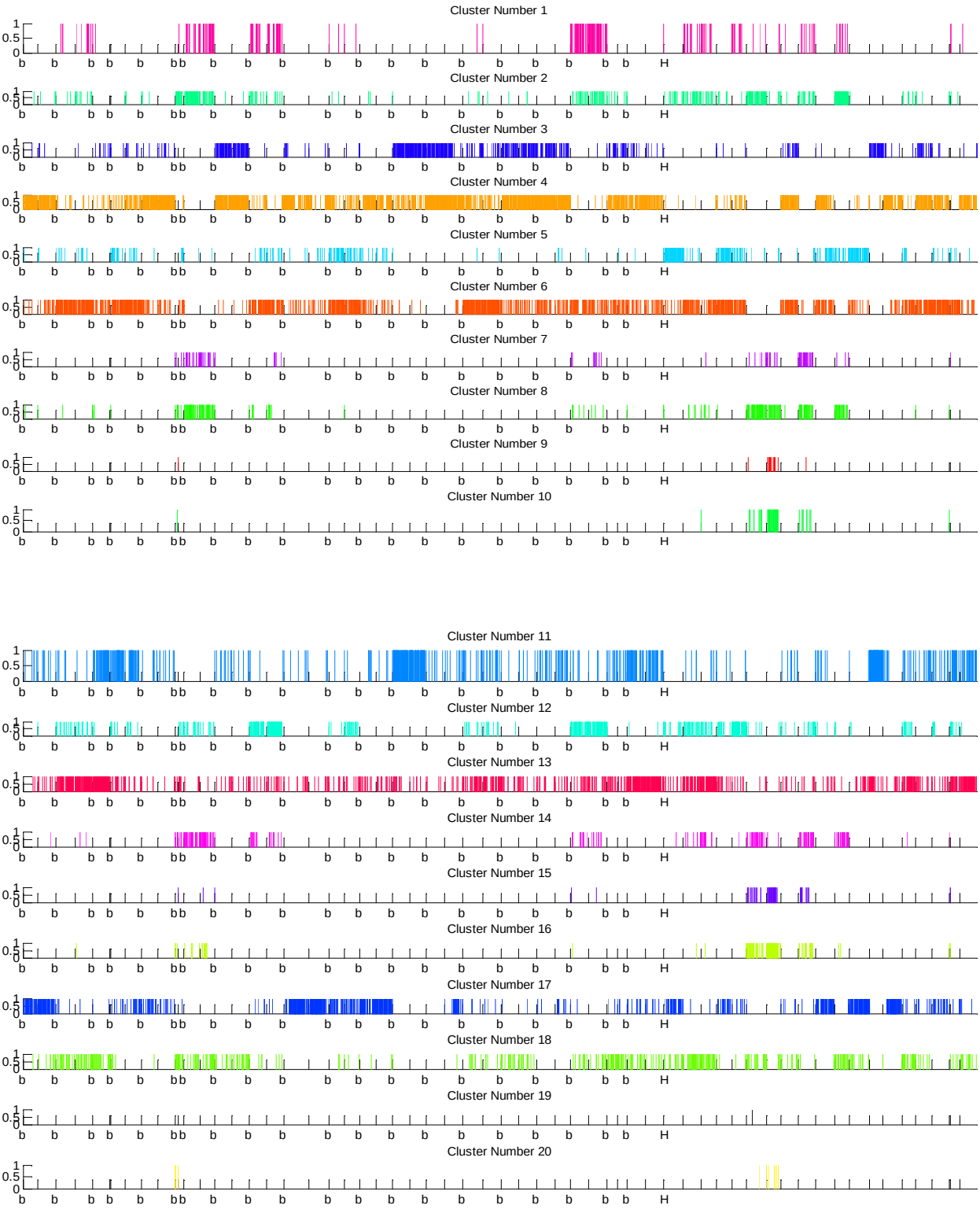

Figure 66: Data from patients who show some hypoglycaemic patterns classified them in 20 clusters.

5.2.3 Program 3: Monitoring & CalCAP exercises together

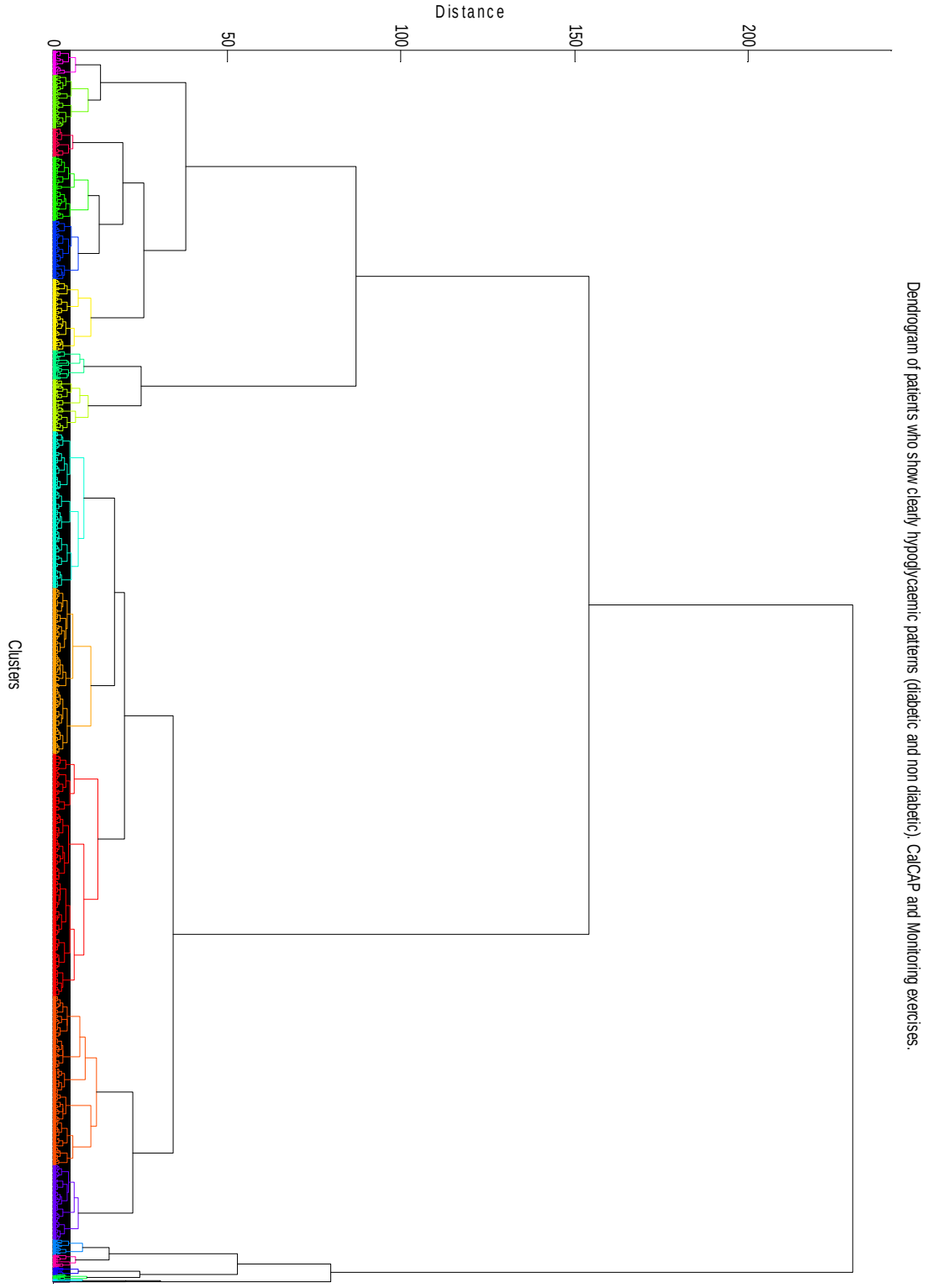

Figure 67: Dendogram for patients who show clearly hypoglycaemic patterns for CalCAP and monitoring exercises together.

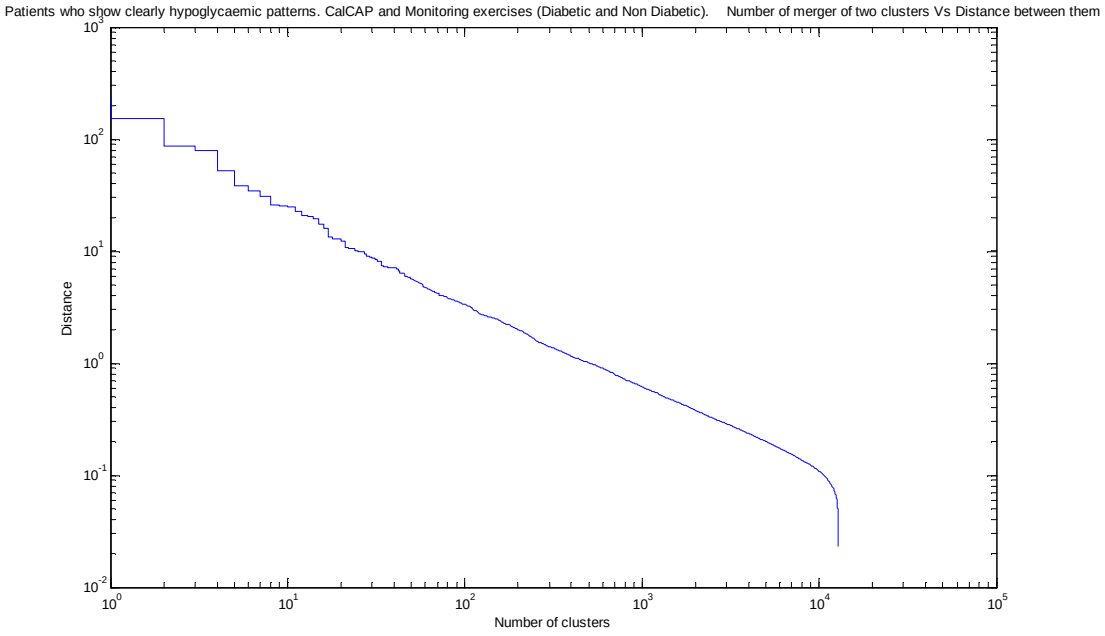

Figure 68: Clusters VS distance between them, using logarithmic scale in both axis.

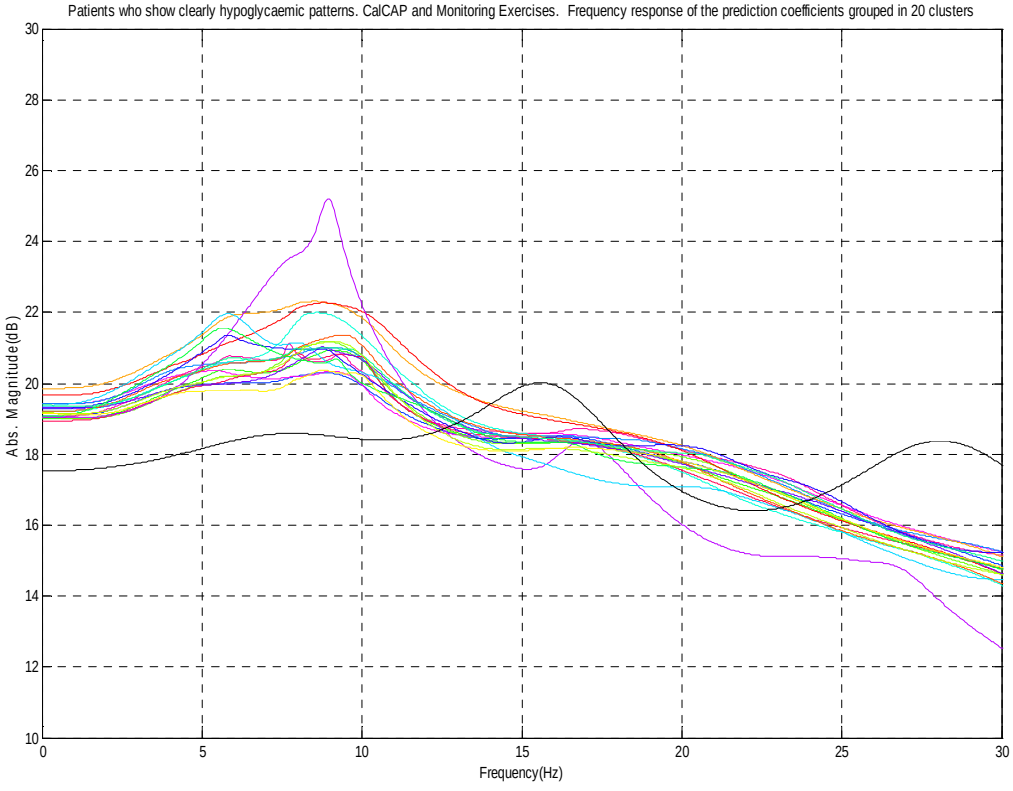

Figure 69: Frequency response of the prediction coefficients from patients' data that show clearly hypoglycaemic patterns. Data has been grouped in 20 clusters.

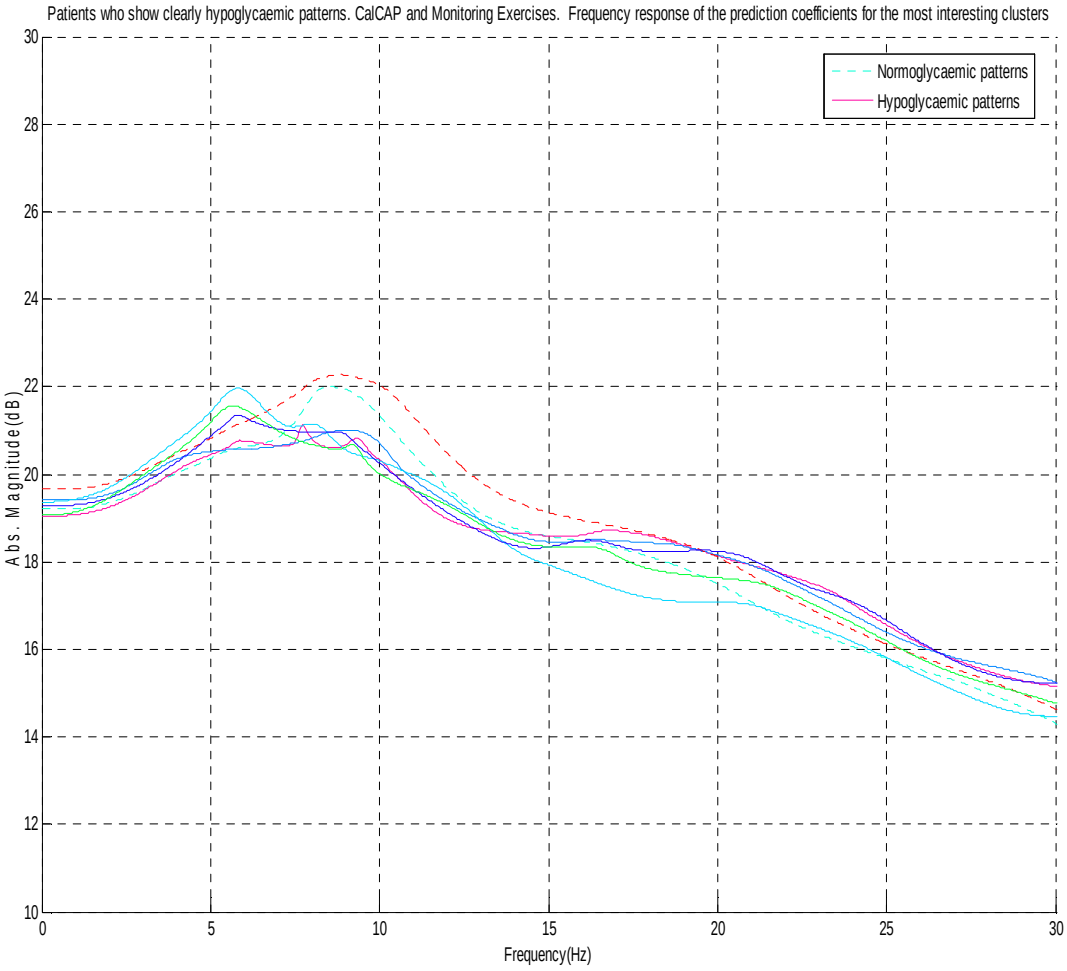

Figure 70: Frequency response of the prediction coefficients for the most interesting clusters. Dashed lines represents patterns that occur during all the sequences, and solid lines represent hypoglycaemic patterns.

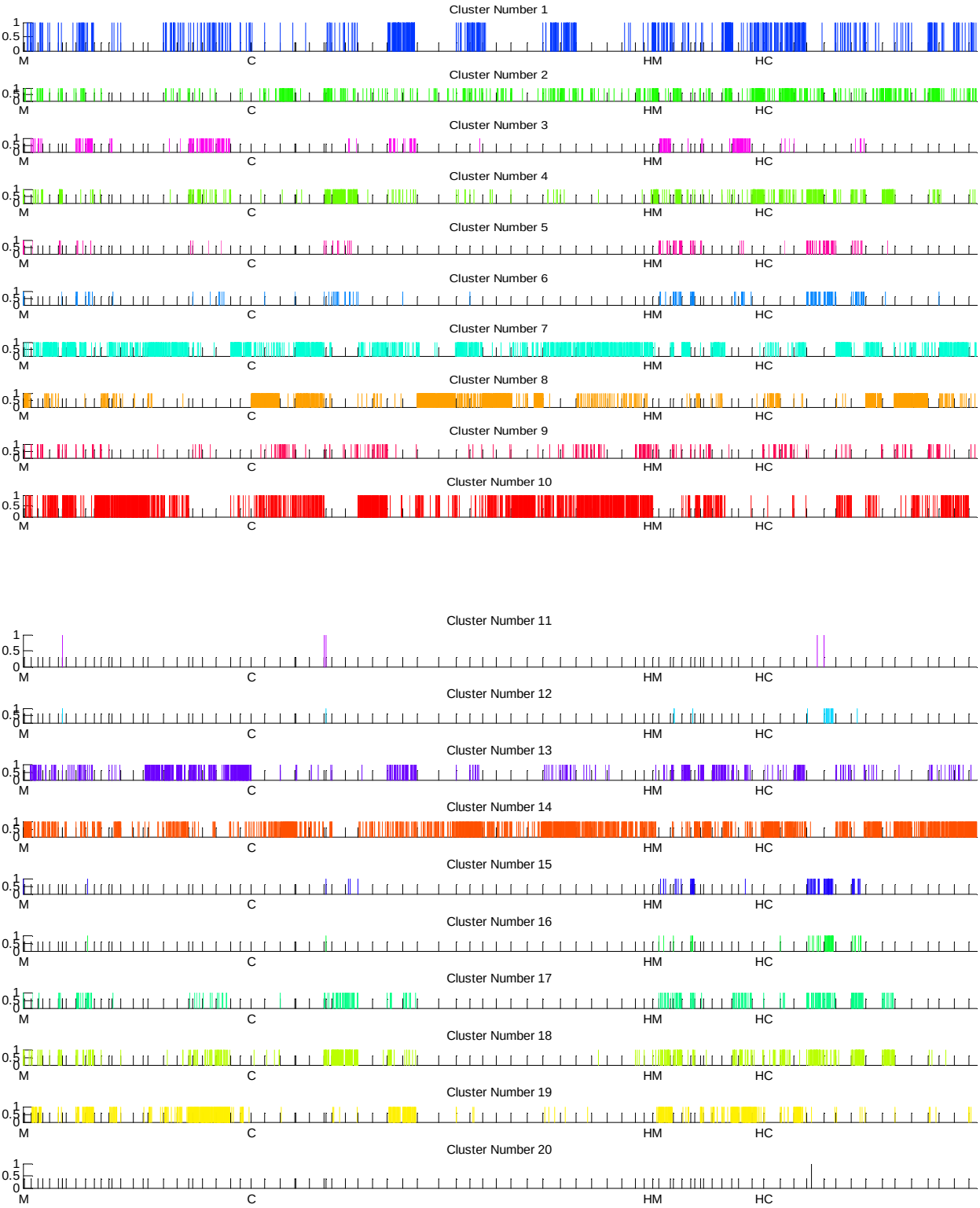

Figure 71: Data from patients who show clear hypoglycaemic patterns, grouped in 20 clusters.

5.2.4 Results for groups of patients

5.2.4.1. Results of program 1: Monitoring

- The results from this program mainly capture the interpatient variability mentioned in the first chapters. It can be appreciated that the frequency components of some patients during hypoglycaemia, corresponds to other patients' frequency components during normoglycaemia. Meaning that the interindividual differences in resting activity (alpha rhythm) differ more from person to person than from normo- to hypoglycaemia for the same person.
- However, it is also showed very slightly, some common patterns that occur more often during hypoglycaemia (See cluster 5 and cluster 6 from Figure 60).
- The Prediction Coefficients (PCs) frequency response figure (Figure 59) mostly reflects the frequency of the resting alpha rhythm (8-12Hz).

5.2.4.2. Results of program 2: CalCAP

- As in the previous case, interpatient variability is strongly showed.
- A few of common hypoglycaemic patterns are also presented in clusters 9, 10, 15 and 16 from Figure 65.
- The frequency response of PCs for clusters 9, 10, 15 and 16 (hypoglycaemic patterns) and clusters 4, 6 and 13 (patterns that occur in every sequence and every patient) has been represented. (See Figure 64)
  The dashed lines corresponds to clusters 4, 6 and 13 and the solid lines correspond to the hypoglycaemic clusters mentioned above.

  It can not be appreciated huge differences between normoglycaemic and hypoglycaemic clusters, but it seems it starts to appear lower high frequency (>7Hz) and higher content of 4 to 7 Hz activity, as it will be shown in a clearer way in the next program.

5.2.4.3. Results of program 3: Monitoring&CalCAP exercises together

- Besides the interpatient variability found in the two previous programs, common patterns can be more strongly appreciated here. All the clusters corresponding to the third main branch of the dendrogram (if the tree was cut keeping 3 different clusters) show frequency components that occur more often during hypoglycaemia sequences (see Figure 66). These clusters are (from left to right in the dendrogram): cluster number 6, 5, 15, 16, 12, 11, 20 (see Figure 70).
- It has been plotted the PC frequency response of the clusters: 5, 6, 12, 15 and 16 representing hypoglycaemic patterns, together with the clusters 7 and 10 that correspond to clusters that occur during all sequences (normoglycaemia and hypoglycaemia). (See Figure 69)
  The dashed lines corresponds to clusters 10 and 7 and the solid lines represent hypoglycaemic patterns.

It seems to be less high frequency (>7Hz) in the clear hypoglycaemic patterns and a correspondent higher content of 4 to 7 Hz activity.

Important for all the programs: Even when these patterns stand out in the common clustering, it is important to emphasize that they do not seem to appear for more than a few patients.

6. Probabilistic classifier

6.1 Introduction

In the field of machine learning, the goal of classification is to group items that have similar feature values, into groups by taking an input vector x and assigning it to one of *K* discrete classes C_k where *k=1, …, K*. The input space is thereby divided into decision regions whose boundaries are called decision boundaries or decision surfaces [16].

In this research, it has been considered a linear model for classification, therefore the decision surfaces are linear functions of the input vector, and hence are defined by 11 dimensional hyperplanes within the 12-dimensinoal input space, due to the fact that each set of features contain 12 Autocorrelation coefficients.

Bayes decision theory is a fundamental statistical approach to the problem of pattern classification. This approach is based on the assumption that the decision problem is posed in probabilistic terms, and that all of the relevant probability values are known [19].

The classification was based on: the feature vector x, extracted for each 2 seconds segment; the number of classes; the a priori probability and a statistical description for each class. The a posterior probability $p(C_k|\mathrm{x})$ for the object belonging to the class C_k given the feature vector x could be calculated from Bayes' theorem [2,19]:

$$P(C_k|\bar{x}) = \frac{p(\bar{x}|C_k)P(C_k)}{p(\bar{x})}$$

where

$$p(\bar{x}) = \sum_{k=1}^{K} p(\bar{x}|C_k)P(C_k)$$

Calculating $p(C_k|\mathrm{x})$ for all classes *k=1, …, K* and choosing the class with the largest probability will give a minimum probability for mis-classifications.

It has been assumed that the probability density function $p(\mathrm{x}|C_k)$ for each class was a multi-dimensional normal distribution:

$$p(\bar{x}|C_k) = \frac{1}{(2\pi)^{d/2}|\bar{\Sigma}_k|^{1/2}} \exp\left[-\tfrac{1}{2}(\bar{x}-\bar{\mu}_j)^t \bar{\Sigma}_k^{-1}(\bar{x}-\bar{\mu}_j)\right]$$

Thus, each class centroid was determined by the mean value vector $\bar{\mu}$ and the size and shape by the covariance matrix $\bar{\Sigma}$.

Minimum rate classification can be achieved by use of a set of discriminant functions $g_k(\bar{x}), k = 1,...,K$. The classifier is said to assign a feature vector x to class C_k if

$$g_i(\bar{x}) > g_j(\bar{x}) \qquad \text{for all } j \neq i.$$

For the minimum-error-rate case, we can simplify things further by taking $g_i(\bar{x}) = P(C_i \mid \bar{x})$, so that the maximum discriminant function corresponds to the maximum a posteriori probability.

A simple case arises when the covariance matrices for all of the classes are chosen to be identical. Geometrically, this corresponds to the situation in which the samples fall in hyperellipsoidal clusters of equal size and shape. In this case, the classification problem leads to this simple set of linear discriminant functions, reduced from the basic expression for the a posteriori probability and using expansion of the quadratic form of the Mahalanobis distance as measure:

$$g_i(\bar{x}) = \bar{W}_{1i}{}^t \cdot \bar{x} + \bar{W}_{0i}$$

$$\textit{where} \qquad \bar{W}_{1i} = \bar{\Sigma}^{-1} \cdot \bar{\mu}_i$$

$$\textit{and} \qquad \bar{W}_{0i} = -\tfrac{1}{2} \cdot \bar{\mu}_i{}^t \cdot \bar{\Sigma}^{-1} \cdot \bar{\mu}_i$$

Along with the assumption of a common covariance matrix $\bar{\Sigma}$ for all classes, the a priori probability during this study was equally distributed. The statistical parameters ($\bar{\mu}_i$ and $\bar{\Sigma}_i$) for each class are the maximum likelihood estimates, and the common covariance matrix $\bar{\Sigma}$ was calculated as the pooled covariance matrix [2, 19].

6.2 Methodology

In order to develop a general classifier for our group of patients, patients who showed hypoglycaemic patterns during the most complete analysis (program 3: Monitoring and CalCAP exercises together) were identified, see Figure 70, page 68. The diabetic patient number 6 and the diabetic patient number 9 showed hypoglycaemic patterns during both exercises. The diabetic patient number 7 did it during CalCAP and the non diabetic patient number 18 did it during monitoring. The diabetic patient number 3 showed more slightly hypoglycaemic patterns during monitoring.

From all of them, patient 6 was selected to develop the classifier due to his/her strong hypoglycaemic patterns presented during the most general program.

After grouping patient six's data in 10 clusters (also called classes) (for monitoring and CalCAP together), as it was done for every individual patients, a set of four "super classes" has been defined.

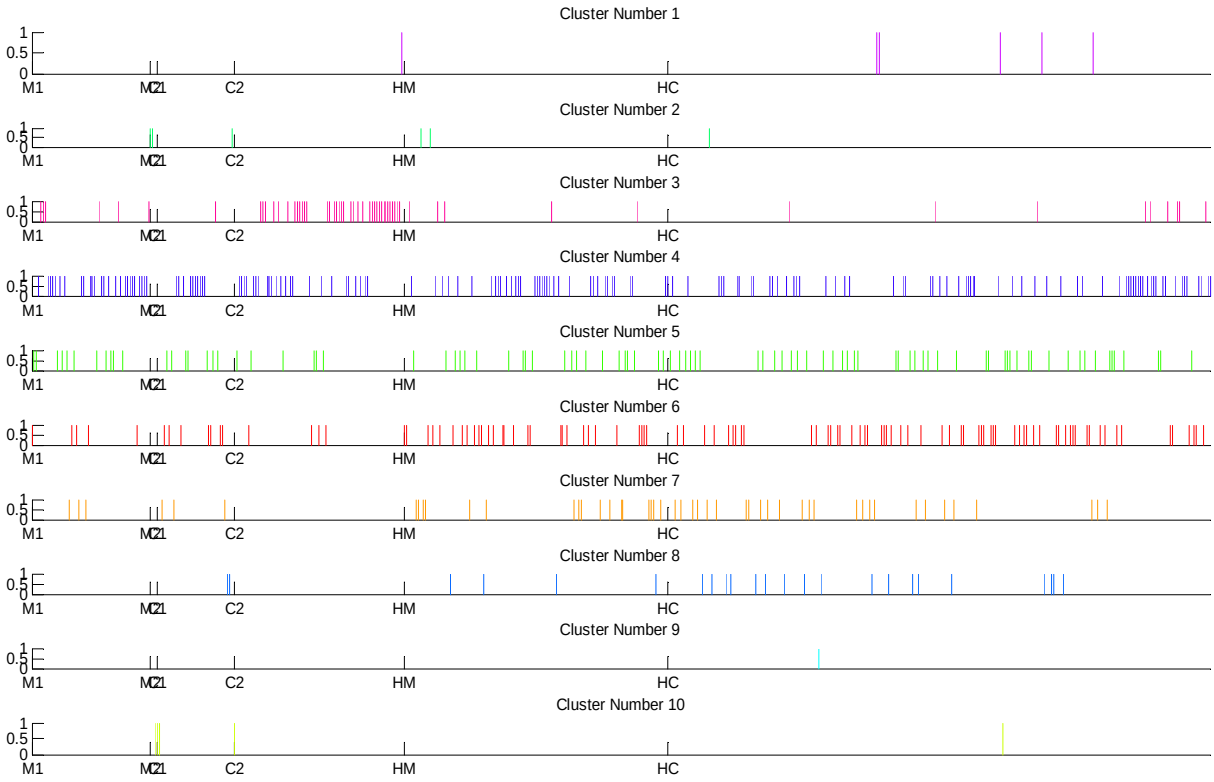

Figure 72: Data from diabetic patient number 6 grouped in 10 clusters.

The four classes were created as follows.

1. "Normoglycaemia": Represents patterns that occur more often during baselines 1 and baselines 2. Colour blue was assigned to this class.
2. "Everywhere": Represents patterns occurring during all sequences. Colour green was assigned to this class.
3. "Hypoglycaemia": Represents patterns that occur more often during the hypoglycaemic sequences. Colour red was assigned to this class.
4. "Other": Represents artefacts not removed, or clusters containing a few patterns. Colour black was assigned to this class.

These classes consist of:

1. "Normoglycaemia": Consist of patterns from cluster number 3.
2. "Everywhere": Consist of clusters number 4, 5 and 6.
3. "Hypoglycaemia": Consist of clusters number 7 and 8.
4. "Other": Consist of clusters number 1, 2, 9 and 10.

At this point, having the a posteriori probability (probability for a class belonging some feature vector) of diabetic patient number 6, it can be calculated the a priori probability for the other persons (probability of belong to one class given the feature vector). For this aim,

the mean vector $\bar{\mu}_i$ of each class has been calculated, as well as the pooled covariance matrix $\bar{\bar{\Sigma}}$.

To calculate the pooled covariance matrix, first the covariance matrices for each class have been calculated. Secondly, multiplying each matrix by the number of elements of each class and adding all matrices, another matrix is obtained. Dividing the resulting matrix by the total number of elements, a pooled covariance is achieved. It is a weighted average of the individual covariance matrices.

Using the mean vector and the pooled covariance matrix, the weights of each class were calculated using the formulation of the previous section. With these weights the linear transformation of the feature vector is calculated for each class, obtaining the discriminant function for each class. For each input vector, the ten different discriminant functions are calculated, and the input vector is assigned to the class with highest discriminant function value, which was later classified in four super classes according to different blood glucose states.

With this method, the data from all the patients was classified in four classes representing four blood glucose level states.

As it was expected, the classification for the diabetic patient number 6 represents each state of the patient pretty clearly.

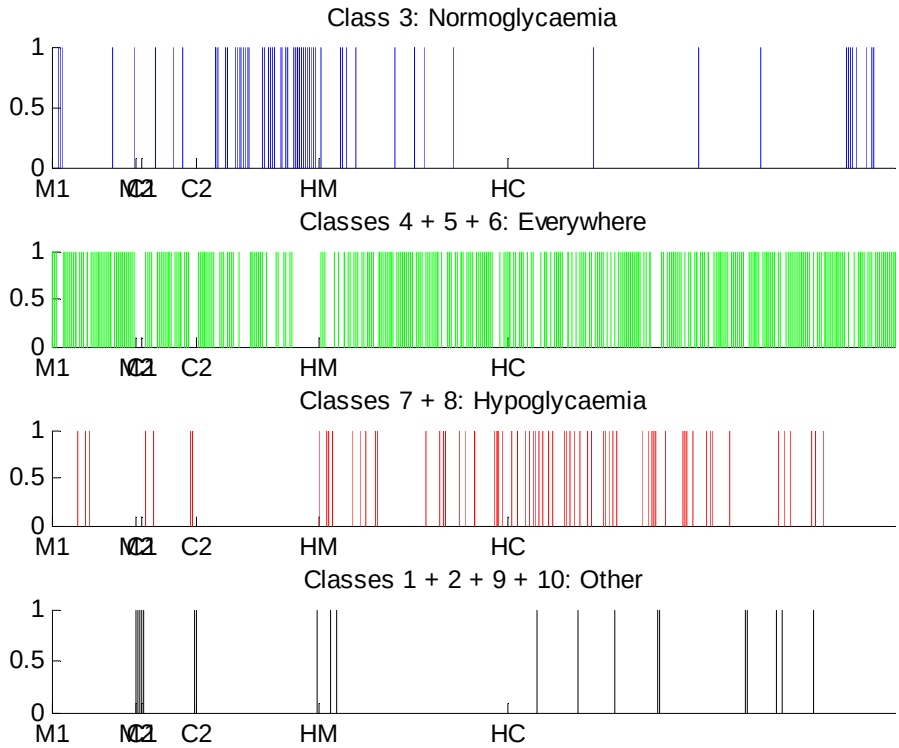

Figure 73: Data from diabetic patient number 6 classified in 4 "superclasses". This patient is the one used to develop the classifier.

6.3 Results

After classify every patient and visualize every figure obtained for each one, some conclusions can be extracted. Obviously the classifier works for the diabetic patient number 6 (the one used to create it) and it works somehow for the diabetic patients 3 and 9. And maybe it works slightly for diabetic patient number 7 and non diabetic patients 17 and 18.

The patients that somehow have a good response to the classifier correspond to the patients that showed hypoglycaemic patterns in the program 3, where the clustering for every person for the 2 exercises together (monitoring and CalCAP) was performed. The patients that showed hypoglycaemic patterns in the program 3 were diabetic patients: 3, 6, 7 and 9. And non diabetic patient: 18. It can be said that the classifier works for them (for ones better than for others). These are the corresponding figures obtained after classifying their data for these patients (Figures for all patients can be visualized in the additional DVD).

Diabetic patient number 3:

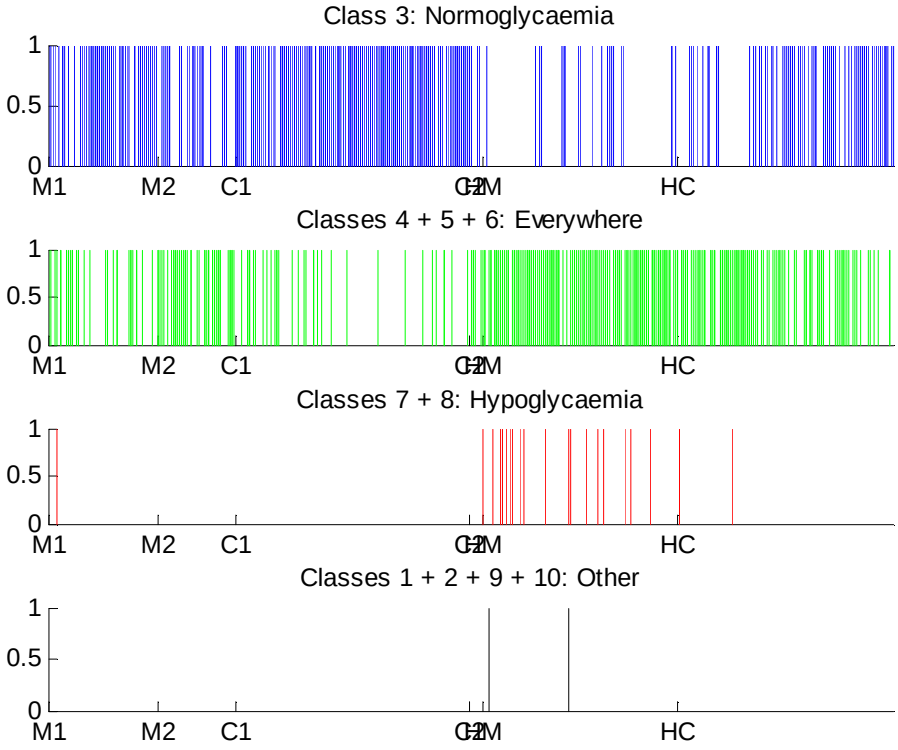

Figure 74: Data from diabetic patient number 3 after classification in 4 classes.

Diabetic patient number 7:

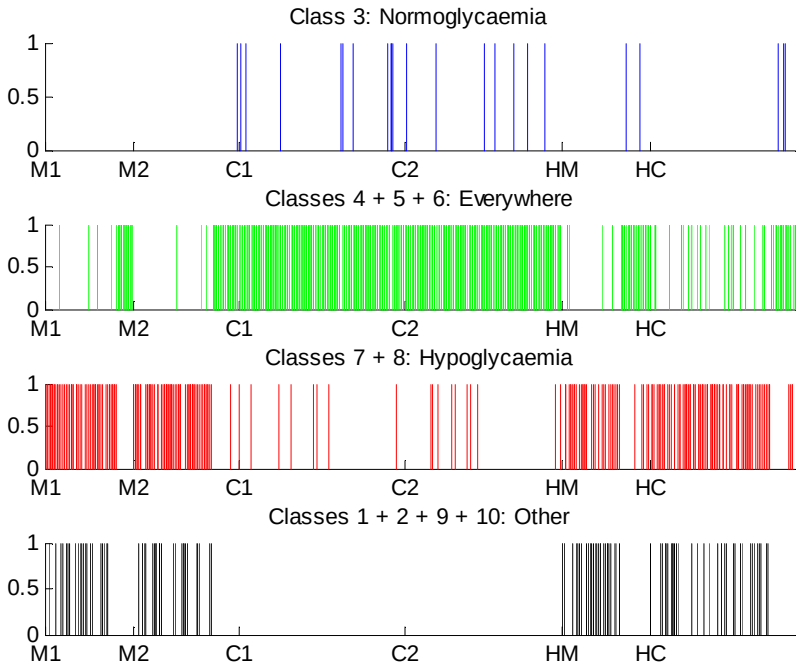

Figure 75: Data from diabetic patient number 7 after classification in 4 classes.

Diabetic patient number 9:

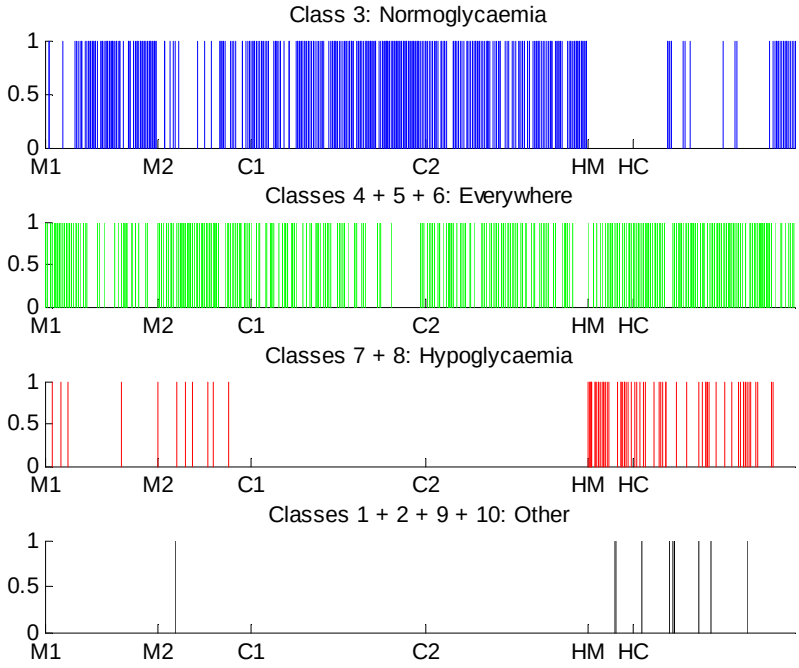

Figure 76: Data from diabetic patient number 9 after classification in 4 classes.

Non diabetic patient number 18:

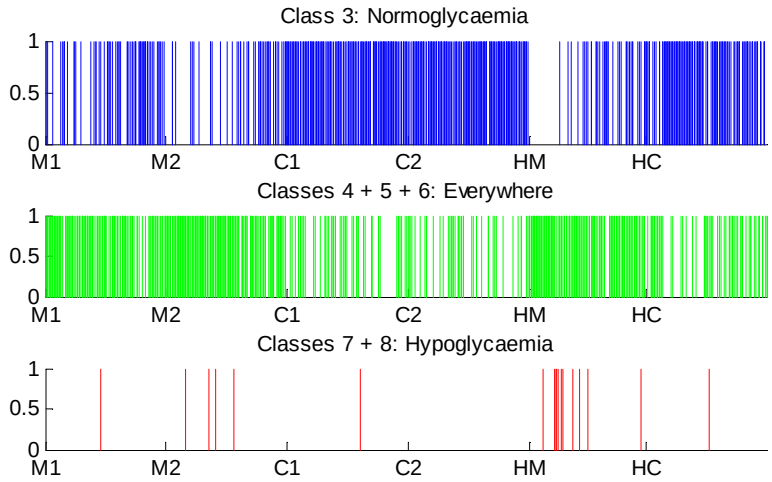

Figure 77: Data from non diabetic patient number 18 after classification in 4 classes (there is no data belonging to the fourth class).

The main conclusion extracted is that interpatient variability is more important that the differences between normo and hypoglycaemia. It can be easily seen that patterns representing a normoglycaemic state for one patient represent hypoglycaemic patterns for other patients, as it occurs for the diabetic patient number 5:

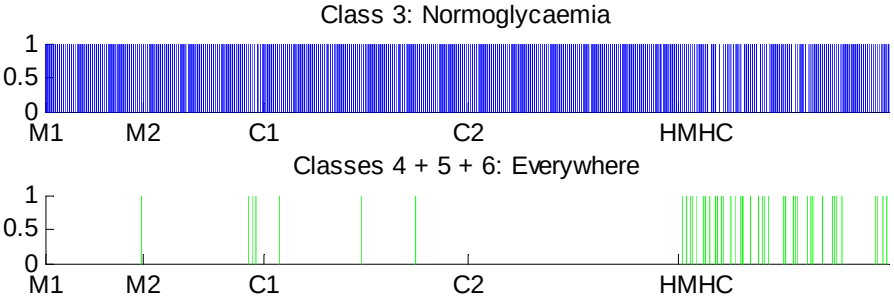

Figure 78: Data from diabetic patient number 5 after classification in 4 classes. Hypoglycaemic patterns for this patient correspond to normoglycaemic patterns for diabetic patient number 6 (used to develop the classifier).

For this reason, a classifier can be developed for a small group of patients with similar characteristics, but a separate study of each patient is needed. For each patient, typical hypoglycaemic patterns can be found and thereby a method to find out when hypoglycaemia takes place for each individual patient can be developed, as well as for a small group of patients if after study it similar hypoglycaemic characteristics are found out.

7. Conclusions and future work

7.1 Conclusions

For most of the persons, individual characteristic patterns occurring principally during the hypoglycaemic sequence were found out. For these patients, it has been observed higher content of 8 to 12 Hz activity (alpha rhythm) during hypoglycaemia and the highest frequency activity seems to occur from 6 to 10-12 Hz, instead of from 8 to 12 as it occurs during normoglycaemia.

Diabetic patients develop more hypoglycaemic patterns than non diabetic patients. The reason may be their many several past hypoglycaemic events. When performing a demanding mental task during hypoglycaemia, diabetic patients seem to be more affected, and the differences between diabetic and non diabetic patients are more marked.

The main and most important result obtained is the interpatient variability. The frequency patterns of some patients during hypoglycaemia correspond to other patients' patterns during normoglycaemia. This means that the interindividual differences differ more from person to person, than from normo- to hypoglycaemia for the same person.
Even when hypoglycaemic patterns stand out in the common clustering, it is important to emphasize that they do not seem to appear for more than a few patients.

When a group of patients is studied together, it seems to be less high frequency (>7Hz) in the clear hypoglycaemic patterns and a correspondent higher content of 4 to 7 Hz activity.

Due to the strong interpatient variability found out during all this research, a classifier can be developed for a small group of patients with similar characteristics, but a separate study of each patient is needed. For each patient, typical hypoglycaemic patterns can be found and thereby a method to find out when hypoglycaemia takes place for each individual patient can be developed, as well as for a small group of patients with similar hypoglycaemic characteristics.

These results and conclusions can be classified them according to the goals proposed at the beginning of this research.

Table 15: Results classified according to the goals proposed at the beginning of this research.

Set of goals	Set of results
Find out the differences between normoglycaemia and hypoglycaemia state for each patient during monitoring exercise.	For most of the persons, individual characteristic patterns occurring principally during the hypoglycaemic sequence were found out. For these patients, it has been observed higher content of 8 to 12 Hz activity (alpha rhythm) during hypoglycaemia and the highest frequency activity seems to occur from 6 to 10-12 Hz, instead of from 8 to 12 as it occurs during normoglycaemia. For a group of patients: The main and most important result obtained is the interpatient variability. The frequency patterns of some patients during hypoglycaemia correspond to other patients' patterns during normoglycaemia. This means that the interindividual differences differ more from person to person, than from normo- to hypoglycaemia for the same person.
If the differences exist, analyze how hypoglycaemia affects to brain functioning for each different patient and extract conclusions and results.	
Find out how hypoglycaemia affects to brain functioning for the CalCAP exercise for each patient.	The results are the same than for the monitoring exercise, but it can be added that diabetic patients develop more hypoglycaemic patterns than non diabetic patients. The reason may be their many several past hypoglycaemic events.
Find out if there is some relation between hypoglycaemic patterns for different exercises (CalCAP and monitoring).	Individually: When performing a demanding mental task during hypoglycaemia (CalCAP), diabetic seems to be more affected, and the differences between diabetic and non diabetic patients are more marked. For a group of patients it can be said that hypoglycaemic patterns stand out in the common clustering, but it is important to emphasize that they do not seem to appear for more than a few patients. For these patients it seems to be less high frequency (>7Hz) in the clear hypoglycaemic patterns and a correspondent higher content of 4 to 7 Hz activity.
Analyze if our conclusions are valid for a group of patients and if it is possible to do some generalization.	Due to the strong interpatient variability found out during all this research, a classifier can be developed for a small group of patients with similar characteristics, but a separate study of each patient is needed. For each patient, typical hypoglycaemic patterns can be found and thereby a method to find out when hypoglycaemia takes place for each individual patient can be developed, as well as for a small group of patients with similar hypoglycaemic characteristics.

7.2 Future work

In the artefact removal part, the optimum threshold to decide which data is valid and which one it is not, should be different for each person to achieve a better generalization. For this aim, it is purpose to build a filter using an adaptive threshold for each patient, depending on the energy of each person at each moment.

Referring to the whole data processing and analysis in this research, the same analysis can be carried out for the other 2 exercises that have not been studied here, and it can be checked that the AEP exercise is expected to show similar results to the monitoring exercise's results and the AQT exercise is expected to behave like CalCAP.

The same study can be done for different channels, to analyze how hypoglycaemia is presented in different parts of the brain.

From the raw data, it can be extracted intermediate sequences (sequences when blood glucose level decreases and during recovery), because there are many intermediate states that have not been studied here, and they can form new classes represented by other colours in the classifier.

In this research, patients' state of the brain has been study while they are sitting and performing specific tasks. A next close step to this research would be to run the developed classifier in the entire raw data, which contains more artefacts and movements, and analyze its performance for a more similar situation to real life.

An important future work must be to analyze the differences existing between the blood glucose level and how the brain is affected for each patient. This analysis can be seen from different angles: it can be studied how brain is affected for the same blood glucose level for each patient (different patients might have different brain behaviour to the same blood glucose level), or it can be studied, for patients who show similar brain patterns, which blood glucose level have at that moment. Since each patient's brain is affected in a different way, the goal would be to try to find these boundaries between patients' brain behaviour and blood glucose levels for each patient. And then, try to do some generalization for patients with similar characteristics.

7B. Conclusiones y trabajo futuro

7.1 Conclusiones

Para la mayoría de las personas, han sido encontrados patrones individuales característicos que ocurren principalmente durante la secuencia hipoglucémica. Para estos pacientes, se ha observado un contenido más alto de actividad en el rango de 8 a 12 Hz (ritmo alfa) durante hipoglucemia y la actividad frecuencial más alta parece ocurrir de 6 a 10-12 Hz, en lugar de 8 a 12 como ocurre durante normoglucemia.

Los pacientes diabéticos desarrollan más patrones hipoglucémicos que las personas no diabéticas. La razón podrían ser los diversos eventos hipoglucémicos que han sufrido en el pasado. Cuando se encuentran desarrollando una tarea que demanda concentración mental durante hipoglucemia, los pacientes diabéticos parecen estar más afectados, y las diferencias entre diabéticos y no diabéticos son más marcadas.

El resultado principal y más importante obtenido es la variabilidad entre pacientes. Los patrones frecuenciales de algunos pacientes durante hipoglucemia corresponden a patrones de otros pacientes durante normoglucemia. Esto significa que hay más diferencias en los patrones frecuenciales de una persona a otra, que de normo a hipoglucemia para la misma persona. Incluso cuando en el *clustering* más general encontramos patrones hipoglucémicos, es importante resaltar que sólo aparecen para unos pocos pacientes.

Cuando un grupo de pacientes es estudiado junto, parece haber menos altas frecuencias (>7Hz) en las funciones que representan claros patrones de la hipoglucemia y un contenido más alto de actividad en el rango de 4 a 7 Hz.

Debido a la gran variabilidad entre pacientes encontrada durante toda esta investigación, un clasificador puede ser construido para un pequeño grupo de pacientes con similares características, pero es necesario realizar un estudio separado de cada paciente. Para cada paciente, se pueden encontrar patrones hipoglucémicos típicos y por lo tanto, se puede desarrollar un método para averiguar cuándo tiene lugar la hipoglucemia para cada paciente, al igual que para un pequeño grupo de pacientes con características hipoglucémicas similares.

Estos resultados y conclusiones pueden ser clasificados de acuerdo a los objetivos propuestos al principio de este estudio.

Tabla 16: Resultados clasificados de acuerdo a los objetivos propuestos al inicio de este estudio.

Conjunto de objetivos	Conjunto de resultados
Descubrir las diferencias entre los estados de normoglucemia e hipoglucemia para cada paciente durante el ejercicio de monitorización.	Para la mayoría de las personas, han sido encontrados patrones individuales característicos que ocurren principalmente durante la secuencia hipoglucémica. Para estos pacientes, se ha observado un contenido más alto de actividad en el rango de 8 a 12 Hz (ritmo alfa) durante hipoglucemia y la actividad frecuencial más alta parece ocurrir de 6 a 10-12 Hz, en lugar de 8 a 12 como ocurre durante normoglucemia.
Si estas diferencias existen, analizar cómo afecta al funcionamiento del cerebro la hipoglucemia para cada paciente y extraer resultados y conclusiones.	Para un grupo de pacientes: El resultado principal y más importante obtenido es la variabilidad entre pacientes. Los patrones frecuenciales de algunos pacientes durante hipoglucemia corresponden a patrones de otros pacientes durante normoglucemia. Esto significa que hay más diferencias en los patrones frecuenciales de una persona a otra, que de normo a hipoglucemia para la misma persona.
Descubrir cómo afecta al funcionamiento del cerebro la hipoglucemia para el ejercicio CalCAP para cada paciente.	Los resultados son los mismos que para el ejercicio de monitorización, pero se puede añadir que los pacientes diabéticos desarrollan más patrones hipoglucémicos que las personas no diabéticas. La razón podrían ser los diversos eventos hipoglucémicos que han sufrido en el pasado.
Averiguar si hay alguna relación entre patrones hipoglucémicos de diferentes ejercicios (CalCAP y monitorización).	Individualmente: Cuando se encuentran desarrollando una tarea que demanda concentración mental durante hipoglucemia, los pacientes diabéticos parecen estar más afectadas, y las diferencias entre diabéticos y no diabéticos son más marcadas. Para un grupo de pacientes, se puede decir que encontramos patrones hipoglucémicos en el *clustering* más general, pero es importante resaltar que sólo aparecen para unos pocos pacientes. Para estos pacientes, parece haber menos altas frecuencias (>7Hz) en las funciones que representan claros patrones de la hipoglucemia y un contenido más alto de actividad en el rango de 4 a 7 Hz.
Analizar si nuestras conclusiones son válidas para un grupo de pacientes y si es posible hacer alguna generalización.	Debido a la gran variabilidad entre pacientes encontrada durante toda esta investigación, un clasificador puede ser construido para un pequeño grupo de pacientes con similares características, pero es necesario realizar un estudio separado de cada paciente. Para cada paciente, se pueden encontrar patrones típicos hipoglucémicos y por lo tanto, se puede desarrollar un método para averiguar cuándo tiene lugar la hipoglucemia para cada paciente, al igual que para un pequeño grupo de pacientes con características hipoglucémicas similares.

7.2 Trabajo futuro

En la parte de eliminación de artefactos, el umbral para decidir qué datos son válidos y cuales no, debería ser diferente para cada persona para conseguir una generalización mejor. Con este propósito, se propone construir un filtro usando un umbral adaptativo a cada paciente, dependiendo de la energía de cada persona en cada momento.

Con respecto a todo el análisis y procesamiento de datos entero de este estudio, se puede hacer el mismo análisis para los otros dos ejercicios que no han sido estudiados, y se puede comprobar que para el ejercicio AEP se espera encontrar resultados similares al ejercicio de monitorización, y para el ejercicio de AQT se espera encontrar un comportamiento similar al ejercicio CalCAP.

Se puede hacer el mismo estudio para diferentes canales, y analizar así cómo se presenta la hipoglucemia en diferentes partes del cerebro.

Se pueden extraer secuencias intermedias de los datos brutos (secuencias en las que el nivel de azúcar en la sangre decrece y durante la recuperación), porque dichos estados intermedios no han sido estudiados aquí, y pueden formar nuevas clases representadas con otros colores en el clasificador.

En esta investigación, el estado del cerebro de los pacientes ha sido estudiado mientras están sentados y realizando específicas tareas. Un paso próximo a esta investigación sería ejecutar el clasificador desarrollado para todos los datos brutos, que contienen más artefactos y movimientos, y analizar su comportamiento para una situación más parecida a la vida real.

Un trabajo futuro importante, debería ser analizar las diferencias existentes entre el nivel de azúcar en la sangre y cómo el cerebro es afectado para cada paciente. Este análisis puede ser visto desde diferentes ángulos: Se puede estudiar cómo el cerebro es afectado para el mismo nivel de azúcar en la sangre para cada paciente (distintos pacientes pueden tener comportamientos cerebrales diferentes para el mismo nivel de azúcar en la sangre), o puede ser estudiado, para pacientes que tienen características cerebrales similares, qué nivel de azúcar en la sangre poseen en ese momento. Como el cerebro de cada paciente es afectado de forma diferente, el objetivo sería tratar de encontrar esas fronteras entre el comportamiento del cerebro de los pacientes y su nivel de azúcar en la sangre individualmente. Y entonces, intentar hacer alguna generalización para pacientes con características similares.

8. References

[1] Peter Noesgaard Andreasen.
"Monitoring anaesthesia by EEG analysis with self-organizing feature map".
Alborg Universitetscenter. Institut for Elektroniske Systemer. Afdeling for medicinsk informatik. Juni 1992.

[2] Carsten E. Thomsen
"Hierarchical Cluster Analysis and Pattern Recognition Applied to the Electroencephalogram".
Aalborg University, Department of Medical Informatics & Image Analysis, Institute of Electronic Systems. 1992.

[3] "Methods of Information in Medicine". 1/94 Vol.33
Schattauer, 1994.
Editors: J.H. van Bemmel, A.Rosenfalck, N.Saranummi.

[4] I. Bendtson, J. Gade, A.M. Rosenfalck, C.E. Thomsen, G. Wildschiodtz and C.Binder.
"Nocturnal electroencephalogram registrations in Type 1 (insulin-dependent) diabetic patients with hypoglycaemia"
1991

[5] C.E. Thomsen and P.F. Prior
"Quantitative EEG in assessment of anaesthetic depth: comparative study of methodology".
British Journal of Anaesthesia, 1996.

[6] C.E. Thomsen, A. Rosenfalck and K. Norregaard Christensen
"Assessment of anaesthetic depth by clustering analysis and autoregressive modelling of electroencephalograms".
Computer Methods and Programs in Biomedicine, 1991.

[7] Toney Allman
"Diabetes" from "Genes & Disease" series.
Chelsea House Publishers, 2008.

[8] Philip E. Cryer, MD
"Hypoglycemia in Diabetes. Pathophysiology, Prevalence, and Prevention".
American Diabetes Association, 2009.

[9] John M. Stern. Jerome Engel, Jr., editor
"Atlas of EEG patterns".
Lippincott Williams & Wilkins, 2005.

[10] http://www.tbirecoverycenter.org/images/brainmap.jpg

[11] Pamela F. Prior, M.D., M.R.C.P. and Douglas E. Maynard, Ph.D., M.Phil
"Monitoring Cerebral Function. Long-Term Monitoring of EEG and Evoked Potentials."
Elsevier, 1986

[12] http://www-users.cs.york.ac.uk/~fisher/mkfilter/trad.html

[13] L. Hyllienmark, J. Maltez, A. Dandenell, J. Ludvigsson, T. Brismar
"EEG abnormalities with and without relation to severe hypoglycaemia in adolescents with type 1 diabetes".
Diabetologia, 2005.

[14] Inan Güler and Elif Derya Übeyli
"Multiclass Support Vector Machines for EEG-Signals Classification"
IEEE Transactions on information technology in biomedicine. vol.11, no.2. March 2007.

[15] Pasan Hapuarachchi
"Feature selection and artifact removal in sleep stage classification"
Waterloo, Ontario, Canada, 2006.

[16] Christopher M. Bishop
"Pattern Recognition and Machine Learning".
Cambridge, February 2006.

[17] Alan V. Oppenheim
"Discrete-Time Signal Processing"
Prentice-Hall, 1989.

[18] Matlab Help Navigator

[19] Richard O. Duda and Peter E. Hart
"Pattern Classification and Scene Analysis"
John Wiley and Sons, 1973.

[20] Wild S, Roglic G, Green A, Sicree R, King H.
"Global prevalence of diabetes: estimates for the year 2000 and projections for 2030".
Diabetes Care **27** (5): 1047–53. doi:10.2337/diacare.27.5.1047. PMID 15111519.
http://care.diabetesjournals.org/cgi/content/full/27/5/1047.
May 2004.

9. Appendix

Content of the additional DVD:

- Matlab codes and Matlab figures.
- Figures of every single patient.
- Figures of groups of patients analyzed.
- Figures after classify data from every patient.
- Electronic version of the Master Thesis report.
- Important: Only data from diabetic patients belonging to channel 1 has been included in the programs.

10. Presupuesto

Gastos generales

- Cánula intravenosa en una vena ante cubital en ambos antebrazos.
- Casco EEG y dos cables ECG precordiales conectados a un grabador digital de EEG (Cadwell, Kennewick, Washington, USA).
- Auriculares conectados a un ordenador que libera estímulos auditivos..
- Unidad de presión sanguínea ambulatoria.
- Videocámara para grabar los experimentos.
- Software para visión de videos.
- Disco duro para almacenación de datos.
- Insulina y alimentos azucarados.

Gastos sufragados por

- Hillerød Hospital (Dinamarca)
- Rigshospitalet (Copenhague, Dinamarca)
- University of Copenhagen (Dinamarca)

Madrid, Septiembre de 2009

El Ingeniero Jefe de Proyecto

Fdo.: María Riesco García
Ingeniero Superior de Telecomunicación

11. Pliego de condiciones

Este documento contiene las condiciones legales que guiarán la realización, en este proyecto, de un sistema de detección de hipoglucemia. En lo que sigue, se supondrá que el proyecto ha sido encargado por una empresa cliente a una empresa consultora con la finalidad de realizar dicho sistema. Dicha empresa ha debido desarrollar una línea de investigación con objeto de elaborar el proyecto. Esta línea de investigación, junto con el posterior desarrollo de los programas está amparada por las condiciones particulares del siguiente pliego.

Supuesto que la utilización industrial de los métodos recogidos en el presente proyecto ha sido decidida por parte de la empresa cliente o de otras, la obra a realizar se regulará por las siguientes:

Condiciones generales

1. La modalidad de contratación será el concurso. La adjudicación se hará, por tanto, a la proposición más favorable sin atender exclusivamente al valor económico, dependiendo de las mayores garantías ofrecidas. La empresa que somete el proyecto a concurso se reserva el derecho a declararlo desierto.

2. El montaje y mecanización completa de los equipos que intervengan será realizado totalmente por la empresa licitadora.

3. En la oferta, se hará constar el precio total por el que se compromete a realizar la obra y el tanto por ciento de baja que supone este precio en relación con un importe límite si este se hubiera fijado.

4. La obra se realizará bajo la dirección técnica de un Ingeniero Superior de Telecomunicación, auxiliado por el número de Ingenieros Técnicos y Programadores que se estime preciso para el desarrollo de la misma.

5. Aparte del Ingeniero Director, el contratista tendrá derecho a contratar al resto del personal, pudiendo ceder esta prerrogativa a favor del Ingeniero Director, quien no estará obligado a aceptarla.

6. El contratista tiene derecho a sacar copias a su costa de los planos, pliego de condiciones y presupuestos. El Ingeniero autor del proyecto autorizará con su firma las copias solicitadas por el contratista después de confrontarlas.

7. Se abonará al contratista la obra que realmente ejecute con sujeción al proyecto que sirvió de base para la contratación, a las modificaciones autorizadas por la superioridad o a las órdenes que con arreglo a sus facultades le hayan comunicado por escrito al Ingeniero Director de obras siempre que dicha obra se haya ajustado a los preceptos de los pliegos de condiciones, con arreglo a los cuales, se harán las modificaciones y la valoración de las diversas unidades sin que el importe total pueda exceder de los presupuestos aprobados. Por consiguiente, el número de unidades que se consignan en el proyecto o en el presupuesto, no podrá servirle de fundamento para entablar reclamaciones de ninguna clase, salvo en los casos de rescisión.

8. Tanto en las certificaciones de obras como en la liquidación final, se abonarán los trabajos realizados por el contratista a los precios de ejecución material que figuran en el presupuesto para cada unidad de la obra.

9. Si excepcionalmente se hubiera ejecutado algún trabajo que no se ajustase a las condiciones de la contrata pero que sin embargo es admisible a juicio del Ingeniero Director de obras, se dará conocimiento a la Dirección, proponiendo a la vez la rebaja de precios que el Ingeniero estime justa y si la Dirección resolviera aceptar la obra, quedará el contratista obligado a conformarse con la rebaja acordada.

10. Cuando se juzgue necesario emplear materiales o ejecutar obras que no figuren en el presupuesto de la contrata, se evaluará su importe a los precios asignados a otras obras o materiales análogos si los hubiere y cuando no, se discutirán entre el Ingeniero Director y el contratista, sometiéndolos a la aprobación de la Dirección. Los nuevos precios convenidos por uno u otro procedimiento, se sujetarán siempre al establecido en el punto anterior.

11. Cuando el contratista, con autorización del Ingeniero Director de obras, emplee materiales de calidad más elevada o de mayores dimensiones de lo estipulado en el proyecto, o sustituya una clase de fabricación por otra que tenga asignado mayor precio o ejecute con mayores dimensiones cualquier otra parte de las obras, o en general, introduzca en ellas cualquier modificación que sea beneficiosa a juicio del Ingeniero Director de obras, no tendrá derecho sin embargo, sino a lo que le correspondería si hubiera realizado la obra con estricta sujeción a lo proyectado y contratado.

12. Las cantidades calculadas para obras accesorias, aunque figuren por partida alzada en el presupuesto final (general), no serán abonadas sino a los precios de la contrata, según las condiciones de la misma y los proyectos particulares que para ellas se formen, o en su defecto, por lo que resulte de su medición final.

13. El contratista queda obligado a abonar al Ingeniero autor del proyecto y director de obras así como a los Ingenieros Técnicos, el importe de sus respectivos honorarios facultativos por formación del proyecto, dirección técnica y administración en su caso, con arreglo a las tarifas y honorarios vigentes.

14. Concluida la ejecución de la obra, será reconocida por el Ingeniero Director que a tal efecto designe la empresa.

15. La garantía definitiva será del 4% del presupuesto y la provisional del 2%.

16. La forma de pago será por certificaciones mensuales de la obra ejecutada, de acuerdo con los precios del presupuesto, deducida la baja si la hubiera.

17. La fecha de comienzo de las obras será a partir de los 15 días naturales del replanteo oficial de las mismas y la definitiva, al año de haber ejecutado la provisional, procediéndose si no existe reclamación alguna, a la reclamación de la fianza.

18. Si el contratista al efectuar el replanteo, observase algún error en el proyecto, deberá comunicarlo en el plazo de quince días al Ingeniero Director de obras, pues transcurrido ese plazo será responsable de la exactitud del proyecto.

19. El contratista está obligado a designar una persona responsable que se entenderá con el Ingeniero Director de obras, o con el delegado que éste designe, para todo relacionado con ella. Al ser el Ingeniero Director de obras el que interpreta el proyecto, el contratista deberá consultarle cualquier duda que surja en su realización.

20. Durante la realización de la obra, se girarán visitas de inspección por personal facultativo de la empresa cliente, para hacer las comprobaciones que se crean oportunas. Es obligación del contratista, la conservación de la obra ya ejecutada hasta la recepción de la misma, por lo que el deterioro parcial o total de ella, aunque sea por agentes atmosféricos u otras causas, deberá ser reparado o reconstruido por su cuenta.

21. El contratista, deberá realizar la obra en el plazo mencionado a partir de la fecha del contrato, incurriendo en multa, por retraso de la ejecución siempre que éste no sea debido a causas de fuerza mayor. A la terminación de la obra, se hará una recepción provisional previo reconocimiento y examen por la dirección técnica, el depositario de efectos, el interventor y el jefe de servicio o un representante, estampando su conformidad el contratista.

22. Hecha la recepción provisional, se certificará al contratista el resto de la obra, reservándose la administración el importe de los gastos de conservación de la misma hasta su recepción definitiva y la fianza durante el tiempo señalado como plazo de garantía. La recepción definitiva se hará en las mismas condiciones que la provisional, extendiéndose el acta correspondiente. El Director Técnico propondrá a la Junta Económica la devolución de la fianza al contratista de acuerdo con las condiciones económicas legales establecidas.

23. Las tarifas para la determinación de honorarios, reguladas por orden de la Presidencia del Gobierno el 19 de Octubre de 1961, se aplicarán sobre el denominado en la actualidad “Presupuesto de Ejecución de Contrata” y anteriormente llamado ”Presupuesto de Ejecución Material” que hoy designa otro concepto.

Condiciones particulares

La empresa consultora, que ha desarrollado el presente proyecto, lo entregará a la empresa cliente bajo las condiciones generales ya formuladas, debiendo añadirse las siguientes condiciones particulares:

1. La propiedad intelectual de los procesos descritos y analizados en el presente trabajo, pertenece por entero a la empresa consultora representada por el Ingeniero Director del Proyecto.

2. La empresa consultora se reserva el derecho a la utilización total o parcial de los resultados de la investigación realizada para desarrollar el siguiente proyecto, bien para su publicación o bien para su uso en trabajos o proyectos posteriores, para la misma empresa cliente o para otra.

3. Cualquier tipo de reproducción aparte de las reseñadas en las condiciones generales, bien sea para uso particular de la empresa cliente, o para cualquier otra aplicación, contará con autorización expresa y por escrito del Ingeniero Director del Proyecto, que actuará en representación de la empresa consultora.

4. En la autorización se ha de hacer constar la aplicación a que se destinan sus reproducciones así como su cantidad.

5. En todas las reproducciones se indicará su procedencia, explicitando el nombre del proyecto, nombre del Ingeniero Director y de la empresa consultora.

6. Si el proyecto pasa la etapa de desarrollo, cualquier modificación que se realice sobre él, deberá ser notificada al Ingeniero Director del Proyecto y a criterio de éste, la empresa consultora decidirá aceptar o no la modificación propuesta.

7. Si la modificación se acepta, la empresa consultora se hará responsable al mismo nivel que el proyecto inicial del que resulta el añadirla.

8. Si la modificación no es aceptada, por el contrario, la empresa consultora declinará toda responsabilidad que se derive de la aplicación o influencia de la misma.

9. Si la empresa cliente decide desarrollar industrialmente uno o varios productos en los que resulte parcial o totalmente aplicable el estudio de este proyecto, deberá comunicarlo a la empresa consultora.

10. La empresa consultora no se responsabiliza de los efectos laterales que se puedan producir en el momento en que se utilice la herramienta objeto del presente proyecto para la realización de otras aplicaciones.

11. La empresa consultora tendrá prioridad respecto a otras en la elaboración de los proyectos auxiliares que fuese necesario desarrollar para dicha aplicación industrial, siempre que no haga explícita renuncia a este hecho. En este caso, deberá autorizar expresamente los proyectos presentados por otros.

12. El Ingeniero Director del presente proyecto, será el responsable de la dirección de la aplicación industrial siempre que la empresa consultora lo estime oportuno. En caso contrario, la persona designada deberá contar con la autorización del mismo, quien delegará en él las responsabilidades que ostente.