Theses by compendium of publications

Interuniversity Doctoral Program in Bioinformatics (UAB, UPC, UdG, UdL, UOC, UVic-UCC, URV, UB)

Data Science Applied to Chronic Fatigue Syndrome

PhD candidate:
Marcos Lacasa Cazcarra

Thesis supervisors:
Dr. José Alegre Martín
Dr. Jordi Casas Roma

June 2024

Acknowledgements

In 2020, I first came into contact with Dr Alegre at the Vall d'Hebrón Hospital in Barcelona by chance through a mutual friend, Dr Antonio Segura. Naturally, I mentioned my interest in confirming whether machine learning could improve people's lives. Without any expectation of financial reward for his time and dedication, Dr Alegre made it his own. The commitment has been absolute and the determination energetic and unwavering. Unfortunately, the only return was the knowledge that the work was right. The achievement of the objectives was optimal. The quality of the work was exemplary. There is no expectation of recognition for the dedication and enthusiasm that goes into every protocol, every article, every extra visit to one more patient. This thanks, which may seem hollow, formal and obligatory, is intended to make those reading these lines realise that all the initiative put into this thesis has been despite everything and thanks to us. Therefore, So I cannot be grateful that a brilliant mind wants to share his talent with society against all odds. Rather, I must emphasize it.

Supervising doctoral students, as well as everything related to research, is poorly paid at best. Dr Jordi Casas has respected, valued, and improved this work. With respect, prudence, and know-how, qualities to be recognized in these times of noise, shouting, and loss of the essential. You do not have to shout to be heard. You don't have to attack to be respected. You don't have to impose to lead. I hope this will be one of many doctoral projects that I will consider directing for the good of the doctoral student.

The Open University of Catalonia has all the tools to make doctoral students feel competitive: funding for the publication of articles, access to scientific journals, and support for international dissemination. So, this work has been completed thanks to the environment that the university has provided us with.

Finally, after reading several letters from patients diagnosed with Chronic Fatigue Syndrome who have told me of their suffering and fears, I would like to express my final intention for this work: to give hope. That research will continue and, if possible, with resources. As far as I am concerned, this is what will happen.

Tarragona, June 2024.

Dedicatory note

To Ariadna and Sonia, for my new project: to make up for all the hours that I have spent on this work.

Abstract

La encefalomielitis miálgica o síndrome de fatiga crónica (EM/SFC) es un proceso orgánico, debilitante y multifacético. El inicio heterogéneo y la presentación clínica con comorbilidades adicionales dificultan su diagnóstico. No hay evidencias de pruebas diagnósticas ni biomarcadores que determinen, por sí solos, su diagnóstico.

Las pautas de tratamiento actuales se centran en el manejo de los síntomas. Se desconoce el mecanismo de acción causante. Las líneas de investigación son heterogéneas y se requiere un consenso en aunar recursos para la búsqueda de resultados sólidos. Existe una necesidad urgente de realizar ensayos clínicos para identificar tratamientos eficaces.

Un elemento limitante es la falta de medidores de eficacia objetivos, y éstos, se basan actualmente en cuestionarios subjetivos que contesta el propio paciente. Una de las conclusiones más importantes es la necesidad del uso de medidas objetivas. En esta investigación se aporta 2 biomarcadores que pueden ser utilizados en este fin: el consumo pico de oxígeno en la prueba de esfuerzo y el resultado del test CPT3 para medir el deterioro cognitivo.

El consumo de oxígeno en la prueba de esfuerzo

Se demuestra que la medida de consumo pico de oxígeno en la prueba de esfuerzo determina el estado físico del paciente afecto EM/SFC. Además se puede realizar un cribado mediante el uso de un modelo basado en inteligencia artificial que predice qué pacientes requieren de una prueba de esfuerzo por un estado físico limitante.

El test CPT3 para la detección del deterioro cognitivo

Se demuestra que el uso del test CPT3 mide el deterioro cognitivo del paciente afecto a EM/SFC y puede ser un medidor de eficacia de tratamientos en futuros ensayos clínicos.

Transferencia de conocimiento

Se diseña una aplicación que ofrece un análisis multidisciplinar, y predice el riesgo físico de un paciente afecto EM/SFC requiriendo las respuestas del cuestionario SF-36 como datos de entrada. Esta aplicación no sólo favorece la detección precoz del deterioro físico y la sugerencia de la derivación a una unidad especializada, sino que favorecería una detección mayor del síndrome y aportaría conocimiento a la comunidad médica sobre este campo.

Resum

L'encefalomielitis miàlgica o síndrome de fatiga crònica (EM/SFC) és un procès orgànic, debilitant i multifacètic. L'inici heterogeni i la presentació clínica amb comorbiditats addicionals en dificulten el diagnòstic. No hi ha evidències de proves diagnòstiques ni biomarcadors que en determinin, per si sols, el diagnòstic.

Les pautes de tractament actuals se centren en el maneig dels símptomes. Es desconeix el mecanisme d'acció causant. Les línies de recerca són heterogènies i es requereix un consens a unir recursos per a la cerca de resultats sòlids. Hi ha una necessitat urgent de fer assajos clínics per identificar tractaments eficaços.

Un element limitant és la manca de mesuradors d'eficàcia objectius, i aquests es basen actualment en qüestionaris subjectius que contesta el mateix pacient. Una de les conclusions més importants és la necessitat de fer servir mesures objectives. En aquesta investigació s'aporten 2 biomarcadors que es poden utilitzar en aquesta finalitat: el consum pic d'oxigen a la prova d'esforç i el resultat del test CPT3 per mesurar el deteriorament cognitiu.

El consum d'oxigen a la prova d'esforç

Es demostra que la mesura de consum pic d'oxigen a la prova d'esforç determina l'estat físic del pacient afecte EM/SFC. A més, es pot fer un cribratge mitjançant l'ús d'un model basat en intel·ligència artificial que prediu quins pacients requereixen una prova d'esforç per un estat físic limitant.

El test CPT3 per a la detecció del deteriorament cognitiu

Es demostra que l'ús del test CPT3 mesura el deteriorament cognitiu del pacient afecte a EM/SFC i pot ser un mesurador d'eficàcia de tractaments en assaigs clínics futurs.

Transferència de coneixement

Es dissenya una aplicació que ofereix una anàlisi multidisciplinària i prediu el risc físic d'un pacient afecte EM/SFC requerint les respostes del qüestionari SF-36 com a dades d'entrada. Aquesta aplicació no només afavoreix la detecció precoç del deteriorament físic i el suggeriment de la derivació a una unitat especialitzada, sinó que afavoriria una detecció més gran de la síndrome i aportaria coneixement a la comunitat mèdica sobre aquest camp.

Abstract

Myalgic Encephalomyelitis, or Chronic Fatigue Syndrome (ME/CFS), is a chronic, debilitating, multi-faceted disease. Its heterogeneous onset and clinical presentation with additional co-morbidities make diagnosis difficult. There is no evidence of diagnostic tests or biomarkers alone to establish a diagnosis.

Current treatment guidelines focus on symptom management. The causative mechanism is unknown. Research lines are heterogeneous. A consensus is needed to pool resources and find solid results. Clinical trials are urgently needed to identify effective treatments.

One limiting factor is the lack of objective measures of efficacy. These are currently based on subjective questionnaires answered by the patients themselves. The need to use objective measures is one of the main conclusions. This research provides two biomarkers that can be used for this purpose: peak oxygen consumption in the stress test and the CPT3 test result to measure cognitive impairment.

Exercise test oxygen consumption

It is shown that the physical status of the ME/CFS patient can be determined by measuring the peak oxygen consumption in the exercise test. Furthermore, screening can be performed using an artificial intelligence-based model that predicts patients requiring stress testing due to limiting physical status.

The CPT3 Test for the detection of cognitive impairment

The CPT3 test has been shown to measure cognitive impairment in ME/CFS patients and may be used to measure treatment efficacy in future clinical trials.

Knowledge Transfer

An application using the SF-36 questionnaire as input data will be developed that provides multidisciplinary analysis and predicts the physical risk of a ME/CFS patient. This application not only favors the early detection of physical deterioration and the suggestion of referral to a specialized unit, but would also favor a greater detection of the syndrome and provide knowledge to the medical community in this field.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This section contextualizes concepts commonly used in recent scientific publications dealing with data science applied to health or bioinformatics. As the interest that this thesis may arouse is both for practitioners interested in chronic fatigue syndrome and for data scientists or biostatisticians interested in the methods, I would like to approach both the development and the conclusions of this research from both sensitivities. It is important to study the particularities of the syndrome in order to understand the methods used. But also the knowledge of the methodological context from machine learning and health sciences.

1.1 About data science and bioinformatics

The term **artificial intelligence** (AI) was associated in 1988 with the behavior of a machine, which, if a human behaves in the same way, is considered intelligent [1]. A consensus definition is currently being sought due to the many nuances that the term AI encompasses. [2].

AI is a basic tool in interdisciplinary research fields that transcend academia. AI is associated with a category of algorithms known as machine learning, including deep learning. These algorithms rely on vast amounts of data to deliver results that have led to innovation across all sectors. It is transforming our society by changing how people interact with each other and how they interact with the world [3].

The development of AI has been made possible by the use of large classified databases and an unprecedented improvement in the technology of informatics services. In healthcare, it impacts clinicians through rapid and accurate interpretation of medical records, tests, biomarkers, and images; healthcare systems through improved workflow; and patients through access to medical information systems that make complex data easier to understand [4].

Machine learning is defined as a set of methods that can automatically detect patterns in data and use them to predict future data or make other types of decisions under conditions of uncertainty. The accepted typology is [5]:

- Supervised: The data used is characterised by attributes, it is also called labeled. The problems it solves are usually classification or regression problems.
- Unsupervised: The data is not labeled. This is a much less well-defined problem, as it does not indicate what kind of patterns to look for, and there is no obvious error metric to use. An example would be discovering homogeneous groups, given an n-dimensional data set. Graph analysis would also be included in this section.
- Reinforcement learning: A machine learning training method based on rewarding and punishing an agent in its decision-making. In general, a reinforcement learning agent is able to sense and interpret its environment, take actions and learn through trial and error. An example is automated driving.

When it comes to processing natural data in its original form, these machine-learning techniques have limitations. **Deep learning** is a set of methods that use raw data (no longer just rows and columns of characterised numbers) and automatically discover patterns for detection or classification. [6]. Image recognition and classification were among the first models based on this technology. Neural networks are a set of algorithms inspired by the communication mechanism of the biological neuron. Their learning ability makes them both adaptive and sophisticated algorithms [7].

Bioinformatics has recently been defined as an interdisciplinary field concerned with the development and application of algorithms that analyze biological data to investigate the structure and function of biological polymers (carbohydrates, nucleic acids, and proteins) and their relationships to living systems [8]. And we define biological data as any record of a person that is properly obtained. This data can refer to anything that is included in the patient's medical history, laboratory tests, diagnostic images and any information that has been collected in any medium that refers to the patient. Legal standards are followed in the collection of any biological data by any institution, whether or not it is part of the patient's medical record. Thus, collecting and using such data in any research must comply with regulations.

The greater accuracy of predictions and the use of more sophisticated and ambitious algorithms is due to a significant improvement in the computing power of servers at the service of artificial intelligence. In bioinformatics, two areas are developing significantly: diagnostic imaging and protein and gene sequencing. In the last five years, they have evolved to the point where paradigms have changed. And thanks to the continuing evolution of technology in the service of data science, this is likely to continue at a rapid pace in the coming years.

1.2 Biomarkers

The term **biomarker** refers to a broad category of medical signs that can be measured accurately and reproducibly. A biomarker provides clinically relevant information. The validity of a biomarker refers to the need to characterize its effectiveness as an endpoint that could replace clinically meaningful endpoints [9]. Recent advances in biomarker analysis and bioinformatics have opened up new possibilities in precision and predictive medicine [10].

Biomarkers can be used clinically for the detection, diagnosis, or monitoring of disease activity and can be used to guide molecularly targeted therapy or to assess therapeutic response [11]. Validation of the use of biomarkers for clinical trials, particularly in humans, is required. Reliability (the ability to replicate tests and obtain the same results under the same measurement conditions) is the most important consideration in biomarker development. Reliability requires the ability to measure error using statistical methods, and indices have been developed to estimate reliability. Other parameters must also be evaluated, such as reproducibility, sensitivity, and specificity. Finally, interlaboratory variability must be measured [12]. The application of biomarkers in developing new drugs requires a general description of the development of a specific biomarker [13].

Genomics focuses on universally detecting genes; if it focuses on detecting messenger RNA (mRNA), we will talk about transcriptomics. If the goal is to analyze proteins, it is proteomics, and if it is to analyze metabolites, it is metabolomics. The main source of information for analysis in a given biological sample must be non-selective and unbiased [14].

The **genome** is the total DNA of a cell or organism. The human genome contains 3.2 billion bases and about 40,000 protein-coding genes. A genomic analysis aims to measure differences in DNA sequence between individuals and the expression of thousands of genes simultaneously to look for abnormal chromosomal patterns. They are particularly interesting when related to diseases [14].

The **transcriptome** is the total mRNA of a cell and plays a fundamental role in protein synthesis in a process called translation. The main goal of transcriptomics is to discover how the transcripts of a living cell, tissue or organism are affected by disease or environmental factors (such as drugs, hormones, etc.). Non-coding RNA is another very important aspect of transcriptomics. These functional elements play an important role in the pathogenesis of various diseases and their response to treatment. Transcriptomics complements genomics and bridges the gap between genomics and proteomics in the era of precision medicine [15].

The databases used for transcriptomics are microarrays and RNA sequencing (*RNA-seq*). *RNA-seq* assays represent an increasing proportion of new gene expression experiments and are emerging as a better option than microarrays. Integrative analyses of gene expression require the combination of these types of data. This task is of utmost importance for rare diseases or poorly understood biological processes and organisms, where all available assays are needed to discover robust signatures or biomarkers. Therefore, for future functional genomic and transcriptomic experiments, effective strategies for combining data from both platforms, perhaps in a way that leverages both the advantages of *RNA-seq* and the abundance of microarray data, will be critical [16].

The result of any microarray assay, regardless of the technology, is to provide a measurement for each gene in a way that answers the most important question: do the genes in the sample have a higher or lower level of expression than the samples that we want to compare? The recent application of DNA microarrays in clinical research has been a very important step towards developing more complex markers based on multigene signatures (information about the activity of a specific group of genes). A machine learning supervised classification analysis of whether these signatures are associated with a disease in any

of their variations is called **biomarker detection** [17].

The **proteome** is defined as the set of all proteins expressed in a cell, tissue or organism. Proteomics aims to characterize the flow of information within the cell through protein pathways and networks. The goal is to understand the functional relevance of proteins. Proteomic research requires the analysis of more than 100,000 proteins. It is still difficult to accurately detect all proteins. Proteomics is a fundamental line of research in biomarker discovery because proteins are likely to be ubiquitously involved in disease and disease response [14].

Metabolomics can be defined as the study of global metabolite profiles in a system (cell, tissue or organism) under given conditions [18]. Metabolomics has some theoretical advantages over other omics approaches. Because the metabolome is the end product of gene transcription, changes in the metabolome will be amplified relative to changes in the transcriptome or proteome. The metabolome is more closely related to the phenotype of the biological system that is being studied. Approximately 5,000 metabolites have been defined, but the analysis is more diverse and includes many different biological molecules, which makes it more physically and chemically complex than previous analyses [14].

Many of the pattern recognition strategies currently used in metabolomics are based on unsupervised machine learning. **Clustering** methods are used to assess how similar a set of samples are to each other based on their metabolomic profiles. Supervised learning methods use some patterns that are labeled because a relationship has already been defined. The goal of supervised learning is to find a model that correctly associates the inputs (classified samples) with the targets (mathematical relationship with the samples under study).

In summary, bioinformatics has become an essential tool in translational research (understood as the synergy between clinical and basic research knowledge) [19]. The ultimate goal is to accelerate the discovery of biomarkers. The transfer of this knowledge to society applies to the individualized clinical management of patient's health and should facilitate the development of new personalized treatments.

1.3 Explanatory Memorandum

This research has been developed under the auspices of the Unit Specialising in Myalgic Encephalomyelitis or Chronic Fatigue Syndrome (ME/CFS) of the Vall d'Hebron Hospital in Barcelona, whose research lines are attached to the Research Institute of the same hospital (VHIR, Vall d'Hebron Research Institute).

Led by Dr Alegre, the ME/CFS unit has built up a database of patients diagnosed with ME/CFS over the last twenty years, with more than 3,000 records. This database is an order of magnitude larger than those used in state-of-the-art reference works in the field, which gives an idea of its importance.

No consensus on the definition of biomarkers can objectively and unambiguously indicate whether a patient has ME/CFS or not. This work has shown that the interpretation of the patient's clinical history determines the diagnosis. The questionnaires listed in this research represent a reliable tool for assessing the health status of the patient with ME/CFS.

We therefore ask whether applied data science can help improve the lives of people with ME/CFS, both those who have been diagnosed and those who have symptoms but have not yet been diagnosed.

1.4 Objectives

The objectives defined at the beginning of this research are listed below:

1. **Search for new biomarkers.**
   It is necessary to define new biomarkers to measure the efficacy of treatments in future clinical trials. The fact of being able to repeatedly evaluate the patient affected by ME/CFS at different times would already be considered a very important improvement, since an objective and accurate measure of their clinical evolution would be available.

2. **The creation of models capable of predicting the severity of the pathology in patients diagnosed with ME/CFS.**
   Relating the database, especially questionnaire responses to objective tests makes it possible to obtain robust results in supervised machine learning. These advances have produced the findings presented in the scientific publications of this thesis.

3. **The transfer of knowledge through bioinformatics applications could help the medical community advance in the study of pathology.**
   Offering both the scientific community and patients an easily accessible application provides some key points defined in this thesis, such as: education of the primary care physician in ME/CFS, improvement in the detection and evolution of the affected patient, and detection of the patient at high risk of morbidity. An early detection of symptoms in physical or cognitive deterioration could affect the course of the disease and improve its evolution.

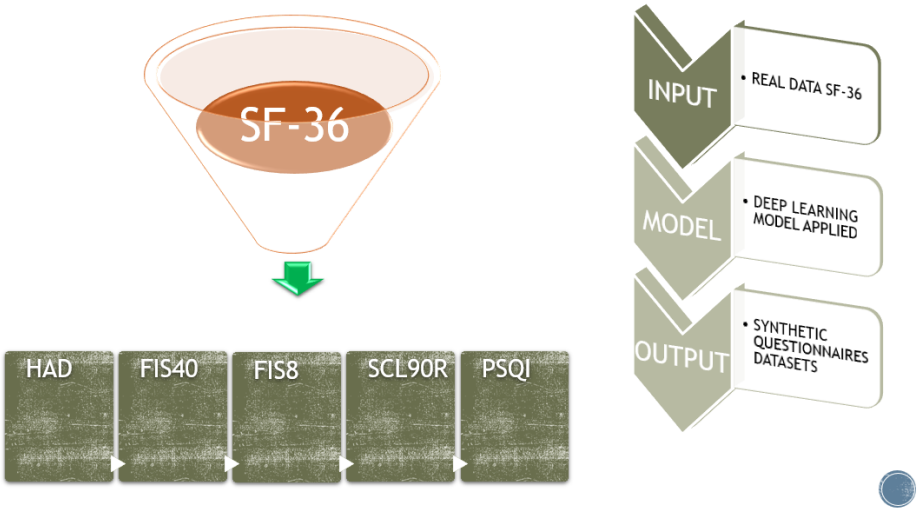
1.5 Knowledge transfer

The generation of scientific knowledge demonstrated through scientific publications has, as part of its objective, its transfer to society, so that it results in improvements for people. Numerous studies have highlighted the difficulty involved in this [20] [21].

As part of the tasks of this research, we have collaborated with companies specialized in the development of applications to evaluate a computer application that develops the main conclusions of the studies presented in this compendium in order to substantiate the generation of knowledge, precisely so that it will benefit people. For more information on the use of questionnaires in section 5.1.

Specifically, this application allows the following:

Figure 1.1: Process and generation of information



The generation of information from the app is based on obtaining the answers to the SF-36 questionnaire, and using the model generated in this research, a study is offered on the results, not only the SF-36 but also of 5 other questionnaires. The responses can be collected from a web or mobile application.

- The application (app) can be viewed by scanning QR codes, or direct access, which can be displayed in health centers, or through digital media. The app system will be web-based, so it will be accessible from any device. The SF-36 questionnaire can be completed directly without registration, but in order to access the results the user must provide a series of basic data (e.g. name and email). In this way, the results can be shared with a health professional if the patient wishes.
- The system will use an AI algorithm developed in the publication of section 6.1.3, which will receive the answers of the SF-36 questionnaire as input parameters and return, within seconds, a series of output parameters, corresponding to the following health information
  - Anxiety and depression provided by the HAD questionnaire.
  - Physical status and its cognitive impairment provided by the FIS40 and FIS8 studies.
  - State of psychological impairment as measured by the SLC90R questionnaire.
  - State of sleep dysfunction as measured by the PSQI questionnaire.
  - Health status and its prediction of peak O_2 consumption in a hospital stress test.
- The system will have a web access for registered users, where they will be able to register health centers, and the centers will be able to register authorized health personnel. The physician will be able to view the results of those tests that have been shared with him. In the center, an administrator role will be able to see all the tests managed by the assigned professionals. Each test will have a unique code, which will allow the patient to share the code with the physician, who will be able to find it in the system, in a secure way, and without sharing personal data.
- The user, to visualize his results, must create an access to the system (authorized according to European security standards). By logging in with your credentials, you will be able to view your

result and your unique code, which you can share with your trusted physician so that he/she can view and evaluate the result.

- It will be possible to access an environment of recommendations and useful information for the patient suffering from chronic fatigue syndrome.

Chapter 2

ME/CFS Review

2.1 Clinical definition

Fatigue is a common symptom in everyday medical consultations. It is defined as a feeling of exhaustion and weakness and difficulty in performing sustained physical and/or mental activity [22]. Fatigue that persists or recurs over more than six months is called chronic fatigue (CF). Its presence may be due to a psychiatric disorder, an organic disease; it may be idiopathic, or it may be chronic fatigue syndrome (CFS) [23].

ME/CFS is a well-defined clinical entity that involves severe post-exertional fatigue that causes severe disability in patients, significantly interfering with their work and daily living tasks. It affects young adults, predominantly women [24]. In addition to fatigue and intolerance to physical exercise, these patients are accompanied by characteristic symptomatology of inflammatory, muscular, sleep dysfunction, and impaired cognitive functions [25].

ME/CFS is a complex chronic clinical condition that lacks biomarkers that define the diagnosis by blood tests [26]. There are 20 clinical and consensus research definitions of ME/CFS in the literature [27], which include: Canadian consensus criteria [28], Fukuda [29], Holmes [30], International criteria [31], Oxford [32], among others. The US Institute of Medicine appointed a committee in 2015 that drafted new criteria for ME/CFS and renamed it Systemic Exertion Intolerance Disease (SEID). With inconsistent ME/CFS criteria, it has been challenging to conduct definitive studies leading to a new understanding of pathophysiology, new diagnostic tests, and treatment methods.

As a result of not having diagnostic biomarkers, many physicians have been unconvinced of the definition of ME/CFS. As a result, these patients have been misunderstood and told that they had no real physical illness and that it was “all in their head”. The Institute of Medicine (IOM) and the National Academy of Medicine (NAM) recognize that diagnosing ME/CFS is not medical.

The report *"Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness"*

[33], described ME/CFS as "a syndrome characterized by profound fatigue, cognitive dysfunction, sleep abnormalities, autonomic manifestations, pain, and other symptoms that worsen with exertion" of any kind. ME/CFS can severely affect patients' ability to lead an everyday life.

ME/CFS case definitions are evaluated in terms of sensitivity (ability to identify patients with ME/CFS correctly) and specificity (ability to exclude patients who do not have ME/CFS). The SEID criteria are similar to the 1994 Fukuda CFS criteria [29] and the 1991 Oxford CFS criteria [32], which include other major psychiatric pathologies in their criteria. There are "detrimental implications for research in interpreting epidemiological, etiological and treatment aspects" of patients with ME/CFS [34].

2.2 Diagnosis

ME/CFS continues to cause significant morbidity worldwide, and it is estimated that 84% to 91% of people with ME/CFS symptoms remain undiagnosed due to the lack of diagnostic biomarkers [34]. The evolution of diagnostic criteria can be summarized in the following agreements:

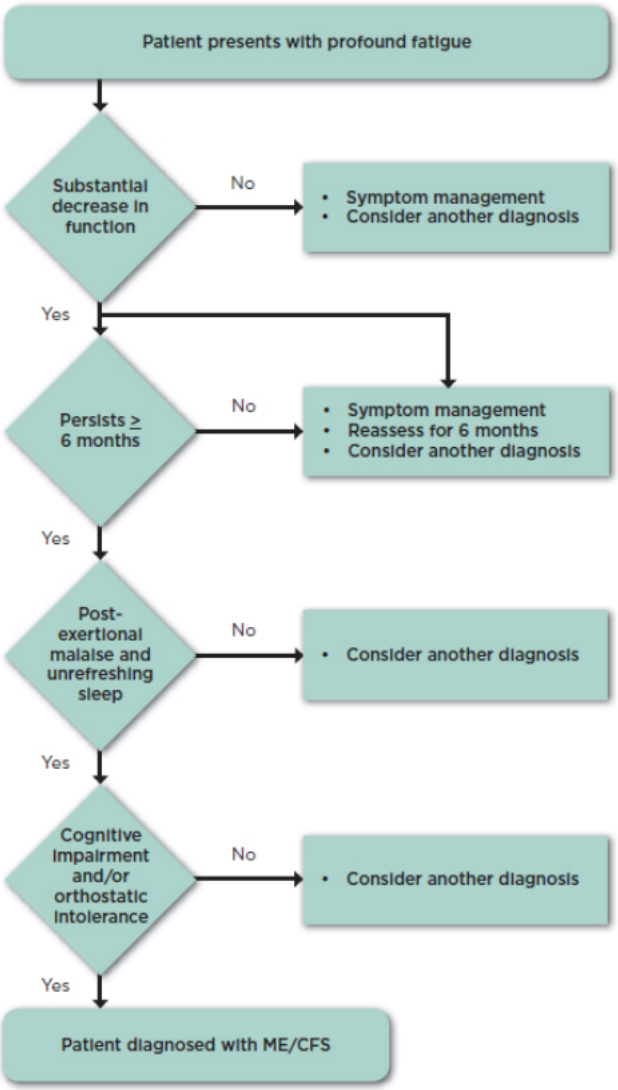
1. In 1994, the Atlanta Center for Disease Center (CDC) Diagnostic Criteria proposed by Fukuda [29].
2. A new definition of ME/CFS created to exclude psychiatric cases was proposed in 2003 through the Canadian ME/CFS case consensus document [35], which complements the CDC criteria and allows us to study the symptomatic blocks (neurological, muscular, cognitive, neurovegetative and immunological).
3. In 2011, these criteria were updated, proposing post-exertional fatigue as a mandatory criterion [31].
4. In 2016, a consensus was published to review the evidence of the variety of diagnostic criteria and recommend a new terminology to refer to this entity by the Institute of Medicine in the USA, calling it systemic intolerance syndrome [36].

2.3 Clinical Diagnostic Criteria Worksheet

To diagnose ME/CFS, the patient must present the following symptomatic clinical criteria:

1. Disabling fatigue, post-exertional malaise, sleep disturbance, pain, two neurocognitive symptoms, and at least one symptom from two categories: autonomic, neuroendocrine, or immunologic.
2. Fatigue and other symptoms must persist or relapse for at least six months in adults or three months in children and adolescents. A provisional diagnosis may be possible.
3. Another disease cannot explain the symptoms.

Figure 2.1: Diagnostic algorithm



The current "Centers for Disease Control and Prevention" (CDC) diagnostic criteria for ME/CFS are also known as Systemic Exertion Intolerance Disease (SEID) criteria. The Centers for Disease Control and Prevention is a U.S. government agency dedicated to epidemiology and public health. It operates under the auspices of the Department of Health and Human Services. Source: Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness, Institute of Medicine of the National Academies.

Clinical criteria and description of symptoms:

1. **Pathological fatigue**: A significant degree of new onset, unexplained, persistent, or recurrent physical and/or mental fatigue that substantially reduces activity levels and is not the result of sustained exertion and is not relieved by rest.
2. **2. Post-exertional malaise**: Worsening symptoms of mild exertion or normal activity are followed by malaise, loss of physical and mental stamina and/or worsening of other symptoms. Recovery is delayed, taking more than 24 hours.
3. **sleep disturbance**: Sleep is not restful.
4. **Pain**: Pain is generalized, migratory or localized: myalgia, arthralgia (without signs of inflammation) and/or headache.
5. **Cognitive symptoms**: Impaired concentration, short-term memory or word retrieval, hypersensitivity to light, noise or emotional overload. Confusion, disorientation, slowed thinking, muscle weakness and/or ataxia.

In addition, at least one symptom from two of the following categories must be detected:

a) Autonomous:
Orthostatic intolerance, neurally mediated hypotension (NMH), postural orthostatic tachycardia (POTS), dizziness, extreme pallor, palpitations, exertional dyspnea, urinary frequency, irritable bowel syndrome (IBS), and nausea.

b) Neuroendocrine:
Low body temperature, cold extremities, sweating, intolerance to heat or cold, reduced tolerance to stress, and other symptoms worsen with stress, weight change, or abnormal appetite.

c) Immune:
Recurrent flu-like symptoms include sore throat, tender lymph nodes, fevers, and new sensitivities to foods, drugs, odors, or chemicals.

2.3.1 Symptom characteristics

The most common onset is sudden, but it may be gradual. Symptoms may vary from day to day or during the same day. Relapses and remissions are common. Reactivating symptoms after exertion may be immediate or delayed by 24 hours or more.

2.3.2 Exclusionary diseases

Many other diseases have symptoms that mimic ME/CFS. Active pathologic processes that could explain the main symptoms of fatigue, sleep disturbances, pain, and neurocognitive dysfunction should be ruled

out by anamnesis, physical examination, and medical tests. Some common exclusionary diseases are listed below:

- Anemia, autoimmune diseases such as rheumatoid arthritis, lupus, heart disease, endocrine disorders such as type I diabetes mellitus, Addison's disease, uncontrolled thyroid disease, menopause, infectious diseases such as tuberculosis, HIV/AIDS, or chronic hepatitis.
- Intestinal diseases such as celiac disease or Crohn's disease. Neurological disorders such as multiple sclerosis, Parkinson's disease, and myasthenia gravis.
- Primary psychiatric disorders and substance abuse. Melancholic and/or bipolar major depression are exclusionary psychiatric processes.

2.3.3 Non-excludable diseases

Some comorbid entities often appear in association with ME/CFS, such as:

- Allergies, fibromyalgia (FM), irritable bowel syndrome (IBS) and multiple chemical sensitivity (MCS). Any medical condition that has been adequately treated and is under control. Any isolated physical abnormality or laboratory test that is insufficient to diagnose an exclusionary condition.
- ME/CFS and FM are often closely associated and should be considered superimposed syndromes. A comorbid condition may precede the onset of ME/CFS by many years but then become associated with it. If the patient presents with unexplained and prolonged fatigue but an insufficient number of symptoms to meet the criteria for ME/CFS, the disease should be classified as idiopathic.
- Any physician can diagnose the condition by following standard medical procedures. A complete medical history should be taken; similar fatigue-related illnesses should be ruled out; typical signs and symptoms of the disease should be observed; all medications the patient is currently taking, and any other factors that may influence the severity or persistence of the fatigue should be noted. ME/CFS is also associated with comorbid health conditions. These comorbidities vary in prevalence and severity but are higher than in patients without ME/CFS [37].

2.4 Prevalence

Research on the prevalence of the syndrome has predominantly focused on meta-analysis due to the variety of this type of data and the complexity of the calculations. Published data indicate a prevalence of between 0.89 and 1.14 % among the US population [38]. Additional data include an analysis of adolescents in the United Kingdom with an estimated prevalence range of between 1.47 % and 2.99 % [39]. In contrast to these studies, recent research in China estimates the prevalence among the Chinese population to be 12.54 % [40]. The prevalence in Australia was estimated at 0.79 % [41]. In a survey based

on telephone interviews conducted in Wichita (US), middle-aged women constituted 70 % of patients with ME/CFS [42].

This syndrome is found in people of all ages and races. It has been reported in one case as young as four years old. It is more common in adolescents than in younger children. Cases may occur in clusters or sporadically [43].

The large discrepancy in prevalence statistics is due to different definitions of ME/CFS. The more specific the exclusion criteria in the definition, the lower the number of patients diagnosed with ME/CFS. This estimates that 836,000 and 2.5 million Americans may have ME/CFS, according to the IOM-SEID [36]. The prevalence in a recent European meta-analysis was estimated to be between 0.1%-2.2% [44].

2.5 Pathophysiology

No specific cause of ME/CFS has been found, and no specific diagnostic test is clinically available. The pathophysiological consequences of ME/CFS are multisystemic [26]. Although viral infections have long been considered the main trigger for disease onset, a precise mechanism of pathogenesis has not yet been defined [45] several reviews in the literature address general and specific issues related to ME/CFS. Through reviews of the scientific literature, we address how immune dysfunction, hormonal imbalance, genetics/epigenetics, and cognitive changes affect patients with ME/CFS, provide insight into the emerging role of non-coding RNAs and gut microbiome alterations in disease pathogenesis, and the possible relationship between the newly coined "persistent COVID" and chronic fatigue.

2.5.1 Inflammation

Patients diagnosed with ME/CFS experience symptoms related to immunological changes, such as high susceptibility to infections, prolonged recovery times, chronically tender and swollen lymph nodes, and a feeling of fever [26]. It is still unclear whether ME/CFS is an inherently low-grade inflammatory disease or is only accompanied by systemic inflammation [45]. The results suggest a common link between the central nervous system, infection, and the immune system. Current data from various sources support the model that ME/CFS has a propensity to overproduce pro-inflammatory cytokines, along with misregulation of anti-inflammatory cytokines. ME/CFS patients appear to have a specific immune dysfunction profile with increased basal activation of lymphoid subsets but suppression of specific immune responses, particularly Th1-driven (antiviral and antitumor responses) [46]. Another study suggests that natural killer (NK) cell cytotoxic activity may be a suitable biomarker for diagnosing ME/CFS [47]. The activated state of the immune system is also indicated by an increase in the biomarker neopterin, which is released by monocytes and macrophages, and a high concentration of acute phase reactants [48]. The body's ability to fight infections decreases with impaired NK cell function. The more severely the function of these cells is impaired, the worse the patient's symptoms of ME/CFS become and the more likely patients are to get recurrent infections due to immune suppression [49].

As inflammatory pathways seem to be altered in general, a possible explanation for the disease symptoms could be an altered intestinal barrier [50]. ME/CFS and irritable bowel syndrome (IBS) often co-occur [51]. Another possible explanation for the pronounced immune response could be autoimmunity.

Patients with ME/CFS show a more pronounced immune response to exercise than healthy sedentary controls. Many of these immune changes are related to the post-exertional malaise in ME/CFS, which is one of the main features of the syndrome [52]. In addition, patients appear to undergo increased oxidative stress more rapidly and for longer after exercise than healthy controls, and their antioxidant response is delayed and reduced [53].

2.5.2 Genetic and epigenetic alterations

Although most studies report an association between ME/CFS and one or a few polymorphisms, it should be noted that, being a multifactorial disease, a varied genetic contribution is more likely to explain predisposition and inheritance than a single variation. According to a recent publication [54], it is concluded that genetic studies are the best way to understand the etiology of ME/CFS due to the causal nature of genetic associations. A large genome-wide association study focused on uncovering the biomolecular mechanisms of ME/CFS is urgently needed because no study on the genetics of ME/CFS has yet seen repeated and replicated results. A genome-wide association study (GWAS) is a research approach used to identify genomic variants statistically associated with a risk of a particular disease or trait. The method involves studying the genomes of many people, looking for genomic variants that occur more frequently in those with a specific disease or trait than those without the disease or trait. Once such genomic variants are identified, they are typically used to search for nearby variants directly contributing to the disease or trait. Although recruiting thousands of people with ME/CFS, particularly severely affected individuals who are homebound or bedridden, is a challenging task, it will be essential to GWAS using their samples if disease mechanisms are to be understood [54].

2.5.3 Cognitive symptoms and depression

Sleep disturbances, depression, anxiety, and mood swings are often found in and characterize ME/CFS. Melancholic major depression and bipolar major depression (MD) may be exclusion criteria for ME/CFS. However, while the two conditions show some similar symptoms, they can still be distinguished. For example, in depressed individuals, fatigue is associated with apathy, whereas in patients with ME/CFS it is associated with intense frustration with their condition [55]. Every ME/CFS evaluation should include a mental status examination to identify abnormalities in mood, intellectual function, memory, and personality changes. Special attention should be paid to acute depressive, anxious, or self-destructive thoughts. It is concluded that MD and ME/CFS can be defined as comorbid. But an antidepressant treatment in patients without a history could be detrimental [55]. The paper on cognitive dysfunction and ME/CFS appended in the compendium is useful here [56].

2.5.4 Hormonal imbalance

Dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis appears to be associated with ME/CFS. Scientific evidence suggests a mechanism by which genetic and environmental factors may serve to create an endocrine environment that can affect immune and vascular processes and, therefore, presumably trigger and maintain the symptoms experienced by those diagnosed with ME/CFS. Despite the heterogeneity of symptoms in these patients and the evidence for multifactorial pathogenesis, the hormonal imbalance has been shown to be directly related to some of the symptoms present in ME/CFS, such as debilitating fatigue, difficulty concentrating, and sleep disturbance [57]. A popular hypothesis is the so-called "allostatic load state", in which the neuroendocrine system responds to a stressor (allostatic state) to restore the physiological set point (homeostasis). When this mechanism fails, allostatic overload occurs, and how the body deals with the stressor perpetuates the stress and chronicity of the condition. This situation may be due to a dysfunction of the HPA axis. However, clear evidence to support this suggestion is still lacking and further studies are needed to understand the role of neuroendocrinology in the pathogenesis of ME/CFS [58].

2.5.5 Dysbiosis and intestinal permeability

In chronic inflammatory conditions, such as ME/CFS, microbiome dysbiosis or imbalance occurs that favors the multiplication of other pathogens, and many microbes, viruses, and their corresponding proteins and metabolites may modulate and/or alter metabolic, immune, and neurologic reactions [59]. For example, NK cell dysfunction has been found in ME/CFS and their activity is known to modulate the bacterial microbiome [60]. Similarly, lactic acid producers have been found to increase NK cytotoxicity, and tryptophan produced by the microbiome interacts with aryl hydrocarbon receptors (AHR), microglia, and astrocytes [61] and in the genesis of autonomic dysfunction in ME/CFS. The microbiome would also be important in regulating blood pressure through Olfr78 and Gpr41 receptors [62]. The integrity of the gut-brain axis is of great importance for ME/CFS, as the vagus nerves can transmit and modify information about gut dysbiosis, and intracellular pathogens would travel through various pathways, such as nanotubules, reaching glial cells, dendritic cells, and organelles such as mitochondria [63].

Emerging evidence linking alterations in gut microbiota composition to clinical symptoms reported by patients with ME/CFS suggests a potential therapeutic role for microbiome-targeted interventions, such as supplementation with prebiotics and/or probiotics. Restoring intestinal homeostasis may help alleviate patients' health problems and improve their quality of life. The Compendium includes this approach [64].

2.5.6 Non-coding RNAs

It has been suggested that plasma may be a satisfactory parameter to determine miRNAs as biomarkers in ME/CFS. The data show that miR-127-3p, miR-142-5p and miR-143-3p may be potential plasma biomarkers for diagnosing ME/CFS. However, further studies are now needed to validate these findings

in a larger cohort to determine their diagnostic power [65]. The pathways in which each miRNA exerts its activity are still unclear, but several altered miRNAs have been identified in patients with ME/CFS [66]. The differential expression of circulating miRNAs in severely affected ME/CFS patients was measured. Eleven miRNAs associated with ME/CFS were discovered and validated. Based on these distinct miRNA signatures, the machine learning algorithm classified ME/CFS patients into four groups related to symptom severity. These findings may provide a basis for developing a new noninvasive test to diagnose ME/CFS patients. These miRNA signatures and clusters could eventually be used to predict the response to drug treatments and even to identify individuals for whom such treatments may be beneficial. In addition, possible mechanisms, yet to be validated, by which each of the miRNAs may play a role in the pathogenesis and etiology of ME/CFS are suggested [67]. The above-mentioned studies provide a basis for developing a comprehensive diagnostic and prognostic program that includes both metabolic and molecular analytes, such as miRNA or lncRNA, for diagnosing and selecting the most appropriate treatment.

2.5.7 ME/CFS and COVID-19

A proportion of COVID-19 survivors fail to return to their pre-existing condition and report persistent debilitating ME/CFS-like symptoms several months after resolution of acute COVID-19 infection. Increasing reports of persistent malaise and ME/CFS-like exhaustion can leave patients with physical weakness and emotional disturbances. Compounded by the psychological toll of the pandemic experienced by the entire population, persons recovering from COVID-19 may be at even greater risk for depression, anxiety, post-traumatic stress disorder, and medication use disorder [68]. This chronic post-viral syndrome has been termed "prolonged COVID" or "post-acute COVID-19 syndrome" and has been reported to affect patients regardless of the severity of acute infection [69].

As a member of the research team of the project *Characterization of post-exertional fatigue, cognitive dysfunction and dysautonomia in post-COVID-19 syndrome: a new emerging paradigm within post-viral syndromes*, which is currently under development, I note the relevance that the results of the research may offer. Its main objective is to evaluate the clinical similarities and/or differences, risk factors, pharmacological and non-pharmacological treatments, and stratification of post-exertional fatigue, cognitive impairment, and dysautonomia in patients with post-COVID-19 syndrome from the experience of other post-viral fatigue syndromes such as ME/CFS.

2.5.8 Status of biomarkers

There is no consensus on a biomarker for ME/CFS; however, there is substantial biological evidence that ME/CFS results in a widespread impairment of the immune system. The medical and scientific community accepts biomarkers with high sensitivity, specificity, and selectivity for detecting disease targets. They are effective and accessible. In addition, evidence such as high genetic and physiological reproducibility of results is paramount. The following is a summary of findings from publications addressing the search for biomarkers in ME/CFS. Study design, definition of outcome markers, quality and poten-

tial to define diagnostic biomarkers are heterogeneous. Due to the consistency observed over time, this review also supports using NK cells as an appropriate model to study the pathological mechanism of ME/CFS. The review is followed by a list of possible associations, and one of the findings is that these systems do not function independently and may be part of a complex integrative network; there is a need for a multidisciplinary approach to the study of these systems. Multidisciplinary research and uniform protocols are needed for ME/CFS biomarker research [70]. In this study, biomarkers are analyzed. A classification of those that I consider relevant is proposed.

Genetic biomarkers

- Sirt1/eNOS modulator microRNAs: miR-21, miR-34a, miR-92a, miR-126, miR-200c [71]. Expression miR-21, miR-17-5p, miR-10a, miR-103, miR-152, miR-146a, miR-106, miR-223 and miR-191. Significant values for NK cells [72].
- miRNAs (hsa-miR-127-3p, hsa-miR-142-5p and hsa-miR-143-3p that were differentially expressed and significantly increased in ME/CFS based on RT-qPCR ($p < 0.05$) [65].
- Increased gene expression of purinergic and cellular modulators (p= 0.04) and mediators of nociception and stress were positively associated with ME/CFS ($p = 0,02$) [73].
- Gene expression for the following: ASIC3, P2X4, and P2X5, for sensory receptors α-2A, β-1, β-2, and COMT and IS genes for IL10 and TLR4 was increased in ME/SFC up to 0.5 at 48 hours ($p < 0,05$) [74].
- ME/CFS patients had significantly lower target cell lysis by NK cells than HCs. Eleven SNPs were identified in TRP ion channel genes (TRPC4, TRPC2, TRPM3 and TRPM8). Five of these SNPs were associated with TRPM3 ($p < 0,05$). There were 14 SNPs associated with nicotinic and muscarinic genes, including CHRNA2, CHRNA3, CHRNB4, CHRNA5 and CHRNE ($p < 0,05$). There were six SNPs associated with CHRNA3. Sixteen genotypes were identified from the SNPs ($p < 0,05$) [75].
- Seventy-eight SNPs were identified in nicotinic and muscarinic acetylcholine receptor genes in B cells of ME/CFS patients compared to HC. Thirty-five of these SNPs were identified in mAChM3. Other SNPs identified were: nAChR delta, nAChR alpha 9, TRPV2, TRPM3, TRPM4, mAChRM3 2 and mAChRM5 ($p < 0,05$). From these SNPs, nine genotypes were identified ($p < 0,05$) [75].
- The mRNA levels were significantly different in ME/CFS patients compared to the control group in 12 genes. Quantitative real-time PCR validated these changes in nine of these genes. These genes encoded: granzyme in activated T cells or natural NK cells, energy regulators (ATP5J2, COX5B and DBI), proteasome subunits (PSMA3 and PSMA4), a putative inhibitor of protein kinase (HINT), GTPase (ARHC) and signal transducers and activators of transcription 5A (STAT5A) ($p < 0,05$). A follow-up investigation with microarrays showed that 9 genes effectively classified 79 % of patients with ME/CFS compared to others without ME/CFS experiencing fatigue. In contrast, real-time PCR was able to classify 94 % of ME/CFS patients effectively [76].

- Patients with ME/CFS had elevated mRNA in P2X4, TRPV1, CD14 and all adrenergic receptors compared to control groups after exercise ($p =< 0.05$). Patients with ME/CFS and comorbid fibromyalgia showed elevated mRNA in ASIC3 and P2X5 ($p < 0.05$) [77].

Immunological Biomarkers

- ME/CFS patients had significantly increased expression of IL-10, IFN-g, TNF-a, CD4+CD25+T cells, FoxP3, and VPACR2 compared to HC. The cytotoxic activity of NK cells, particularly CD56bright NK cells, was decreased in the patients with ME/CFS. Granzyme A and granzyme K expression was decreased, and perforin levels were elevated in ME/CFS patients. CD8+ T-cell cytotoxic activity and NK phenotypes were also decreased in patients ($P \leq 0,05$) [78].

- NKCC significantly decreased at all time points: baseline (T1), six months (T2), and 12 months (T3) in ME/CFS patients compared to HC ($p < 0,05$). Differences between each time point were also significantly different (p < 0, 001). ME/CFS patients had fewer CD56bright CD16-NK cells at T1 and T2 time points ($p < 0,05$). After mitogenic stimulation, there were significant increases i n IL- 10, IFN-y and TNF-a at T1. IL-10 and IL-17A were significantly decreased at T2. IL-2 increased in T3 in ME/CFS patients ($p < 0,05$) [72].

- A significant decrease of CD57 was observed in lymphocytes from ME/CFS patients. This decrease was significant in both the percentage of cells expressing CD57 (p = 0, 024) and the amount of CD57 molecule expressed per cell (p <0.001) in T lymphocytes. In non-T lymphocytes, only the amount of CD57 per cell significantly differed in ME/CFS patients compared to HC ($p = 0,007$) [79].

- The following cytokines were elevated in ME/CFS patients: LTα, IL-1α, IL-1β, IL-4, IL-5, IL-6 and IL-12 and the following cytokines were decreased in ME/CFS patients: IL-8, IL-13 and IL-15 ($p < 0,05$). The following cytokines showed good biomarker potential based on the area under the curve: IL- 5, LTα, IL-4 e IL-12 ($p < 0,05$) [80].

- NKCC was significantly lower in patients with ME/CFS (p < 0, 001). There was a significantly higher percentage of CD2+ lymphocytes positive for DPPIV/C26 in patients with ME/CFS ($p = 0,001$), although the levels of DPPIV/C26 expressed on T cells and NK cells were lower ($p < 0,001$) [81].

- MCP-1 was significantly increased in ME/CFS patients compared to HCs, while other immune markers such as IL-1β, Il-4, IL-6, TNF-α, TGF-β1, TGF-β2, TGF-β3, IL-10 e IL17 were significantly lower ($p < 0,01$) [82].

- Patients with ME/CFS had higher levels of anti-VEB-dUTPase antibodies ($p < 0,001$) and anti-humanUTPase antibodies ($p = 0,0241$) compared to the control groups [83]. In post-giardiasis ME/CFS patients, NK cell levels were significantly lower than in the control groups. There was a negative correlation between abdominal and fatigue symptoms and NK cell levels ($p < 0,05$) [84].

- Patients with moderate ME/CFS had a significant increase in CD8+CD45RA effector memory T cells, expression of the signaling lymphocyte activation molecule in NK cells, 2DL5A on CD4+ T

cells and BTLA4+ on central memory CD4+T cells, as well as a significant reduction of LFA-1 on central memory CD8+T cells, KLRG1 on total CD8+T cells, KLRG1 on naïve CD4+T cells, and CD56dimCD16-CD2+ and CD18+CD2+ NK cells ($p < 0,05$). Patients with severe ME/CFS had a significant increase of CD18+CD11c on CD56dimCD16-NK cell phenotype and a significant reduction of NKp46 on CD56brightCD16dimNK cells ($p < 0,05$) [85].

- CD62L iNKT expression increased over time in patients with moderate ME/CFS. CD56bright NK receptors differed significantly in patients with severe ME/CFS ($p = 0,004$). At 6 months, naïve CD8+ T cells ($p = 0,041$), CD8-CD4- ($p = 0,024$) and CD56-CD16- ($p = 0,030$) iNKT phenotypes,$\gamma\delta$2T cells ($p = 0,035$) and effector memory subsets were significantly increased in patients with severe ME/CFS. Patients with severe ME/CFS also had significantly reduced CD56-brightCD-16dim NKG2D, CD56dimCD16- KIR2DL2/DL3, CD94-CD11a-$\gamma\delta$1T and CD62L+CD11a-$\gamma\delta$1T cells at six months [86].

- IL-17A levels are significantly higher in ME/CFS patients compared to HC ($p < 0,01$) [73].

- The following cytokines were significantly higher in ME/CFS patients than in HCs: TNF-α, IL-12, IL-4, IL-12, IL-1β and IL-25. IL-17F and CXCL8 were significantly lower than in HCs ($p < 0,05$). IL-17F and CXCL8 were significantly lower compared to HC ($p < 0.05$) [87].. Serum IL-1 ($p < 0,001$), TNFα ($p < 0,001$), neopterin ($p < 0,001$), lysozyme ($p < 0,001$) and plasma PMN- elastase ($p = 0,007$) were significantly higher in ME/CFS patients than in HC [86].

- There was a significant association with the plasma proteome of immunoglobulin heavy variable region 3-23/30 and ME/CFS (p $<$ 0, 001). ME/CFS patients with coexisting IBS showed a significant association with immunoglobulin λ constant region 7 ($p < 0,001$), and ME/CFS patients without IBS showed an association with immunoglobulin κ variable region 3-11 ($p < 0,001$) [88].

- ME/CFS patients showed significantly lower values of regulatory T cells (CD4+CD25++(high) FOXP3+) and higher values of NKT-like cells (CD3+CD16+/-CD56+) compared to HCNKCD69 and NKCD56 were significantly higher and NKG2C were significantly lower in ME/CFS patients compared to HC ($p < 0,05$) [89].

- ME/CFS exhibited significantly lower antibody-dependent cell-mediated cytotoxicity (ADCC) compared to HC (AUC: 77 %, $p < 0,02$). Related relatives also showed significantly lower ADCC [90].

- Lower concentrations of sCD26 were only found in female patients with ME/CFS, and diagnostic adequacy could not be confirmed. However, an association was observed between low sCD26 concentrations and higher levels of autoantibodies against α1-adrenergic and M3-muscarinic acetylcholine receptors. In infection-triggered ME/CFS, sCD26 was positively associated with activated T cells (r: 0.153-0.316; p $<$0.05), liver enzymes, creatine kinase (r: 0.247; p :$<$ 0, 001) and lactate dehydrogenase (r: 0.173; p : 0, 014) and negatively with interleukin-1 β (r: -0,182; $p : 0,018$) [91].

- The subgroups of ME/CFS patients had higher levels of antibodies against microtubule-associated proteins 2 (MAP2) ($p = 0,03$) and single-stranded DNA ($p = 0,04$). However, no significantly elevated levels of autoantibodies were observed in the entire ME/CFS cohort [92].

Metabolomic biomarkers

- Glutamine ($p = 0,002$) and ornithine ($p < 0,05$) levels in the blood of ME/CFS patients were significantly lower than those of HCs. Correlation analysis showed associations with glucogenic aminoacids and metabolites involved in the urea cycle [93].
- The following mitochondrial proteins, aconitate hydratase ($p < 0.05$) and ATP synthase subunit beta ($p < 0,01$), were up-regulated in ME/CFS patients compared to control twins [94].
- Patients with ME/CFS had significantly reduced plasma levels of antioxidant cap- antioxidant ($p = 0,009$) and increased levels of lipoperoxide ($p = 0,021$), inflammatory cytokines, FGF21 and NT-proBNP ($p < 0,05$). Significant positive correlations were observed between NT- proBNP and IL-1b ($p = 0,04$) and IL6 ($p = 0,01$) compared to HC [95].
- Lymphocyte mortality rate, mitochondrial respiratory function and TORC1 activity showed a sensitivity of more than 90% for diagnostic differentiation of ME/CFS patients compared to HC ($p < 0,001$) [96].
- Patients with ME/CFS and IBS had more abundant unclassified *Alistipes* and less *Faecalibacterium* compared to HC ($p < 0,05$). Patients with ME/CFS without IBS had more unclassified Bacteroides and less *Bacteroides Vulgatus* ($p < 0,05$). Patients with ME/CFS had fewer metabolic pathways associated with unsaturated fatty acid biosynthesis and more atrazine degradation pathways, independent of IBS comorbidity ($p < 0,05$) [97].
- Patients with ME/CFS had significantly lower levels of phosphatidylcholine ($p = 0,017$), choline, and carnitine ($p = 0,017$) than the control groups. Patients with ME/CFS and IBS had significantly higher levels of triglycerides ($p = 0,004$) and ceramide ($p = 0,021$) [98].
- IL-17A ($p = 0,018$), la FABP-2 ($p = 0,002$), and 3-Hydroxykynurenine ($p = 0,037$) were increased in ME/CFS patients compared to those in the control group, whereas kynurenine ($p = 0,012$) and serotonin ($p = 0,045$) were lower in ME/CFS patients. Changes in kynurenine and 3- hydroxykinurenine were associated with higher kynurenic acid/kynurenine and 3- hydroxykinurenine/kynurenine ratios [99].

Endovascular biomarkers

- Patients with ME/CFS had significantly higher extracellular vesicle (EV) counts and EVs were also significantly smaller compared to the control group ($p < 0.05$). Blood creatine kinase (CK) was significantly lower compared to HC ($p < 0.05$). The zeta potential was significantly different in patients with ME/CFS compared to those in the control group, where ME/CFS patients had more negative values regardless of whether EVs were isolated in the presence or absence of proteinase κ ($p < 0,05$). Blood CK values and physical characteristics of plasma EVs (including count, size, and zeta potential) miRNAs were significantly associated with severe ME/CFS ($p < 0,05$) [100].
- Compared to the control group, the fraction enriched in extracellular vesicles (EVs) was significantly higher in ME/CFS patients ($p = 0,007$), and EVs were significantly smaller ($p = 0,014$) [101].

- The number of circulating EVs was significantly higher in ME/CFS patients than in the control group ($p < 0,001$). The number of circulating EVs correlated with serum c-reactive protein levels ($p = 0,442, p = 0,0007$). Variations in Talin-1, Filamin-A, and 14-4-4 family proteins were identified by proteomic analysis in ME/CFS patients ($p < 0,05$) [102].

- Significant differences were observed in three biological variables after exercise: M wave, thiobarbituric acid reactive substances (TBARS) and resting CD26-expression. These variables were correlated with each other. Health-related quality of life was negatively correlated with exercise-induced increase in TBARS ($r : 0.570, p < 0.001$) and positively correlated with CD26-expression ($r : 0.486, p < 0,01$). The pain component of the **SF-36** correlated negatively with CD26 expression ($r : 0.618, p < 0,001$). Increased TBARS and decreased M-wave were the highest, and CD26 expression level was the lowest in those who referred an infectious onset [103].

- Plasma neuropeptide Y (NPY) was significantly higher in ME/CFS patients compared to the control group ($p < 0.001$). NPY had significant associations with various subjective measures, including levels of perceived stress and depression ($p < 0,05$) [80].

- **Random Forest Analysis** (RFA) indicated that serum activin B ($p < 0.001$) and associated markers 24-hour urinary creatinine clearance ($p = 0,02$) and serum urea ($p = 0,002$) were significantly higher in ME/CFS patients compared to the control group [104].

- Serum dehydroepiandrosterone sulfate (DHEAS) was significantly lower in ME/CFS patients compared to those in the control group ($p = 0,003$). There was no significant difference in IGF1 or IGFBP3/IGF1 ratio [105].

- GDF15 levels in patients with severe ME/CFS were significantly elevated ($p = 0,01$) and positively correlated with fatigue scores ($p = 0,026$). Circulating levels of GDF15 were stable at two different time points (over seven months) in mild/moderate patients [106].

- Patients with severe ME/CFS had significantly lower mean serum creatine kinase values compared to patients with control group and non-severe ME/CFS ($p < 0,001$) [107].

- Male patients with ME/CFS had significantly higher levels of hexosyl-ceramides (HexCer), monounsaturated plasma phospholipids (PL) and saturated triglycerides compared to the control group ($p < 0,05$). In ME/CFS patients, the levels of total phosphatidylethanolamine (PE), PE with omega-6 arachidonic acid and total HexCer ($p < 0,05$). Oxylipins derived from omega-6 linoleic acid were significantly higher in male patients with ME/CFS compared to males with SC ($p < 0,05$). **Principal component analysis** showed that most PCs and some PE, PI, and SM species were negatively associated with headache severity and fatigue; this correlation was not sex-specific ($p < 0,05$). Lower correlations of oxylipins and ethanolamides were associated with headache, fatigue, and cognitive difficulties. This correlation was sex-dependent ($p < 0,05$) [108].

- The mean concentration of α-melanocyte-stimulating hormone was significantly higher in patients with ME/CFS compared with the control group ($p = 0,02$). There was a negative correlation between disease duration and plasma concentration of alpha-melanocyte-stimulating hormone ($p =$

$0,04$). In patients with diagnosis $\leq$ 5 years, there was a significant difference in α- melanocyte-stimulating hormone concentration compared to the control group ($p < 0,01$) [109].

- Endothelial dysfunction, as measured by flow-mediated dilatation and post-occlusive reactive hyperemia, was significantly lower in ME/CFS patients compared to the control group ($p = 0,005$ y $p = 0,003$, respectively) [110].

2.6 Design of studies on biomarkers

The following is an example of the design of some of the studies referenced. It is an example for most of those presented in list 2.5.8.

The first study indicated that the analysis of micro RNAs [71], in particular, MiRNA-21, which is associated with 16 diseases as diverse as some oncological or cardiac diseases [111]. It is also associated with autoimmune diseases. The study indicates that a sample size of n = 87 was analyzed, with the control group being a healthy population of n = 29, mostly men. This study shows that there is an overexpression of MiRNA-21 in this sample.

If MIR-127-3p is analyzed in [65] it is known to be associated with an increase in IL - 10 (a marker studied and referenced in the exposed relationship) which can lead to inflammatory responses or develop various lymphomas. The sample size is n = 40 with the control group of healthy participants n = 20 with no information other than age to analyze the participants' phenotype. MIR- 127-3p is also associated as a biomarker for frontotemporal dementia [112] or gastric cancer [113].

Finally, a study proposing NK as a biomarker in [78] is analyzed without wanting to extend it. The sample size is n = 145 with a control group of healthy patients n = 50. A significant gender difference is observed, with 70% of records in affected patients being female and only 57% in controls. The results show a lower number of NK in affected patients. It has been linked to several autoimmune diseases.

2.6.1 Conclusions

Definition of the design of the participants

The following limitations were detected in all the studies:

1. The control group is made up of healthy participants.
2. There is heterogeneity in the group of patients affected by ME/CFS; the sex ratio is not considered, and their clinical history is not reported.
3. Since the biomarkers cover a variety of diseases and do not provide a discriminating factor, the conclusions always have the limitation of confirming the initial hypothesis.

Discriminant design proposal

If the aim is to define biomarkers characteristic of the population affected by ME/CFS, consideration should be given to defining the groups to be compared in such a way that there is homogeneity in order to seek discriminating results.

For example, autoimmunity in ME/CFS and its characterization would be investigated. For this purpose, the target population would be defined as follows:

1. A homogeneous population is defined based on the clinical history of patients with ME/CFS who have answered the SF-36 questionnaire and which correspond to the group of patients with in-hospital stress test results in the Webber classification in groups C and D.
2. The control group is defined as patients with cardiovascular disease who are classified by a hospital stress test according to the Webber classification in groups C and D.

Thus all patients will have an autoimmune characterization, and the analysis should focus on:

- Compare autoimmune biomarkers.
- Compare other biomarkers.

In this way, we would expect common results in the analysis of immunological biomarkers with homogeneous patients (belonging to group C and D of Webber's classification) and we would look for other types of biomarkers to characterize the ME/CFS population.

If this type of analysis is performed by comparing different groups with diseases that respond to the same pattern, discriminating biomarkers would be detected to help define a characteristic phenotype of ME/CFS.

2.7 Long-term forecast

The prognosis for ME/CFS is variable. Patients may show some improvement during the first five years of the illness and usually stabilize at a level of health below their level of functioning before the illness. The majority of patients never return to their pre-morbid level of health or functioning. [114] [115].

In a review of 14 studies, the conclusion was that [116]:

- Five percent of patients recovered with a range [0 %-31 %)].
- Forty percent of patients improved during follow-up with a range [8%-63%].
- Between 8% and 30% returned to work.

- Between el 5% and 20% of patients reported a worsening of symptoms.

They may have mild disability (able to work normal hours or work from home), or moderate or severe disability (confined to bed) [117]. Patients often have good days (remissions) and bad days during the disease, called crises [118]. Some patients may have a gradual worsening of the disease. Risk factors for disease severity are [119] [120]:

- The severity of the disease at the time of onset.
- The standard of early disease management. Late diagnosis or overexertion in the early stages of the disease is likely to lead to physical deterioration of the patient.
- Having a mother with the disease.
- Comorbid diagnosis of fibromyalgia.

Death certificates for patients with ME/CFS are generally documented as if the patient died of some other co-morbid illness, since the cause of death is not listed as ME/CFS on the death certificate. Therefore, it is impossible to determine the mortality rate by reviewing death certificates because patients with ME/CFS are not reflected on death certificates. One study found suicide, heart disease, and cancer to be the most common causes for death in ME/CFS patients, and the average age at death from these causes was significantly lower than the national average. The overall death rate in people with ME/CFS was not significantly different from the standard death rate [121] [122].

2.8 Economic costs

ME/CFS is a debilitating chronic illness. It places an enormous burden on those affected and their caregivers and the healthcare system. It is estimated that between 836,000 to 2.5 million Americans suffer from ME/CFS, resulting in an annual financial cost ranging from 17 to 24 billion US$ per year. Unemployment rates among those with the disease range from 35% to 69%. Individual income losses are approximately $20,000 per year per household [33].

The economic costs to patients, their families and society are enormous. Annual direct medical costs per patient with ME/CFS range from $2,342 in a community-based sample (previously undiagnosed) to $8,675 in a tertiary sample (previously diagnosed) [123]. Seventy-three percent of those affected were unable to work [124]. In another study, 25% were bedridden [125].

The 2005 and 2010 Canadian Community Health Survey documented that patients with ME/CFS are significantly impaired compared to Canadians with other chronic diseases such as cancer and heart disease. Patients with ME/CFS reported high levels of permanent disability, a need for help with activities of daily living, and a high number of medical consultations (more than 10 per year). A substantial proportion reports a loss of income and productivity of $20,000 per patient, with many reporting an annual household income of less than $15,000 [119] [123].

Why do physicians struggle to diagnose ME/CFS when it is presented to them? A first hypothesis would be inadequate training: Most medical schools do not include ME/CFS in their undergraduate programs. Medical textbooks are not up to date in this area of medicine. Physicians, both general practitioners and specialists, have not been taught about ME/CFS in a formal and systematic way, as was done in the past for emerging diseases such as HIV/AIDS. As a result of the lack of professional training in ME/CFS, it takes many years for patients to receive a diagnosis. There is a need for information, visibility, and awareness campaigns about the syndrome.

2.9 The treatments

In the absence of a specific treatment for the condition, the primary role of the clinician is to confirm the diagnosis, educate the patient about the importance of avoiding overexertion and mental stress, "pacemaker" therapy, and the use of medications to manage symptoms as necessary and appropriate for the patient. It is important to follow up periodically, to establish follow-up times, as well as to monitor progress and consider the possible development of new diagnoses and comorbidities. "Pacing" is the interruption of physical or mental activity with periods of rest before significant levels of fatigue or symptom exacerbation are reached or expected after the patient has been diagnosed with exertion. As a general rule of thumb, although flexibility should be exercised to reflect the particular circumstances, needs, and conditions of each patient, it is recommended that activity be maintained at two-thirds of the duration and intensity expected (based on past experience) to produce post-exertional symptoms [24].

Recent studies suggest that cognitive behavioral therapy may have a role in the treatment of ME/CFS. It may have long-term benefits in chronic fatigue [126], but with little evidence, it should be used with caution to avoid frustration [127]. It should be kept in mind that this i s supportive and not curative therapy [128].

American recommendations for palliation of ME/CFS symptoms have been published during 2021 [129]. Some disease experts are evaluating the selective use of antivirals and immunomodulators in ME/CFS and have seen promising results in some patients. Specialized consultation may be helpful in identifying and treating these or other aspects of the disease. Clinicians should also be aware of any potential impact of excipients in drugs. The following are symptoms that are recommended for medical management [129]:

- Drugs for orthostatic intolerance
- Drugs to improve sleep.
- Drugs for cognitive impairment and fatigue
- Drugs for pain.
- Drugs for immune dysfunction.
- Drugs under suspicion of small intestinal bacterial overgrowth.

2.10 Open clinical trials and research lines

The following is a list of studies in progress in which I participate as a member of the multidisciplinary team:

1. Pivotal, single-center, randomized, double-blind, comparative clinical trial to assess the improvement of fatigue, pain, quality of life, sleep, depression-anxiety and neurovegetative alteration after the administration of melatonin and zinc versus placebo in patients with chronic fatigue syndrome associated with COVID19 infection (CFIDS). Funding source: Fundacion Mutua Madrileña.
2. Study of the neuropsychological profile in adult patients with Chronic Fatigue Syndrome using the computerized software CPT-3 (Conners Continuous Performance Test).
3. Longitudinal study of the microbiota in adult patients with chronic fatigue syndrome. Project code PR(AG)257/2021.
4. Study of physical exercise intolerance (PEI) in adult patients with chronic fatigue syndrome (CFS) by performing a two-day cardiopulmonary exercise stress test (CPET). Project code PR-AG9 01/2020.
5. Characterization of post-exertional fatigue, cognitive dysfunction and dysautonomia in post-COVID-19 syndrome: a new emerging paradigm within post-viral syndromes.
6. Pivotal, single-center, randomized, double-blind, comparative clinical trial to assess the improvement of fatigue, pain, quality of life, sleep, depression-anxiety and neurovegetative alteration after the administration of melatonin and zinc versus placebo in patients with chronic fatigue syndrome.

Chapter 3

Patient management in the specialized unit of Vall d'Hebron Hospital

3.1 Protocols of diagnosis

The accredited chronic fatigue unit of the Vall d'Hebrón Hospital in Barcelona, led by Dr. José Alegre Martín, has maintained a registry of patients diagnosed with ME/CFS from January 2008 to date. The study population are patients older than 18 years and meet the inclusion criteria, which coincide with those of diagnosis of ME/CFS according to the criteria of Fukuda 1994 [130] and Carruthers 2003 [35], Chronic Fatigue (CF) associated with cancer survivor [131], CF associated with immuno-inflammatory diseases [36], CF associated with other infectious diseases [132] and idiopathic CF [133].

Each record contains information on socio-demographics (age, sex, marital status, profession and educational level) and clinical variables, such as the family history of CFS, FM, auto-immune diseases, rheumatologic diseases, thyroid diseases, characteristics of fatigue including the apparent trigger-physical trauma, intoxication, stressful life event, surgical intervention, infectious process, transfusion, pregnancy-childbirth, bariatric surgery, not specified), a form of onset of fatigue (sudden, gradual, insidious), its evolutionary course (improvement and/or worsening), and the age of onset and time of evolution.

The presence of non-restorative sleep, recurrent headache, and muscular, cognitive, neurological, neurovegetative and immunological symptomatology was recorded [28]. Also somatic comorbid phenomena, such as fibromyalgia (FM), temporomandibular joint syndrome, dry syndrome, thyroiditis and hypothyroidism, endometriosis, multiple chemical hypersensitivities, ligamentous hyperlaxity, shoulder tendinopathy, degenerative and/or mechanical vertebral disease, plantar fasciitis, epicondylitis, shoulder carpal tunnel, dyslipidemia, and hypovitaminosis D [134]. And comorbid psychological phenomena, such as generalized anxiety, depression, panic attacks, dysthymia, personality, and adaptive disorder.

The pharmacological treatment prescribed for each patient was also recorded, specifically: antidepressants (tricyclics, duals, SSRIs), anticonvulsants, opioids (minor, major), NSAIDs, analgesics, anxiolytics,

whether they underwent cognitive-behavioral therapy and programmed physical exercise therapy, in addition to a multidisciplinary treatment program.

In the study of the ME/CFS patient, after diagnosis and assessment of comorbid phenomena, it is essential to quantify and evaluate fatigue, quality of life or anxiety and/or depression psychopathology using a battery of clinically self-administered questionnaires. There are currently few specialized ME/CFS units in the world, which also have a relatively low number of well-documented cases and a lack of publicly available data compared to other disorders.

As already discussed, there are no commercially available diagnostic tests, specific laboratory biomarkers, or specific drugs approved by the Federal Drug Association (FDA) for ME/CFS syndrome [134]. Therefore, each subject to be diagnosed with ME/CFS must undergo an evaluation according to Fukuda criteria and a procedure that each unit has established using validated self-administered questionnaire batteries. The self-administered questionnaires used by this unit are:

- The fatigue impact scale FIS40 [135] and FIS8 [136].
- Sleep quality using the PSQI questionnaire [137].
- General quality of life such as the Short Form Health Survey (SF-36) questionnaire [138].
- The Symptom Checklist-90-revised psychological inventory (SCL 90 R) [139].
- The Hospital Anxiety and Depression Scale (HAD) [140].

No consensus exists on the number and type of questionnaires, so not every unit has the same number of questionnaires per person. As a result, it is complex for ME/CFS units to have a large number of records of the questionnaires that are needed to efficiently carry out large longitudinal and multi-center studies of patients with this pathology using the latest advances in data analysis, such as **machine learning** techniques [141].

3.2 Use of questionnaires

There have been positive evaluations of the usefulness of using general health status questionnaires in randomized trials. Whether or not the results are consistent with the primary outcomes, it would be important to collect and report such data systematically [142]. The SF-36 was originally developed as a general health survey instrument. It has been translated into over 50 countries as part of the International Quality of Life Assessment Project and has become the most widely validated and used generic QOL (quality of life) instrument. It has broad applications for population health surveys, comparing relative disease burdens, and distinguishing between the health benefits of different interventions [143].

The SF-36, which is not specific to fatigue, was found to be the most widely used measure in a review of randomized clinical trials in ME/CFS. The SF-36, especially the physical functioning subscale, has been consistently used as the primary efficacy measure, often in combination with other specialized

fatigue measures such as the Checklist Individual Strength (CIS) or the FIS-40. Questionnaires are further discussed in the scientific literature in Section 5.1.

Chapter 4

ME/CFS population-based dataset

4.1 Definition

On October 18, 2006, the research project entitled "Population-based registry of patients with chronic fatigue syndrome," presented by Dr. José Alegre Martín, was submitted to the clinical research ethics committee and the Vall d'Hebrón Hospital research commission in Barcelona. After its approval, the cases diagnosed with ME/CFS have been recorded by the specialized unit of the same hospital. The following is a descriptive analysis of the variables collected. The results are updated as of November 2023. The patients come predominantly from Catalonia (89%) and the rest come from other autonomous communities. This analysis updates the statistical analysis of 824 cases from 2011 [144] of the same project.

4.2 Statistical analysis

This prospective cross-sectional study included 89.3% women and 10.7% men. Data from the SF-36, HAD, FIS8, FIS40, FIS40, SCL 90 R and PSQI questionnaires have been obtained and recorded from 2008 to 2023. See the final records in Table 4.1. Patients were eligible if they were 18 years of age or older and had a confirmed diagnosis of ME/CFS according to the 1994 CDC/Fukuda definition [130] and provided signed written informed consent and ethics committee approval. These inclusion criteria have been updated as detailed in section 3.1.

Sixty percent of the patients reported that they were not employed. Twenty-six percent are working and 14% have never worked. There is an increasing percentage of ME/CFS patients who end up with permanent and absolute disability.

The following is a summary of the records of the questionnaires used in the study detailed in Table 4.1. Throughout the research presented in this paper, these datasets were used. The summary statistics

are relevant for several purposes:

- There are no databases with a similar volume. The results presented below are of interest to corroborate or qualify the conclusions of studies based on meta-analysis of the scientific literature.
- Information on more than 100 variables is collected from each patient. It is still in the analysis phase, and several works in development are expected to shed light on the syndrome.

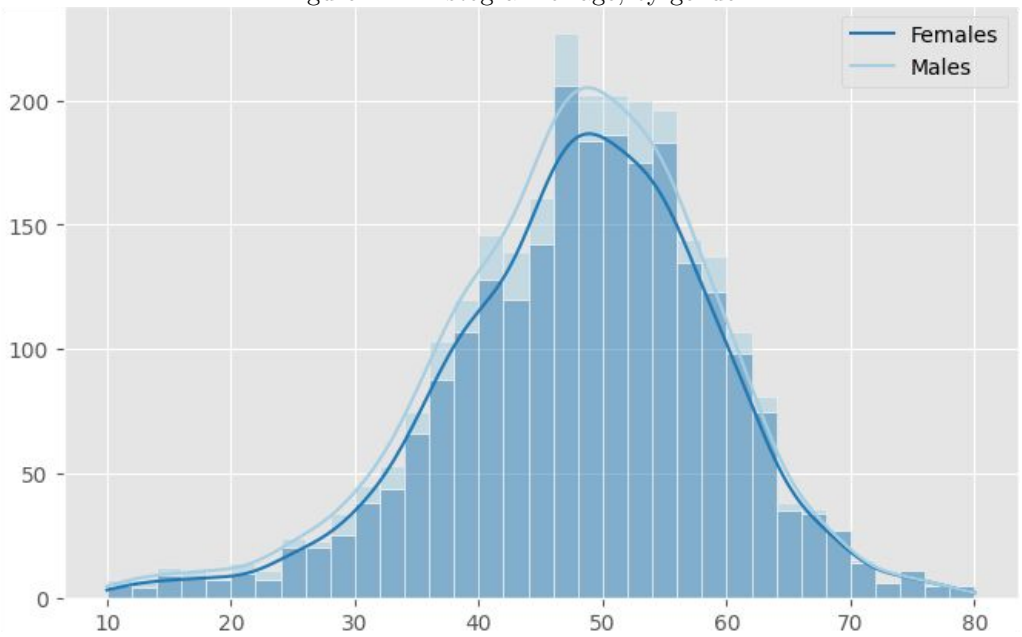
Table 4.1: Statistical summary of questionnaires collected

Questionnaire	Records	Questions	Subscales	Total Value	Response Range
SF-36	2.346	36	10	NO	{1,2,3,4,5,6}
HAD	2.339	14	2	SI	{0,1,2,3}
FIS8	2.057	8	0	SI	{0,1,2,3,4}
FIS40	2.362	40	3	SI	{0,1,2,3,4}
SCL 90 R	2.361	90	12	NO	{0,1,2,3,4}
PSQI	1.959	34	7	SI	{0,1,2,3}

Data available for each questionnaire and characteristics of the questionnaires. From left to right, the title of each column means ***Records****: number of forms available.* ***Questions****: number of questions per questionnaire.* ***Subscales****: number of subscales defined. Usually, a total summary value would be composed of the sum of the subscales in the HAD, FIS40, and PSQI questionnaires. In the case of SF-36, "global" values are created using mathematical formulas.* ***Total value****: The questionnaire offers a single summary value.* ***Response range****: Possible response values for each question.*

The study's relevant statistical data are presented below. Graphs are used as much as possible to help us understand the information and facilitate attention to the relevant information. The figure caption expands the information in the graph.

Figure 4.1: Histogram of age, by gender

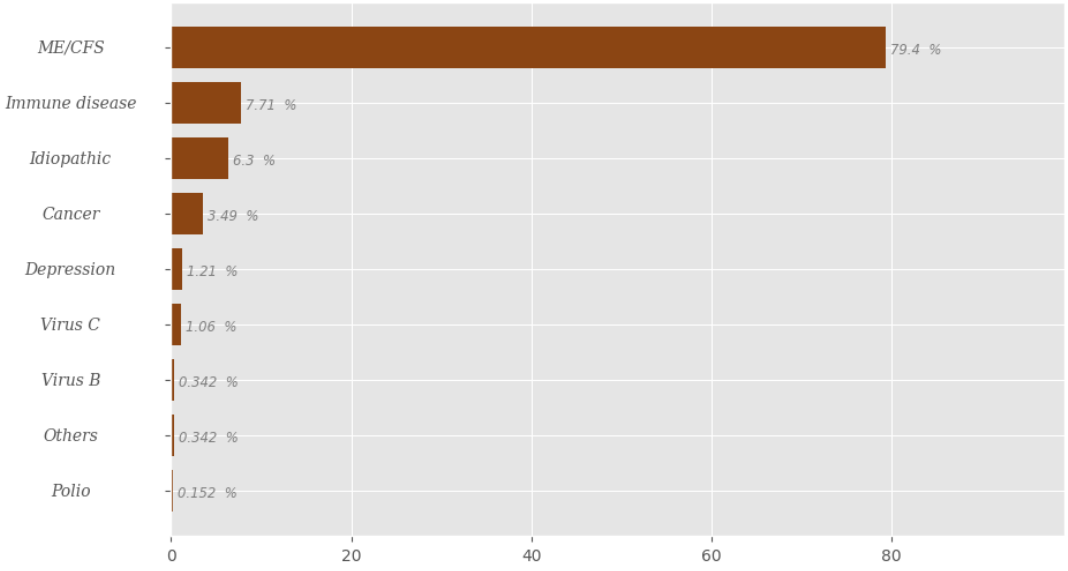


Histogram of age, taking into account that 89.3 % of the sample are women and the remaining 10.7 % are men. The age distribution in both sexes does not show significant statistical differences being the mean and standard deviation data for men and women respectively $\mu = 48.20$, $\sigma = 10.77$ *y* $\mu = 43.77$, $\sigma = 11.99$.

Figure 4.2 shows the details of the diagnoses of the patients who followed the inclusion criteria of the registry, and which forms associated with immunological diseases are noted in detail:

- Sjögren's syndrome ($n = 5$), multiple sclerosis ($n = 68$), inflammatory bowel disease ($n = 82$), systemic lupus erythematosus ($n = 28$), rheumatoid arthritis and other inflammatory rheumatic diseases ($n = 52$) and 6 patients affected by idiopathic thrombopenic purpura.
- In cancer survivors, breast ($n = 88$), prostate ($n = 10$), colon ($n = 8$), lymphoma ($n = 21$) and 9 other types.
- In those related to chronic infections 13 hepatitis B virus, 18 hepatitis C, 6 human immunodeficiency virus and 12 poliomyelitis.
- Of the psychiatric illnesses, bipolar major depression was reported in 35 cases and psychotic major depression in 10 cases.
- Of the remainder, heart failure (n = 9), chronic obstructive pulmonary disease (n = 12) and two cases of paroxysmal nocturnal hemoglobinuria were reported.

Figure 4.2: Histogram according to diagnosis



Diagnoses of patients who have followed the inclusion criteria of the registry and that forms associated with immunological diseases are noted.

Figure 4.3 shows the highest level of education at the time of diagnosis and interprets non-compulsory secondary education as the High School and intermediate and higher training cycles or previously Vocational Training.

Figure 4.3: Histogram according to academic education

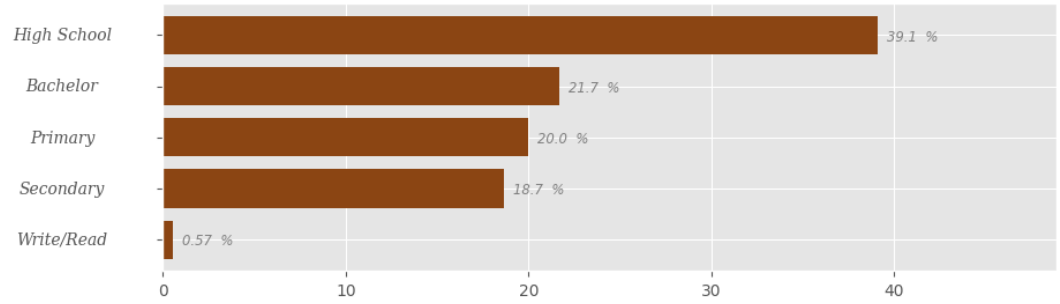


figure 4.4 shows information on the type of work performed at the time of registration. Specialized work requires more academic training. In this classification, half of the patients recognize their profession as non-specialized, considering the rest of the options.

Figure 4.4: Histogram by profession

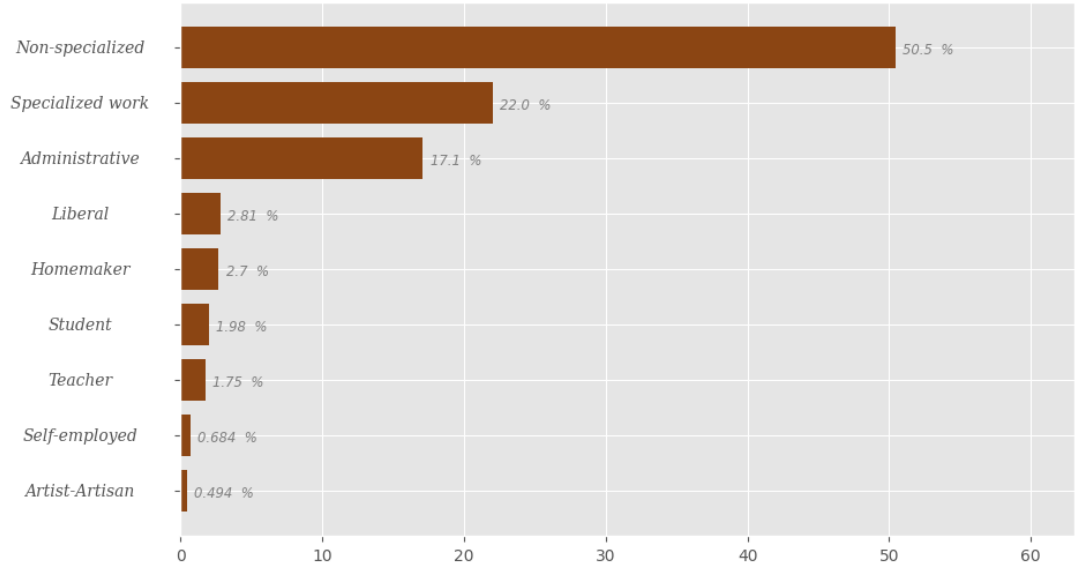
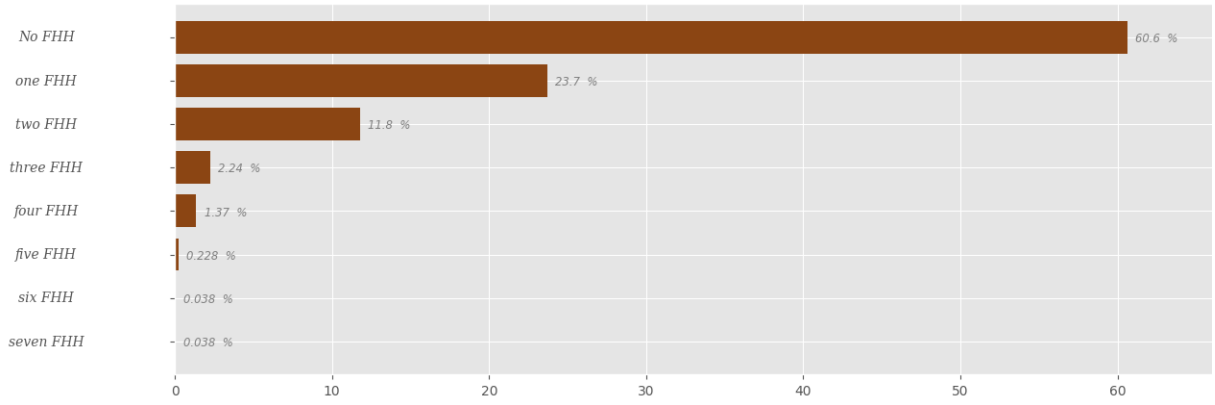


figure 4.5 shows whether the patient reports an FHH and, if so, the number. In this case, several relatives up to the second degree of consanguinity may suffer from one of the diseases detailed in Figure 4.6. Information is collected on whether any FHH is known in the following cases:

- According to the degree of consanguinity: it is recorded in father or mother, children, and siblings.
- Depending on the background, the following are asked: ME/CFS, Fibromyalgia (FM), Immune diseases, rheumatologic diseases, and thyroid diseases.

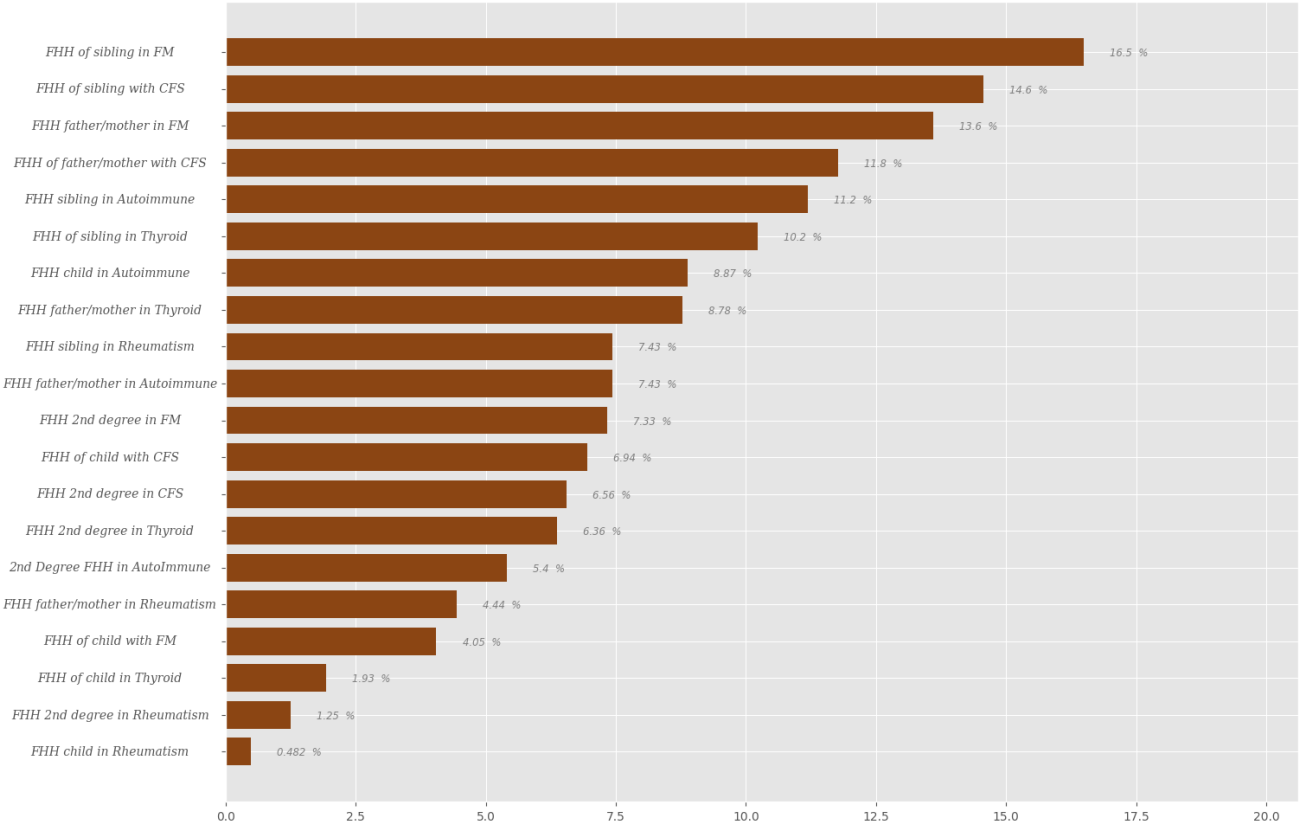
Figure 4.5: Family Health History (FHH



Forty percent of the patients state that they know of a close relative who has suffered or is suffering from one of the diseases. One refers to knowing of a case, and so on.

Figure 4.6 analyzes the results of 40% of patients who report a FHH of one of the described types. Thus, when it is indicated that 16.5 %, it is with respect to 40 % of those with a known FHH.

Figure 4.6: Family history by type.



Family health history (FHH) is recorded according to 5 diseases and according to the degree of co-blood relationship. From the prefixes, ***FHH****:* ***SFC:****FHH of ME/CFS.* ***FM****: FHH in fibromyalgia.* ***Autoinmune****: FHH of autoimmune diseases.* ***Rheumatism****: FHH in rheumatologic diseases.* ***Tiroides****: FHH in thyroid diseases. The degree of family consanguinity is registered '****father****': Family ancestor of the father or mother, and successively, '****brother****', '****2ndDegree****'.*

The values represent the percentage of patients who report some type of FHH. The first row should be read: FHH of sibling in fibromyalgia represents 16.5% of the total FHH recorded. For a correct interpretation of Figure 4.6, it should be taken into account that of the 40% of the records where a FHH has been noted.

Of those with a FHH:

- Forty percent are recorded as CFS.
- Forty-one percent are registered as FM.
- Thirty-three percent are recorded as autoimmune disease.

- 14% are recorded for rheumatologic disease.
- Thyroid disease accounts for 19%.

These data are analyzed with a **recommendation algorithm** [145]. This involves analyzing the frequency and its relationship with that associated with the previous data, from which the first rows of the result are incorporated in Table 4.2.

Table 4.2: Results of the recommendation algorithm for family history data.

Antecedent	Consequent	Frec1	Frec2	*Support*	*Confidence*
FHHfm'	FHHsfc_child'	1.6%	2.7%	1.4%	85.7%
FHHsfc_fathers'	FHHfm_fathers'	4.6%	5.4%	3.5%	75.4%
FHHsfc_silibing'	FHHfm_silibing'	5.7%	6.5%	3.8%	66.9%
FHHfm_fathers'	FHHsfc_fathers'	5.4%	4.6%	3.5%	65.2%
FHHsfc_2ndDegree'	FHHfm_2ndDegree'	2.6%	2.9%	1.6%	63.2%
FHHfm_silibing'	FHHsfc_silibing'	6.5%	5.7%	3.8%	59.1%
FHHfm_2ndDegree'	FHHsfc_2ndDegree'	2.9%	2.6%	1.6%	56.6%
FHHsfc_child'	FHHfm_child'	2.7%	1.6%	1.4%	50.0%

Explanation of variables. ***Antecedent****: The detailed antecedent occurs.* ***Consequent****: When the consequent occurs, knowing that the antecedent has occurred.* ***Frec1****: Probability (prevalence) that the antecedent occurs* ***Frec 2****: Probability (prevalence) that the consequent occurs.* ***Support****: Probability that the antecedent will occur simultaneously with the consequent.* ***Confidence****: The conditional probability that if the antecedent occurs, the consequent occurs. For a value of Confidence = 1 it means that whenever the first occurs, the second occurs.*

The reading of Table 4.2 follows, in its first row:

- FHHfm child: FHH of the child (FHH) suffering from fibromyalgia (fm). In the registry it occurs 1.6% of the time.
- FHHsfc child: FHH of the child (FHH) suffering from ME/CFS. In the registry, it occurs 2.7% of the time.
- The occurrence of both happens 1.4% of the time, but when the antecedent happens, 85.7% of the time the consequent happens.

Considering that the records analyzed are 40% of the total, it follows that FHH is usually associated with ME/CFS and FM.

Figure 4.7 records the personal history of the patient at the time of diagnosis of ME/CFS, whether he/she has been previously diagnosed with episodes of fatigue, pain, autoimmune diseases, or has been

diagnosed with any of the psychological pathologies described. Fatigue and, to a lesser extent, pain are the most common episodes.

Figure 4.7: Personal history

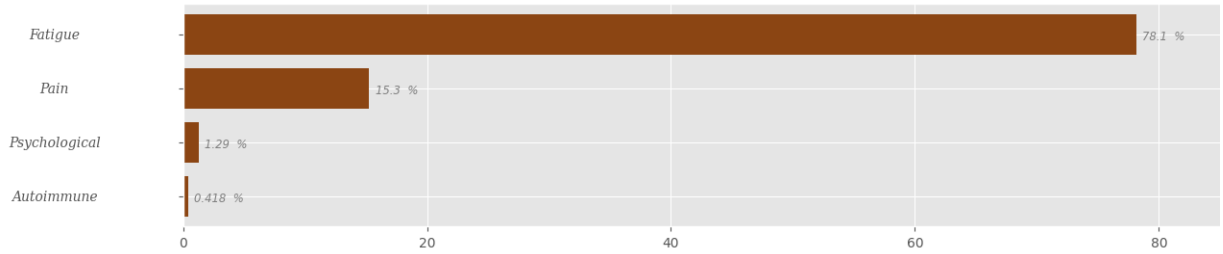


figure 4.8 shows graphically the records related to the time of fatigue detection, both the onset and other relevant data.

Among patients with gradual onset (accounting for 78% of cases), 89% reported worsening during the course of the disease. As for the apparent trigger, the percentages were 19% stressful life event, 11% infection, 10% childbirth, and 9% physical trauma or surgery.

Figure 4.8: Disease onset form

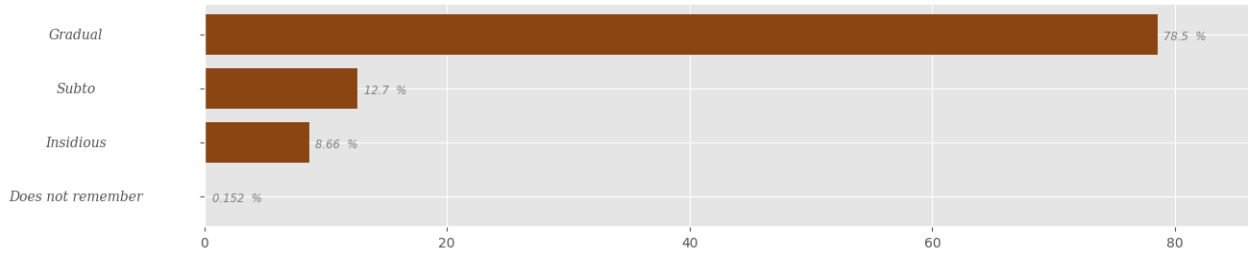
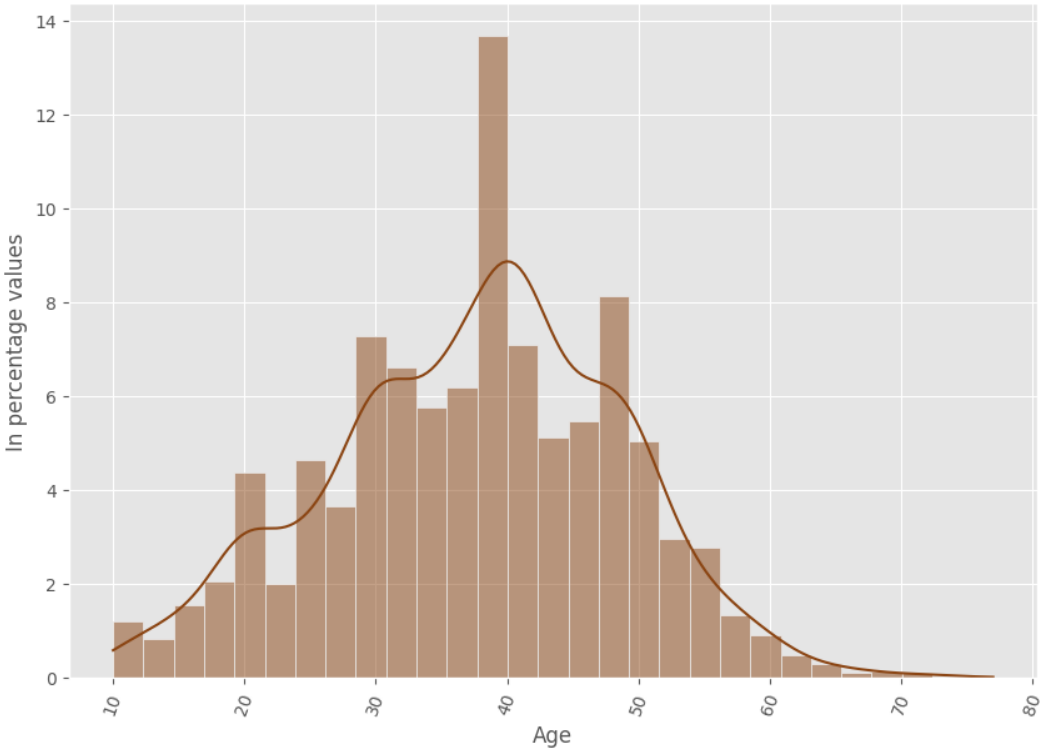


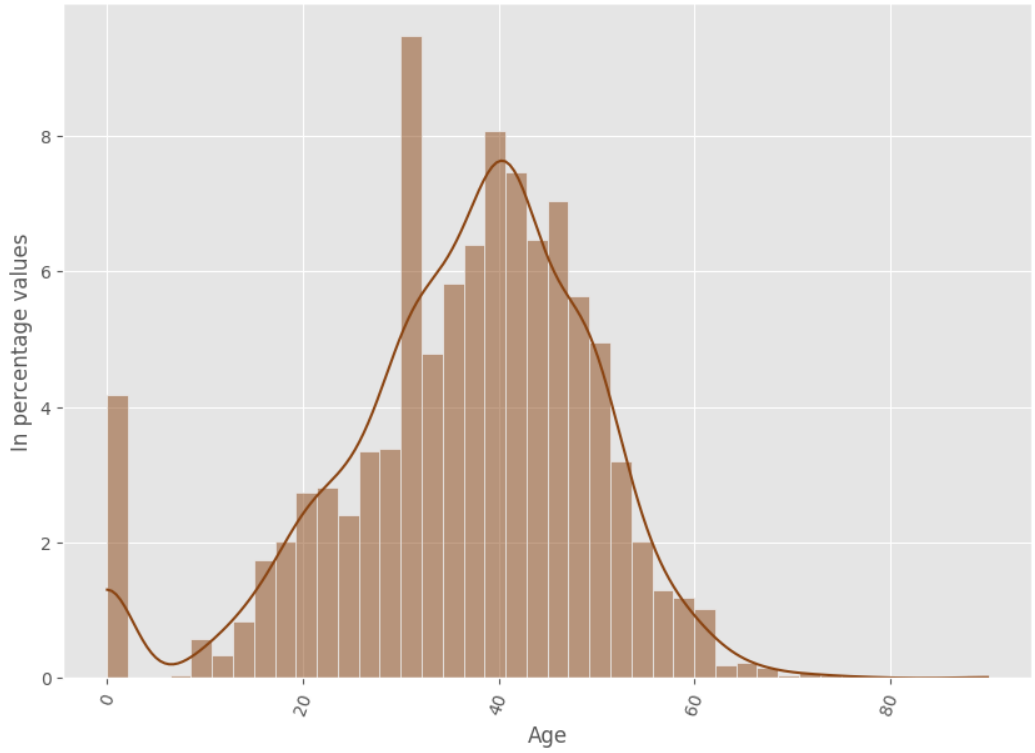
Figure 4.7, fatigue and pain have been common antecedents in patients diagnosed with ME/CFS. These aspects, such as the age of onset, time of evolution, and intensity, are analyzed. It should be noted that the mean evolution of fatigue and pain was 10 years at the time of diagnosis.

Figure 4.9: Age at onset of fatigue



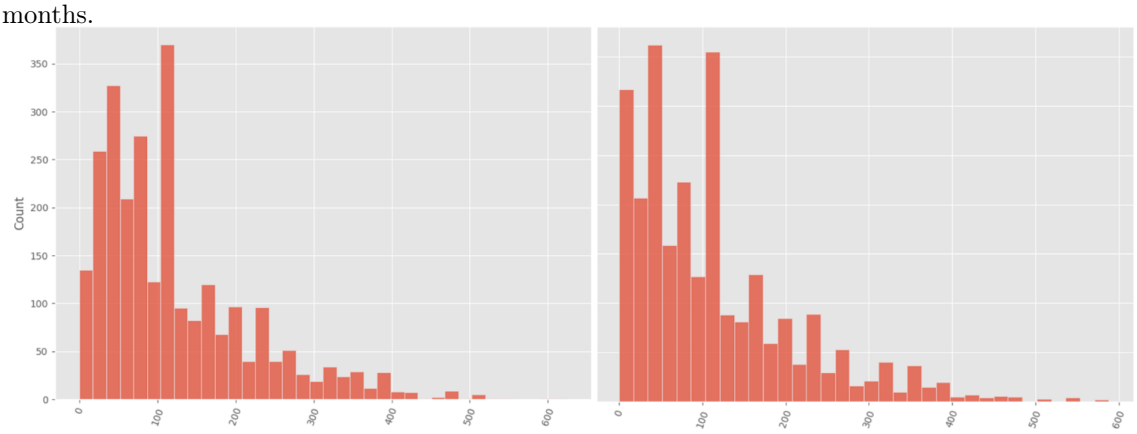
The mean age of fatigue onset was $\mu = 37.5$ *years with a standard deviation* $\sigma = 11.13$.

Figure 4.10: Age at onset of pain



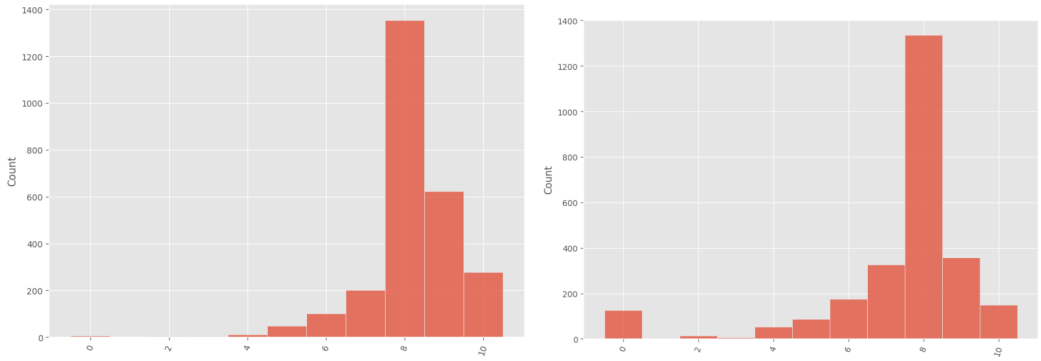
The mean age at onset of pain was $\mu = 36.4$ *years with a standard deviation* $\sigma = 13.3$.

Figure 4.11: Histograms of fatigue evolution time (left) and average pain evolution time (right) measured in months.



The mean time of evolution of fatigue was $\mu = 121.5$ *5 months and* $\sigma = 97$ *and the mean time of evolution of pain was* $\mu = 113.6$ *and* $\sigma = 98$. *A similarity is observed in both statistics, and both the time of evolution of fatigue and pain show similar* μ *y* σ *values.*

Figure 4.12: Histograms of fatigue and pain values according to the visual analog scale (VAS



The Visual Analog Scale (VAS) measures the intensity of pain or fatigue felt by the patient in a way that is reproducible to the investigator. For zero values, it indicates no or less intensity, and for values close to ten, it indicates greater intensity. The patient scores the intensity, which is measured in centimeters or millimeters. In both cases, the histograms are very similar, with values close to 8.

Figure 4.13 shows the variables analyzed on the symptomatology referred to the type of sleep, of which the following were collected when asked if the sleep was not restful:

- If the sleep was superficial, if there were frequent nightmares, if there was insomnia, hypersomnia, daytime sleepiness (somnolence).
- The associated phenomena were also recorded: restless legs syndrome (legs), sleep paralysis syndrome (sleep paralysis), snoring and sleep apnea syndrome. The number in percentage is expressed as a percentage of the total number of symptomatology records.

Figure 4.13: Relationship of sleep symptomatology

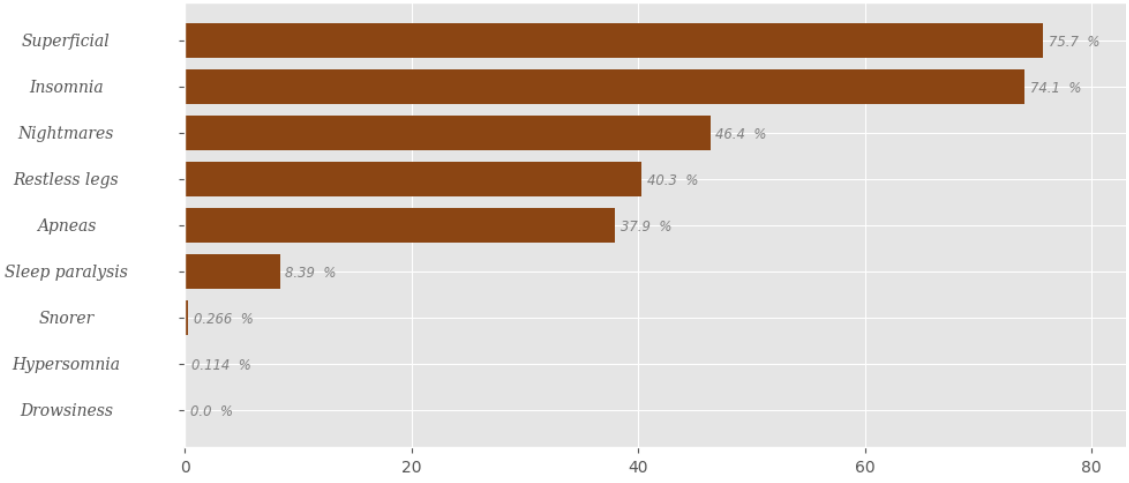


Figure 4.14 shows the descriptive statistics of muscle symptomatology, which included generalized pain (Generalized), muscle weakness (Weak), post-exercise fatigue (Post-exercise), difficulty in fine movements

(Movements), marked muscle contractures (Contractures), myofascial syndrome (Myofascial), myoclonic (Myoclonic) and falls due to loss of tone (Falls). The values express the percentage of registered patients reporting these symptoms.

Figure 4.14: Muscle symptomatology relationship

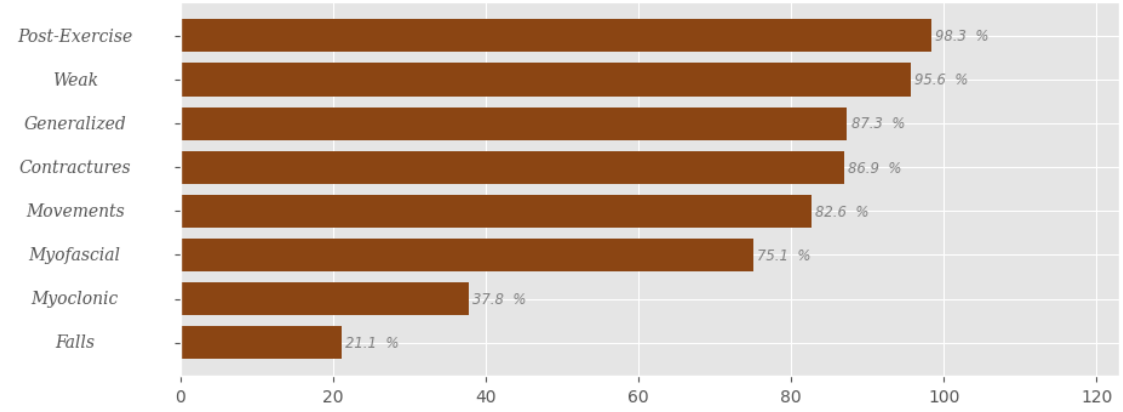


Figure 4.15 shows the descriptive statistics on cognitive symptomatology. The following variables are recorded: impaired concentration (concentration), impaired recent memory (memory), impaired task planning (tasks), impaired calculation (calculation), reading difficulty (reading), confusion with frequent forgetfulness (forgetfulness), temporospatial disorientation (spatial), episodes of nominal aphasia (aphasia), auditory-visual agnosia (visual). Values are expressed as a percentage of the total number of records.

Figure 4.15: Relationship of cognitive symptomatology

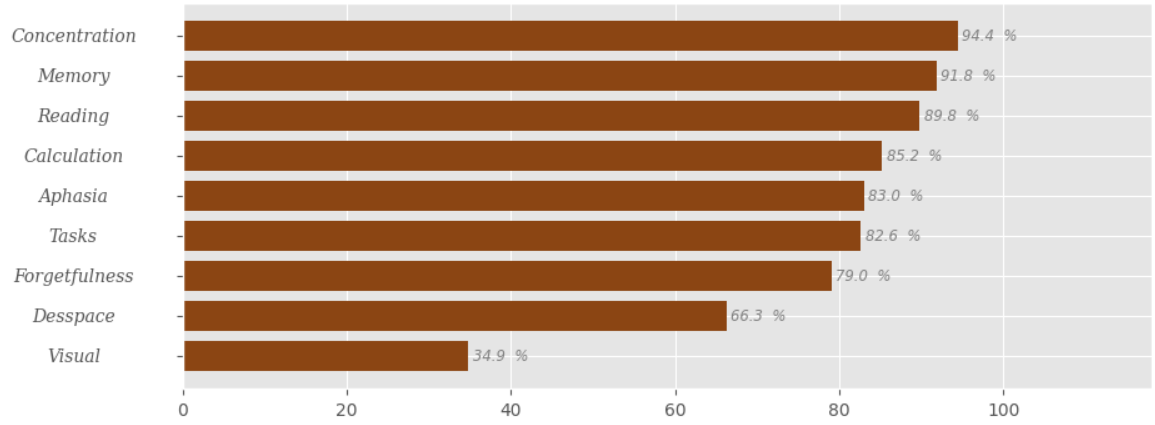


Figure 4.16 shows the descriptive statistics on neurological symptomatology. The following are recorded:

The existence of ataxia and/or dyssymmetry (ataxia), sensory hypersensitivity (sensory), and, in the previous affirmative case, if the cause is light (light), noise (noise), or odors (odor). Visual disturbances (visual) were also recorded, and it is specified whether it is blurred vision (vision), light spots (luminance),

or amaurosis (amaurosis). Finally, motor incoordination with or without falls (falls) is recorded. The values show the incidence in the registration sample, but in the case of value specificity (as in the case of sensory hypersensitivity and visual disturbances) the extent of each of them is recorded.

Figure 4.16: Neurological symptomatology relation

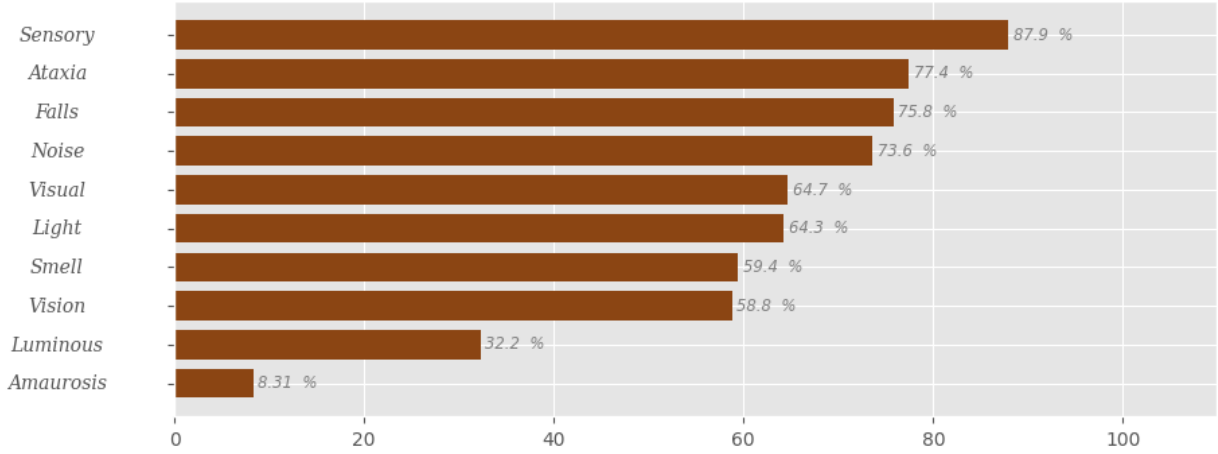


Figure 4.17 shows the statistics on neurovegetative symptomatology. The following symptoms are collected:

Dizziness or head instability (dizziness), vertigo (vertigo), episodes of orthostatic hypotension (orthostatic hypotension), lipothymias, syncope, frequent palpitations (palpitation), tremors (tremor), profuse sweating (sweating), altered bowel rhythm (bowel), altered urination (urination), decreased libido/anorgasmia/impotence (libido), visual accommodation difficulties (visual). Values indicate the incidence of each symptom in relation to the sample.

Figure 4.17: Relationship of neurovegetative symptomatology

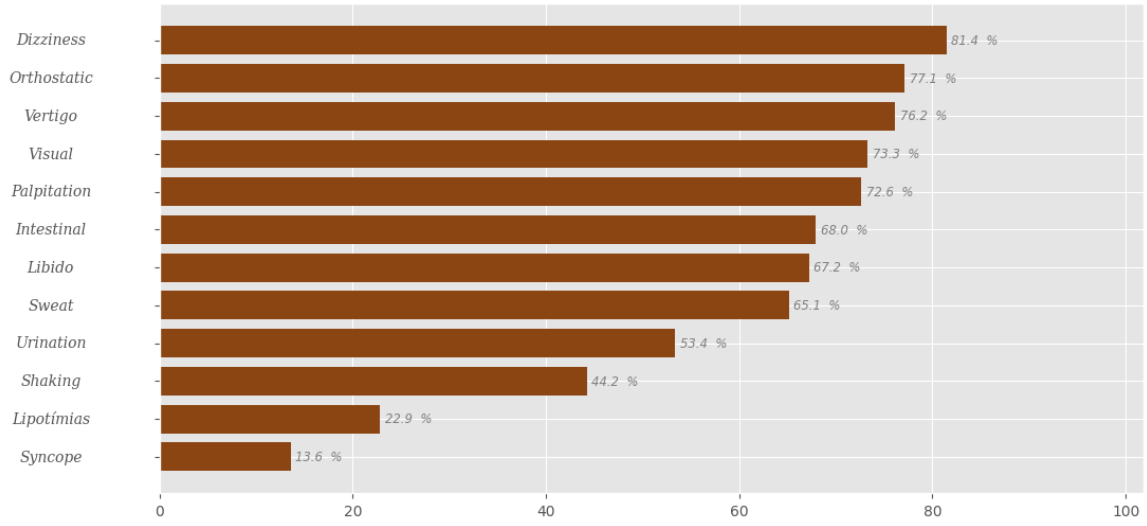
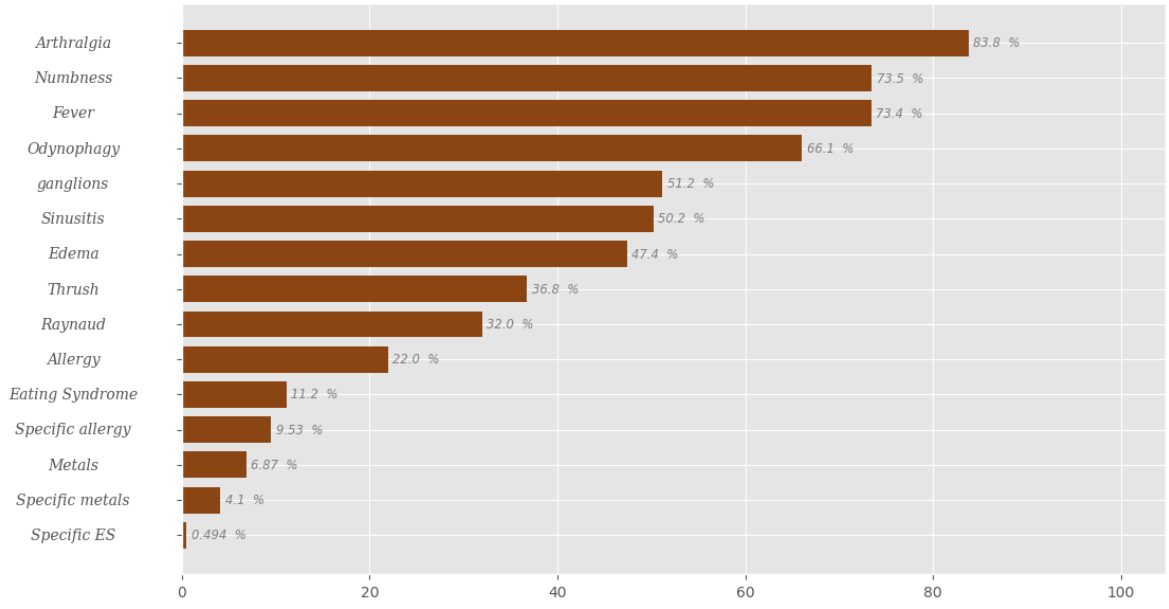


figure 4.18 shows the descriptive statistics on immunological symptoms. The following are collected:

Recurrent febrile fever, recurrent odynophagia (odynophagia), painful lymph nodes, Raynaud phenomenon (Raynaud) sensation of numbness and coldness in some areas of the body, such as fingers and toes, in response to cold temperatures or stress, generalized morning numbness (numbness), migratory arthralgias (arthralgia), multiple drug allergies (allergy), food intolerance (Eating Syndrome), multiple metal allergies (metal) and if collected if the patient specifies the type of metal, history of sinusitis (sinusitis), facial edema, cold sores (thrush), herpes and candida. The values refer to the incidence of these symptoms in the total sample.

Figure 4.18: Relationship of immunological symptomatology



Of the comorbidities, data on fibromyalgia, regional myofascial syndrome, degenerative spine disease, shoulder tendinopathy, epicondylitis, trochanteritis, patellar chondropathy, plantar fasciitis, carpal tunnel syndrome, ligamentous hyperlaxity were collected in Figure 4.19. Data on vitamin D and bone metabolism were collected on vitamin D depletion, osteopenia, and osteoporosis. Metabolic syndrome and vascular risk data were collected on hypercholesterolemia, hypertriglyceridemia, diabetes mellitus, peripheral insulin resistance, arterial hypertension, increased abdominal perimeter, overweight, obesity, peripheral arteriopathy, ischemic heart disease, cerebral infarction, and vascular risk. Data on dry syndrome, endometriosis, goiter, hypothyroidism or thyroiditis, hyperthyroidism, bronchial asthma, chronic urticaria, and other immune phenomena were collected. Regarding hypersensitivities, data were collected on multiple chemical hypersensitivity, electromagnetic hypersensitivity, intolerance to gluten, lactose, or fructose, and intolerance to drugs, pollen, or others. Regarding psychopathology, it was recorded whether there was an adaptive disorder, dysthymia, personality alteration, anxiety and panic disorder, or hyperactivity. Finally, data were collected on tinnitus, bilateral neurosensory hyperacusis, repeated unilateral infections, chronic inter syndical cystitis, irritable bladder syndrome, chronic prostatitis, mitral valve prolapse, or cancer.

Figure 4.19: Frequent comorbidities collected in the study and their prevalence.

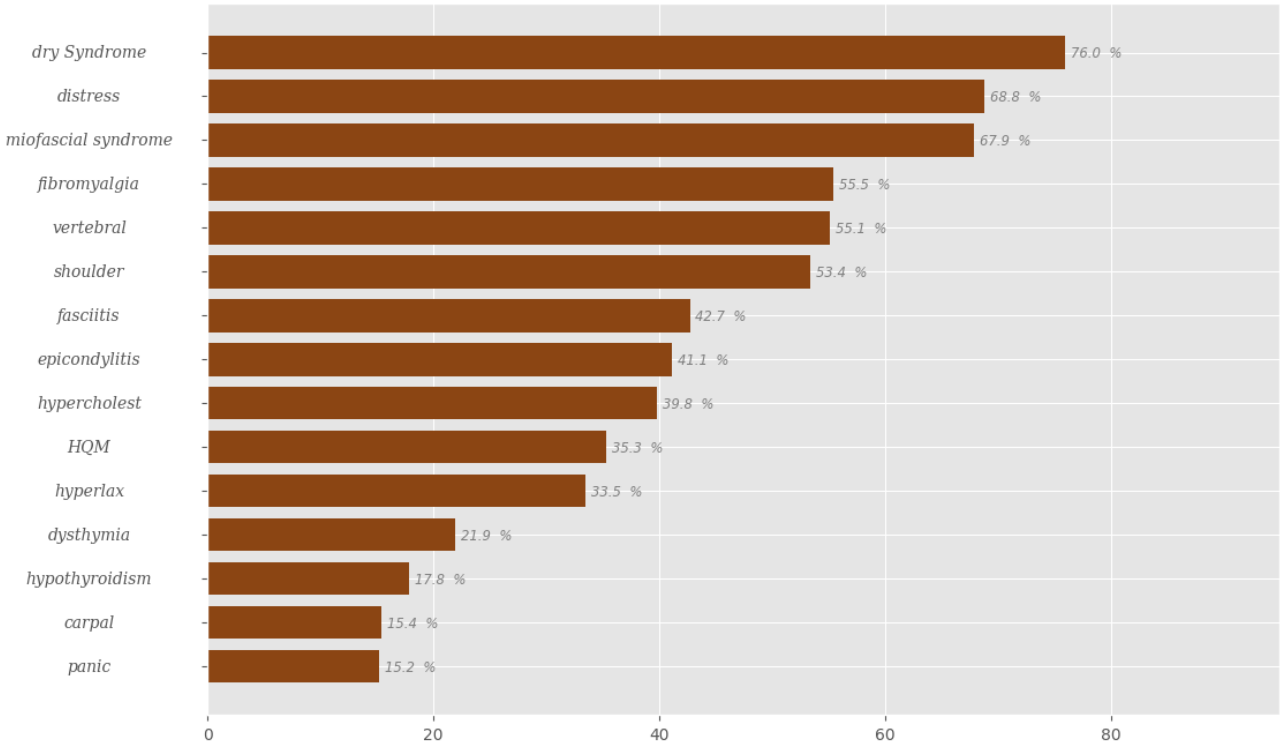
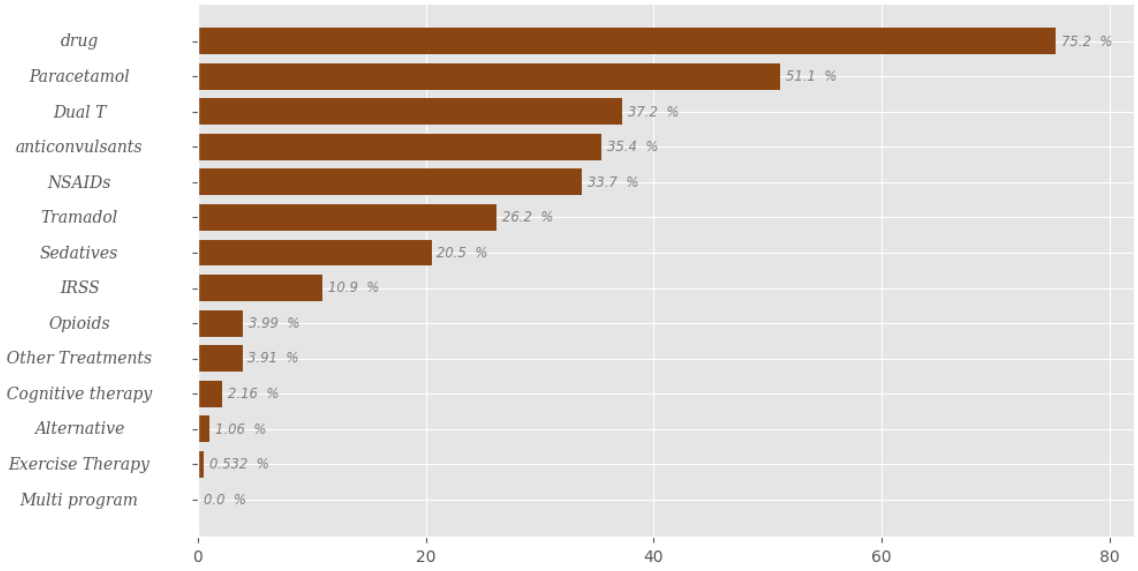


Figure 4.20 shows the treatments prescribed at the time of the registry. Seventy-five percent of the patients reported pharmacological treatments. Antidepressant treatments are specifically included, being tricyclic, dual, and selective serotonin inhibitors (such as Citalopram (Celexa), Escitalopram (Lexapro),

Fluoxetine (Prozac), Paroxetine (Paxil, Pexeva) or Sertraline (Zoloft)). Pharmacological treatments for pain are specifically Tramadol, Paracetamol (including Ibuprofen), and major opioids (characterized by having no analgesic ceiling, i.e., the degree of analgesia grows almost unlimitedly with the dose and could reach a very high analgesic ceiling were it not for their adverse effects. These are morphine, fetanyl, oxycodone and methadone). Non-steroidal anti-inflammatory drugs (aspirin, ketoprofen, celecoxib, diclofenac or ketorolac) and anxiolytics or sedatives. Non-pharmacological treatments include cognitive-behavioral therapy, programmed physical exercise, multidisciplinary treatment programs or other alternative treatments.

Figure 4.20: Recorded treatments



Chapter 5

State of the Art

5.1 Use of questionnaires in clinical longitudinal trials

The usefulness of self-administered questionnaires on general health status has been positively evaluated in randomized trials. It would be necessary to systematically collect and report such data, regardless of whether the results agree with the primary outcomes or not [142].

A review of the scientific literature is presented below to contextualize the use of self-administered questionnaires to measure effectiveness in different contexts related to ME/CFS. The titles of the studies relate to the research field and the type of questionnaire in use. The fields are diverse, as can be seen. Where appropriate, global data or subscales are usually used as numerical variables.

STUDY AND SCALES STUDIED

- Multidisciplinary rehabilitation treatment versus cognitive behavioral therapy for patients with chronic fatigue syndrome: a randomized controlled trial [146].
  - SF-36 Gloval Values
- Health-related quality of life in patients with chronic fatigue syndrome: group cognitive behavioral therapy and graded exercise versus usual treatment. A randomized controlled trial with one year of follow-up [147].
  - SF-36 Global Values
- Cognitive behavioral therapy in chronic fatigue syndrome: a randomized controlled trial of an outpatient group program [148].
  - SF-36 Global Values
- A Preliminary Placebo-Controlled Crossover Trial of Fludrocortisone for Chronic Fatigue Syndrome [149].

  - SF-36 Global Values

- Treatment and management of chronic fatigue syndrome/myalgic encephalomyelitis: all roads lead to Rome [134].

  - SF-36 Global Values

- Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a randomised trial [150].

  - SF-36 Physical Function

- Guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome (GETSET): a pragmatic randomized controlled trial [151].

  - SF-36 Physical Function

- Effectiveness of a group-based self-management program for people with chronic fatigue syndrome: a randomized controlled trial [152].

  - SF-36 Physical Function

- Randomised Controlled Trial of Cognitive Behaviour Therapy Delivered in Groups of Patients with Chronic Fatigue Syndrome.

  - SF-36 Physical Function

- Cognitive Behaviour Therapy for Chronic Fatigue Syndrome in Adults: Face to Face versus Telephone Treatment - A Randomized Controlled Trial [153].

  - SF-36 Physical Function

- Implementing a minimal intervention for chronic fatigue syndrome in a mental health centre: a randomized controlled trial [154].

  - SF-36 Physical Function

- Cytokine Inhibition in Patients With Chronic Fatigue Syndrome [155].

  - SF-36 Physical Function

- Nurse led, home based self help treatment for patients in primary care with chronic fatigue syndrome: randomised controlled trial [156].

  - SF-36 Physical Function

- Dietary intervention in chronic fatigue syndrome [157].

  - SF-36 Physical Function

- Randomised controlled trial of patient education to encourage graded exercise in chronic fatigue syndrome [158].

  - SF-36 Physical Function

- Implementing a minimal intervention for chronic fatigue syndrome in a mental health centre: a randomized controlled trial [154].
  - SF-36 Physical and Social Function
- Effectiveness of Distant Healing for Patients with Chronic Fatigue Syndrome: A Randomised Controlled Partially Blinded Trial (EUHEALS) [159].
  - SF-36 Mental Function
- Effect of Melatonin Plus Zinc Supplementation on Fatigue Perception in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Randomized, Double-Blind, Placebo-Controlled Trial [160].
  - SF-36 + FIS 40
- Effect of Dietary Coenzyme Q10 Plus NADH Supplementation on Fatigue Perception and Health-Related Quality of Life in Individuals with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Prospective, Randomized, Double-Blind, Placebo-Controlled Trial [160].
  - SF-36 + FIS 40
- Yeast Beta-Glucan Supplementation with Multivitamins Attenuates Cognitive Impairments in Individuals with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Randomized, Double-Blind, Placebo-Controlled Trial [141].
  - SF-36 + FIS 40
- Physical activity measures in patients with myalgic encephalomyelitis/chronic fatigue syndrome: correlations between peak oxygen consumption, the physical functioning scale of the SF-36 questionnaire, and the number of steps from an activity meter [161].
  - SF-36 + VO2 CPET
- Does Oral Coenzyme Q10 Plus NADH Supplementation Improve Fatigue and Biochemical Parameters in Chronic Fatigue Syndrome? [162].
  - FIS - 40

The clinical trial of the only approved treatment for ME/CFS in one country.

- A Double-Blind, Placebo-Controlled, Randomized, Clinical Trial of the TLR-3 Agonist Rintatolimod in Severe Cases of Chronic Fatigue Syndrome [163].

Two hundred thirty-four patients diagnosed with ME/CFS were analyzed to evaluate the efficacy of Rintatolimod. The method was to analyze intra- and inter-group results by comparing baseline and week 40 of the Placebo/Rintatolimod treatment.

- Primary Endpoint.

  - Exercise tolerance.

- Secondary objectives.
  - Vitality subscale (SF-36).
  - Global values (SF-36).
  - Karnovsky (KPS).

The results in the primary objective were positive, with a significant difference in increased exercise tolerance, but not in the SF-36 results. The safety of the drug was questioned, and it is approved for use in Argentina. The study is from 2012, and the FDA rejected its approval in 2013, as will be discussed in 8.1.

5.2 The use of machine learning in ME/CFS. The diagnostic tests of image

Standard functional MRI analysis attempts to compare univariate regions of brain activation at rest or during a task between patients with a diagnosis of ME/CFS and control groups. However, multivariate classification methods have become increasingly popular to identify patterns of brain activity that may differentiate disease physiology [164]. Machine learning algorithms such as logistic regression, **support vector machines** (SVM), and random forests, combined with feature selection, can be applied to clustered voxel data (the cubic unit that makes up a three-dimensional MRI) to identify patterns of brain regions that may be indicative of disease [165]. The hypothesis is that acute physical stress, such as a light exercise task, combined with implementing a machine learning algorithm, will identify a dominant pattern of behavior during functional MRI scans of individuals with ME/CFS performing the n-back reminiscence paradigm [166].

Working memory (WM), refers to the structures and processes used for temporarily storing and manipulating information in the face of continuous processing and distractors. A popular assessment of WM in the imaging literature is the $N-back$ task. In this task, participants are presented with a series of stimuli, and their decision regarding each stimulus is analyzed to determine whether it matches the N previously presented items. Researchers have shown that processing load is systematically altered, as expressed by changes in accuracy and reaction time (RT), by varying the value of N. The psychometric properties of the $N-back$ task as a measure of WM have been little studied despite its widespread use in neuroimaging. Furthermore, little is known about individual differences in $N-back$ performance and their relationship to individual differences in WM on other measures of cognitive ability. [167].

Some findings suggest attentional dysfunction and possible distraction by sensory processing in pain and interoception. Differential activation after exercise may indicate objective changes associated with post-exercise discomfort involving the frontal and lateral temporal nodes of the default mode network, sensory hypervigilance, and attention using the left rolandic operculum, the visual network and ventral

attentional network, and the basal ganglia in the affective. However, a small number of analyzed cases require further study for the evaluation of diagnostic methods of the disease using deep learning techniques based on diagnostic imaging such as MRI [168].

Chapter 6

Consistency of contributions

The following is a list of the publications that are part of the compendium and their contribution to the objectives. The common thread running through the papers is the use of the population-based registry of patients with ME/CFS described in Chapter 4. The models generated throughout the research have facilitated the findings presented in the papers. These models have demonstrated high accuracy due to the adequate number of records that make up the database. This fact is different from similar research and highlights the results obtained.

6.1 List of publications

6.1.1 Publication 1:

Unsupervised Cluster Analysis Reveals Distinct Subtypes of ME/CFS Patients Based on Peak Oxygen Consumption and SF-36 Scores.

In this work, the authors demonstrate the relationship between responses to the self-administered SF-36 questionnaire and peak oxygen consumption in the in-hospital stress test. These findings align with the two primary objectives of the present thesis. On the one hand, a data science-based model is proposed, which predicts oxygen consumption in a hospital stress test of an MS/CFS-affected patient using a questionnaire response. On the other hand, an application is offered so that, from primary care, a physician can detect this fact and can refer the patient to a specialized to a specialized hospital unit on a preferential basis.

Both facts are relevant, and help predict the severity of the pathology and detect the physical deterioration of the patient affected by ME/CFS early on.

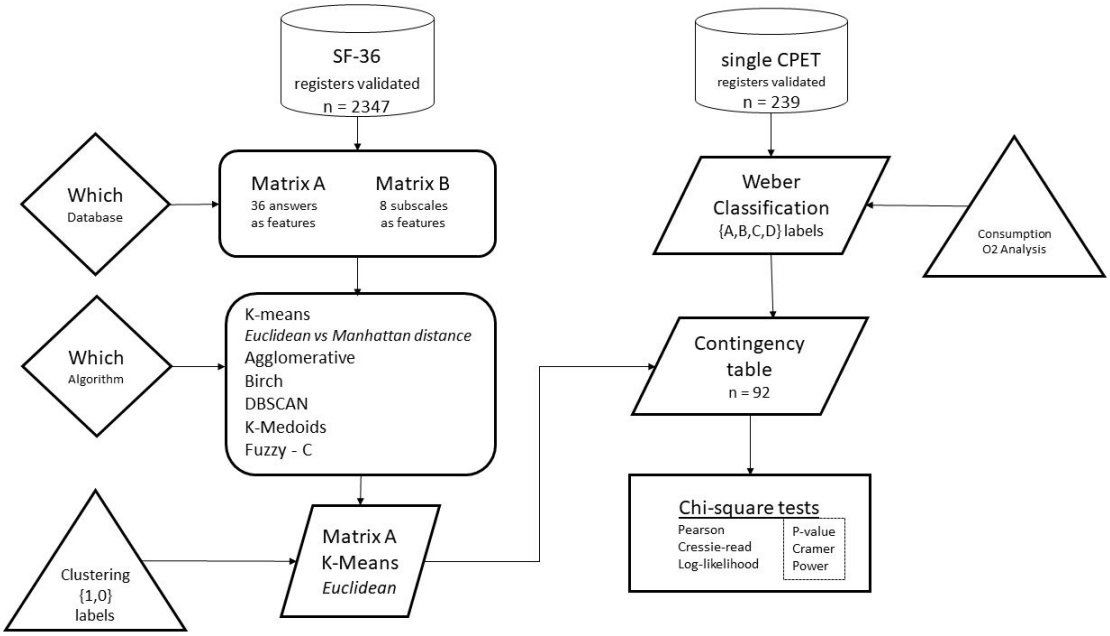
The same model has been reproduced in another research center in another paper, which is still under peer review. This confirms the results. The conclusions of both papers would confirm the use of the

machine learning model as a predictor of the health status of the ME/CFS patient, as well as the use of peak oxygen consumption in the exercise test as a biomarker.

Applied data science

In this publication the whole mathematical and algorithmic argument was provided as supplementary material, which can be consulted in the appendix of this thesis. An explanation of the strategy followed is given here, and I show the scheme in figure 6.1 without substituting in detail the documentation provided.

Figure 6.1: Flow diagram



The flowchart follows the BMPN notation.

The objective was to relate the results of a questionnaire to the results of a hospital stress test. In other words, to relate what the patient feels through a series of questions to his or her physical fitness as measured objectively in a hospital environment. The SF-36 questionnaire datatypes have been used as a reference in the investigations consulted. However, a comparison is made between the two alternatives: using the raw answers or the data types. Each response calculates a datatype so that one matrix can be calculated by a linear application of the other (see in more detail in the appendix). This fact observes the characterization of each patient as a vector formed by the different dimensions of each matrix. Since we intend to create two groups of unlabeled vectors, we are required to use **clustering** algorithms.

This article uses linear algebra theory and its applications and a complete study of the use of **clustering** algorithms that demonstrate that the raw response matrix provides better information than the decatypes matrix, in contrast to what has been studied in the scientific literature so far.

6.1.2 Publication 2:

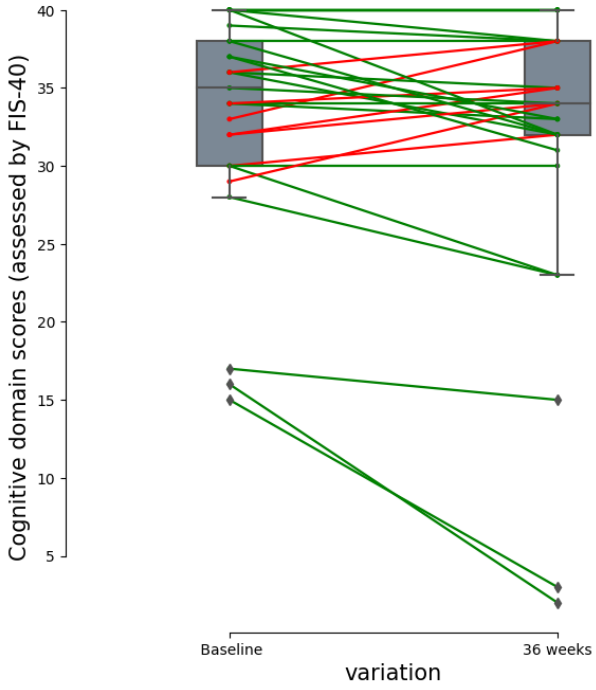
Yeast Beta-Glucan Supplementation with Multivitamins Attenuates Cognitive Impairments in Individuals with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Randomized, Double-Blind, Placebo-Controlled Trial.

It shows the efficacy of a nutritional product under study using the results of the answers to questionnaires used as a primary endpoint. This fact is relevant as it confirms the use of questionnaires for prospective clinical trials of drug efficacy. It supports the relevance of the questionnaires and their routine clinical practice and reinforces the importance of the findings in the other publications.

Applied data science

The results of this research publication are based on inferential statistics. Work is in progress on future publications with the new biochemical data to be analyzed together. Visualization of the relevant results makes it easier for the reader to understand the conclusions, and is a relevant field in data science. Figure 6.1 is discussed.

Figure 6.2: Paired data analysis



A two-dimensional graph is shown, with the abscissa axis represented by the time of data collection and the ordinate axis by the values of the variable. Each record is identified with a point and the difference between the two measurements, indicating in color whether it is reduced (green) or increased (red), providing the box plot for a correct interpretation of the statistical data provided in the study.

6.1.3 Publication 3:

A synthetic data generation system for questionnaires for myalgic encephalomyelitis/chronic fatigue syndrome.

The synthetic data model offers several services. On the one hand, it makes the generation of valid synthetic data for syndrome research available to the scientific community. This fact is relevant due to the scarce volume of records of this type of database in specialized international units. On the other hand, it proposes an application that details information from up to seven different questionnaires mentioned in section 5.1, using the answers of only one. This type of application can help to stratify the severity of the syndrome once it has been diagnosed according to the diagnostic criteria.

Applied data science

The problem solved by this work is one of prediction. The models tested were regression and classification models, the latter giving the best results. **XGBOOST** and neural networks have provided better accuracy. Information on the process is provided in the publication, and this section, I will add some details specialized in **deep learning**.

About the use of the "class-weight" parameter.

The final objective is to create a model that, knowing the answers to the SF-36 questionnaire, can predict the answers to the rest of the questionnaires. Thus, the prediction is calculated question by question. The order is established using the graph analysis explained in the article; in this case, we start with the first question of the HAD questionnaire. Of the 4 possible answers, not all have the same probability of being answered in our database, thus producing an imbalance. Of the 2321 HAD questionnaires collected, the first question was answered:

- 66 option 1 with 0% success rate.
- 886 option 2 with 69% success rate.
- 807 option 3 with 51% success rate.
- 562 option 4 with 61% success rate.

The model did not predict any first-choice response, and yet it could still have a good accuracy ratio. Detection of this fact is critical in this type of problem. In the following code, the expected weight of the prediction is defined, and the model will "punish" the errors proportionally so that it corrects itself.

```
from sklearn.utils import class_weight

class_weights = class_weight.compute_class_weight('balanced', np.unique(y), y)
class_weights = dict(enumerate(class_weights))
```

Tras el uso de este parámetro y un mejor ajuste, los resultados fueron:

- 66 option 1 with 80% success rate.
- 886 option 2 with 81% success rate.
- 807 option 3 with 76% success rate.
- 562 option 4 with 82% success rate.

Below is the code of the model used attached:

```
def model( activation = 'relu', init = 'glorot_normal', regularizer = 'l2'):
    optimizer = Adam(lr=0.0001, beta_1=0.97, beta_2=0.995, epsilon=1e-08)
    # create model
    model = Sequential()
    model.add(Dense(400,
                    input_shape=(X_train.shape[1],),
                    activation=activation,
                    kernel_regularizer = regularizer,
                    kernel_initializer = init)
              )

    model.add(Dense(200, activation=activation, kernel_regularizer =regularizer,
        kernel_initializer = init))
    model.add(Dropout(0.5))
    model.add(Dense(100, activation=activation, kernel_regularizer =regularizer,
        kernel_initializer = init))
    model.add(Dense(ohe.shape[1], activation = 'softmax'))
    model.compile(loss='categorical_crossentropy', optimizer=optimizer, metrics=metrics)
    return model
```

And when training the model, the "class-weights" parameter is added in the "fit" method,

```
estimator = KerasClassifier(build_fn=model, nb_epoch=4000, batch_size=64, verbose=0)
es = EarlyStopping(monitor='val_recall', mode='max', patience = 400, verbose=0)
mc = ModelCheckpoint('best_model_{col}.h5'.format(col = 'tunning'),
monitor='val_recall', mode='max', verbose=0, save_best_only=True)
h = estimator.fit(X_train, y_train, validation_data = (X_test, y_test),
    epochs = 4000, batch_size = 64,
    class_weight=class_weights, callbacks = [es, mc], verbose = 0)
```

Each response has a value associated with it to calculate each questionnaire's different datatypes and provide the interpretation of the questionnaire. In preparing a synthetic data generator, the anonymity of the real data must be preserved, but the statistical characteristics of the real data must be preserved. On this basis, the model's errors are calculated, and it is found that 95% of the errors differ by 1. As a result, we obtain artificial records with a statistical structure corresponding to the original one.

6.1.4 Publication 4:

The Conners Continuous Performance Test CPT3™: Is it a reliable marker to predict neurocognitive dysfunction in Myalgic encephalomyelitis/chronic fatigue syndrome?

The results analyzed from the data science point out relevant information: it postulates the result of an objective test as a biomarker for measuring cognitive impairment in patients affected by MS/CFS. This fact and other sub-analyses that are being carried out show a line of action for the early detection of cognitive impairment and are part of the primary objectives of this thesis. A work has been developed based on a subanalysis where it is demonstrated that the time of evolution of the disease influences cognitive impairment and it is expected to be submitted to a journal for evaluation.

Chapter 7

A copy of the publications of the Compendium

Clinical Therapeutics 45 (2023) 1228–1235

Contents lists available at ScienceDirect

Clinical Therapeutics

journal homepage: www.elsevier.com/locate/clinthera

Original Research

Unsupervised Cluster Analysis Reveals Distinct Subtypes of ME/CFS Patients Based on Peak Oxygen Consumption and SF-36 Scores

Marcos Lacasa[1,*], Patricia Launois[5], Ferran Prados[1,2,3,4], José Alegre[5], Jordi Casas-Roma[1]

[1] *e-Health Center, Universitat Oberta de Catalunya, Barcelona, Spain*
[2] *Center for Medical Image Computing, University College London, London, United Kingdom*
[3] *National Institute for Health Research Biomedical Research Centre at UCL and UCLH, London, United Kingdom*
[4] *Queen Square MS Center, Department of Neuroinflammation, UCL Institute of Neurology, Faculty of Brain Sciences, University College London, London, United Kingdom*
[5] *Myalgic Encephalomyelitis / Chronic Fatigue Syndrome Unit, Division of Rheumatology, Vall d'Hebron Hospital Research Institute Universitat Autònoma de Barcelona, Barcelona, Spain*

ARTICLE INFO

Keywords:
Biomarker
Chronic fatigue syndrome
Cardiopulmonary exercise test
Clustering
K-means

ABSTRACT

Purpose: Myalgic encephalomyelitis, commonly referred to as chronic fatigue syndrome (ME/CFS), is a severe, disabling chronic disease and an objective assessment of prognosis is crucial to evaluate the efficacy of future drugs. Attempts are ongoing to find a biomarker to objectively assess the health status of (ME/CFS), patients. This study therefore aims to demonstrate that oxygen consumption is a biomarker of ME/CFS provides a method to classify patients diagnosed with ME/CFS based on their responses to the Short Form-36 (SF-36) questionnaire, which can predict oxygen consumption using cardiopulmonary exercise testing (CPET).

Methods: Two datasets were used in the study. The first contained SF-36 responses from 2,347 validated records of ME/CFS diagnosed participants, and an unsupervised machine learning model was developed to cluster the data. The second dataset was used as a validation set and included the cardiopulmonary exercise test (CPET) results of 239 participants diagnosed with ME/CFS. Participants from this dataset were grouped by peak oxygen consumption according to Weber's classification. The SF-36 questionnaire was correctly completed by only 92 patients, who were clustered using the machine learning model. Two categorical variables were then entered into a contingency table: the cluster with values {0,1} and Weber classification {A, B, C, D} were assigned. Finally, the Chi-square test of independence was used to assess the statistical significance of the relationship between the two parameters.

Findings: The results indicate that the Weber classification is directly linked to the score on the SF-36 questionnaire. Furthermore, the 36-response matrix in the machine learning model was shown to give more reliable results than the subscale matrix ($p - value < 0.05$) for classifying patients with ME/CFS.

Implications: Low oxygen consumption on CPET can be considered a biomarker in patients with ME/CFS. Our analysis showed a close relationship between the cluster based on their SF-36 questionnaire score and the Weber classification, which was based on peak oxygen consumption during CPET. The dataset for the training model comprised raw responses from the SF-36 questionnaire, which is proven to better preserve the original information, thus improving the quality of the model.

Introduction

Myalgic encephalomyelitis (ME), commonly referred to as chronic fatigue syndrome (CFS), is a serious, complex, chronic, multisystem illness of unknown etiology. It is also known as post-viral fatigue syndrome as it is often triggered by a persistent viral infection. Research on the prevalence of the disease has focused predominately on meta-analysis due to the variety of this type of data and the complexity of calculations. ME/CFS is characterized by unexplained and persistent post-exertional fatigue that is not relieved by rest, is exacerbated by physical and mental exertion, and shows other core symptoms such as cognitive, immunometabolic, autonomic, and neuroendocrine dysfunc-

Abbreviations: ME/CFS, Myalgic encephalomyelitis chronic fatigue syndrome; CPET, cardiopulmonary exercise test; VO2 peak, peak oxygen consumption; VO2 VT1, oxygen consumption at the anaerobic threshold; RPE, rate of perceived exertion; RER, respiratory exchange ratio.

* Corresponding author: ADaS Lab - e-Health 5Center, Universitat Oberta de Catalunya, Rambla del Poblenou, 156, 08018 Barcelona. (SPAIN)
E-mail address: mlacasaca@uoc.edu (M. Lacasa).

https://doi.org/10.1016/j.clinthera.2023.09.007
Accepted 9 September 2023

tion.[1] Published data indicates a prevalence of between 0.89 and 1.14% among the population of the United States.[2] Additional data includes an analysis of adolescents in the United Kingdom with an estimated prevalence range of between 1.47% and 2.99%.[3] In contrast to these studies, recent research carried out in China estimates the prevalence among the Chinese population to be 12.54%.[4] The prevalence in Australia was estimated at 0.79%.[5] ME/CFS is a major cause of disability, with many patients feeling unable to fulfil their family responsibilities and forced to limit their social activities. The condition significantly impacts their work (difficulty performing their job) and daily life (for example, climbing stairs, lifting or carrying groceries, moving a table, or pushing a vacuum cleaner).[6] Patients also have characteristic inflammatory and muscular symptoms, sleep disturbances, and altered cognitive function.[7] Muscular symptoms include pain, generalized muscle weakness, post-exertional fatigue, and neurological (mental fatigue, impaired cognition, psychomotor slowing, disturbed sleep, hypersensitivity to noise, light, and odours, headache, pain, paresthesia and severe dysautonomia).[8] Neurocognitive symptoms have also been reported (slow reaction time, indicating the likelihood of an ADHD-like pattern of functioning).[9]

The most widely used international definition for ME/CFS is the Fukuda criteria, developed in 1994 by the Center for Disease Control and Prevention (CDC) in Atlanta, Georgia, USA[10]. As some symptoms in the Fukuda criteria overlap with depression, it has been criticized for its lack of specificity.[11] In contrast, the Canadian Consensus Criteria distinguish patients with ME from those who are depressed, identifying patients who are more physically debilitated and have greater physical and cognitive dysfunction.[12] These criteria were updated in 2011 to include post-exertional fatigue.[13]

Cardiopulmonary exercise testing (CPET) provides information that helps identify abnormal values in cardiac and/or ventilatory responses that rule out ME/CFS, thus differentiating these patients from those likely to be affected by other etiologies causing exercise intolerance. We used a gas analyser to measure peak oxygen consumption (VO2 peak) and anaerobic threshold oxygen consumption (VO2 VT1) during exercise. These measurements determine the degree of deterioration of the patient's functional capacity (simultaneously the cardiovascular and ventilatory response to a known metabolic stress are examined) and early transition to anaerobic metabolism during exercise, respectively. Both can be measured during a single exercise test which provides information on the efficiency of the metabolism in response to physiological exercise,[14] with VO2 peak showing a strong predictive and prognostic value. Patients with chronic heart failure were evaluated with a single CPET using Weber and Janicki's[15] established classification system.

Attempts are still ongoing to identify a reliable, objective biomarker for the health status of ME/CFS patients. Currently, the most widely used technique to measure ME/CFS patient status is internationally validated self-administered questionnaires such as the SF-36 questionnaire. Several studies have attempted to identify patterns of association between these questionnaires and the pathophysiology of specific systems and links to contemporary multidisciplinary molecular pathology, including comparative MRI,[16] exploring symptom co-occurrence using network analysis,[17] and investigating the relationship between oxygen consumption and the physical subscale of the SF-36 questionnaire in retrospective studies.[18] Physicians currently base ME/CFS diagnosis and prognosis on exclusion and subjective clinical interpretation, and to date, a reliable, objective method using accessible inexpensive tests that can identify ME/CFS patients with worse prognosis due to the wide-ranging symptoms accompanying the disease and subjective interpretation of results has not yet been found. Therefore, validating an objective biomarker such as oxygen consumption in CPET would be helpful as a primary outcome of trials developing new treatments and for a deeper understanding of the evolution of the disease.

This aim of this study was twofold: first, to prove that oxygen consumption is a potential biomarker of ME/CFS; and second, to develop a new machine learning-based method to identify patients diagnosed with ME/CFS and at high risk of physical impairment. Our proposed model predicts CPET scores from self-administered SF-36 questionnaires in the primary healthcare setting and assists in the early identification of patients in need of referral to a dedicated hospital unit.

Patients and Methods

Study Population

This study involved two clinical trials, both approved by the ethics committee of the Vall d'Hebron University Hospital, Barcelona, Spain: 1) "Population-based Registry of Chronic Fatigue Syndrome Patients", approved 10/18/2006; and 2) "Study of exercise intolerance in adult patients with chronic fatigue syndrome using a cardiopulmonary exercise stress test (CPET)", approved 09/22/2020.

The 36-Item Short Form Health Survey

This prospective cross-sectional study included 2,522 patients diagnosed with ME/CFS from the Vall d'Hebron University Hospital (Barcelona, Spain), made up of 90.5% women (mean age 48.11±10.31 years) and 9.5% men (mean age 44.41±11.35 years). Data from the Spanish version of the SF-36 questionnaire[19] between 2008 and 2022 was obtained and recorded. Both the Spanish and English versions of the SF-36 can be found in Supplementary Material.

The SF-36 questionnaire consisted of 36 questions requiring the patient to choose one option. The options ranged from 1 to 6, as illustrated in Table 1. Eight subscales were defined and calculated based on the weighted sum of a small number of responses. The score for each question was coded so that higher scores reflected better health.

Net rank *equals the difference between the maximum and minimum theoretical values. The minimum is equal to the number of questions formed by each subscale, as the minimum value of each question is always 1.*

The formula to calculate each subscale value:

$$Subscale = \frac{(Real\ punctuation - minimum)}{Net\ rank} \times 100$$

If the minimum value of each question were zero, the procedure could be proposed as a weighted sum expressed as a percentage. This was essential point as it meant the matrices could be converted to a linear application. Refer to Supplementary Material for further details.

Decoded SF-36 questionnaire

The items and scales on the SF-36 are scored so that a higher score represents better health. For example, the function scales are scored so that a higher score indicates better function, and the pain scale is scored so that a high score suggests the patient is pain-free. After data entry, the items and scales were scored in three steps:

1. Re-encode the 10 items listed in the SF-36 manual.
2. Compute the scale score by summing the items in the same scale (raw scale score).
3. Convert the raw scale scores to a 0-100 scale (converted scale scores).

Modulus

The 36 responses to the SF-36 questionnaire were analysed, with the eight subscales investigated separately. The 36 responses for each participant were collected in columns, and each row represented one participant's anonymized data. In both cases, a 36-value vector (in the case of the 36-response matrix) or an 8-dimension vector (in the case of the subscales) was computed for each participant. For further details, see Supplementary Material.

Table 1
Subscales calculation scheme.

Subscale	Number of questions	Rank punctuations (min-max)	Net rank
Physic Function	10	10-30	20
Physic Rol	4	4-8	4
Body Pain	2	2-12	10
General Health	5	5-25	20
Vitality	4	4-24	20
Social Function	2	2-10	8
Emotional Rol	3	3-6	3
Mental Health	5	5-30	25

For each vector, the modulus was calculated as follows:

$\text{modulus}_i = \sqrt{\Sigma_{j=1}^{d} x_j^2}$

$modulus_i$: *the* i^{th} *register.*

d : *number of dimensions.* 36 *for our model.*

x_j : *the* j^{th} *element of vector.*

Single CPET test dataset

The "Study of exercise intolerance in adult patients with chronic fatigue syndrome using a cardiopulmonary exercise test (CPET)" is a prospective, cross-sectional, enrolled cohort study initiated in 2020. Details of the study are set out below:

- **Objectives**: To evaluate exercise intolerance in adult patients diagnosed with ME/CFS using CPET, taking into consideration cardiovascular, ventilatory, muscular, and metabolic variables to determine functional capacity and metabolic efficiency of exercise.
- **Inclusion criteria**: Participants aged between 18 and 60 and diagnosed (in accordance with the Fukuda[10] and Carruthers[12] criteria) with ME/CFS by the physician of the Central Sensitization Syndromes Unit, Vall d'Hebron Hospital, Barcelona. Participants must have clinical manifestations of the symptomatic muscle group of exercise intolerance, be able to perform the exercise test, and agree to take part in the study by signing the informed consent form.
- **Exclusion criteria**: Patients with contraindications for CPET or who do not provide informed consent to participate in the study.

Two hundred and thirty-nine participants were enrolled and referred to the Cardiopulmonary Rehabilitation Exercise Laboratory to perform CPETs. The study included 85.7% women (mean age 50.15±8.61 years) and 14.3% men (mean age 46.94±9.82 years). Patients were advised to take their regular medication and to avoid strenuous exercise on the test day. They were also asked to fast for two hours before the test and to drink only water. The participants were connected to an electrocardiogram, a pulse oximeter, and a blood pressure cuff to continuously monitor their vital signs during all study phases. They underwent a symptom-limited maximal exercise test using an electronically braked bicycle ergometer (Ergoline GmbH ER 800 S, Bitz, Germany). Exhaled air was collected via a two-way breathing valve attached to a mask that covered the participant's nose and mouth. Respiratory gasses breath-by-breath were analyzed using a Vyntus CPX gas analyzer (Ergospirometer, Vyaire, Hoechberg, Germany) and SentrySuite 3.0 software. The spirometer was calibrated under ambient conditions before testing. Version V-781239 V 06.02 of the technical manual is available on the SentrySuite website. We used cardiopulmonary exercise testing based on work by Dr. Wasserman and colleagues at the University of California, Los Angeles.[20]

CPET Protocol Phases

The CPET tests were performed in the Cardiopulmonary Rehabilitation Exercise Laboratory under the supervision of a physician and a nurse and following the standardized protocols of the Central Sensitization Syndromes Unit of the Vall d'Hebron Hospital, Barcelona. These tests were performed in a hospital setting at no additional cost to the patient. The tests involved the following 3 phases:

Phase 1. Determine baseline cardiovascular and respiratory values as follows:

- Simple spirometry: three consecutive manoeuvres were performed to obtain reproducible data.
- Slow spirometry: one assessment was performed to determine maximum voluntary ventilation (MVV).

Phase 2. Maximum incremental cycle ergometer test: starting with three minutes of rest, followed by two minutes of unloaded pedalling, then adding an incremental load at a rate of 10 W/min (Watts / minutes) while maintaining a pedalling cadence of between 40-50 and 50-60 W/min (depending on the training status and previous fatigue level of the participant) until exhaustion due to muscular fatigue and/or dyspnoea, or according to the operator's medical criteria as electrocardiographic changes or symptoms contraindicating continuation of the test.

Phase 3. Recovery. On cycling completion, recovery from exercise was monitored for 3 minutes and the reasons for stopping CPET recorded. Continuous electrocardiographic monitoring of heart rate (12-lead) and blood pressure was performed every 2 minutes throughout the test. A Borg test was performed at baseline and at maximal exercise.

Maximal criteria test

Gas exchange data were recorded during CPET and recovery phases. The criteria used to determine whether participants had reached maximal physiological effort were as follows: plateau in oxygen consumption with increasing workload; modified rate of perceived exertion (RPE) > 8 (scale of 0-10); respiratory exchange ratio (RER) > 1.1, reaching at least 85% of age-predicted maximal heart rate, or a peak blood lactate > 8 mmol. When two of the three criteria were met, it was considered that the patient had exerted maximal effort.

The key measure sought in this study was the cardiorespiratory fitness of patients with ME/CFS, as determined by peak oxygen uptake (VO2 peak: measured during incremental exercise) and representing the maximum aerobic power during cumulative effort. The VO2 peak was measured in millilitres of oxygen per kilogram of body weight per minute ($mL^{*}Kg^{-1}{}^{*}min^{-1}$). This parameter is described as the maximum energy capacity achieved by aerobic metabolism per unit of time (aerobic capacity) during an incremental CPET. Any pathophysiological situation that impairs oxygen transport from the air to the mitochondria and its utilization during exercise will reduce the predicted values of peak oxygen uptake according to age and sex.[21]

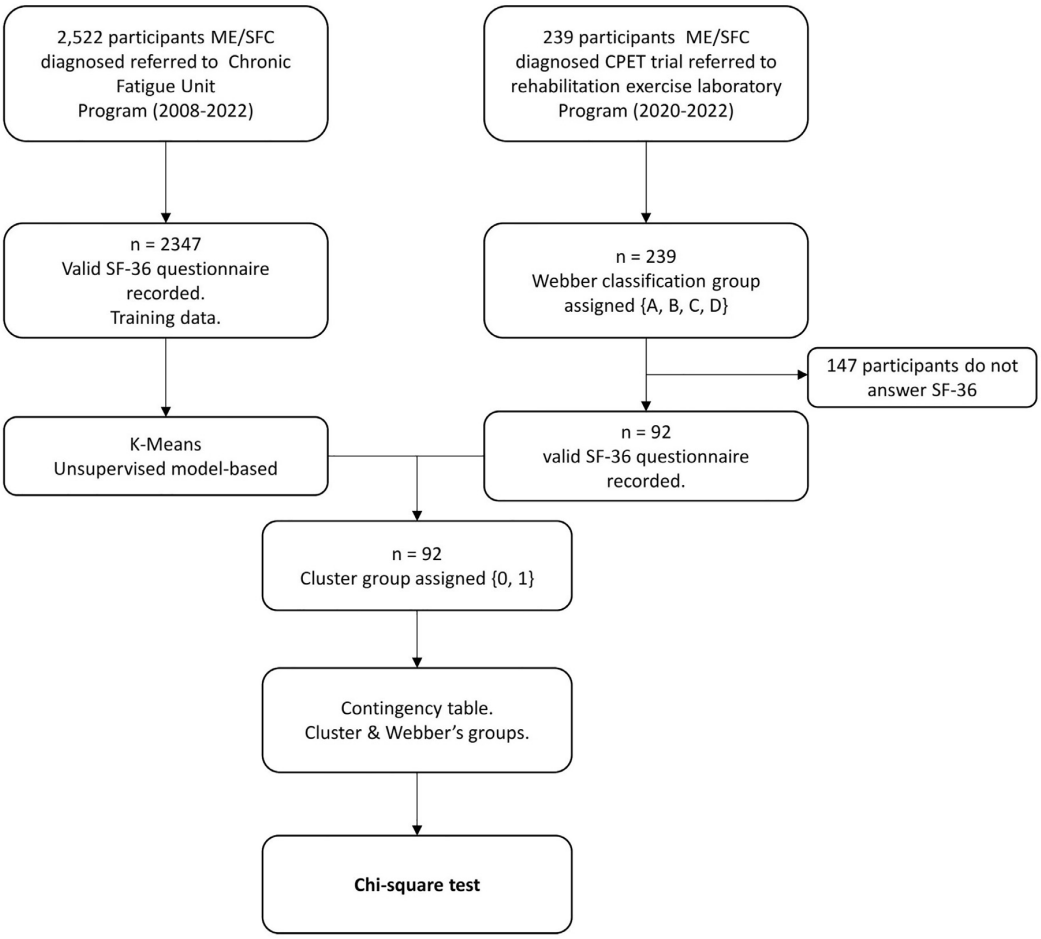


Figure 1. CONSORT diagram.

Patient enrolment procedure

This research presents data from two studies (see Figure 1): the first, "Population-based Registry of Chronic Fatigue Syndrome Patients", was approved 10/18/2006 and is used for automatic patient clusters using machine learning; the second, "Study of exercise intolerance in adult patients with chronic fatigue syndrome using a cardiopulmonary exercise stress test (CPET)", was used as a validation set.

The first study, "Population-based Registry of Chronic Fatigue Syndrome Patients" included SF-36 questionnaires collected from 2522 patients recruited at the Chronic Fatigue Syndrome Unit of the Vall d'Hebron Hospital, Barcelona, between 2008 and 2022, with their informed consent to participate in research. From this initial group, only 2347 of the questionnaires were valid. These 2347 questionnaire responses were used to train the proposed machine learning model to identify two different ME/CFS clusters.

The second study: "Study of exercise intolerance in adult patients with chronic fatigue syndrome using a cardiopulmonary exercise stress test (CPET)" included 239 CPET tests collected between 2020 and 2022 from ME/CFS patients referred to the Cardiopulmonary Rehabilitation Exercise Laboratory from the Central Sensitization Syndromes Unit of the Vall d'Hebron Hospital, Barcelona. The CPET results were collected according to Weber's classification. From the 239 patients, 92 completed both studies (CPET and SF-36 questionnaire). A contingency table of the 92 selected datasets was created and the previously trained clustering model applied, thus classifying each of the 92 patients as belonging to cluster 0 or cluster 1. Finally, as a validation step, the Chi-square test of independence was performed to determine whether there was a relationship between the two categorical variables (Weber's classification and clusters).

Statistical Analysis

Contingency table difference analysis

The creation of the contingency table is based on two categorical variables. The first variable is computed by the previously trained machine learning model, which assigns a cluster value {0, 1} based on all the answers from each input SF-36 questionnaire. The second variable is denoted by a letter {A, B, C, D} contingent on the peak oxygen consumption measured in the CPET.

- Mild to none (A) when Peak VO2 value is > 20 $mL^{*}Kg^{-1}{}^{*}min^{-1}$
- Mild to moderate (B) when Peak VO2 value range is 16-20 $mL^{*}Kg^{-1}{}^{*}min^{-1}$
- Moderate to severe (C) when Peak VO2 value range 10-16 $mL^{*}Kg^{-1}{}^{*}min^{-1}$
- Severe (D) when Peak VO2 value is < 10 $mL^{*}Kg^{-1}{}^{*}min^{-1}$

This is evaluated by groups {A, B, C, D} of patients who performed the exercise test, n=239. The self-administered SF-36 questionnaire was

completed by 92 patients only. All the SF-36 responses from these 92 patients were assigned to clusters {0, 1} in the model. These data were then used to create a contingency table for the 92 participants. In summary, each participant is assigned to a cluster (0 or 1) and given a letter (corresponding to the peak oxygen consumption: A, B, C or D) that defines the two categorical variables in our analysis (Figure 1).

The Chi-square test of independence is a statistical hypothesis test used to determine whether two categorical or nominal variables are likely to be related and can be used when counting the values of two categorical variables. The parameters were analysed for a more in-depth evaluation of the results. The likelihood ratio chi-square tests provide a range of parameterizations that support accurate selections for various distributions and sample sizes. Pearson's and Cressie-Read's chi-squared tests often tend to select overly complex bivariate parameterizations (up to 4 options in Weber's classification). This is because chi-squared statistics are overly inflated when models are fitted to sparse bivariate distributions and minor expected frequencies are used as divisors.[22] We decided to compare different strategies using different values of lambda. Python package Pingouin has used (version 0.5.2) to see if they differed.[23]

Parameters analysed in the contingency table:

- **Lambda**. A measure of association that reflects the proportional reduction in error when the values of the independent variable are used as predictors of the values of a dependent variable. A value equal to 1 indicates that the independent variable is a perfect predictor of the dependent variable. A value equal to 0 indicates that the independent variable has no contribution to the prediction of the dependent variable: Pearson (lambda=1), Cressie-Read (lambda=0.67) and Log-likelihood (lambda=0).
- **dof** is the Chi-square's degrees of freedom, and is calculated using the equation $dof = (r - 1) \times (c - 1)$, where r is the number of rows and c is the number of columns.
- The **p-value** is the probability of obtaining a chi-square equal to or greater than that obtained in the current experiment in the current experiment, given that the null hypothesis is true. Generally, a p-value of 0.05 or greater is considered critical. Anything less indicates significant variances, and the hypothesis must be rejected.
- **Cramer's V** is an effect size measurement for the Chi-square test of independence and measures how strongly two categorical fields are associated. The degree of freedom (dof) is 3. A value of between 0.06 and 0.17 is considered small-medium; up to 0.30, medium-large; and greater than 0.30, large.[28] Cramer's V is based on Pearson's Chi-squared statistic and was published by Harald Cramer.[29]
- The **power of the goodness of fit** (δ) or **Chi-square independence test**. High power means there is a low probability of concluding that no effect exists when there is one. Statistical power depends upon effect size and the sample size.

$$\delta = d \times \sqrt{n}$$

Where d is Cohen's coefficient and n is the number of registers.[24].

Clustering analysis

The clustering analysis was implemented in Python (version 3.7.14) using and comparing the decoded SF-36 answers and subscale matrices. The dimensions were 2347 × 36 and 2347 × 8, respectively. To select the optimal number of clusters, some models were fitted with values in the range.[2,6] for k (Birch and Spectral Clustering) using the elbow method.[25] and the Calinski and Harabasz metric (see Supplementary Materials for further details). Three validation metrics were proposed using the scikit-learn package (version 1.0.2) to evaluate the performance of each tested model when the true labels are unknown:

- Silhouette Coefficient
- Calinski–Harabasz Index[26]
- Davies–Bouldin Index[27]

Results

Matrix Analysis

The two databases for analysis were compared and contrasted to determine which provided the best information for classifying patients with ME/CFS; first, the 36 raw responses, followed by the eight subscales (for all results, see Supplementary Material). No correlation was found between the variables in the two matrices. However, when the modulus of each matrix was compared, this yielded different results, with the modulus of the raw responses showing a different health status to the modulus of the matrix of the subscales. This finding is to argue that using one matrix cannot be used as an equivalent to the other, as the model would give different results for each. A PCA analysis shows that in both matrices, emotional role (ER) and mental health (MH) are correlated and symmetric with those parts of the questionnaire dealing with physical function and vitality. Thus, the differences between the matrices were given by weighting the subscales, which significantly modified a single value of the SF-36 questionnaire. It shows that the subscales cannot be interpreted globally and must be analysed explicitly. This study demonstrates that the matrix of the 36 raw responses used as training data improves the results of the unsupervised machine learning model.

Cluster Analysis

The best-performing model for clustering MS/CFS patients was the K-means model, with the optimal number of clusters being two. The model was trained using the 36 raw answers of the 2,347 validated questionnaires to classify each subject into one of the two labelled groups with values {0, 1}. See Supplementary Material for further details.

Validation group Analysis

A total of 239 validated records were included. The Weber Classification results of the study population that completed the CPET are presented below. Participants' physical data is shown, and the peak oxygen consumed during the test is analysed and labelled as stated above:

- Mild to none (A) when Peak VO2 value is > 20 $mL^{*}Kg^{-1}{}^{*}min^{-1}$
- Mild to moderate (B) when Peak VO2 value range is 16-20 $mL^{*}Kg^{-1}{}^{*}min^{-1}$
- Moderate to severe (C) when Peak VO2 value range is 10-16 $mL^{*}Kg^{-1}{}^{*}min^{-1}$
- Severe (D) when Peak VO2 value is < 10 $mL^{*}Kg^{-1}{}^{*}min^{-1}$

A total of 239 records were collected and validated. The results are shown in Table 2.

Contingency table for single CPET: Weber's classification and clustering analysis

Only 92 records were selected. These corresponded to patients with validated SF-36 questionnaire responses and CPET. A similar analysis to the previous group analysis was performed on each of the 92 patients individually. Table 3 shows the records grouped according to Weber's classification. Note that there are no class D records for the male group as all collected values for the VO2 Peak variable are greater than 10.

Following this, the 92 selected patients were classified into two clusters using the previously trained machine learning model. Table 4 shows the results of the data analysed according to the cluster assigned by the machine-learning-based model.

Table 2
Weber Classification Results (n=239). Following Weber's classification, class was defined according to the VO2 value of a single CPET. Weber's classification stratifies patients based on peak VO2 and anaerobic threshold to define functional exercise capacity[15]: Mild to none (A) when Peak VO2 value is > 20; Mild to moderate (B) when Peak VO2 value range is 16-20; Moderate to severe (C) when Peak VO2 value range is 10-16; and Severe (D) when Peak VO2 value is < 10. **n**: Number of participants in each subgroup. Arithmetic means and standard deviation (in brackets) are shown for other parameters. Age in years, weigh in kg, height in cm and VO2 peak in $mL^{*}Kg^{-1}{}^{*}min^{-1}$.

Gender	Class	n	Age	Weight	Height	VO2 peak
Female	A	35	47.37 (10.07)	56.85 (8.46)	160.48 (6.80)	24.56 (3.89)
	B	50	49.48 (9.23)	65.36 (13.92)	161.01 (6.02)	17.32 (1.01)
	C	95	51.47 (7.78)	73.05 (14.39)	160.75 (6.26)	13.13 (1.60)
	D	25	50.40 (7.51)	74.68 (17.06)	163.24 (5.26)	8.54 (1.09)
Male	A	15	47.0 (12.10)	78.02 (11.30)	174.26 (8.38)	25.19 (3.65)
	B	8	42.75 (9.13)	82.25 (11.75)	174.50 (13.58)	17.58 (1.29)
	C	9	50.22 (6.18)	88.88 (17.60)	177.11 (6.62)	12.76 (1.78)
	D	2	48.50 (2.12)	77.45 (16.19)	177.5 (0.70)	8.80 (0.84)

Table 3
Weber Classification Contingency Table Analysis (n=92). Weber's classification defined the class according to the VO2 value of a single CPET. Weber's Classification stratifies patients based on peak VO2 and anaerobic threshold to define functional exercise capacity[15]: Mild to none (A) when Peak VO2 value is > 20; Mild to moderate (B) if Peak VO2 value range is 16-20; Moderate to severe (C) when Peak VO2 value range is 10-16; and Severe (D) when Peak VO2 value is < 10. **n**: Number of participants in each subgroup. Arithmetic means and standard deviation (in brackets) are shown for other parameters. Age in years, weigh in kg, height in cm and VO2 peak in $mL^{*}Kg^{-1}{}^{*}min^{-1}$. **Modulus**: subgroup means modulus value.

Gender	Class	n	Age	Weight	Height	VO2 peak	Modulus
Female	A	7	46.85 (10.25)	60.31 (11.15)	159.85 (9.92)	23.15 (0.8)	12.61 (1.89)
	B	22	51.54 (9.59)	61.85 (12.35)	161.18 (4.44)	17.51 (0.8)	13.98 (2.35)
	C	44	51.59 (7.15)	72.79 (14.92)	160.87 (5.88)	12.87 (1.54)	12.20 (2.46)
	D	8	51.87 (8.44)	75.0 (22.4)	160.87 (4.64)	8.12 (1.44)	11.73 (1.99)
Male	A	2	51.0 (7.07)	85.0 (14.12)	169.0 (12.73)	24.45 (2.19)	17.17 (1.61)
	B	5	44.6 (5.31)	86.24 (13.51)	177.60 (12.76)	17.18 (1.07)	13.65 (2.3)
	C	2	47.0 (7.07)	71.5 (12.02)	181.0 (1.41)	14.1 (0.84)	13.47 (1.23)

Table 4
Clustering Classification Contingency Table Analysis (n = 92). Cluster: number defined by the Euclidean K-means result with matrix A. Cluster 0 corresponds to better health status because of the higher modulus. Similarly, group one corresponds to patients with worse health status. n: Number of participants in each subgroup. Arithmetic means and standard deviation (in brackets) are shown for other parameters. Age in years, weight in kg, height in cm and VO2 peak in $mL^{*}Kg^{-1}{}^{*}min^{-1}$. **Modulus**: mean of modulus value of the subgroup.

Gender	Cluster	n	Age	Weight	Height	VO2 peak	Modulus
Female	0	32	53.12 (7.87)	64.96 (16.63)	159.75 (6.45)	16.50 (3.52)	15.06 (1.38)
	1	49	49.94 (8.28)	71.57 (14.57)	161.77 (5.17)	13.28 (3.94)	11.12 (1.59)
Male	0	8	47.50 (5.45)	83.37 (13.97)	178.00 (5.17)	18.56 (4.63)	14.79 (2.23)
	1	1	39.00 (-)	77.20 (-)	164.00 (-)	16.50 (-)	11.26 (-)

Table 5
Chi-square test results. All parameters defined in the method section.

Test	Lambda	Chi2	dof	*p*-value	Cramer	Power
Pearson	1.00	19.06	3	0.000266	0.46	0.96
Cressie-read	0.67	19.57	3	0.000208	0.46	0.97
Log-likelihood	0.00	22.29	3	0.000005	0.49	0.98

Finally, Chi-square tests were used to determine whether there is a statistically significant difference between the observed and expected values. Three tests were performed to analyse the independence of the variables, as shown in Table 5. It should be emphasised that the three tests (Pearson, Cressie-read and Log-likelihood) are positively correlated, as a p-value<0.01 indicates a clear relationship between the categories analysed (clusters and Weber's Classification). This implies that the model correctly predicts the CPET scores with a precision equal to or higher than 99%. As a general guideline for consistency, the observed and expected contingency tables should not have cells with frequencies less than 5.

Discussion

The purpose of this study was to determine whether there is a relationship between the results of self-administered SF-36 questionnaires and CPET oxygen consumption values. CPET results were analysed, focusing on peak VO2 as a determinant of functional capacity and stratified according to Weber's classification.[15] The one-day test has been examined in other pathologies such as cardiac and pulmonary,[30] and more recently in long COVID-19 syndrome.[31] Previous studies have observed decreased cardiovascular response and increased global and maximal heart rate in ME/CFS patients using the one-day exercise test, which may be pathology-specific, resulting in early-onset fatigue, dysfunctional exercise capacity, inconsistent response, or lack of motivation.[32] In contrast, other studies[33] suggest that reduced exercise capacity may be related to autonomic dysfunction, as ME/CFS patients have difficulty reaching their age-predicted maximal heart rates. This may be one of the reasons why their physical performance is impaired.[33] Data analysis shows that ME/CFS patients have lower cardiorespiratory fitness levels than healthy control subjects,[34] and another recent study revealed that results from an analysis of various factors obtained after a single CPET could be used as biomarkers for diagnosing ME/CFS.[35] Although the exact mechanisms associated with low exercise capacity in patients with ME/CFS have not been determined, a single CPET can be a reliable and accessible test for patients that provides objective physiological data on their response to exercise.

It has been suggested that ME/CFS could have an autoimmune etiology, as antibodies against beta2-adrenergic receptors (β2AdR) and muscarinic acetylcholine receptors (M3 AChR and M4 AChR) have been

identified in symptoms such as cognitive deficits, autonomic dysregulation, and immune activation.[36] To assess this, patients must undergo laboratory testing, which is considerably more costly to the healthcare system than filling in a SF-36 questionnaire. Furthermore, laboratory testing is not considered an optimal method for diagnosing a cognitive or physical impairment. The procedure that we propose, however, is based on analysing answers to a SF-36 questionnaire, which is easily accessible in primary care and has the advantage that only those patients with specific results (classified as cluster 1 in the proposed model) would be evaluated for referral to a dedicated unit.

An earlier study demonstrated that low VO2 max values are directly correlated with a subscale of the SF-36 questionnaire.[18] The differences in peak VO2 values between the male and female populations observed in our study are in line with those found in previous studies,[37,38] in that women presented lower peak VO2 values than men on the first CPET and a worse classification on the Weber scale. This is confirmed in the clustering analysis shown in Table 4.

Results of the study were based on a chi-square test of independence (see Table 5) using three tests with different lambda values and indicate clearly that worse physical condition on exertion corresponds to a worse response to the SF-36 questionnaire (p-value < 0.001, for the three tests). These results suggest that a CPET can be used as a biomarker to measure oxygen consumption and objectively assess the health status of patients diagnosed with ME/CFS. The relationship between the SF-36 and CPET results of this study may have important implications for clinical practice. The SF-36 questionnaire is readily available for all ME/CFS patient assessment services (both primary care and dedicated units). Using the questionnaire would help predict oxygen consumption and could initiate a referral from primary care to a dedicated unit to begin early multidisciplinary evaluation and measure patient progress after treatments. These results are consistent with other studies discussed in this paper and highlight the relationship between responses to quality of life questions the SF-36 questionnaire and CPET results from a machine learning perspective.

Limitations

The main limitation of our study is that it is performed in only one hospital centre with an unbalanced gender population. It would be appropriate to contrast our results with a larger number of participants from multiple centres.

Conclusions

Low oxygen consumption on a CPET could be considered a diagnostic biomarker in patients with ME/CFS. Findings of our study reveal a relationship between the SF-36 questionnaire and Weber's classification. Clustering datasets from health questionnaires such as the SF-36 should be performed on raw response data, which preserves the original information and improves the quality of the model. Adopting this procedure in primary care by analysing responses to a SF-36 questionnaire may help early referral of potential patients with ME/CFS to a dedicated unit (accredited by the Department of Health of the Generalitat de Catalunya requires a group of professionals, internists, nurses, physiatrists, psychologists and physiotherapists, who carry out multidisciplinary work, with the application of pharmacological and non-pharmacological treatments, such as programmed physical exercise and cognitive-behavioral therapy). Further research could focus on confirming the results presented in a multi-centre scenario.

Ethics Disclosure

Protection of human and animal subjects. We confirm that the procedures this study were in accordance with the regulations established by the Clinical Research and Ethics Committee and the World Medical Association's Declaration of Helsinki. The research protocols were approved by the Research Ethics Committee of the Vall d'Hebron University Hospital for both studies: 1) "Population-based Registry of Patients with Chronic Fatigue Syndrome", approved 18/10/2006; and 2) PR-AG9 01/2020 "Study of exercise intolerance (IEF) in adult patients with chronic fatigue syndrome (CFS) using a cardiopulmonary exercise stress test (CPET)", approved 22/09/2020.

Confidentiality of data. We confirm we have followed the protocols of their work centre on publication of patient data, right to privacy and informed consent. The authors obtained the informed consent of all patients and/or subjects mentioned in the article.

Authorship. We confirm that all named authors of the manuscript have participated in the conception and design of the study; collection, analysis, and interpretation of the data; and the submitted manuscript has been written, revised and approved by all of us.

Declaration of Conflicting Interests. We wish to confirm there are no potential conflicts of interest concerning this article's research, authorship, and/or publication.

Funding

This research received no specific grant from public, commercial, or non-profit funding agencies.

Acknowledgments

We wish to thank Amelia Marquino for her dedication and excellent work managing and maintaining the database. We are also grateful to the Fundación Mútua Madrileña for financing the creation of the database through a grant awarded in 2006 for the "Population-based Registry of Patients with Chronic Fatigue Syndrome" project.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.clinthera.2023.09.007.

References

1. Morris G, Puri BK, Walker AJ, Maes M, Carvalho AF, Walder K, et al. Myalgic encephalomyelitis/chronic fatigue syndrome: From pathophysiological insights to novel therapeutic opportunities. *Pharmacological Research*. 2019:104450. doi:10.1016/j.phrs.2019.104450.
2. Lim E-J, Ahn Y-C, Jang E-S, Lee S-W, Lee S-H, Son C-G. Systematic review and meta-analysis of the prevalence of chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME). *Journal of Translational Medicine*. 2020. doi:10.1186/s12967-020-02269-0.
3. Norris T, Collin SM, Tilling K, Nuevo R, Stansfeld SA, Sterne JA, et al. Natural course of chronic fatigue syndrome/myalgic encephalomyelitis in adolescents. *Arch Dis Child*. 2017;102:522–528. doi:10.1136/archdischild-2016-311198.
4. Wu Q, Gao J, Bai D, Zhong Y, Yang Z, Jiang X. Prevalence of chronic fatigue syndrome in China: a meta-analysis. Youjiang Med J.
5. Lloyd AR, Hickie I, Boughton CR, Spencer O, Wakefield D. Prevalence of chronic fatigue syndrome in an Australian population. *Med J Aust*. 1990;153:522–528. doi:10.5694/j.1326-5377.1990.tb126191.x.
6. Committee on the Diagnostic Criteria for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome, Board on the Health of Select Populations, Institute of Medicine. Background. National Academies Press (US); 2015. Available: https://www.ncbi.nlm.nih.gov/books/NBK284897/.
7. Maes M, Twisk FN. Why myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) may kill you: disorders in the inflammatory and oxidative and nitrosative stress (IO&NS) pathways may explain cardiovascular disorders in ME/CFS. *Neuro Endocrinol Lett*. 2009;30:677–693. Available. https://www.ncbi.nlm.nih.gov/pubmed/20038921.
8. Wirth KJ, Scheibenbogen C, Paul F. An attempt to explain the neurological symptoms of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *J Transl Med*. 2021;19:471. doi:10.1186/s12967-021-03143-3.
9. Fernández-Quirós J, Lacasa-Cazcarra M, Alegre-Martín J, Sanmartín-Sentañes R, Almirall M, Launois-Obregón P, et al. The Conners Continuous Performance Test CPT3TM: Is it a reliable marker to predict neurocognitive dysfunction in Myalgic encephalomyelitis/chronic fatigue syndrome? *Front Psychol*. 2023;14:1127193. doi:10.3389/fpsyg.2023.1127193.
10. Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Ann Intern Med*. 1994;121:953–959. doi:10.7326/0003-4819-121-12-199412150-00009.

11. Christley Y, Duffy T, Martin CR. A review of the definitional criteria for chronic fatigue syndrome. *J Eval Clin Pract*. 2012;18:25–31. doi:10.1111/j.1365-2753.2010.01512.x.
12. Carruthers BM, Jain AK, De Meirleir KL, Peterson DL, Klimas NG, Martin Lerner A, et al. Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *Journal of Chronic Fatigue Syndrome*. 2003:7–115. doi:10.1300/j092v11n01_02.
13. Carruthers BM, van de Sande MI, De Meirleir KL, Klimas NG, Broderick G, Mitchell T, et al. Myalgic encephalomyelitis: International Consensus Criteria. *J Intern Med*. 2011;270:327–338. doi:10.1111/j.1365-2796.2011.02428.x.
14. Stevens S, Snell C, Stevens J, Keller B. VanNess JM. Cardiopulmonary Exercise Test Methodology for Assessing Exertion Intolerance in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *Front Pediatr*. 2018;6:242. doi:10.3389/fped.2018.00242.
15. Weber KT, Janicki JS. Cardiopulmonary exercise testing for evaluation of chronic cardiac failure. *Am J Cardiol*. 1985;55:22A–31A. doi:10.1016/0002-9149(85)90792-1.
16. Asprusten TT, Sletner L, Wyller VBB. Are there subgroups of chronic fatigue syndrome? An exploratory cluster analysis of biological markers. *J Transl Med*. 2021;19:48. doi:10.1186/s12967-021-02713-9.
17. Kujawski S, Słomko J, Newton JL, Eaton-Fitch N, Staines DR, Marshall-Gradisnik S, et al. Network Analysis of Symptoms Co-Occurrence in Chronic Fatigue Syndrome. *Int J Environ Res Public Health*. 2021;18. doi:10.3390/ijerph182010736.
18. van Campen CLMC, Rowe PC, Visser FC. Validation of the Severity of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome by Other Measures than History: Activity Bracelet, Cardiopulmonary Exercise Testing and a Validated Activity Questionnaire: SF-36. *Healthcare (Basel)*. 2020;8. doi:10.3390/healthcare8030273.
19. Alonso J, Prieto L, Antó JM. The Spanish version of the SF-36 Health Survey (the SF-36 health questionnaire): an instrument for measuring clinical results. *Med Clin*. 1995;104:771–776. Available. https://www.ncbi.nlm.nih.gov/pubmed/7783470 .
20. Wasserman K, Hansen JE, Sue DY, Stringer WW, Whipp BJ. Principles of exercise testing and interpretation: including pathophysiology and clinical applications. *Med Sci Sports Exercise*. 2005;37:1249.
21. Mezzani A. Cardiopulmonary Exercise Testing: Basics of Methodology and Measurements. *Ann Am Thorac Soc*. 2017;14:S3–S11. doi:10.1513/AnnalsATS.201612-997FR.
22. Moses T, Holland PW. A comparison of statistical selection strategies for univariate and bivariate log-linear models. *Br J Math Stat Psychol*. 2010;63:557–574. doi:10.1348/000711009X478580.
23. Vallat R. Pingouin: statistics in Python. *J Open Source Softw*. 2018;3:1026. doi:10.21105/joss.01026.
24. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd Edition. Routledge; 1988. doi:10.4324/9780203771587.
25. Bengfort B, Gray L, Bilbro R, Roman P, Deziel P, McIntyre K, et al. Yellowbrick v1.5. 2022. doi:10.5281/zenodo.7013541
26. Kozak M. A Dendrite Method for Cluster Analysis" by Caliński and Harabasz: A Classical Work that is Far Too Often Incorrectly Cited. *Communications in Statistics - Theory and Methods*. 2012;41:2279–2280. doi:10.1080/03610926.2011.560741.
27. Halkidi M, Batistakis Y, Vazirgiannis M. On Clustering Validation Techniques. *J Intell Inf Syst*. 2001;17:107–145. doi:10.1023/A:1012801612483.
28. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math*. 1987;20:53–65. doi:10.1016/0377-0427(87)90125-7.
29. Novikov A. PyClustering: Data Mining Library. *J Open Source Softw*. 2019;4:1230. doi:10.21105/joss.01230.
30. Goulart C da L, Dos Santos PB, Caruso FR, Arêas GPT, Marinho RS, Camargo P de F, et al. The Value of Cardiopulmonary Exercise Testing in Determining Severity in Patients with both Systolic Heart Failure and COPD. *Sci Rep*. 2020;10:4309. doi:10.1038/s41598-020-61199-5.
31. Ferreira EVM, Oliveira RKF. Mechanisms of exercise intolerance after COVID-19: new perspectives beyond physical deconditioning. *J Bras Pneumol*. 2021;47:e20210406. doi:10.36416/1806-3756/e20210406.
32. Inbar O, Dlin R, Rotstein A, Whipp BJ. Physiological responses to incremental exercise in patients with chronic fatigue syndrome. *Med Sci Sports Exerc*. 2001;33:1463–1470. doi:10.1097/00005768-200109000-00007.
33. De Becker P, Roeykens J, Reynders M, McGregor N, De Meirleir K. Exercise capacity in chronic fatigue syndrome. *Arch Intern Med*. 2000;160:3270–3277. doi:10.1001/archinte.160.21.3270.
34. Zambolin F, Duro-Ocana P, Faisal A, Bagley L, Gregory WJ, Jones AW, et al. Fibromyalgia and Chronic Fatigue Syndromes: A systematic review and meta-analysis of cardiorespiratory fitness and neuromuscular function compared with healthy individuals. *PLoS One*. 2022;17:e0276009. doi:10.1371/journal.pone.0276009.
35. Pifarré F, Rosselló L, Hileno R, Palmi J, Bañeres L, Planas A, et al. The use of oxygen as a possible screening biomarker for the diagnosis of chronic fatigue. *Apunts Sports Medicine*. 2022;57:100379. doi:10.1016/j.apunsm.2022.100379.
36. Gravelsina S, Vilmane A, Svirskis S, Rasa-Dzelzkaleja S, Nora-Krukle Z, Vecvagare K, et al. Biomarkers in the diagnostic algorithm of myalgic encephalomyelitis/chronic fatigue syndrome. *Front Immunol*. 2022;13:928945. doi:10.3389/fimmu.2022.928945.
37. van Campen C (linda) MC, Rowe PC, Visser FC. Two-Day Cardiopulmonary Exercise Testing in Females with a Severe Grade of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Comparison with Patients with Mild and Moderate Disease. *Healthc Pap*. 2020;8:192. doi:10.3390/healthcare8030192.
38. van Campen CLMC, Visser FC. Comparing Idiopathic Chronic Fatigue and Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) in Males: Response to Two-Day Cardiopulmonary Exercise Testing Protocol. *Healthcare (Basel)*. 2021:9. doi:10.3390/healthcare9060683.

Clinical Therapeutics 45 (2023) 1228–1235

Contents lists available at ScienceDirect

Clinical Therapeutics

journal homepage: www.elsevier.com/locate/clinthera

Original Research

Unsupervised Cluster Analysis Reveals Distinct Subtypes of ME/CFS Patients Based on Peak Oxygen Consumption and SF-36 Scores

Marcos Lacasa[1,*], Patricia Launois[5], Ferran Prados[1,2,3,4], José Alegre[5], Jordi Casas-Roma[1]

[1] e-Health Center, Universitat Oberta de Catalunya, Barcelona, Spain
[2] Center for Medical Image Computing, University College London, London, United Kingdom
[3] National Institute for Health Research Biomedical Research Centre at UCL and UCLH, London, United Kingdom
[4] Queen Square MS Center, Department of Neuroinflammation, UCL Institute of Neurology, Faculty of Brain Sciences, University College London, London, United Kingdom
[5] Myalgic Encephalomyelitis / Chronic Fatigue Syndrome Unit, Division of Rheumatology, Vall d'Hebron Hospital Research Institute Universitat Autònoma de Barcelona, Barcelona, Spain

SUPPLEMENTARY MATERIAL

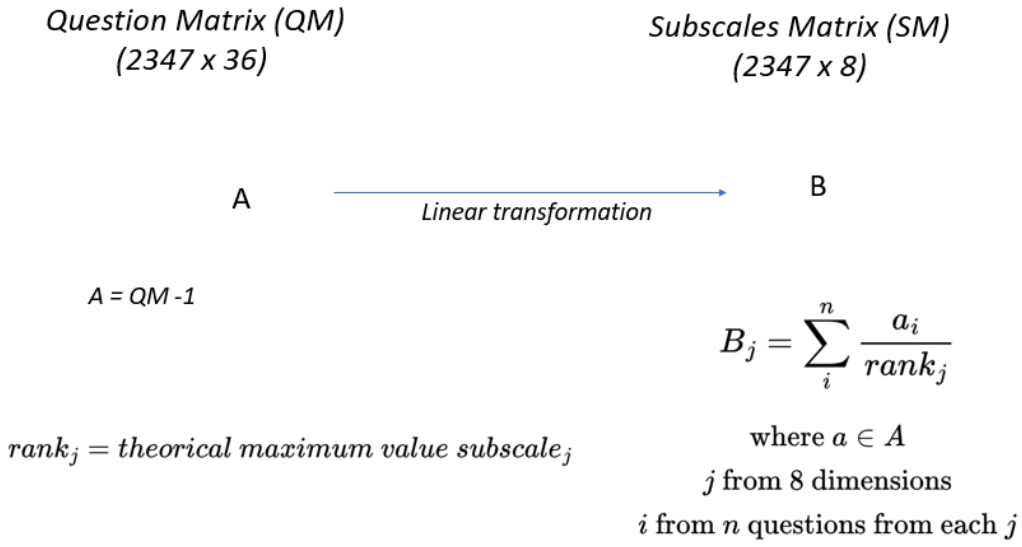
Mathematical arguments based on the work performed and technical specifications on the use of artificial intelligence algorithms

Methods

Linear application of ordered sets

Responses on the SF-36 range from 1 to 6, as shown in **Table 1**, with higher scores corresponding to better health. A subscale is a set of questions related to a specific condition. Therefore, we can define a set of ordered responses for each individual as an ordered 36-dimensional vector. Similarly, subscales define an ordered 8-dimensional vector for each individual. **Figure S1** details the process of computing the set of 8-dimensional vectors of the subscales (matrix B) from the set of 36-dimensional vectors of the responses of SF-36 questionnaires (matrix A) through a linear application. Note that matrix A is rescaled to be in the range between 0 and 5, i.e., the minimum value is 0 instead of 1.

Figure S1.- Diagram of the transition from matrix A to B through a linear application.



While maintaining its properties, this linear application does not guarantee bijectivity since two different 36-dimensional vectors of matrix A could generate the same 8-dimensional vector in matrix B. However, it is relevant to underline that this transformation maintains the properties of the linear application.

The clustering analysis

The clustering analysis was implemented in Python (version 3.7.14). The decoded SF-36 answers and subscale matrices were used and compared. The dimensions were 2347x36 and 2347x8, respectively. To select the optimal number of clusters, some models were fitted with values in the range [2,6] for k (Birch and spectral clustering) by the elbow method. (Bengfort et al. 2022) and the Calinski and Harabasz metric (see **Stable 1**). Three validation metrics are proposed using the scikit-learn package (version 1.0.2) to evaluate the performance of each tested model when the truth labels are unknown:

- Silhouette Coefficient
- Calinski–Harabasz Index (Kozak 2012)
- Davies–Bouldin Index (Halkidi, Batistakis, and Vazirgiannis 2001)

STable 1.- Clustering Algorithms tested

Algorithm	Parameters optimized	Package used
K-means	• Initializer: K-Means++. This method was used to find out optimal initial centers. • Metric distance: Euclidean and Manhattan have been compared. • Optimized using Silhouette score (Rousseeuw 1987) by 200 runs.	Pyclustering (v 0.10.1.2) (Novikov 2019)
Agglomerative	• Links: Centroid, single, complete, and average links were tested, and the maximum silhouette score was chosen	pyclustering
Birch	• Branching_factor and threshold. • Optimized by grid-search using maximum silhouette score.	scikit-learn
DBSCAN	• Eps (The maximum distance between two samples, for one to be considered as in the neighborhood of the other).	scikit-learn

STable 1.- Clustering Algorithms tested

Algorithm	Parameters optimized	Package used
	• Minimum samples. • Left size (this can affect the speed of the construction and query and the memory required to store the tree) using maximum silhouette score.	
K-MEDOIDS	• Initializer: k-medoids++. • Optimized using Silhouette score by 200 runs.	scikit-learn
Fuzzy-C	• Initializer: K Means++. Farthest Centre Candidate option.	pyclustering

Principal Components Analysis

As mentioned before, the set of SF-36 questionnaires consists of 2,347 records or individuals. For each question, we examine the normality and correlation between all pairs of variables. In addition, we use Principal Component Analysis (PCA) to summarize and visualize the interaction between the characteristics. Once the standardization is complete, the analysis is complete and is illustrated by the factor map in **Figure S7**:

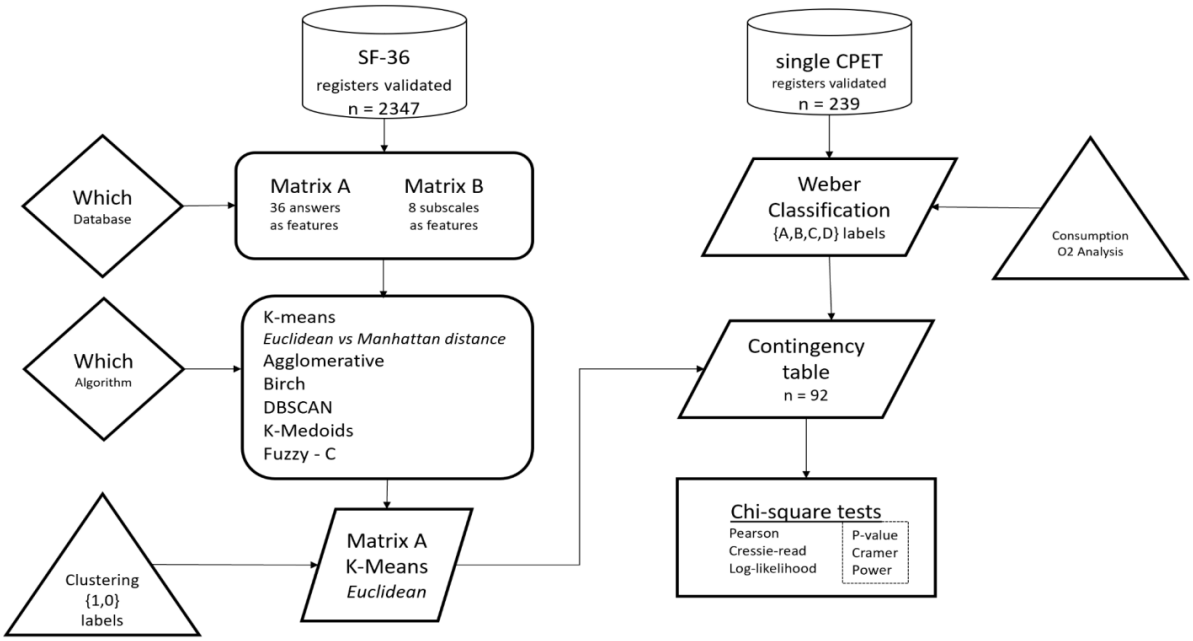
- Eigenvalues and variances for each feature.
- Plot analysis, where positively correlated vectors are grouped in the same quadrant, and negatively correlated vectors are positioned on opposite quadrants.
- Vectors far from the origin (i.e., the center of the coordinate system) are well represented on the factor map.

The cosine square (cos^2) method is computed to measure the quality of each feature, where a high value indicates a good representation of the variable on the principal component and is close to the circumference of the correlation circle (radius equal to 1). This analysis is performed using R (v 4.2.1) with packages factoextra (v 1.0.7) and factominer (v 2.6) (Kassambara 2017).

Contingency table difference analysis

A contingency table was created for patients who completed the single CPET test and the SF-36 questionnaire. We selected 92 patients who met the criteria **(Figure S2)** from 239 single CPET registries. The contingency table is constructed using the clustering labels and Weber's classification based on VO2 levels.

Figure S2.- Steps for the contingency table.

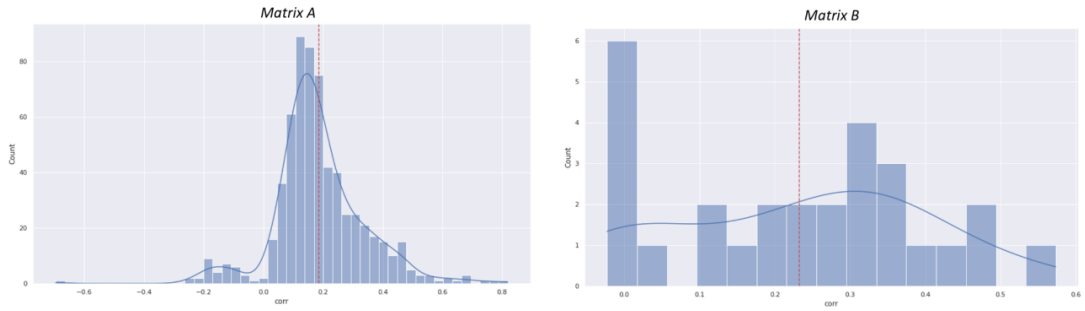


Results

Matrix Analysis

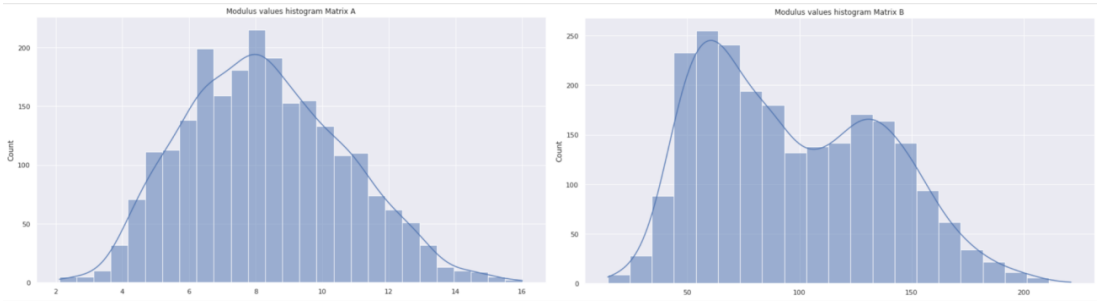
The d'Agostino and Pearson tests are used to analyze the normality of each SF-36 question response and its aggregated subscales (D'agostino & Pearson, 1973). Pearson's linear correlation is calculated between all pairs of variables, and the histogram of the results is shown in **Figure S3**, where the maximum value is 0.8 in the RE1-RE2 pair corresponding to the emotional role. None of the distributions fits a normal distribution.

Figure S3.- Histograms of the series of pairs of linear correlations of variables by matrix. The number of correlations is x/2 (x-1), being x the number of features of each matrix. The red dotted line represents the mean value.



Each data set can be interpreted as a vector, and as mentioned above, it is an ordered set of vectors that represent the state of health of the patient. The modulus of each vector is calculated to create a new ordered set of values. The histograms of these ordered sets are shown below in **Figure S4**.

Figure S4.- Histograms of vector's modulus by matrix.



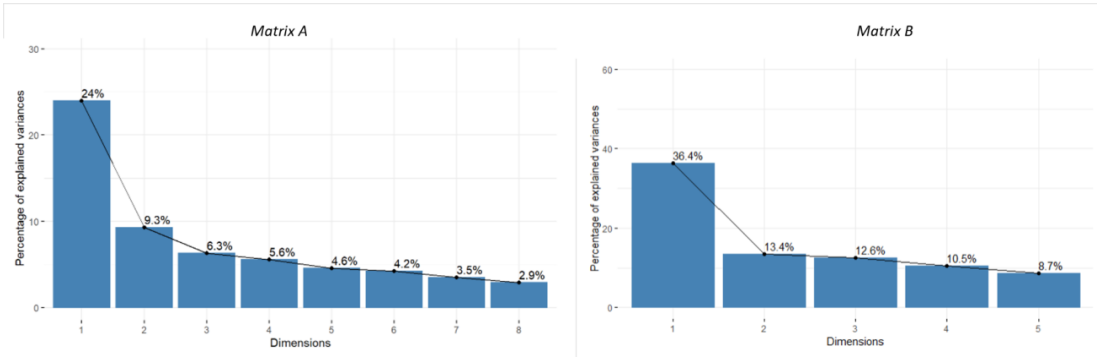
The modulus histograms are not equivalent to matrix A and matrix B. The order of each register was analyzed by percentile performance. If the order is similar, the percentiles would be similar as well. We analyzed the difference in the percentiles of each record in the two matrices, giving an error margin of ± 0.1. We noted that the records with the higher margin of error, 40.14% of the records, indicate a change in the percentile < 0.1 in both matrices, which concludes that the order changes in both matrices. Furthermore, a very differentiated histogram structure is observed, showing a distribution compatible with normality in matrix A and a significant

concentration in lower values (patients with worse health status) between values of 30-70 and less when the modulus value is more significant than 100 in matrix B.

PCA analysis

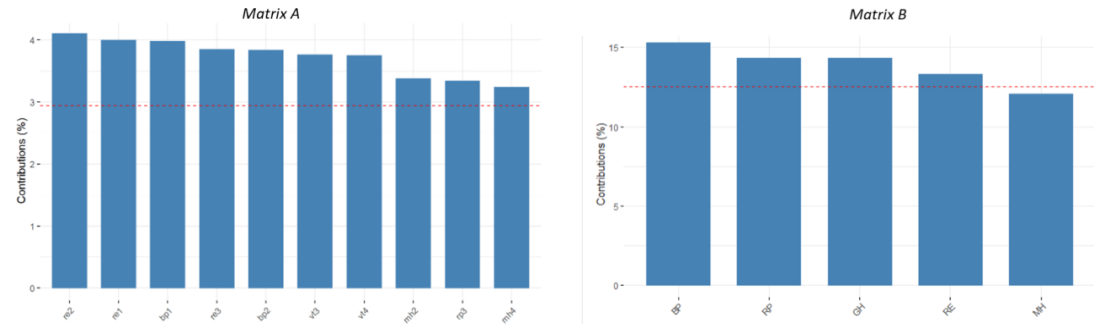
The eigenvalues measure the amount of variation retained by each principal component. The variance contributed by each dimension is calculated with a PCA of 8 components for matrix A and five components for matrix B, as shown in **Figure S5**.

Figure S5.- Percentage of explained variances by matrix.



Matrix A needs eight dimensions or components (reduced from 36) to explain 60.4% of the variance, and matrix B needs five dimensions (reduced from 8) to explain 81.6%. The contributions of these variables for the variability in a given principal component are expressed in percentages. Variables that are correlated with any dimension are the most important in explaining the variability in the dataset. The contribution of variables is shown in **Figure S6**.

Figure S6.- Contribution of variables by Matrix. The contribution is calculated for an 8-dimensional PCA in matrix A and a 5-dimensional PCA in matrix B. The red dashed line on the graph above indicates the expected average contribution. In matrix A, each column represents the name of each response. In matrix B, each column represents the name of each subscale.



Critical variables have larger contributions. Significant differences are observed between the matrices in the representativeness of the variance of the dimensions. In Matrix A, the variables related to the emotional role (ER) have the most significant influence. On the other hand, in matrix B, somatic pain (BP) contributes the most, with two variables (bp1, bp2) among those that contribute the most in matrix A. The physical part significantly influences Matrix B, with the scales of Body Pain (BP) and Physical Role (PR).

The correlation plot of variables shows the relationship between all pairs of variables, with positively correlated variables grouped in the same quadrant and negatively correlated variables in opposite quadrants. The distance between the variables and the origin measures the quality of the variables, where a greater distance from the origin of the coordinates implies better quality. The cos^2 value indicates the goodness of the variable's representativeness. In **Figure S7**, we depict the most relevant vectors, i.e., the vectors with $cos^2 > 0.45$.

Figure S7.- Correlation plot of variables. Note that the display takes the two main dimensions of each matrix, which contribute 33.3% of the variance in the case of matrices A and 39.8% in the case of matrices B. The color indicates the contribution to the variance. The color indicates the contribution to the variance. The redder the color, the more significant the variance contributed. Each column represents the name of each response in Matrix A. Each column represents a subscale name in Matrix B.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 27 February 2023
DOI 10.3389/fpsyg.2023.1127193

OPEN ACCESS

EDITED BY
Fabrizio Stasolla,
Giustino Fortunato University,
Italy

REVIEWED BY
Khalida Akbar,
Durban University of Technology,
South Africa
Maurel Sara,
Autonomous University of Barcelona,
Spain

*CORRESPONDENCE
Judith Fernández-Quirós
✉ judith.fquiros@vallhebron.cat

SPECIALTY SECTION
This article was submitted to Neuropsychology, a section of the journal Frontiers in Psychology

RECEIVED 19 December 2022
ACCEPTED 07 February 2023
PUBLISHED 27 February 2023

CITATION
Fernández-Quirós J, Lacasa-Cazcarra M, Alegre-Martín J, Sanmartín-Sentañes R, Almirall M, Launois-Obregón P, Castro-Marrero J, Rodríguez-Urrutia A, Navarro-Sanchis JA and Ramos-Quiroga JA (2023) The Conners Continuous Performance Test CPT3™: Is it a reliable marker to predict neurocognitive dysfunction in Myalgic encephalomyelitis/chronic fatigue syndrome? *Front. Psychol.* 14:1127193. doi: 10.3389/fpsyg.2023.1127193

The Conners Continuous Performance Test CPT3™: Is it a reliable marker to predict neurocognitive dysfunction in Myalgic encephalomyelitis/chronic fatigue syndrome?

Judith Fernández-Quirós[1,2]*, Marcos Lacasa-Cazcarra[3], Jose Alegre-Martín[1], Ramón Sanmartín-Sentañes[1], Miriam Almirall[1], Patricia Launois-Obregón[1], Jesús Castro-Marrero[1], Amanda Rodríguez-Urrutia[2,4,5,6], Jose A. Navarro-Sanchis[2,4] and J. Antoni Ramos-Quiroga[2,4,5,6]

[1]Myalgic Encephalomyelitis/Chronic Fatigue Syndrome Unit, Division of Rheumatology, Vall d'Hebron Hospital Research Institute Universitat Autònoma de Barcelona, Barcelona, Spain, [2]Department of Mental Health, Hospital Universitari Vall d'Hebron, Barcelona, Spain, [3]e-Health Center, Universitat Oberta de Catalunya, Barcelona, Spain, [4]Group of Psychiatry, Mental Health and Addictions, Vall d'Hebron Research Institute (VHIR), Barcelona, Spain, [5]Biomedical Network Research Centre on Mental Health (CIBERSAM), Barcelona, Spain, [6]Department of Psychiatry and Forensic Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain

Introduction: The main objective is to delimit the cognitive dysfunction associated with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) in adult patients by applying the Continuous Performance Test (CPT3™). Additionally, provide empirical evidence on the usefulness of this computerized neuropsychological test to assess ME/CFS.

Method: The final sample (*n* =225; 158 Patients/67 Healthy controls) were recruited in a Central Sensitization Syndromes (CSS) specialized unit in a tertiary hospital. All participants were administered this neuropsychological test.

Results: There were significant differences between ME/CFS and healthy controls in all the main measures of CPT3™. Mainly, patients had a worse indicator of inattentiveness, sustained attention, vigilance, impulsivity, slow reaction time, and more atypical T-scores, which is associated with a likelihood of having a disorder characterized by attention deficits, such as Attention Deficit Hyperactivity Disorder (ADHD). In addition, relevant correlations were obtained between the CPT3™ variables in the patient's group. The most discriminative indicators of ME/CFS patients were Variability and Hit Reaction Time, both measures of response speed.

Conclusion: The CPT3™ is a helpful tool to discriminate neurocognitive impairments from attention and response speed in ME/CFS patients, and it could be used as a marker of ME/CFS severity for diagnosing or monitoring this disease.

KEYWORDS

neuropsychological test, cognitive impairments, continuous performance test, CPT3™, neurocognitive dysfunction, Central Sensitization Syndromes, chronic fatigue syndrome

Introduction

Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) is a complex multisystem disease that presents a chronic course with periods of symptomatic exacerbation frequently related to acute stress. This condition predominantly affects women and is a severe functional disorder (Afari and Buchwald, 2003; Prins et al., 2006). The prevalence is estimated to be between 0.2 and 2.6% of the general population (Reyes et al., 2003; Nacul et al., 2011). The nuclear symptoms are chronic central fatigue (>6 months) and post-exertional malaise with a recovery time longer than 24 h of idiopathic origin. The international diagnostic criteria were established by the Centres for Disease Control (CDC) in Atlanta (Georgia) in 1994 (Fukuda et al., 1994). Furthermore, the heterogeneity of symptoms inME/CFS is studied in clusters from 2003 as a complement to the CDC's criteria (neurological, muscle, cognitive, neurovegetative, and immunological; Fukuda et al., 1994; Carruthers et al., 2003). In 2011, these criteria were updated, and post-exertional exhaustion was proposed as a disease hallmark (Carruthers et al., 2011). ME/CFS is associated with different comorbid phenomena (anxiety-depressive disorders, fibromyalgia (FM), sicca syndrome, regional myofascial pain syndrome, plantar fasciitis, degenerative or mechanical disk disease, and tendinopathy of the shoulder; Ruiz et al., 2011 Castro-Marrero et al., 2016) that are more prevalent in ME/CFS patients than in non-CFS individuals (Rivera et al., 2006). At present, the most widely accepted hypothesis for the pathogenesis of ME/CFS characterizes it as a genetic-based process with different triggering factors and subsequent neuroimmunology and immunoinflammatory dysfunction (Bassi et al., 2008).

Fatigue is the most common symptom associated with chronic diseases and is experienced as really distressing. Nature fatigue is a subjective state with both physical and psychological elements, and there is a lack of effective treatments for it. New methods are being developed to quantify fatigue and are increasing in clinical settings (Swain, 2000). Fatigue may occur due to a physical or psychological event, or fatigue may cause a physical event. The concept of fatigue appears to be defined by overlapping terms of cognitive or mental fatigue; however, the description for each slightly varies. For example, cognitive fatigue is said to occur when cognitive performance decreases by engaging in tasks requiring sustained activity (Morrow et al., 2015). Mental fatigue has been defined as a subjective feeling of tiredness and inertia that occurs during extended periods of demanding cognitive activity (Badin et al., 2016). Fatigue has also been defined as physiological fatigue, which is described as muscle weakness that may occur due to exercise (Prinsen et al., 2015). There is insufficient evidence examining the relationship between fatigue and cognitive impairments in patients with chronic illnesses. More studies are needed to examine the potential relationships between these two symptoms to develop effective treatments for individuals living with NCDs (Chronic Non-Communicable Diseases; Menzies et al., 2021).

Cognitive deficits are common in ME/CFS patients and limit their quality of life and psychological well-being. The discrepancy between self-reported common cognitive deficits and results obtained by objective neuropsychological tests has been reported (Rasouli et al., 2019). This disparity could be explained by several reasons, mainly methodological (Cockshell and Mathias, 2010), suggesting establishing an appropriate method to avoid it. Patients with higher levels of fatigue, pain, and depression reported greater subjective cognitive difficulties, and those with greater pain were associated with lower objective working memory function. ME/CFS patients primarily had psychomotor speed and attention problems, measured by objective neuropsychological tests (Rasouli et al., 2019). Other studies confirm that cognitive deficits in ME/CFS depend mainly on compromised attention, memory, and reaction time, but motor functioning, vocabulary, reasoning, and global deficits are absent (Cockshell and Mathias, 2010).

Cognitive dysfunction has been described in patients with depression and anxiety, and it is difficult to determine whether this is due to psychopathological comorbidity or fatigue. Slowing information processing speed occurs, especially in complex tasks requiring sustained attention. However, the underlying mechanisms of the manifested cognitive dysfunction remain unclear (Cvejic et al., 2016), advising a need for appropriate neuropsychological assessment tools in this complex disease capable of overcoming such limitations. One study showed that the dysfunction of information processing speed is independent of depressive symptoms in ME/CFS, suggesting that attentional deficits may be primary to memory problems, which could show an underlying neurological basis for the attentional dysfunction in these patients (Santamarina-Perez et al., 2014). Unfortunately, there are currently no commercially available diagnostic tests, specific lab biomarkers, or targeted FDA-approved ME/CFS drugs (Castro-Marrero et al., 2017). For all these arguments, and given the disparity of the published studies, there is no consensus on the results regarding cognitive dysfunction in ME/CFS and its repercussions. Therefore, more research is needed on this symptom in ME/CFS and the detection of the underlying mechanisms. The limitations found in previous studies on this topic must be considered to improve their research and obtain more specific neuropsychological assessment tools to evaluate it (Beaumont et al., 2012). General dysfunction in attention and information processing speed have been suggested as reasons for ME/CFS memory complaints (Deluca et al., 2004; Dickson et al., 2009).

Considering the evidence of neuropsychological dysfunction in ME/CFS patients, this work is carried out using a computerized test that allows comprehensive measurement of different cognitive domains of interest in this disease. Therefore, this is the first study in ME/CFS patients using the CPT3™. This neuropsychological test has been widely used to objectively and comprehensively assess cognitive functioning in patients diagnosed with Attention-Deficit Hyperactivity Disorder (ADHD; Fasmer et al., 2016; Berger et al., 2017).

CPT3™ allows evaluating the participant's performance in just 14 min, contemplating the "fatigue" variable (the core symptom of this disease) and other multiple cognitive measures. It is not only interesting because of the novelty of introducing this software tool in patients with ME/CFS, but also because it allows the evaluation of multiple indicators of cognitive functioning such as attentional capacity and some executive functions (inhibition and information processing speed) trying to find a differential clinical profile between patients and healthy controls (Kao and Thomas, 2010).

Method

Sample

The study subjects were composed of two groups, healthy controls and patients diagnosed with ME/CFS. Patients with a diagnosis of

ME/CFS have been recruited consecutively from the Central Sensitization Syndromes (SCC) specialized unit at the Vall d'Hebron University Hospital in Barcelona from July 2021 to March 2022. Inclusion criteria were fulfilled the Fukuda et al. (1994) and Carruthers criteria for ME/CFS(2), age between 18 and 65 years, the subjective clinical manifestation of cognitive impairment according to Carruthers et al. (2003, 2011), understanding and acceptance of participation in the study and signed written informed consent. The research protocol was approved by the ethics committee with this registration code(PR(AG)257/2021). The patients underwent an extensive medical examination by specialist physicians belonging to the Central Sensitization Syndromes (CSS) experienced Unit of the Rheumatologist Department at this hospital. The exclusion criteria were difficulty in understanding and/or completing the self-reported questionnaires, the presence of severe unstable psychiatric disorders (such as psychotic episodes, major depressive episodes, manic episodes, and anorexia nervosa), history of neurological disorders with cognitive impairment (such as severe head-brain trauma), presence of another systemic disease with cognitive alterations and presence of the disease entity Mild Cognitive Impairment (MCI). All patients diagnosed with ME/CFS included in the study were recorded on a data collection sheet in encrypted form and subsequently included in a database. The results obtained from the sample through the CPT3™ software were recorded in an excel table. The minimum sample calculation to demonstrate a statistically significant difference between the mean of the measurements of the cognitive domain variable, information processing speed (HRT), according to Student's *t*-test, is 92 patients with a confidence level equal to 0.95. Pearson's coefficient, which measures the variance to the mean, is 20%.

Of the total initial sample, four patients were excluded because of an associated severe psychiatric or neurological comorbidity at that precise moment that could account for some of the ME/CFS symptoms (three had a major depressive episode, and one patient with severe head-brain trauma). In addition, six patients were also excluded because they withdrew on their own initiative, indicating fear of fatigue from the mental overexertion of the test.

Neuropsychological evaluation

In the first session, the participants were individually examined by the ME/CFS specialist physician and were given questionnaires to fill in. The perception of disabling fatigue, sleep problems, and health-related quality of life using self-administered questionnaires: the fatigue impact scale FIS40 (Fisk et al., 1994), FIS8 (Fisk and Doble, 2002), PSQI (Verster et al., 2008), the Short Form Health Survey (SF-36; Alonso et al., 1995), the Symptom Checklist-90-revised (SCL 90 R) psychological inventory (McGregor et al., 1997) and hospital anxiety and depression scale (HAD; Castresana et al., 1995). The interviews were conducted by two internists and one rheumatologist physician skilled in diagnosing and treating this syndrome.

The neuropsychological evaluation was performed in a second appointment. It consisted of a single standardized individualized examination in which cognitive functioning was assessed using the latest version of the 3rd Conners Continuous Performance Test CPT3™ (Kao and Thomas, 2010). The participants had never been administered the CPT3™ to avoid familiarity. They were required to respond when any letter except "X" appeared on the monitor. CPT3™ was performed from 9:00 am to 2:00 pm, always in the same medical office under standard temperature, noise, and lighting conditions, and with mobile turned off to avoid interference. It was decided to establish this time slot for the evaluation due to several reasons (for example, attentional, and cognitive performance is not as optimal at certain times of the day, and fatigue accumulates as the day progresses both in the clinical population and in healthy controls as a general rule). The evaluation protocol was not modified for any patient.

CPT3™ is a computerized task of continuous execution based on performance to assess primarily attention-related impairments in individuals aged 8 years and older. The administration time is 14 min, during which the subject must maintain attention to perform this task. There are six blocks of trials, with three sub-blocks, each consisting of 20 trials. Each stimulus appears on the screen with varying time frequencies (inter-stimulus intervals from 1, 2, and 4 s), which allows for comparison of the subject's attention and response speed according to the time intervals. This software makes it possible to compare the changes in performance that may be experienced in separate segments (block measurement), thus checking whether the level of vigilance fluctuates during this simple task. Its duration allows for controlling the "fatigue effect" on cognitive performance, a key aspect to consider in patients diagnosed with CFS. Before starting, some instructions must be given. Individuals are seated in front of a computer. Each participant is required to respond (pressing the space bar on the keyboard) when any letter (target stimuli), except the letter "X" (non-target stimuli), appears on the screen. It is essential to warn them that they must keep responding until the end of the test to obtain the computer-generated reports describing the respondent's performance in detail. Once the subjects have understood the instructions, they perform the test. Each participant's final report was recorded in an Excel spreadsheet. Subsequently, the neurocognitive profile of the whole sample was studied by analyzing all the results obtained (Kao and Thomas, 2010).

Main features of the CPT3™

The CPT3™ test offers already standardized and raw scores to determine not only the overall performance of the evaluation but also specify different types of attention deficits (e.g., inattentiveness, impulsivity, sustained attention, vigilance) and response speed, allowing a comprehensive assessment. So, multiple features are measured in The CPT3™.

A brief description of the main measures used in this study is exposed below:

- Response speed: to obtain information performance about motor reaction time and information processing speed measured by the software (HRT, HRTSD, HRT block change, …).
- Impulsivity: is an indicator of the response inhibition capacity of the evaluated and includes a faster than normal HRT and a higher than average rate of commissions and/or perseverations.
- Inattentiveness: scores about focused attention. This indicator relates to poor detectability, a high percentage of omissions and commissions, a slow HRT, and high levels of inconsistency in response speed.
- Sustained attention: is defined as the respondent ability to maintain attention as the administration progresses. A decrease

in sustained attention across time is captured by atypical slowing in the respondent's HRT and by increases in omissions and commissions in later blocks of the administration.

- Vigilance: relates to the respondent's performance at varying levels of stimulus frequency and is defined by the respondent's ability to maintain a performance level even when the task rate is slow. It is captured by changes in the respondent's HRT, as indicated by the variables HRT ISI.
- Detectability: discrimination between non-targets and targets.
- Omissions: are missed targets and are generally an indicator of inattentiveness.
- Commissions: are incorrect responses to non-targets. High commission error rates may indicate either inattentiveness or impulsivity, depending on the respondent's HRT.
- Perseverations are responses made in less than 100 ms following the presentation of a stimulus. Perseverations may be related to impulsivity or an extremely liberal response style.
- Hit reaction time (HRT): the mean response speed, measured in milliseconds, for all non-preservative responses made during the entire administration. An atypically slow HRT may indicate inattentiveness but may also result from a conservative response style. So, HRT is also affected by response style.
- HRT standard deviation (HRT SD): measures the consistency of response speed for the entire administration. A high HRT SD indicates a greater response speed inconsistency, sometimes indicative of inattentiveness.
- Variability: is a measure of response speed consistency; however, Variability is a "within respondent" measure (i.e., the amount of variability the respondent showed in 18 separate sub-blocks of the administration with the overall HRT SD score). High response speed variability indicates that the respondent's attention and information processing efficiency varied throughout the administration.
- HRT block change: the slope of change in HRT across the six blocks of the administration. A positive slope indicates decelerating HRT as the administration progressed; a negative slope indicates accelerating HRT; a flat slope indicates no change in HRT.
- HRT Inter-Stimulus Interval (ISI) change: is the slope of change in reaction time across the three ISIs (1, 2, and 4 s). A positive slope indicates decelerating HRT at longer intervals, whereas a negative slope indicates accelerating HRT at longer intervals.

Based on the respondent's pattern and scores in each attentional dimension, the software identifies the presence and severity of the kinds of attention problems the respondent is most likely having. Therefore, CPT3™ provides in the outcome report the likelihood of having a disorder characterized by attention deficits, such as ADHD (Kao and Thomas, 2010).

In this research, a categorical dichotomous variable (YES/NO; obtained from the assessment report) is analyzed to indicate whether those evaluated have this probability of having an attentional disorder (YES indicates moderate to high possibility).

Statistical method

Categorical features

For each feature, a 2×2 contingency table is set up to perform the chi-squared test of independence for the groups defined as group C (control group of healthy patients) and group P (diagnosed group of ME/CFS patients). The significance threshold has been defined as a p-value <0.05. Cramer's V is an effect size measurement for the chi-square test of independence. It measures how strongly two categorical fields are associated. The degree of freedom (df) is 1, and if Cramer's value >0.1 is considered small-medium, >0.30 is considered medium-large, and > 0.50 is considered significant (Kim, 2017). Cramer's V was based on Pearson's chi-squared statistic and was published by Harald Cramer in 1946 (Doob, 1946).

Numerical features

The Shapiro–Wilk test and D′Agostino's K-squared test are two of the most commonly used hypothesis tests to analyze normality. In both tests, the null hypothesis is that the data comes from a normal distribution. The p-value of these tests indicates the probability of obtaining data like the observed data if they came from a population with a normal distribution with the same mean and variance as the observed data. The threshold is a p-value <0.05 as sufficient evidence to reject normality. The purpose is to verify the conditions of parametric methods for using the t-test. For features with a non-normal distribution, the test of independence used is the Mann–Whitney U-test for independent samples. The analysis is continued by calculating the arithmetic means of each variable and the graphical calculation using the box plot analysis.

Relationship graph between features

The relationships between the features have been analyzed using graph theory. A graph is a collection of nodes (also called vertices) joined together in pairs by edges (undirected) or arcs (directed; Newman, 2018). The graph structure allows us to capture the pattern of interactions between the nodes (individuals or entities). Graph (or network) analysis is used to study relationships between individuals to discover knowledge about global and local structures. The study of structure networks helps to decide the optimal order (Hagberg et al., 2008). In this work, the graph nodes are defined as all features, and the edges are defined as moderate or strong correlations between nodes (features). The correlation between two features is represented by $corr(i,j)$, and Spearman correlation is defined as moderate or strong if $corr(i,j) \geq 0.5$ (Suchowski, n.d.) in case of direct correlation. It has been created an $edge(i,j)$ if $abs(corr(i,j)) \geq 0.5$ in order to include direct and indirect correlation. This analysis used the values of the P-group diagnosed by ME/CFS.

Random forest algorithm

Random forest is a widespread algorithm for classification and offers the importance of feature values. All classified samples from the two defined groups are used to obtain a model. This model classifies future samples according to both categorical and numerical CPT3 scores. The metric used was accuracy. Features importance are computed as the mean and standard deviation of accumulation of the impurity decrease within each tree. It informs which feature has greater power to classify (Liaw and Wiener, n.d.).

Results

The sample size's value was based on a t-student analysis for two populations, ME/CFS patients (P) and healthy controls (C).

Distributions of CPT3™ variables were examined before analysis. Sample Analysis is reported in Table 1.

There are 17 features analyzed, six categorical and 11 numerical. For categorical features, the Chi-Squared test of independence is used related to C (Control group) or P (Diagnosed group). Tests of independence (Chi-Square, T-Student, and Mann–Whitney) indicate if there are significant differences between the two populations of the study. The results are shown in Tables 2, 3.

As a result, a pattern of attention deficit is obtained in the ME/CFS group when recording as a categorical dichotomous variable (YES/No) the probability of having a disorder characterized by attention deficits, such as ADHD (YES, including moderate to high likelihood of having it). With T-Student (if normality), the detectability feature is significant. With Mann–Whitney *U* Test (not normality), all variables are significant except HRT Blo C and HRT ISI Ch N, which are not.

The categorical variables are dichotomous, with values {1,2}, the plots showing whether significant differences exist due to belonging to one group or another. The graph in Figure 1 shows the percentage of the value of each variable according to the group. For example, it can be seen that in the first graph, the D.AT variable (attention deficit) is positive (value 2) in 30% of the control group and over 70% in the group of diagnosed patients. In this and the inattentiveness (INAT) variable, the differences are marked and reflected in the subsequent Chi-Square test of independence.

Box plots show for continuous variables the differences between quartiles between the different groups, including outliers. It is interesting to note the differences in some variables, although their significance was calculated using an independence test, such as omission (Omission N), perseverance (Persever N) or variability (Variabil N), are very evident in Figure 2.

Spearman's rho value is calculated for the continuous variables to define the graph illustrating the relationship between them. Each edge of the graph will correspond to a value greater than 0.5 from Table 4. The *p*-value is added to check the significance of the rho value.

As a result, three variables are found to be unrelated. In comparison, the five variables on the left of the graph show a complete regular graph as a pentagon, i.e., all five variables are related. The thicker the edge, the higher the correlation, and the thinner and redder the values close to 0.5.

Features importance random forest based

The classification algorithm is run in a Python environment (v 3.7.14), and the library used is sklearn (v 1.1). Random forest is a Supervised Machine Learning Algorithm that is used widely in classification and regression problems. This model classifies if a patient is in a control group or not. The database is divided into 75% to train the model and provide a generalization to predict, based on the values of the continuous and categorical variables, whether it belongs to group P or group C. The remaining 25% is kept for testing the model, i.e., the model is run and compared to see how well the model fits the sample. The model offers 77.19% accuracy, which means the 25% test dataset predicted with 77.19% accuracy. After dependence analysis for continuous

TABLE 1 Sample analysis.

	C (Control group)	P (CFS Diagnosed group)
n (225)	67	158
Age mean (std)	44.82 ± 14.26	51.40 ± 8.11
Females	53 (79.1%)	145 (91.77%)
Males	14 (20.9%)	13 (8.23%)

TABLE 2 Results from numerical features from CPT3™ by groups.

Feature	Control group (C)		Diagnosed group (P)		Independent test
	Mean	STD	Mean	STD	*p*-Value
t-student test (for normal distribution variables)					
Res.Styl N	47.82	8.05	48.95	11.21	0.39644
Detectab N	46.94	8.67	57.30	11.68	0.00000
Mann–Whitney *U* Test (for not normal distribution variables)					
Omission N	46.54	3.60	55.28	13.85	0.00000*
Commiss N	48.64	8.76	57.11	12.18	0.00000*
Persever N	48.63	5.93	58.81	16.25	0.00045*
HRT N	47.33	12.22	53.36	13.24	0.00011*
HRT SD N	46.54	7.56	61.65	14.70	0.00000*
Variabil N	45.01	6.48	55.61	11.73	0.00000*
HRT Blo C N	49.76	8.30	48.03	13.38	0.457403*
HRT ISI Ch N	50.42	9.04	50.63	14.23	0.949235*

Mean values according to the type of patient in the P group (CFS patient) or the Control group. The independence test was performed using Student's *t*-test if the values correspond to a normal distribution or Mann–Whitney *U* test otherwise and marked with an asterisk "*." For values with a *p*-value of less than 0.05, the hypothesis that group membership predetermines the outcome of the variable studied is considered to be accepted.

TABLE 3 Chi-square test of independence results.

Feature	Chi square value	*p*-Value	Cramer's*
Attention deficit	43.14	0.00000	0.44
Inattentiveness	50.45	0.00000	0.47
Impulsivity	6.27	0.01225	0.17
Sustained attention	3.87	0.04920	0.13
Vigilance	12.33	0.00045	0.23
Gender	6.00	0.01430	0.16

*Attention deficit, inattentiveness, impulsivity, sustained attention, and vigilance was all significant. Gender was also found to be significant. Cramer, greater than 0.30, in Inattentiveness (0.47 Cramer's *), Attention Deficit (0.43 Cramer's *), and Vigilance (0.23 Cramer's *) are significant (medium relation if Cramer >0.30 with df = 1). Significant differences indicate that the samples come from different populations. Consequently, it can be inferred that the significant variables predict the group the participant belongs to.

10.3389/fpsyg.2023.1127193

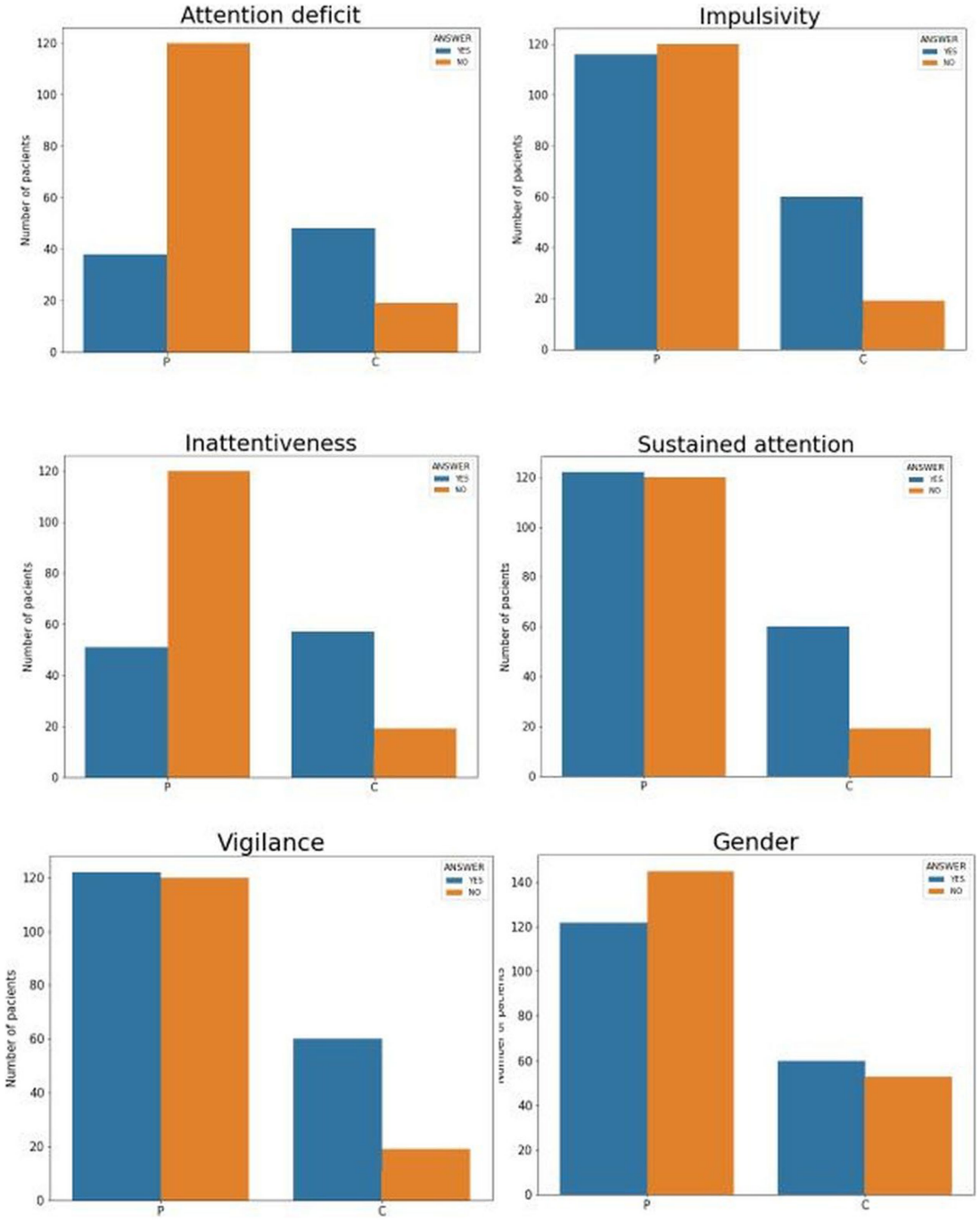


FIGURE 1
Graph of the contingency tables by groups and normalized, where each bar corresponds to the percentage of each value by groups. Each drawing represents for each dichotomous variable the group differences, with C being the control group and P the diagnosed group. The following acronyms mean the six dichotomous variables: DAT, attention deficit; INAT, inattentiveness; IMPUL, impulsivity; AT.SOST, sustained attention; VIGIL, vigilance; and finally, Gender.

features, the most important were HRT N, Detecta N, and Variability. Both were found to be able to classify ME/CFS patients shown in Figure 3.

A threshold of 60 is identified in the variable Variability N. For the values above, it is rare to find patients in the control group, which could indicate a risk of cognitive impairment for patients diagnosed

10.3389/fpsyg.2023.1127193

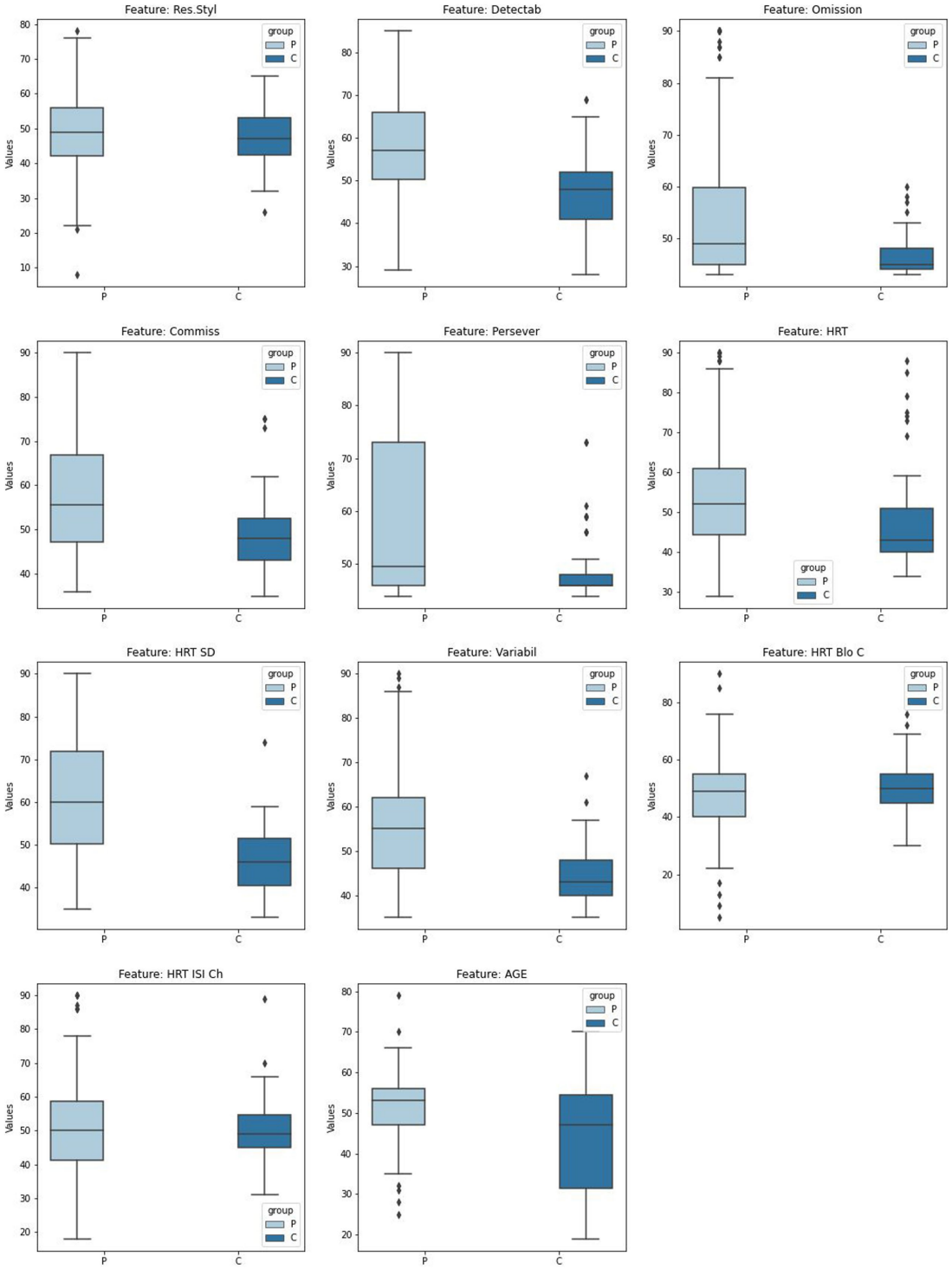


FIGURE 2
Box plot numerical features analysis by groups. Box plots for continuous variables by groups. The horizontal axis shows the groups, and the vertical axis shows the variable analyzed.

with ME/CFS. The results of the random forests showed that the CPT3™ could discriminate between presentations, mainly for the inattentive presentation.

Figure 4 shows the importance of each variable. HRT N and VARIABILITY N were the critical variables. In contrast, other variables related to HRT, such as HRT by Block and HRT ISI

TABLE 4 Rho values > abs(0.5) for variables with high correlation in group P (diagnosed ME/CFS).

Features		Rho	*p*-Value
Detectab N	Commiss N	0.820382	0.00000000
HRT SD N	Variabil N	0.786083	0.00000000
Detectab N	Omission N	0.776174	0.00000000
Omission N	Variabil N	0.766189	0.00000000
Res.Styl N	HRT N	0.722694	0.00000000
Omission N	HRT SD N	0.691635	0.00000000
Detectab N	Variabil N	0.689658	0.00000000
Detectab N	Persever N	0.673743	0.00000000
Detectab N	HRT SD N	0.614872	0.00000000
Omission N	Persever N	0.609529	0.00000000
Persever N	Variabil N	0.590245	0.00000000
Persever N	HRT SD N	0.531106	0.00000000
Res.Styl N	Omission N	0.521658	0.00000000
HRT N	HRT SD N	0.501458	0.00000000
Res.Styl N	Commiss N	−0.563669	0.00000000

The undirected graph is constructed with the data from Table 4 using the rho value itself as the weight on each edge, and the result is shown in Figure 2.

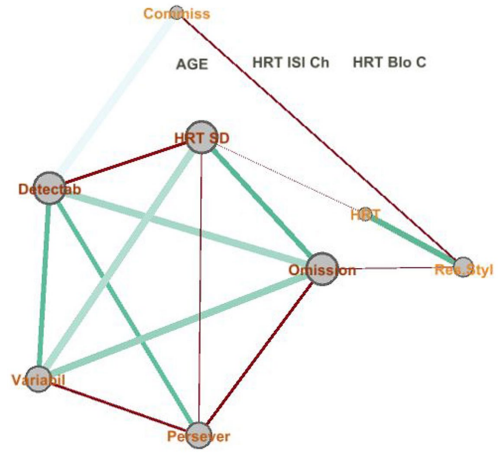


FIGURE 3
Network relationship between numerical features. The graph represents each continuous variable as a node, indicating the size of the node as a function of degree (the number of edges). Both the size and color depend on Spearman's rho value.

Change, had less importance in discriminating ME/CFS presentation.

Moreover, the results of the CPT3™ showed that a relevant number of ME/CFS patients were associated with a likelihood of having a disorder characterized by attention deficits, such as ADHD.

Discussion

Together with comprehensive diagnostic interviews, neuropsychological assessments could give the patient the "gold standard" for diagnosing ME/CFS. To the best of our knowledge, this is the first study assessing cognitive dysfunction in ME/CFS with CPT3™. It has been found that CPT3™ could detect cognitive impairments in all their main attentional measures (inattentiveness, impulsivity, sustained attention, and vigilance) in our ME/CFS individuals compared to healthy controls. More specifically, patients had a worse indicator of inattention, sustained attention, vigilance, impulsivity, and slow reaction time, showing the likelihood of having an ADHD-like pattern of functioning. In addition, relevant correlations were obtained between CPT3™ measures in this ME/CFS group. Accordingly, the adoption of the CPT3™ is supported as a specific test in ME/CFS aimed at analyzing the cognitive domains (attention and response speed) that seem to be key in the cognitive dysfunction of ME/CFS.

This work is in line with another previous study that assessed different types of attentional impairments in ME/CFS but using another brief and simple instrument (The Toulouse-Piéron Test) that allows the measurement of maintained attention—concentration, and resistance to monotony, and also, evaluates the multidimensional domain of attention, classified into different types such as arousal attention (alertness/activation), focused attention (detection of a stimulus) and sustained attention (attention to a stimulus or task for a prolonged time). These types of attention problems compromise various neuroanatomic structures, pathways, neurotransmitters, and their receptors. Their results support the reliability of maintained attention as a biomarker of ME/CFS, and attention deficit is a significant disability in patients affected by central fatigue. This neurocognitive dysfunction points to the neural networks involved in attention and focuses the pathological substrate in areas like the anterior cingulate cortex, lateral ventral prefrontal cortex, basal ganglia, or locus coeruleus (Murga et al., 2021).

Our results are in agreement with other neurocognitive research in ME/CFS reflecting attentional impairments on cognitive performance in these patients (Murga et al., 2021). Following these current findings, it would be quite interesting to monitor these brain areas with advanced techniques, such as fMRI, PET-scan, etc., to establish the key neurological bases that could be involved in CFS.

A neurocognitive profile in ME/CFS has not been described only because of the heterogeneity of the symptoms, but also for other shortcomings such as the lack of specific neuropsychological tests to evaluate it. CPT3™ represents a reliable alternative for assessing attention disorders in ME/CFS and allows a comprehensive measurement of multiple cognitive variables quickly. As a computerized test in the form of a computer game, the administration is more practical and stimulating for these patients, and the elaboration and quantification of scores are accurate. The Conners Continuous Performance Test (CPT) is widely used in clinical practice for its usefulness in the study of attention in various pathologies such as Attention Deficit Hyperactivity Disorder (ADHD; Newcorn et al., 2001; Baggio et al., 2020), although a previous study reported that the CPT3™ might be sensitive only to some of the core deficits of ADHD, but not hyperactivity. This instrument is not considered specific for ADHD, and although CPT may not differentiate between psychiatric and neurological disorders that result in executive dysfunctions (Baggio et al., 2020), it can be used in other diseases.

FIGURE 4
Scatter plot showing the discriminant threshold. Scatter plot of Variabil N (vertical axe) and HRT N (horizontal axe) where number 1 (red color) corresponds to P group, and number 0 (green color) corresponds to C group.

CPT3™ makes corrections for age and gender, so published results have already been corrected for biases identified in several studies (Newcorn et al., 2001; Conners et al., 2003; Seidman et al., 2005; Burton et al., 2010; Ramtekkar et al., 2010). T-scores are relative to age group and gender, with sex-independent consultation being possible as an option. Although the study design did not take into account a gender or age balance in the sample, and this should be taken into account in future studies, the gender variable was surprisingly significant in the Chi-square test. Furthermore, despite the fact that there is also a gender imbalance, especially in the group of diagnosed patients, the scores are in line with other studies that report on the influence of gender on chronic fatigue (Lim et al., 2020).

CFS/ME patients have more difficulty discriminating between targets and non-targets, and this poor detectability indicates inattention. A very unusual number of omission errors may indicate clinical impairment, fatigue, poor understanding of instructions, or a lack of motivation to respond with full effort. The results of the Mann–Whitney *U* test show that all variables are significant except HRT Block Change (HRT Blo C, meaning that the slope of HRT change in the six test blocks) and HRT ISI change [HRT ISI Ch N; indicates that the HRT change's slope in the three ISIs (1, 2, and 4 s)]. A T-score of 60–69 in HRT is classified as slow response, and a T-score of 60–69 in Variability is interpreted as under-average performance. Slower reaction times and high inconsistency of response speed may be associated with the inattentive profile. High response speed variability indicates that the respondent's attention and information processing efficiency varied throughout the administration (Kao and Thomas, 2010).

Indicators that measure response speed act as good discriminants between both groups. Hit Reaction Time (HRT) is the average response speed of correct responses for the entire administration, measured in milliseconds, for all non-perseverative responses made during the entire administration. An atypically slow HRT may indicate inattention, especially when error rates are high, but may also result from a very conservative response style. Variability, like HRT SD, is a measure of response speed consistency; however, Variability is a "within-respondent" measure (i.e., the amount of variability the respondent exhibited in 18 separate sub-blocks of the administration relative to their overall HRT SD score). Although Variability is a different measure than HRT SD, the two measures typically produce comparable results, and both are related to inattention. High response speed variability indicates that the respondent's attention and processing efficiency varied throughout the administration. These scores mean that the response speed in the ME/CFS significantly differed from the response speed in the control group. The high variability of the response speed and the slow reaction time obtained indicate that ME/CFS suffer dysfunctions in the efficiency of information processing and commit more errors of omission, commission, and perseverance than healthy controls, presenting not only more attention problems, but being slower and with greater inconsistency in response speed when performing this test. The high variability of response speed indicates that the attention and processing efficiency of the ME/CFS patients varied throughout the administration. Figure 5 suggests that this is a discriminant element in differentiating ME/CFS from healthy people. The Hit Reaction Time Standard Deviation (HRT SD) measures the consistency of the speed of response to target stimuli

across the administration. A high HRT SD in CFS/ME patients also indicates a greater inconsistency in response speed. It is sometimes indicative of inattention, which could suggest that ME/CFS patients were less engaged and processed stimuli less efficiently than healthy subjects during some parts of the CPT3™ administration. Overall, the ME/CFS group has more-atypical T-scores, which is associated with a high likelihood of having a disorder characterized by attention deficits, such as ADHD. However, assessors should keep in mind that other psychological and/or neurological conditions with attention-disrupting symptoms may also generate atypical scores (and thus a high or very high probability estimate). A previous study suggested that ADHD may be common in ME/CFS patients and is associated with a more severe psychopathology clinical profile (Sáez-Francàs et al., 2012).

For all the above, CPT3™ represents a reliable alternative to objectively and comprehensively assess attentional deficits in CFS. However, it also allows the assessment of other executive function domains, such as information processing speed and inhibition, providing the reaction time measures which can be considered discriminant indicators of ME/CFS patients. This tool makes it possible to assess the severity of cognitive dysfunction in this population and could improve the diagnosis and/or monitor these patients.

It should be noted that for a proper interpretation of CPT3™ test results in patients diagnosed with CFS, it would be necessary to investigate the clinical history further to assess the individualized performance of each measured dimension. The impact of medication use on cognitive functioning has not been explored in-depth, and it would be interesting to consider it in future studies. Even though the CPT lacks ecological validity (Baggio et al., 2020) as it does not adequately simulate the difficulties patients may experience in their daily life (i.e., it is free of external distractions that are likely to impair the patient's real-life performance and it is a rather short task to represent overall performance), ME/CFS patients presented altered CPT3™ parameters when compared to healthy people. Most patients were not referred from primary care but belonged to a specialized unit in a tertiary care hospital. They probably have more mental and non-mental comorbidities, more severe symptoms, and more years of

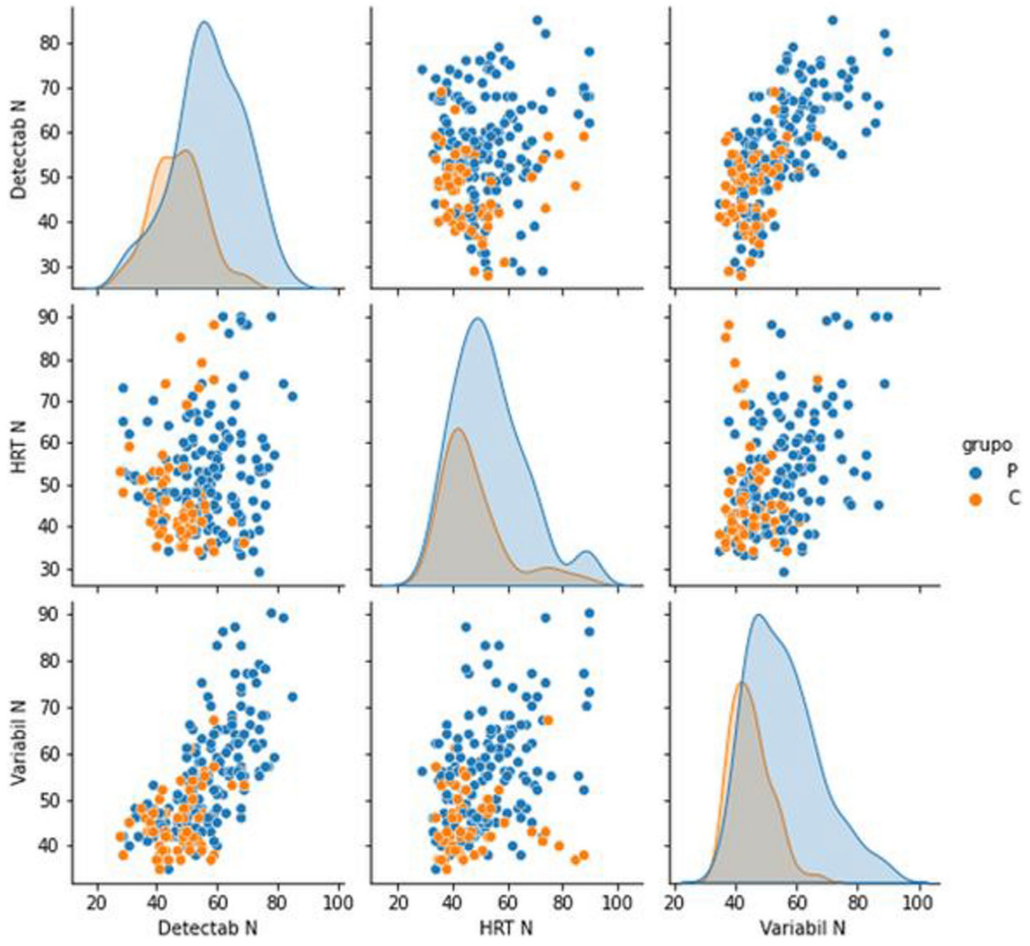


FIGURE 5
Pair plot analysis of the most significant features to discriminate between groups. Histograms as a function of the group and scatter plots of the combinations between the three chosen variables.

diagnosis. It should be considered to analyze the relationship between the results obtained through the CPT3™ with other variables, such as the severity of symptoms, associated comorbidities, and the time of diagnosis of ME/CFS, evaluating the changes over time in these participants. Using the test–retest after an estimated suitable time could be convenient for contemplating the evolution of cognitive functioning in these patients. Furthermore, it would be very interesting to analyze the relationship between the degree of fatigue, the psychopathological symptoms, and the cognitive performance of the CPT3™ in these patients.

Finally, it would be highly recommended in future studies to contemplate an analysis of how the presence of comorbid psychopathological symptoms in ME/CFS influences the cognitive performance obtained in the CPT3™, since it is a test that does not discriminate the presence of these factors. Furthermore, it would be interesting to replicate this study by including another group with severe psychiatric disorders like major depressive disorder.

Limitations

All the patients come from the specialized chronic fatigue unit of the Vall D'Hebrón University Hospital in Barcelona. Patient data from other CFS-specialized units with similar protocols could enrich this work. Further research is needed to determine differences in ME/CFS patients to validate this work. CPT3™ response values have also not been evaluated based on controlled medication. There is no information on the influence of COVID-19 infection on the results in either group.

Conclusion

CPT3™ adequately identifies ME/CFS in this clinical sample of adult participants compared to healthy controls. This profile of cognitive dysfunction could be related to other pathophysiological phenomena of CFS, and its determination could be key to elucidating the underlying basis and providing empirical evidence for the usefulness of this computerized neuropsychological test. Additionally, fatigue is the most common symptom associated with NCDs. There is insufficient evidence examining the relationship between fatigue and cognitive impairments in patients with other chronic diseases in which the symptom fatigue appears (e.g., major depressive disorder, fibromyalgia, postcovida, etc.). Further studies using CPT3™ could be useful to examine possible relationships between these two symptoms.

This study shows that a relevant number of ME/CFS patients had CPT3™ values compatible with a likelihood of having a disorder characterized by attention deficits, suggesting the possible existence of an underlying neurological basis for attentional dysfunction among the study patients. Furthermore, it was shown that two measures of response speed (hit reaction time and variability) act as good discriminants between both groups. Taken together, CPT3™ is a helpful neuropsychological instrument for discriminating cognitive impairments in attention and response speed in CFS. Therefore, the results obtained here will allow us to justify the use of the CPT3™ in other investigations that explore the cognitive functioning of these patients, and even this computerized test could be considered as a possible candidate marker for ME/CFS. As this is the only study reporting CPT3™ scores in adult ME/CFS patients, future studies are needed to compare this test across ME/CFS centers.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of the Vall d'Hebron University Hospital with the registration code PR (AG)257/2021. The patients/participants provided their written informed consent to participate in this study.

Author contributions

All authors have participated in the conception and design of the study, and in the collection, analysis and interpretation of the data, as well as in the writing, revision and approval of the submitted manuscript.

Conflict of interest

JR-Q was on the speakers' bureau and/or acted as a consultant for Janssen-Cilag, Novartis, Shire, Takeda, Bial, Shionogi, Sincrolab, Novartis, BMS, Medicine, and Rubió Uriach, Technofarma, and Raffo in the last 3 years. He also received travel awards (air tickets + hotel) for taking part in psychiatric meetings from Janssen-Cilag, Rubió Shire, Takeda, Shionogi, Bial, and Medice. The Department of Psychiatry chaired by him received unrestricted educational and research support from the following companies in the last 3 years: Janssen-Cilag, Shire, Oryzon, Roche, Psious, and Rubió AR-U has received travel awards (air tickets + hotel) for taking part in annual psychiatric meetings from Lundbeck and has acted as a speaker at various training courses financed by Organon, Janssen-Cilag, and Lundbeck.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Afari, N., and Buchwald, D. (2003). Chronic fatigue syndrome: a review. *Am. J. Psychiatry* 160, 221–236. doi: 10.1176/appi.ajp.160.2.221

Alonso, J., Prieto, L., and Antó, J. M. (1995). The Spanish version of the SF-36 health survey (the SF-36 health questionnaire): an instrument for measuring clinical results. *Med. Clin.* 104, 771–776.

Badin, O. O., Smith, M. R., Conte, D., and Coutts, A. J. (2016). Mental fatigue: impairment of technical performance in small-sided soccer games. *Int. J. Sports Physiol. Perform.* 11, 1100–1105. doi: 10.1123/ijspp.2015-0710

Baggio, S., Hasler, R., Giacomini, V., El-Masri, H., Weibel, S., Perroud, N., et al. (2020). Does the continuous performance test predict ADHD symptoms severity and ADHD presentation in adults? *J. Atten. Disord.*, 840–848. doi: 10.1177/1087054718822060

Bassi, N., Amital, D., Amital, H., Doria, A., and Shoenfeld, Y. (2008). Chronic fatigue syndrome: characteristics and possible causes for its pathogenesis. *Israel Med. Assoc. J.* 10, 79–82.

Beaumont, A., Burton, A. R., Lemon, J., Bennett, B. K., Lloyd, A., and Vollmer-Conna, U. (2012). Reduced cardiac vagal modulation impacts on cognitive performance in chronic fatigue syndrome. *PLoS One* 7:e49518. doi: 10.1371/journal.pone.0049518

Berger, I., Slobodin, O., and Cassuto, H. (2017). Usefulness and validity of continuous performance tests in the diagnosis of attention-deficit hyperactivity disorder children. *Arch. Clin. Neuropsychol.* 32, 81–93. doi: 10.1093/arclin/acw101

Burton, L., et al. (2010). Effects of gender and personality on the Conners continuous performance test. *J. Clin. Exp. Neuropsychol.* 32, 66–70. doi: 10.1080/13803390902806568

Carruthers, B. M., Jain, A. K., de Meirleir, K. L., Peterson, D. L., Klimas, N. G., Lerner, A. M., et al. (2003). Myalgic encephalomyelitis/chronic fatigue syndrome. *J. Chron. Fatig. Syndr.* 11, 7–115. doi: 10.1300/j092v11n01_02

Carruthers, B. M., van de Sande, M. I., de Meirleir, K. L., Klimas, N. G., Broderick, G., Mitchell, T., et al. (2011). Myalgic encephalomyelitis: international consensus criteria. *J. Intern. Med.* 270, 327–338. doi: 10.1111/j.1365-2796.2011.02428.x

Castresana, C., Pérez, A. G.-E., and de Rivera, J. L. G. (1995). Hospital anxiety and depression scale y psicopatología afectiva. *Anal. Psiquiatr.* 2, 126–130.

Castro-Marrero, J., Sáez-Francàs, N., Segundo, M. J., Calvo, N., Faro, M., Aliste, L., et al. (2016). Effect of coenzyme Q10 plus nicotinamide adenine dinucleotide supplementation on maximum heart rate after exercise testing in chronic fatigue syndrome – a randomized, controlled, double-blind trial. *Clin. Nutr.* 35, 826–834. doi: 10.1016/j.clnu.2015.07.010

Castro-Marrero, J., Sáez-Francàs, N., Santillo, D., and Alegre, J. (2017). Treatment and management of chronic fatigue syndrome/myalgic encephalomyelitis: all roads lead to Rome. *Br. J. Pharmacol.* 174, 345–369. doi: 10.1111/bph.13702

Cockshell, S. J., and Mathias, J. L. (2010). Cognitive functioning in chronic fatigue syndrome: a meta-analysis. *Psychol. Med.* 40, 1253–1267. doi: 10.1017/S0033291709992054

Conners, C. K., Epstein, J. N., Angold, A., and Klaric, J. (2003). Continuous performance test performance in a normative epidemiological sample. *J. Abnorm. Child Psychol.* 31, 555–562. doi: 10.1023/a:1025457300409

Cvejic, E., Birch, R. C., and Vollmer-Conna, U. (2016). Cognitive dysfunction in chronic fatigue syndrome: a review of recent evidence. *Curr. Rheumatol. Rep.* 18:24. doi: 10.1007/s11926-016-0577-9

Deluca, J., Christodoulou, C., Diamond, B. J., Rosenstein, E. D., Kramer, N., and Natelson, B. H. (2004). Working memory deficits in chronic fatigue syndrome: differentiating between speed and accuracy of information processing. *J. Int. Neuropsychol. Soc.* 10, 101–109. doi: 10.1017/S1355617704101124

Dickson, A., Toft, A., and O'Carroll, R. E. (2009). Neuropsychological functioning, illness perception, mood and quality of life in chronic fatigue syndrome, autoimmune thyroid disease and healthy participants. *Psychol. Med.* 39, 1567–1576. doi: 10.1017/S0033291708004960

Doob, J. L. (1946). '*Mathematical Methods of Statistics*. Harald Cramèr. Princeton, NJ: Princeton Univ. Press, 1946. Pp. xvi 575. (Illustrated.) $6.00', Science, 450.

Fasmer, O. B., Mjeldheim, K., Førland, W., Hansen, A. L., Syrstad, V. E. G., Oedegaard, K. J., et al. (2016). Linear and non-linear analyses of Conner's continuous performance test-II discriminate adult patients with attention deficit hyperactivity disorder from patients with mood and anxiety disorders. *BMC Psychiatry* 16:284. doi: 10.1186/s12888-016-0993-4

Fisk, J. D., Ritvo, P. G., Ross, L., Haase, D. A., Marrie, T. J., and Schlech, W. F. (1994). Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. *Clin. Infect. Dis.* 18, S79–S83. doi: 10.1093/clinids/18.supplement_1.s79

Fisk, J. D., and Doble, S. E. (2002). Construction and validation of a fatigue impact scale for daily administration (D-FIS). *Qual. Life Res.* 11, 263–272. doi: 10.1023/a:1015295106602

Fukuda, K., Straus, S. E., Hickie, I., Sharpe, M. C., Dobbins, J. G., and Komaroff, A. (1994). The chronic fatigue syndrome: a comprehensive approach to its definition and study. *Ann. Intern. Med.* 121, 953–959. doi: 10.7326/0003-4819-121-12-199412150-00009

Hagberg, A., Swart, P. S., and Chult, D. (2008) *Exploring network structure, dynamics, and function using NetworkX*. Los Alamos, NM (United States): Los Alamos National Lab (LANL).

Kao, G. S., and Thomas, H. M. (2010). Test review: C. Keith Conners Conners 3rd edition Toronto, Ontario, Canada: multi-health systems, 2008. *J. Psychoeduc. Assess.* 28, 598–602. doi: 10.1177/0734282909360011

Kim, H.-Y. (2017). Statistical notes for clinical researchers: chi-squared test and Fisher's exact test. *Restor. Dent. Endod.* 42, 152–155. doi: 10.5395/rde.2017.42.2.152

Liaw, A., and Wiener, M. (n.d.). Classification and regression by randomForest. *R news* [Preprint]. Available at: https://www.researchgate.net/file.PostFileLoader.html?id=57d984d196b7e421f8398993&assetKey=AS%3A406267404275720%401473873105695 (Accessed February 13, 2023).

Lim, E.-J., Ahn, Y. C., Jang, E. S., Lee, S. W., Lee, S. H., and Son, C. G. (2020). Systematic review and meta-analysis of the prevalence of chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME). *J. Transl. Med.* 18:100. doi: 10.1186/s12967-020-02269-0

McGregor, N. R., Dunstan, R. H., Butt, H. L., Roberts, T. K., Klineberg, I. J., and Zerbes, M. (1997). A preliminary assessment of the association of SCL-90-R psychological inventory responses with changes in urinary metabolites in patients with chronic fatigue syndrome. *J. Chronic Fatigue Syndr.* 3, 17–37. doi: 10.1300/J092v03n01_03

Menzies, V., Kelly, D. L., Yang, G. S., Starkweather, A., and Lyon, D. E. (2021). A systematic review of the association between fatigue and cognition in chronic noncommunicable diseases. *Chronic Illn.* 17, 129–150. doi: 10.1177/1742395319836472

Morrow, S. A., Rosehart, H., and Johnson, A. M. (2015). 'Diagnosis and quantification of cognitive fatigue in multiple sclerosis', cognitive and behavioral neurology: official journal of the Society for Behavioral and Cognitive. *Neurology* 28, 27–32. doi: 10.1097/WNN.0000000000000050

Murga, I., Aranburu, L., Gargiulo, P. A., Gómez-Esteban, J. C., and Lafuente, J. V. (2021). The maintained attention assessment in patients affected by Myalgic encephalomyelitis/chronic fatigue syndrome: a reliable biomarker? *J. Transl. Med.* 19:494. doi: 10.1186/s12967-021-03153-1

Nacul, L. C., Lacerda, E. M., Pheby, D., Campion, P., Molokhia, M., Fayyaz, S., et al. (2011). Prevalence of myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) in three regions of England: a repeated cross-sectional study in primary care. *BMC Med.* 9:91. doi: 10.1186/1741-7015-9-91

Newcorn, J. H., Halperin, J. M., Jensen, P. S., Abikoff, H. B., Arnold, L. E., Cantwell, D. P., et al. (2001). Symptom profiles in children with ADHD: effects of comorbidity and gender. *J. Am. Acad. Child Adolesc. Psychiatry* 40, 137–146. doi: 10.1097/00004583-200102000-00008

Newman, M. (2018). *Networks. 2nd Edn.* UK: Oxford University Press.

Prinsen, H., van Dijk, J. P., Zwarts, M. J., Leer, J. W. H., Bleijenberg, G., and van Laarhoven, H. W. M. (2015). The role of central and peripheral muscle fatigue in postcancer fatigue: a randomized controlled trial. *J. Pain Symptom Manag.* 49, 173–182. doi: 10.1016/j.jpainsymman.2014.06.020

Prins, J. B., van der Meer, J. W. M., and Bleijenberg, G. (2006). Chronic fatigue syndrome. *Lancet* 367, 346–355. doi: 10.1016/s0140-6736(06)68073-2

Ramtekkar, U. P., Reiersen, A. M., Todorov, A. A., and Todd, R. D. (2010). Sex and age differences in attention-deficit/hyperactivity disorder symptoms and diagnoses: implications for DSM-V and ICD-11. *J. Am. Acad. Child Adolesc. Psychiatry* 49, 217–28.e1–3. doi: 10.1016/j.jaac.2009.11.011

Rasouli, O., Gotaas, M. E., Stensdotter, A. K., Skovlund, E., Landrø, N. I., Dåstøl, P., et al. (2019). Neuropsychological dysfunction in chronic fatigue syndrome and the relation between objective and subjective findings. *Neuropsychology* 33, 658–669. doi: 10.1037/neu0000550

Reyes, M., Nisenbaum, R., Hoaglin, D. C., Unger, E. R., Emmons, C., Randall, B., et al. (2003). Prevalence and incidence of chronic fatigue syndrome in Wichita, Kansas. *Arch. Intern. Med.* 163, 1530–1536. doi: 10.1001/archinte.163.13.1530

Rivera, J., Alegre, C., Ballina, F. J., Carbonell, J., Carmona, L., Castel, B., et al. (2006). Documento de consenso de la Sociedad Española de Reumatología sobre la fibromialgia. *Reumatol. Clín.* 2, S55–S66. doi: 10.1016/s1699-258x(06)73084-4

Ruiz, E., Alegre, J., García Quintana, A. M., Aliste, L., Blázquez, A., and Fernández de Sevilla, T. (2011). Síndrome de fatiga crónica: estudio de una serie consecutiva de 824 casos evaluados en dos unidades especializadas. *Nuevos Arch. Facult. Med.* 211, 385–390. doi: 10.1016/j.rce.2011.02.013

Sáez-Francàs, N., Alegre, J., Calvo, N., Antonio Ramos-Quiroga, J., Ruiz, E., Hernández-Vara, J., et al. (2012). Attention-deficit hyperactivity disorder in chronic fatigue syndrome patients. *Psychiatry Res.* 200, 748–753. doi: 10.1016/j.psychres.2012.04.041

Santamarina-Perez, P., Eiroa-Orosa, F. J., Rodriguez-Urrutia, A., Qureshi, A., and Alegre, J. (2014). Neuropsychological impairment in female patients with chronic fatigue syndrome: A preliminary study. *Appl. Neuropsychol.* 21, 120–127. doi: 10.1080/09084282.2013.771264

Seidman, L. J., Biederman, J., Monuteaux, M. C., Valera, E., Doyle, A. E., and Faraone, S. V. (2005). Impact of gender and age on executive functioning: do girls and boys with and without attention deficit hyperactivity disorder differ neuropsychologically in preteen and teenage years? *Dev. Neuropsychol.* 27, 79–105. doi: 10.1207/s15326942dn2701_4

Suchowski, M. A. (n.d.). An analysis of the impact of an outlier on correlation coefficients across small sample data where rho is non –zero. *Western Michigan UniversityProQuest Dissertations Publishing, Degree Year*, p. 3007026. Available at: https://search.proquest.com/openview/5d1cbf13c930b7358050381ebab41a85/1?pq-origsite=gscholar&cbl=18750&diss=y (Accessed February 13, 2023).

Swain, M. G. (2000). Fatigue in chronic disease. *Clin. Sci.* 99, 1–8. doi: 10.1042/CS19990372

Verster, J. C., Pandi-Perumal, S. R., and Streiner, D. L. (2008). *Sleep and Quality of Life in Clinical Medicine* Springer Science & Business Media.

scientific reports

Check for updates

OPEN

A synthetic data generation system for myalgic encephalomyelitis/ chronic fatigue syndrome questionnaires

Marcos Lacasa[1✉], Ferran Prados[1,2,3,4], José Alegre[5] & Jordi Casas-Roma[1]

Artificial intelligence or machine-learning-based models have proven useful for better understanding various diseases in all areas of health science. Myalgic Encephalomyelitis or chronic fatigue syndrome (ME/CFS) lacks objective diagnostic tests. Some validated questionnaires are used for diagnosis and assessment of disease progression. The availability of a sufficiently large database of these questionnaires facilitates research into new models that can predict profiles that help to understand the etiology of the disease. A synthetic data generator provides the scientific community with databases that preserve the statistical properties of the original, free of legal restrictions, for use in research and education. The initial databases came from the Vall Hebron Hospital Specialized Unit in Barcelona, Spain. 2522 patients diagnosed with ME/CFS were analyzed. Their answers to questionnaires related to the symptoms of this complex disease were used as training datasets. They have been fed for deep learning algorithms that provide models with high accuracy [0.69–0.81]. The final model requires SF-36 responses and returns responses from HAD, SCL-90R, FIS8, FIS40, and PSQI questionnaires. A highly reliable and easy-to-use synthetic data generator is offered for research and educational use in this disease, for which there is currently no approved treatment.

Myalgic encephalomyelitis, commonly called chronic fatigue syndrome (ME/CFS), is a serious, complex, and chronic multisystem illness of unknown etiology, often triggered by a persistent viral infection (for this reason, it is also known as post-viral fatigue syndrome). ME/CFS affects as many as 17 to 24 million people worldwide, and its prevalence is expected to double by 2030[1]. It is characterized by unexplained and persistent post-exertional fatigue that is not relieved by rest. It is exacerbated by physical and mental exertion and other core symptoms such as cognitive, immunometabolic, autonomic, and neuroendocrine dysfunction[2]. It produces severe disability in patients, significantly interfering with their work activity and their daily life tasks[3]. In addition to fatigue, these patients have characteristic inflammatory and muscular symptoms, sleep dysfunction, and altered cognitive functions[4]. The symptomatic muscle blocks symptoms such as pain, generalized muscle weakness, fatigue after physical exertion, neurological symptoms (sensory hypersensitivity, ataxia, dysmetria, visual disturbances, and motor incoordination), neurocognitive symptoms (alterations in memory, concentration, calculation, task planning). The autonomic block (cephalic instability, dizziness, fainting spells, excessive sweating, orthostatic hypotension, tremor or alterations in intestinal rhythm), immunoinflammatory symptoms (low-grade fever, sore throat, recurrent canker sores, polyarthralgia, morning numbness, infections such as herpes or candida) and deficiency symptoms in the production of cellular metabolic energy. Sleep disturbances have been relevant since their description as their clinical entity. In all versions of the different ME/CFS diagnostic criteria, sleep disorders have played a key role, especially the presence of unrefreshing sleep and the importance of the Pittsburgh Sleep Quality Index (PSQI) questionnaire in the assessment of the severity of alterations in sleep quality and its association with fatigue, pain, psychopathology, and neurovegetative dysfunction[5]. ME/CFS, together with the symptomatic complexity that it presents, as a consequence of its multisystemic nature, is associated with different comorbid phenomena such as fibromyalgia, sicca syndrome, myofascial syndrome, psychopathology, ligament hyperlaxity,

[1]ADaS Lab - E-Health Center, Universitat Oberta de Catalunya, Rambla del Poblenou, 156, 08018 Barcelona, Spain. [2]Center for Medical Image Computing, University College London, London, UK. [3]National Institute for Health Research Biomedical Research Centre at UCL and UCLH, London, UK. [4]Department of Neuroinflammation, Queen Square MS Center, UCL Institute of Neurology, Faculty of Brain Sciences, University College London, London, UK. [5]ME/CFS Unit, Division of Rheumatology, Vall d'Hebron Hospital Research Institute Universitat Autònoma de Barcelona, Barcelona, Spain. ✉email: mlacasaca@uoc.edu

fasciitis plantar, degenerative vertebral disease or mechanical, shoulder tendinopathy, multiple chemical sensitivity, epicondylitis, carpal tunnel syndrome, osteoporosis, hypercholesterolemia, hypertriglyceridemia, vascular risk, endometriosis, thyroiditis, with a higher prevalence than that observed in patients not affected by ME/CFS[6].

In the study of ME/CFS, after the diagnosis and assessment of comorbid phenomena, it is essential to quantify and assess fatigue, quality of life, or anxiety/depression psychopathology using a battery of clinically self-administered questionnaires. Today there are few units specialized in ME/CFS in the world, with a relatively low number of duly documented cases and a lack of publicly available data compared with other disorders. Moreover, unfortunately, there are no commercially available diagnostic tests, no specific lab biomarkers, and no targeted FDA-approved drugs for ME/CFS[7]. Therefore, each subject to be diagnosed with ME/CFS must undergo a Fukuda criteria evaluation and procedure that each unit has established using batteries of validated self-administered questionnaires. As stated before, it is important to evaluate the disabling fatigue perception, sleep problems, and health-related quality of life using self-administered questionnaires such as the fatigue impact scale FIS40[8] and FIS8[9], PSQI[10], and Short Form Health Survey (SF-36)[11], Symptom Checklist-90-revised (SCL 90 R) psychological inventory[12], hospital anxiety and depression scale (HAD)[13]. Ongoing placebo-controlled clinical trials to evaluate the clinical benefits of drugs on ME/CFS symptoms[14] have changed some questionnaire scores from baseline to final study as a primary endpoint.

There is no consensus on the number and type of questionnaires that should be carried out, so not all units record the same number per subject. Consequently, it is complex for ME/CFS units to have many records of the questionnaires necessary to efficiently approach large longitudinal and multicenter studies of patients with this pathology using the latest advances in data analysis, such as Machine Learning techniques.

Machine Learning is a particular method of data analytics that automates model building as it relates to the development of models. Over the last years, it has been proven great performance of machine learning supervised algorithms in several clinical applications[15] to diagnose and treat diseases. Supervised learning involves training machine learning-based algorithms using labeled input datasets requiring however, to be efficient and get optimal results, a large number of records are needed. The learning occurs by comparing results with the expected outputs to identify errors and change the model's weights to infer knowledge. There are few publications, and all of them very recent, that refers to the application of different machine learning techniques in ME/CFS: seeking a new biomarker[16], clustering[17], or discovering the relationship between depression and ME/CFS[18], using neural networks seeking omic biomarkers[16] or neural networks classifiers[19]. While they all make important steps forward in understanding ME/CSF, the limited sample size makes generalization and translation of their findings to clinical practice or other datasets difficult. Also, as stated before, when there are no clear biomarkers to follow the evolution of the illness, like in ME/CSF, quality-of-life questionnaires are used to measure it[14]. There are several lines of investigation, such as clustering[20] or finding relations between blood measurements with questionnaires data[21].

Therefore, there is an increasing demand to access large repositories of high-quality health datasets for better and more reliable predictions from supervised machine learning algorithms. Anonymized electronic health records are bought and sold by insurance[22] and clinical groups[23]. However, they are limited in size or content, might be incomplete, and their applications might be restricted. This problem can be overcome using synthetic datasets coming from simulations[24,25]. Synthetic datasets are generated to create data for improving the sample size of existing cohorts or filling in the missing values, preserving privacy while keeping the real data characteristics. Synthetic data generators preserve the statistical properties of the original. However, they do not reveal any information regarding real people and offer several benefits, such as overcoming real data usage restrictions of data sharing and patient consent. There is a need for developing synthetic datasets that would complement real-world data for various reasons[26]: ease of access, cost-efficiency, test-efficiency, patient privacy protection, completeness, and validation capabilities, handling missingness, complex interactions between variables, resulting sensitivity analysis statistics from latest classifiers and graphical modeling and resampling[27]. A common application of synthetic data generation in medicine is image generation simulating diseases. It helps to test and benchmark the performance and accuracy of different algorithms. Some recent applications are in the simulation of skin lesions[28], brain atrophy in aging or Dementia[29], generation of PET MRI scans for Alzheimer's disease[30], tumor generation in the brain[31], or breast cancer[32].

This work aims to generate a robust and reliable synthetic data generator for ME/CFS questionnaires to produce high-fidelity and risk-free health care records, enhance existing public and private ME/CFS datasets for investigation and educational use, and are free of legal, privacy, security, and intellectual property restrictions.

Patients and methods

Dataset. This prospective cross-sectional study includes 2,522 subjects diagnosed with ME/CFS from the Vall d'Hebron University Hospital, Barcelona, Spain, 90.5% females (mean age 48.11 ± 10.31 years) and 9.5% males (mean age 44.41 ± 11.35 years). Data for SF-36, HAD, FIS8, FIS40, SCL 90 R, and PSQI questionnaires has been obtained and recorded from 2008 to 2021. See Table 1 for final records. Patients were eligible to participate if they were 18 years, had a confirmed diagnosis of ME/CFS, met the Fukuda[33] and Carruthers criteria[34], and provided signed written informed consent and ethics committee approval. The data collected were anonymized in a database to which only those designated for the study had access, and in no case was any information known that could reveal or infer the participant's identity.

Relationship graph between questionnaires. Graph theory was used to analyze the relationships between the subscales of each questionnaire. A graph is a collection of nodes (also called vertices) joined together in pairs by edges (undirected) or arcs (directed)[35]. The graph structure allows us to capture the pattern of interactions between the nodes (individuals or entities). Graph (or network) analysis is used to study relationships

Questionnaire	Registers	Questions	Subscales	Total value	Answers' rank
SF 36	2346	36	10	NO	{1,2,3,4,5,6}
HAD	2339	14	2	YES	{0,1,2,3}
FIS8	2057	8	0	YES	{0,1,2,3,4}
FIS40	2362	40	3	YES	{0,1,2,3,4}
SCL 90 R	2361	90	12	NO	{0,1,2,3,4}
PSQI	1959	34	7	YES	{0,1,2,3}

Table 1. Available data for each questionnaire and the questionnaires' characteristics. From left to right, each column title means Registers: number of available forms. Questions: number of questions per questionnaire. Subscales: number of defined subscales. Total Value: The questionnaire has a unique resume value. Answers' rank: Possible answer value for each question.

between individuals to discover knowledge about global and local structures. The study of structure networks helps to decide the optimal order[36].

In this work, the graph nodes are defined as all subscales, and the edges are defined as moderate or strong correlations between nodes (subscales). The linear correlation between two subscales is represented by $corr(i,j)$, and Pearson correlation is defined as moderate or strong if $corr(i,j) \geq 0.5$ [37] in case of direct correlation. An $edge(i,j)$ is defined if $abs(corr(i,j)) \geq 0.5$.

The relation of subscales between each test is related in Table 2. Each subscale has been classified according to the area to which it has been defined and named as the subject. Thirty-eight subscales, six tests, and twelve subjects form the dataset to create the relationship between them to the graph.

The study of the relationships mentioned above should indicate the order to generate our machine learning models. The SF-36 is prevalent and will be used in our model as initial data. The rest of the order will be given by the relationships between the different tests so that those with a stronger relationship are consecutive in the model. The strength of the relationship is measured in terms of the percentage of connections between the test nodes.

$$max\ (rel_i/rel_j)\ for\ each\ i,j \forall i,j \in [1,n] in\ n\ test$$

$$rel_i = number\ of\ nodes\ of\ test\ i\ related\ with\ nodes\ of\ test\ j$$

$$rel_j = number\ of\ nodes\ of\ test\ j$$

Model architecture. Real data of all of six questionnaires are required to train and build the models. First, an input matrix represents the validated answers of a number of patients, where n is the number of validated responses and f the number of questions. That is the first training data. As predicted, it has to be a second questionnaire which the same n and f_1 questions. The model must generate a predicted matrix with the same dimension. The next step has as input matrix the initial matrix concatenated with the last predicted matrix and the second questionnaire response matrix for prediction, as shown in Fig. 1.

Machine learning algorithms. Classification and regression models can be used. The goal is to provide 186 output dimensions that must be calculated step by step. The output is compared with the real data set to validate the model. The results are validated using the t-student test. The strategy is to validate one questionnaire. The next step is concatenating the questionnaire answers matrix as input with the final output with different models. It has tested machine learning and deep learning algorithms step by step. The validation system has measured whether real and synthetic data come from the same populations within t-student statistics. The models tested have been regressors and classifiers. The comparison between XGBoost and Deep Neural Networks (DNN)[38,39] shows that both models offer similar performance in structured data.

Validation metrics. The F1-score can be interpreted as a harmonic mean of precision and recall, where an F1-score reaches its best value at one and worst score at zero. The relative contribution of precision and recall to the F1-score are equal. The formula for the F1 score is:

$$F_\beta = (1+\beta^2)\frac{precision \times recall}{\beta^2\ precision + recall}$$

and operating,

$$F1 = \frac{TP}{TP + (FN + FP)}$$

where TP is the number of true positives, FN is the number of false negatives, and FP is the number of false positives. Better performance means lower FN and FP values, and better precision and recall mean better F1

Test	Subscale	Subject
SF-36	Physic function (PF)	Physic
	Rol physic (RP)	Physic
	Body pain (BP)	Pain
	General health (GH)	General health
	Vitality (VT)	Vitality
	Social function (SF)	Social
	Rol emotional (RE)	Emotional
	Mental health (MH)	Mental
	Physical component score (PCS)	Physic
	Mental component score (MCS)	Mental
HAD	Total anxiety	Anxiety
	Total depression	Depression
	Total HAD	Depression
FIS40	Physic dim	Physic
	Cognitive dim	Cognitive
	Social dim	Social
	Total FIS40	Physic
FIS8	FIS8	Physic
PSQI	Component 1	Sleep quality
	Component 2	Sleep quality
	Component 3	Sleep quality
	Component 4	Sleep quality
	Component 5	Sleep quality
	Component 6	Sleep quality
	Component 7	Sleep quality
	Total PSQI	Sleep quality
SCL 90 R	Somatizations (SOM)	Mental
	Obsessions (OBS)	Mental
	Interpersonal sensitivity (SI)	Mental
	Depression (DEP)	Depression
	Anxiety (ANS)	Anxiety
	Hostility (HOS)	Anxiety
	Phobic anxiety (FOB)	Anxiety
	Paranoid (PAR)	Mental
	Psychoticism (SIC)	Mental
	Severity global index (GSI)	Mental
	Positive symptoms (PST)	Mental
	Symptomatic discomfort Index (PSDI)	Mental

Table 2. Subscales and subject definitions for questionnaires. Test: Each of the analyzed questionnaires. Subscales: Every dimension defined in every questionnaire. Subject: Area that is associated with each subscale.

performance. In imbalanced data, greater accuracy than the F1 score indicates that some labels perform poorly. Recall is defined by the ratio $recall = \frac{tp}{tp+fn}$[40]. Accuracy is defined if $(y, \hat{y})$ as a (sample, predicted), then the fraction of correct predictions over samples is defined as

$$accuracy(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} 1(\widehat{y_i} = y_i)$$

Mean error is defined as the ratio of overall value questionnaire predicted versus comprehensive value sample questionnaire. The series for t-student value is defined as the sum of all answers of each questionnaire variable, predicted, and sample data.

Ethics approval. The authors declare that the procedures followed were by the regulations of the responsible Clinical Research Ethics Committee and by those of the World Medical Association and the Helsinki Declaration. The research protocols were approved by the Ethics Committee of the Vall d'Hebron University Hospital, the first "Population-based Registry of Patients with Chronic Fatigue Syndrome" approved on 18/10/2006.

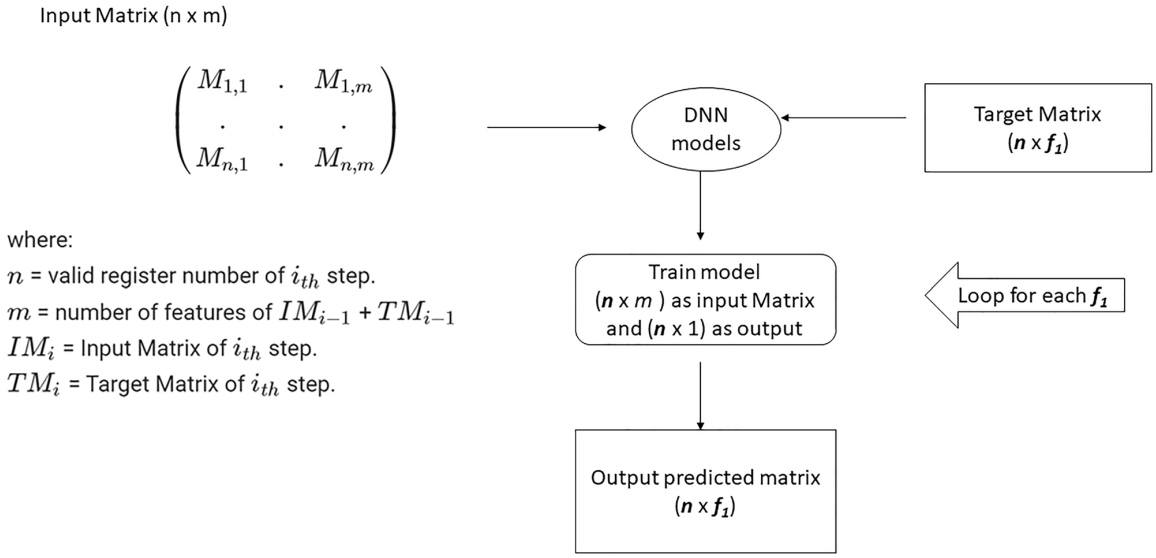


Figure 1. Modeling schema. The model requires a questionnaire answers matrix as an input value. The output is the other five questionnaires. Each questionnaire has a different number of questions and subscales. A simple sum of the number of questions calculates most subscales. For example, the SF-36 input dimension is n × 36, where n is the number of patients who answered the SF-36 questionnaire. The output is n × 186, where 186 is all five questionnaire answers.

Results

Relationship across questionnaires. In our proposed model, an $edge(i,j)$ is defined if $abs(corr(i,j)) \geq 0.5$ which indicates moderate or strong direct and indirect correlation. The 2370 registers were validated, and the Pearson correlation analyzed 38 questionnaire subscales. The subject of each subscale represents networks with each node (9) shown in Fig. 2. Mental, depression, and anxiety are strongly correlated with physical subjects. SF-36 emotional subscales are relational with anxiety, depression, and mental subscales (SCL 90 R and HAD questionnaires). As can be seen, HAD and SCL 90 R are strongly correlated. The node's size is related to the degree of the node, i.e. the number of incident edges.

In supplementary material, the second network analyzes the subscales as a node and the same relationship as an edge. The SF-36 subscales (green) have strong relationships with HAD (magenta) and FIS8, and FIS40 (strong-green and red, respectively). SCL 90 R (brown) has a strong relationship with HAD. Furthermore, PSQI (blue) has no relationship except the total psqi value. The strength of the relationship is measured in terms of the percentage of relationships between the test nodes. The initial test is SF-36, and its nodes have relationships with 100% of HAD's nodes (3 of 3) and only 25% of SCL 90 R (4 of 12). HAD's nodes have a 100% relationship with SCL 90 R's nodes. SCL 90 R has a relationship with the unique FIS8 node, which has relationships with all four FIS40 nodes. The last test with few relations is PSQI. Consequently, the order decided according to aforementioned relationships is: HAD, SCL 90 R, FIS8, FIS40, and PSQI.

Best model selection. A test comparison between XGBoost, Classifier and XGBoost Regressor using SF-36 as training data and HAD as a target with 2321 validated registers, is provided in supplementary material Fig. 3. The hyperparameter defines how our model works[41]. The parameters tuned were *max_depth, gamma, reg_alpha, reg_lambda, colsample_bytree, min_child_weight, subsample, n_estimators* and *eta*. Hyperopt has been used for hyperparameter tuning[41]. Both must be trained for each question; therefore 14 models have to be trained. The order on a set predicted value is {0, 1, 2, 3}, and the trained value is {1, 2, 3}, where in both cases, greater values show worse health status. Regressor predicted rounded to compare between real data. The results of the model are analyzed with XGBoost and the regression and classification are compared. The mean regression error is much higher than the classification error (32.50% vs. 3.16%). Therefore, the regression model is discarded in the following analyses (the results are available in the supplementary material, Table S1). Total connections have been 32,494 (2321 registers × 14 questions HAD questionnaire) and "1" and "2" answers are 67.25% of the total. The model tends to reduce the mean error, so the model predicted 70% more "1" than real and rare predicted, "3" (For more information, see Table S3 in the Supplementary Material).

Imbalanced data occur where one or more class labels have a very high number of observations, and the other has a lower one. The main problem is to increase accurate predictions of the minority class. To consider the skewed distribution of classes of different weights, classes with weights result in a penalty and a minor update of the model coefficients. The model based on the Keras library is more flexible, and for each question, it can be considered as the difference of the unbalanced data. The main difference between the Keras classifier model is

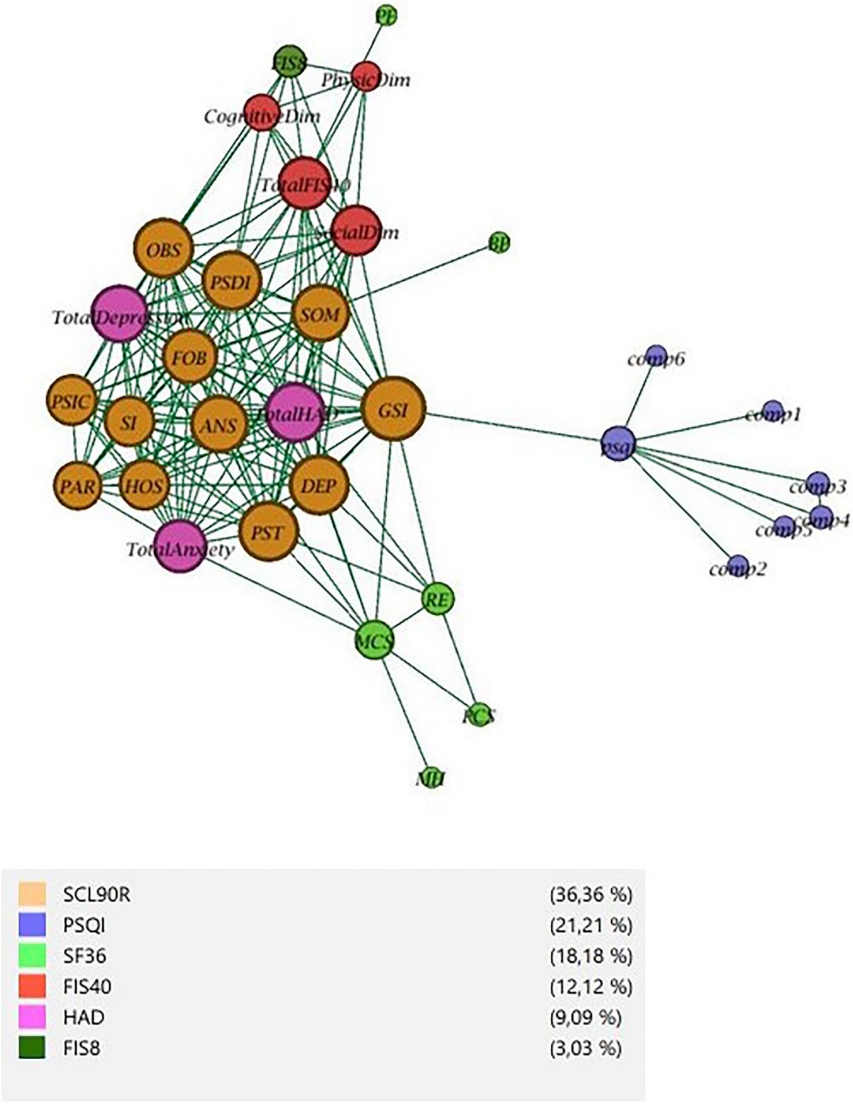


Figure 2. Subscales relationship graph. Color nodes represent the test to which nodes belong. The percentage in the legend represents the number of nodes versus the total.

the usage of the recall value, which helps to reduce the aforementioned problem with imbalanced data (for more information, see Table S3 in the Supplementary Material). For each class, do

$$classWeight_i = \frac{n}{(classes \times count_i)}$$

where n is the number of valid registers, *classes* is the number of classes, and $count_i$ is the support of ith class. Results comparison 1st questions of HAD (for more information, see Table S1–S3 in the Supplementary Material). The answer "0" has 66 (2.8%) support, and the answer "3" has 562 (24.21%) support. Minority-weighted label classes tend to be underrepresented with a low recall rate, 0.00 in the first case. These biases produce worse synthetic quality data for posterior analysis. Table 3 shows the results once corrected by the configuration in our model, improving the results significantly in those responses with low representation.

Model results. Building the model needs five steps, as depicted in Fig. 3. The first step requires an SF-36 questionnaire input matrix with 3019 registers which HAD questionnaire had the same. The output is a HAD

Answers	Precision	Recall	f1-score	Support	Class weights
0	0.62	0.79	0.69	66	8.80
1	0.71	0.83	0.76	886	0.66
2	0.66	0.59	0.62	807	0.72
3	0.83	0.69	0.75	562	1.03
Accuracy			0.71	2321	
Macro avg	0.70	0.73	0.71	2321	
Weighted avg	0.72	0.71	0.71	2321	

Table 3. Keras weighted model results.

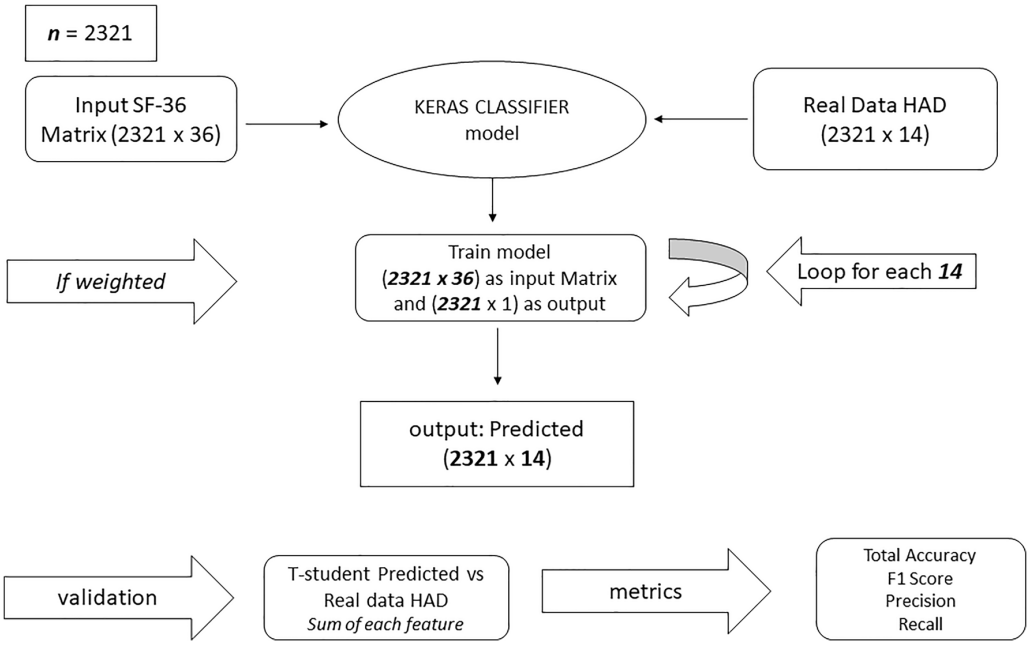


Figure 3. Keras Classifier algorithm schema.

Metrics	Steps models summary				
	STEP 1	STEP 2	STEP 3	STEP 4	STEP 5
Inputs models (dimension)	SF-36 (2321 × 36)	SF-36 + HAD (2314 × 50)	SF-36 + HAD + SCL 90 R (2019 × 140)	SF36 + HAD + SCL 90 R + FIS8 (2019 × 148)	SF-36 + HAD + SCL 90 R + FIS8 + FIS40 (1902 × 188)
Accuracy	0.67	0.78	0.81	0.78	0.78
Precision	0.69	0.78	0.83	0.79	0.80
Recall	0.72	0.76	0.76	0.74	0.76
F1 score	0.70	0.77	0.81	0.76	0.78
Mean error	− 1.35%	− 1.22%	− 2.59%	− 2.81%	− 5.50%
t-student	0.79	0.85	0.67	0.18	0.37
output	HAD	SCL 90 R	FIS8	FIS40	PSQI

Table 4. Final model result. Each question of each step needs different parameters, so it has to train 188 models with other parameters.

Inputs models (dimension)	SF-36 (2321 × 36)	SF-36 + HAD (2314 × 50)	SF-36 + HAD + SCL 90 R (2019 × 140)	SF36 + HAD + SCL 90 R + FIS8 (2019 × 148)	SF-36 + HAD + SCL 90 R + FIS8 + FIS40 (1902 × 188)
Layers	4	3	4	4	4
Dropout	2	2	3	3	2
epochs	4000	3000	3000	3000	3000
Monitor	*Val_recall*	*Val_recall*	*Val_recall*	*Val_recall*	*Val_recall*
Early stopping patience	400	300	300	300	400
Neurons [layer]	[400,400,200,100]	[1500,1500,750]	[1000,1000,500,250]	[1000,1000,500,250]	[500,500,250, 100]
output	HAD	SCL 90 R	FIS8	FIS40	PSQI

Table 5. Steps models summary.

synthetic matrix. The second step requires an input matrix of SF-36 + HAD (synthetic data) and produces synthetic SCL 90 R responses and so on. The results are detailed in Tables 4, 5.

Discussion

Given the SF-36 questionnaire data can create using a new model, synthetic responses from other questionnaires inform the impact of fatigue, psychological phenomena, and sleep dysfunction. The lack of risk-free health data is an issue in ME/SFC hospital units and investigators. This open-source project offers a tool to generate risk-free synthetic data for the health IT and clinical community to use, experiment, and create more synthetic data. The quality based on validation tests did not cover projects or research focused on clinical discovery. Synthetic data can be an alternative to ground truth when data access is restricted and an excellent alternative to machine learning training/testing datasets[26].

The SF-36 includes one multi-item scale that assesses eight health concepts: (1) limitations in physical activities because of health problems; (2) limitations in social activities because of physical or emotional problems; (3) limitations in usual role activities because of physical health problems; (4) bodily pain; (5) general mental health (psychological distress and well-being); (6) limitations in usual role activities because of emotional problems; (7) vitality (energy and fatigue), and (8) general health perceptions and is one of the most used quality life questionnaires used and evaluated[41]. The other five questionnaires used in this work complement most information about the quality of life of ME/SFC patients.

The questionnaires can be answered quickly and are regularly available in primary care and specialized medical consultations. Some applications offer automated analyzed results that inform essential information about patient health conditions.

The graph theory has been used to decide the order of the modeling cascade. Although a deeper analysis of these relationships should be the subject of another, more specific work, in this case, it informs us of the order used in our model. These relationships will characterize our model, which will be more robust with more records analyzed. Our dataset is unusually great in SFC, which becomes robust to our models.

Our synthetic dataset generator applications fill in missing data of real datasets from any other five questionnaires. For those, ME/SFC dataset clinical units with SF-36 questionnaire answers but missing others could build a complete dataset.

Limitations

(1) Single-center trial. (2) Unit of reference in diagnosing and treating CFS/ME, which may be biased towards more severe cases and a longer evolution time than studies in primary care. (3) No information is available on parameters such as the results of the two-day ergometric test for assessing exercise intolerance, a neuropsychological battery for assessing cognitive impairment, and neurovegetative dysfunction, e.g., heart rate variability. (4) That this is a prospective study with cross-sectional data collection. It is not a longitudinal study.

Conclusion

Synthetic patients can be simulated with models of ME/CFS questionnaires data and corresponding standards of care to produce risk-free realistic synthetic healthcare records at scale. An open-source generator offers high-fidelity synthetic data for investigation and educational use, free of legal, privacy, security, and intellectual property restrictions.

Data availability

GitHub is an online platform where researchers and software developers share their work with the scientific community. The following link shares the work described here. The datasets generated and/or analyzed during the current study are available in the SFCSyntheticDataGenerator repository, https://github.com/mlacasa/SFCSyntheticDataGenerator

Received: 22 March 2023; Accepted: 9 August 2023
Published online: 31 August 2023

References

1. Lim, E.-J. *et al.* Systematic review and meta-analysis of the prevalence of chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME). *J. Transl. Med.* **18**, 100. https://doi.org/10.1186/s12967-020-02269-0 (2020).

2. Morris, G. *et al.* Myalgic encephalomyelitis/chronic fatigue syndrome: From pathophysiological insights to novel therapeutic opportunities. *Pharmacol. Res.* **148**, 104450. https://doi.org/10.1016/j.phrs.2019.104450 (2019).
3. Castro-Marrero, J. *et al.* Unemployment and work disability in individuals with chronic fatigue syndrome/myalgic encephalomyelitis: A community-based cross-sectional study from Spain. *BMC Public Health* **19**, 840. https://doi.org/10.1186/s12889-019-7225-z (2019).
4. Maes, M. & Twisk, F. N. Why myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) may kill you: Disorders in the inflammatory and oxidative and nitrosative stress (IO&NS) pathways may explain cardiovascular disorders in ME/CFS. *Neuroendocrinol Lett.* **30**, 677–693 (2009).
5. Castro-Marrero, J. *et al.* Poor self-reported sleep quality and health-related quality of life in patients with chronic fatigue syndrome/myalgic encephalomyelitis. *J. Sleep Res.* **27**, e12703. https://doi.org/10.1111/jsr.12703 (2018).
6. Castro-Marrero, J. *et al.* Comorbidity in chronic fatigue syndrome/myalgic encephalomyelitis: A nationwide population-based cohort study. *Psychosomatics* **58**, 533–543. https://doi.org/10.1016/j.psym.2017.04.010 (2017).
7. Castro-Marrero, J., Sáez-Francàs, N., Santillo, D. & Alegre, J. Treatment and management of chronic fatigue syndrome/myalgic encephalomyelitis: All roads lead to Rome. *Br. J. Pharmacol.* **174**, 345–369. https://doi.org/10.1111/bph.13702 (2017).
8. Fisk, J. D. *et al.* Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. *Clin. Infect. Dis.* **18**(Suppl 1), S79-83. https://doi.org/10.1093/clinids/18.supplement_1.s79 (1994).
9. Fisk, J. D. & Doble, S. E. Construction and validation of a fatigue impact scale for daily administration (D-FIS). *Qual. Life Res.* **11**, 263–272. https://doi.org/10.1023/a:1015295106602 (2002).
10. Buysse, D. J., Reynolds, C. F. 3rd., Monk, T. H., Berman, S. R. & Kupfer, D. J. The pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Res.* **28**, 193–213. https://doi.org/10.1016/0165-1781(89)90047-4 (1989).
11. Alonso, J., Prieto, L. & Antó, J. M. The Spanish version of the SF-36 health survey (the SF-36 health questionnaire): An instrument for measuring clinical results. *Med. Clin.* **104**, 771–776 (1995).
12. McGregor, N. R. *et al.* A preliminary assessment of the association of SCL-90-R psychological inventory responses with changes in urinary metabolites in patients with chronic fatigue syndrome. *J. Chronic Fatigue Syndr.* **3**, 17–37. https://doi.org/10.1300/J092v03n01_03 (1997).
13. Castresana, C., Perez, A. G. -E., de Rivera, J. L. G. Hospital anxiety and depression scale y psicopatología afectiva. Anales de psiquiatría. pp. 126–130. (1995) Available: https://www.academia.edu/download/51823551/95_A138_03.pdf
14. Castro-Marrero, J. *et al.* Effect of dietary coenzyme Q10 plus NADH supplementation on fatigue perception and health-related quality of life in individuals with myalgic encephalomyelitis/chronic fatigue syndrome: A prospective, randomized, double-blind placebo-controlled trial. *Nutrients* https://doi.org/10.3390/nu13082658 (2021).
15. Watson, D. S. *et al.* Clinical applications of machine learning algorithms: Beyond the black box. *BMJ* **364**, l886. https://doi.org/10.1136/bmj.l886 (2019).
16. Kitami, T. *et al.* Deep phenotyping of myalgic encephalomyelitis/chronic fatigue syndrome in Japanese population. *Sci. Rep.* **10**, 19933. https://doi.org/10.1038/s41598-020-77105-y (2020).
17. Słomko, J. *et al.* Autonomic phenotypes in chronic fatigue syndrome (CFS) are associated with illness severity: A cluster analysis. *J. Clin. Med. Res.* **9**, 254. https://doi.org/10.3390/jcm9082531 (2020).
18. Zhang, F. *et al.* Artificial intelligence based discovery of the association between depression and chronic fatigue syndrome. *J. Affect. Disord.* **250**, 380–390. https://doi.org/10.1016/j.jad.2019.03.011 (2019).
19. Hanson, S. J., Gause, W. & Natelson, B. Detection of immunologically significant factors for chronic fatigue syndrome using neural-network classifiers. *Clin. Diagn. Lab. Immunol.* **8**, 658–662. https://doi.org/10.1128/CDLI.8.3.658-662.2001 (2001).
20. Levine, P. H. *et al.* Clinical, epidemiologic, and virologic studies in four clusters of the chronic fatigue syndrome. *Arch. Intern. Med.* **152**, 1611–1616 (1992).
21. Asprusten, T. T., Sletner, L. & Wyller, V. B. B. Are there subgroups of chronic fatigue syndrome? An exploratory cluster analysis of biological markers. *J. Transl. Med.* **19**, 48. https://doi.org/10.1186/s12967-021-02713-9 (2021).
22. Hunter, P. The big health data sale: as the trade of personal health and medical data expands, it becomes necessary to improve legal frameworks for protecting patient anonymity, handling consent and ensuring the quality of data. *EMBO Rep.* **17**, 1103–1105. https://doi.org/10.1525/embr.201642917 (2016).
23. Tate, A. R. *et al.* Exploiting the potential of large databases of electronic health records for research using rapid search algorithms and an intuitive query interface. *J. Am. Med. Inform. Assoc.* **21**, 292–298. https://doi.org/10.1136/amiajnl-2013-001847 (2014).
24. Moniz, L. *et al.* Construction and validation of synthetic electronic medical records. *Online J. Public Health Inform.* https://doi.org/10.5210/ojphi.v1i1.2720 (2009).
25. Weiss, J. C., Page, D. Forest-based point process for event prediction from electronic health records. Machine learning and knowledge discovery in databases. Springer Berlin Heidelberg. 547–562 (2013). https://doi.org/10.1007/978-3-642-40994-3_35
26. Wang, Z., Myles, P. & Tucker, A. Generating and evaluating cross-sectional synthetic electronic healthcare data: Preserving data utility and patient privacy. *Comput. Intell.* **37**, 819–851. https://doi.org/10.1111/coin.12427 (2021).
27. Tucker, A., Wang, Z., Rotalinti, Y. & Myles, P. Generating high-fidelity synthetic patient data for assessing machine learning healthcare software. *NPJ. Digit. Med.* **3**, 147. https://doi.org/10.1038/s41746-020-00353-9 (2020).
28. Qin, Z., Liu, Z., Zhu, P. & Xue, Y. A GAN-based image synthesis method for skin lesion classification. *Comput. Methods Programs Biomed.* **195**, 105568–105616. https://doi.org/10.1016/j.cmpb.2020.105568 (2020).
29. Ravi, D. *et al.* Degenerative adversarial neuroimage nets for brain scan simulations: Application in ageing and dementia. *Med. Image Anal.* **75**, 102257. https://doi.org/10.1016/j.media.2021.102257 (2022).
30. Islam, J. & Zhang, Y. GAN-based synthetic brain PET image generation. *Brain Inform.* **7**, 3. https://doi.org/10.1186/s40708-020-00104-2 (2020).
31. Li, Q., Yu, Z., Wang, Y. & Zheng, H. TumorGAN: A multi-modal data augmentation framework for brain tumor segmentation. *Sensors* https://doi.org/10.3390/s20154203 (2020).
32. Tien, H.-J., Yang, H.-C., Shueng, P.-W. & Chen, J.-C. Cone-beam CT image quality improvement using Cycle-Deblur consistent adversarial networks (Cycle-Deblur GAN) for chest CT imaging in breast cancer patients. *Sci. Rep.* **11**, 1133. https://doi.org/10.1038/s41598-020-80803-2 (2021).
33. Fukuda, K. *et al.* The chronic fatigue syndrome: A comprehensive approach to its definition and study. *Ann. Intern. Med.* **121**, 953–959. https://doi.org/10.7326/0003-4819-121-12-199412150-00009 (1994).
34. Carruthers, B. M. *et al.* Myalgic encephalomyelitis: International consensus criteria. *J. Int. Med.* **2415**, 327–338. https://doi.org/10.1111/j.1365-2796.2011.02428.x (2011).
35. Newman, M. Networks. Oxford University Press; 2018. Available: https://play.google.com/store/books/details?id=YdZjDwAAQBAJ
36. Hagberg, A., Swart, P. S., Chult, D. Exploring network structure, dynamics, and function using NetworkX. Los Alamos National Lab.(LANL), Los Alamos, NM (United States); (2008). Available: https://www.osti.gov/biblio/960616
37. Suchowski, M. A. An analysis of the impact of an outlier on correlation coefficients across small sample data where rho is non-zero. Western Michigan UniversityProQuest Dissertations Publishing, Degree Year. p. 3007026. Available: https://search.proquest.com/openview/5d1cbf13c930b7358050381ebab41a85/1?pq-origsite=gscholar&cbl=18750&diss=y
38. Ferreira, L., Pilastri, A., Martins, C. M., Pires, P. M., Cortez, P. A comparison of AutoML tools for machine learning, deep learning and XGBoost. 2021 International Joint Conference on Neural Networks (IJCNN). pp. 1–8. (2021). https://doi.org/10.1109/IJCNN52387.2021.9534091

39. Park, D. J. *et al.* Development of machine learning model for diagnostic disease prediction based on laboratory tests. *Sci. Rep.* **11**, 7567. https://doi.org/10.1038/s41598-021-87171-5 (2021).
40. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
41. Chen, T., Guestrin, C. XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: Association for Computing Machinery; pp. 785–794 (2016). https://doi.org/10.1145/2939672.2939785

Author contributions

All authors contributed to the study's conception and design. M.L. and J.A performed material preparation, data collection, and analysis. M.L. wrote the first draft of the manuscript, and all authors commented on previous versions. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-40364-6.

Correspondence and requests for materials should be addressed to M.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nutrients

IMPACT FACTOR 5.9

Indexed in: PubMed

CITESCORE 9.0

Article

Yeast Beta-Glucan Supplementation with Multivitamins Attenuates Cognitive Impairments in Individuals with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Randomized, Double-Blind, Placebo-Controlled Trial

Marcos Lacasa, Jose Alegre-Martin, Ramon Sanmartin Sentañes, Luisa Varela-Sende, Joanna Jurek and Jesus Castro-Marrero

Special Issue
Nutrition and Dietary Patterns: Effects on Brain Function

Edited by
Prof. Dr. Roser Granero and Prof. Dr. Gemma Guillazo-Blanch

MDPI

https://doi.org/10.3390/nu15214504

nutrients

MDPI

Article

Yeast Beta-Glucan Supplementation with Multivitamins Attenuates Cognitive Impairments in Individuals with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Randomized, Double-Blind, Placebo-Controlled Trial

Marcos Lacasa [1,†], Jose Alegre-Martin [2], Ramon Sanmartin Sentañes [2], Luisa Varela-Sende [3], Joanna Jurek [2,†] and Jesus Castro-Marrero [2,*,†]

[1] E-Health Center, Universitat Oberta de Catalunya, 08018 Barcelona, Spain; mlacasaca@uoc.edu
[2] Myalgic Encephalomyelitis/Chronic Fatigue Syndrome Research Unit, Division of Rheumatology, Vall d´Hebron University Hospital Research Institute, Universitat Autònoma de Barcelona, 08035 Barcelona, Spain; jose.alegre@vallhebron.cat (J.A.-M.); ramon.sanmartin@vallhebron.cat (R.S.S.); joanna.michalina.jurek@gmail.com (J.J.)
[3] Clinical Research Department, VITAE Health Innovation, Montmeló, 08160 Barcelona, Spain; lvarela@vitae.es
* Correspondence: jesus.castro@vhir.org; Tel.: +34-93-4893000 (ext. 3753)
† These authors contributed equally to this work.

Abstract: This research aimed to examine the potential alleviative effects of beta-glucan administration on fatigue, unrefreshing sleep, anxiety/depression symptoms and health-related quality of life in ME/CFS. A 36-week unicenter, randomized, double-blind, placebo-controlled trial was conducted in 65 ME/CFS patients, who were randomly allocated to one of two arms to receive four capsules each one of 250 mg beta-glucan, 3.75 µg vitamin D3, 1.05 mg vitamin B6, and 7.5 mg zinc (n = 35), or matching placebo including only microcrystalline cellulose as an excipient (n = 30) once daily. The findings showed that the beta-glucan supplementation significantly improved cognitive fatigue (assessed with FIS-40 scores) after the 36-week treatment compared to the baseline ($p = 0.0338$). Taken together, this study presents the novel finding that yeast-derived beta-glucan may alleviate cognitive fatigue symptoms in ME/CFS. Thus, it offers valuable scientific insights into the potential use of yeast beta-glucan as a nutritional supplement and/or functional food to prevent or reduce cognitive dysfunction in patients with ME/CFS. Further interventions are warranted to validate these findings and also to delve deeper into the possible immunometabolic pathomechanisms of beta-glucans in ME/CFS.

Keywords: chronic fatigue syndrome; beta-glucan; zinc; vitamin D_3; vitamin B_6; myalgic encephalomyelitis; mitochondria; non-restorative sleep; quality of life

Citation: Lacasa, M.; Alegre-Martin, J.; Sentañes, R.S.; Varela-Sende, L.; Jurek, J.; Castro-Marrero, J. Yeast Beta-Glucan Supplementation with Multivitamins Attenuates Cognitive Impairments in Individuals with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Randomized, Double-Blind, Placebo-Controlled Trial. *Nutrients* **2023**, *15*, 4504. https://doi.org/10.3390/nu15214504

Academic Editors: Roser Granero and Gemma Guillazo-Blanch

Received: 3 October 2023
Revised: 18 October 2023
Accepted: 23 October 2023
Published: 24 October 2023

1. Introduction

Myalgic encephalomyelitis or chronic fatigue syndrome (ME/CFS) is a complex debilitating disorder affecting approximately 17 million people worldwide with a prevalence ranging between 0.2 and 2.6% [1] and a higher frequency in women [2]. ME/CFS is characterized by unexplained and persistent post-exertional fatigue lasting for at least six months, which worsens with physical or mental activity and is not relieved by rest or sleep [3]. ME/CFS is often accompanied by a cluster of other symptoms such as cognitive impairments, impaired immune function, neuroendocrine alterations, and autonomic dysfunction [4]. The condition represents a considerable public health problem and is a cause of severe disability in society, with a marked impact on professional activities and on social and personal relationships [5,6]. All these symptoms can significantly compromise the quality of life of sufferers, as they limit the ability of ME/CFS patients to perform daily tasks [7,8].

In addition, the sleep disturbances commonly reported in ME/CFS patients have been shown to exaggerate these cognitive symptoms and make them more difficult to control [9,10]. At present, there are no specific tests for diagnosing ME/CFS, and so a comprehensive clinical assessment is essential, especially when evaluating the impact of fatigue, quality of life, anxiety/depression symptoms, and health-related quality of life; this assessment should include consultation of medical records, physical examination, laboratory testing, and neuroimaging techniques, along with validated neuropsychiatric tools [11,12]. Although the exact pathophysiological mechanisms underlying cognitive dysfunction are not yet fully understood, alterations in the autonomic nervous system and cerebral blood flow have been proposed as possible causes [13]. In addition, poor sleep quality in ME/CFS has been shown to worsen mental and muscular fatigue, causing intolerance to physical exercise and presenting clinically as myalgia and insufficient/poor blood flow to the brain with additional disturbances in energy production [10].

Functional alterations of redox metabolism, gene expression, and mitochondrial biomarkers in peripheral blood mononuclear cells (PBMC), aerobic oxidative metabolism, decreased aerobic activity, and deconditioning after physical exercise have also been described in ME/CFS [4,14]. High levels of pro-inflammatory cytokines, such as interleukin-1 (IL-1), tumor necrosis factor (TNF-α), and elastase, as well as high oxidative and nitrosative stress, may have multiple effects; they can inhibit mitochondrial respiration, decrease electron transport chain activity and change mitochondrial membrane potential, increase the mitochondrial membrane permeability, and interfere with the ATP production, finally leading to mitochondrial inactivation [15]. Along with morning fatigue, exercise intolerance, and alterations in concentration and memory, sleep disturbances is a particular salient feature of ME/CFS, in the form of unrefreshing sleep, insomnia, hypersomnia, presence of nightmares and/or bad dreams, restless leg syndrome, and sleep apnea syndrome [16].

In addition, many individuals with ME/CFS report gastrointestinal complaints such as irritable bowel syndrome (IBS), a common functional disorder of the gastrointestinal tract, characterized by abdominal pain or discomfort and altered bowel habit. Although the exact mechanism of the gut disturbances in ME/CFS needs to be explored further, the high frequency of IBS diagnoses in ME/CFS patients suggests that impaired composition of intestinal microbiota along with immune dysfunction and increased inflammation beyond the gut may play a role in the onset and development of the illness [17]. The involvement of a bidirectional link between the gut and the central nervous system (CNS) in ME/CFS, known as the gut–brain axis, was supported by studies showing improvements of various symptoms, including anxiety/depression, sleep quality, and neurocognitive impairments after probiotic supplementation interventions [18,19]. Moreover, the dysbiosis of the gut microbiome observed in ME/CFS was linked to the elevated levels of circulating inflammatory mediators, possibly due to an increase in intestinal permeability that allowed for bacterial translocation [20].

Studies investigating the role of short-chain fatty acids (SCFA) in ME/CFS pathogenesis have suggested that SCFA deficiency, in particular butyrate-producing bacteria, is associated with fatigue severity in these individuals [21], as well in other chronic comorbidity conditions, including IBS [22], cancer [23], multiple sclerosis [24], and type 1 diabetes [25]. In view of these early findings, enhancing butyrate production in the gut may have a beneficial therapeutic effect in ME/CFS. One way to increase SCFA levels is through the supplementation of prebiotics, in form of non-digestible polysaccharides (NPS), which upon fermentation, can selectively stimulate the growth and/or activity of anti-inflammatory bacteria in the gut [25]. Beta-glucan, an example of an NPS found as glucose polymers in yeast, fungi, algae, and cereals, such as oats and barley, can promote the growth of probiotic bacteria and increase the production of SCFA, in particularly butyrate [26,27].

Although many studies have investigated the health benefits of beta-glucan, including the prevention and treatment of chronic gut inflammation and conditions related to metabolic disturbances and neurodegeneration, its use in ME/CFS is still under inves-

tigation [26,28,29]. Experimental pre-clinical studies using murine models have shown that beta-glucan supplementation can reduce microbial translocation, restore immune homeostasis, and decrease fatigue [30]; these early findings are now under investigation in clinical trials (https://clinicaltrials.gov, accessed on 11 May 2023; NCT05726435 and NCT05524688). The accompanying synergistic improvements in terms of blood pressure, decreased confusion, and improved mood can also help to reduce perceived fatigue [30,31].

Furthermore, by promoting the activity of the desired intestinal microbiota and limiting the growth of pathogens, beta-glucan can play an important role in maintaining proper gut function and thus prevent chronic inflammation. For example, the use of a yeast-derived beta-glucan was effective in reducing hyperpermeability in ileal specimens from patients with Crohn's disease [32], and it also reduced chronic abdominal pain associated with altered bowel habits in an experimental IBS murine model [33]. Similar improvements in bloating, flatulence, and abdominal pain were also observed in a clinical intervention with a mixture of beta-glucan, inositol, and digestive enzymes in IBS patients [34]. This finding may be of particular importance in ME/CFS, as the gastrointestinal issues frequently reported by these patients are often associated with psychological distress and increased release of corticotrophin-releasing hormone, which is known to be involved in the stimulation the hypothalamic–pituitary–adrenal (HPA) axis and is associated with major depression and IBS [35,36].

Thus, beta-glucan supplementation is associated with numerous health benefits including improved gut health, enhanced immune function, and reduced fatigue, and may provide a potential therapeutic advantage for ME/CFS. The present randomized placebo-controlled trial aimed to assess the effect of ImmunoVita®, a dietary supplement composed of yeast-derived beta-glucan combined with vitamin D3, vitamin B6, and zinc on fatigue, sleep problems, anxiety/depression, and health-related quality of life in ME/CFS.

2. Materials and Methods

2.1. Participants

This study was conducted in 67 Caucasian ME/CFS patients consecutively recruited from a single outpatient tertiary referral center (ME/CFS Clinical Unit, Vall d'Hebron University Hospital, Barcelona, Spain) from September 2021 to December 2022. Figure 1 shows a flowchart of the participants prior to analysis. Patients were potentially eligible for the study if they were female, aged 18 years or older, and had a confirmed diagnosis of ME/CFS according to the 1994 CDC/Fukuda case definition [37].

Exclusion criteria comprised participation in another intervention within 30 days prior to study inclusion; inability (in the opinion of the investigator) to follow the instructions or to complete the treatment satisfactorily; failure to provide signed informed consent; consumption of certain drugs/supplements that might influence outcome measures in the last 90 days or whose withdrawal might be a relevant problem, anticoagulant treatment, pregnancy or breast-feeding, smoking, alcohol intake or substance abuse, BMI $\geq$ 25 kg/m^2, and hypersensitivity to any of the components of treatments. Patients with missing data from the follow-up visits were considered to have dropped out. The participants signed an informed consent form prior to the study.

2.2. Intervention Protocol

Of the 67 eligible ME/CFS participants screened, two were excluded. The remaining 65 participants were allocated to treatment by an independent investigator not otherwise involved in the intervention, using a list of random numbers generated by a computer program. The participants were randomly assigned in a double-blind fashion in a 1:1 ratio to receive either active treatment (n = 35) or a matching placebo (n = 30) in the form of four capsules daily for nine months. They were instructed to ingest the capsules on an empty stomach, 30 minutes before breakfast and dinner, only with water. The intervention comprised the intake of a food supplement containing beta-glucan, vitamin D3, vitamin

B6, and zinc. The placebo group received the same food supplement containing only microcrystalline cellulose instead of active ingredients.

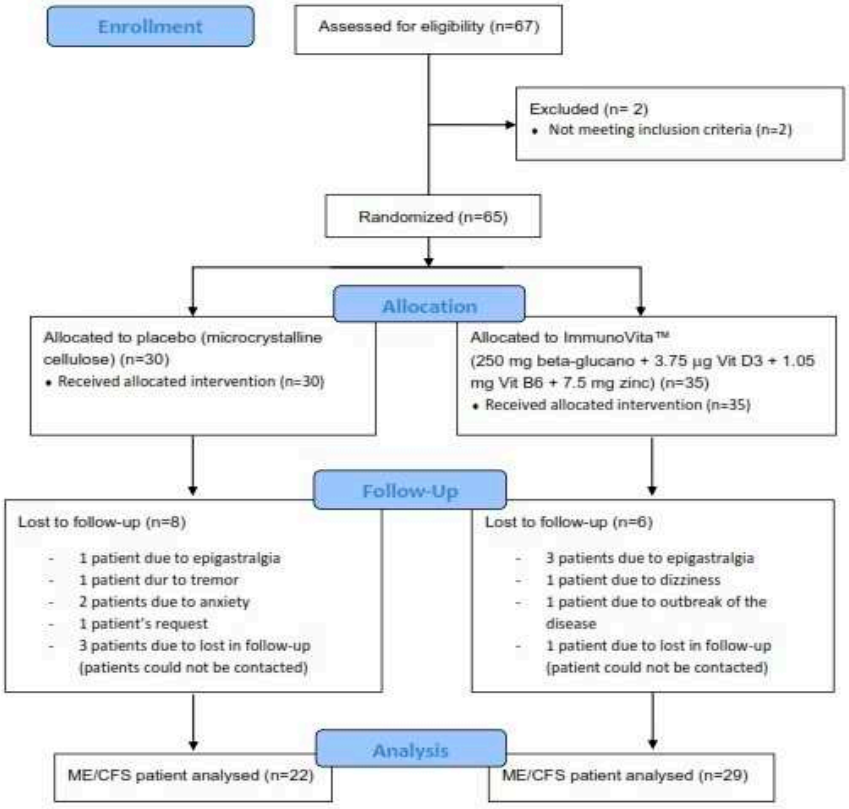


Figure 1. Consolidated Standards of Reporting Trials (CONSORT) flow diagram illustrating the steps of screening, enrollment, assignment, and follow-up of the study participants.

Safety information recorded at all study visits included data on adverse events, vital signs, and the results of the general physical examination. Adverse events, including serious ones, were reviewed throughout the trial by the independent medical monitor, the steering committee, and the independent data and safety monitoring board. Provision was made for investigator-initiated temporary or permanent dose reductions or suspensions due to adverse effects.

During the study, nine subjects dropped out due to adverse events, specifically five in the intervention group (three epigastralgias, one dizziness, and one outbreak of the disease) and four in the placebo group (one epigastralgia, one tremor, and two anxiety episodes). Four patients were lost to follow-up (one in the intervention group and three in the placebo group). Finally, one patient in the placebo group was withdrawn at her own request. The remaining 51 cases of ME/CFS (78%, 22 in the placebo group and 29 in the intervention group) completed all the study protocol procedures and were included in the overall analysis of outcome measures as displayed in Figure 1.

2.3. Testing of Dietary Supplements

Patients randomized to the intervention's experimental group received a daily dose of four capsules composed of 250 mg beta-glucan, 3.75 µg vitamin D3, 1.05 mg vitamin B6, and 7.5 mg zinc, plus microcrystalline cellulose and plant capsule as excipients. The placebo composition was the same capsule with microcrystalline cellulose without any active ingredients. The treatments were identical in terms of size, color, opacity, shape, presentation,

and packaging. All capsules were manufactured and donated by Vitae Health Innovation S.L. (Montmeló, Barcelona, Spain). The study pharmacist recorded all treatments supplied on the medication-dispensing forms along with their original prescription.

2.4. Study Design and Procedures

The trial was a 36-week long, single-center, randomized, double-blind, placebo-controlled study. Clinical visits and trial design of both groups are detailed in Figure 2. After an oral explanation of the study, all participants provided their written consent prior to the commencement of the study, and they received no compensation for their participation.

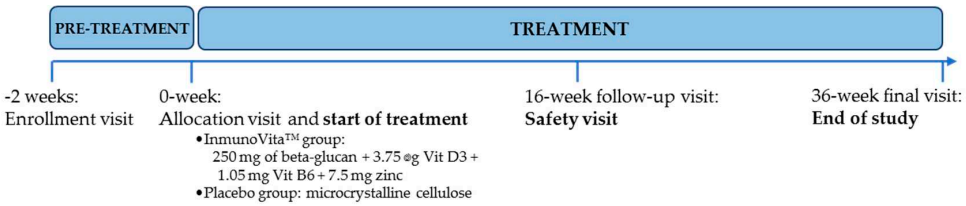


Figure 2. Summary of the study schedule at each visit during the clinical trial.

Patients were evaluated at baseline, at a 16-week follow-up (safety) visit, and then at a 36-week visit by the site investigator. Changes in symptoms were assessed through validated self-report questionnaires completed by participants under the supervision of two trained investigators (J.C.-M. and J.A.). Compliance was checked through medication logs. The use of concomitant medications was tracked at the 36-week visit. The study protocol was reviewed and approved by the local IRB at the participating site (Clinical Research Ethics Committee, Vall d'Hebron University Hospital, Barcelona, Spain, under protocol reference InmunoVitaME-PR(AG)-447-2019, approved on 28 February 2019).

The study protocol was conducted in accordance with the guidelines of the Declaration of Helsinki, the current Spanish regulations on clinical research, and the standards of good clinical practice of the European Union. It also followed the Consolidated Standards of Reporting Trials (CONSORT) guidelines. The current clinical trial was registered on https://clinicaltrials.gov as NCT04301609.

2.5. Primary Endpoint

Fatigue Perception

The primary endpoint was the change in self-reported fatigue perception assessed by using the validated Fatigue Impact Scale (FIS-40) questionnaire from the baseline to the final (week 36) study visit. Briefly, the FIS-40 comprises 40 items divided into three domains that describe how perceived fatigue impacts on cognitive (10 items), physical (10 items), and psychosocial functioning (20 items) over the previous four weeks. Each item is scored from 0 (no fatigue) to 4 (severe fatigue). The total score is calculated by adding together responses from the 40 questions (score range 0–160). Higher scores indicate more functional limitations due to severe fatigue [38].

2.6. Secondary Endpoints

The secondary outcome measures included changes in sleep disturbance, anxiety/depression symptoms, and health-related quality of life (HRQoL) on the validated self-reported questionnaires.

2.6.1. Sleep Quality

Sleep quality was assessed using the self-administered 19-item Pittsburgh Sleep Quality Index (PSQI) questionnaire. Scores are obtained on each of seven domains of sleep quality: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep perturbations, use of sleeping medication, and daytime dysfunction. Each component

is scored from 0 to 3 (0 = no sleep problems and 3 = severe sleep problems). The global PSQI score ranges from 0 to 21 points, with scores of >5 indicating poorer sleep quality [39].

2.6.2. Anxiety and Depression

Severity of anxiety/depression symptoms was assessed using the Hospital Anxiety and Depression Scale (HADS), a validated self-reported tool composed of 14 items (seven related to anxiety symptoms and seven to depression). Each item on the HADS questionnaire is scored from 0–3, and so scores range from 0 to 21; scores of 0–7 are interpreted as normal, 8–10 as mild, 11–14 as moderate, and 15–21 as severe for either anxiety or depression. The total HADS score ranges from 0 (no anxiety or depression) to 42 (severe anxiety and depression) [40].

2.6.3. Health-Related Quality of Life

The 36-item Short Form Health Survey (SF-36) was used to assess HRQoL. The SF-36 is a broadly based self-reported survey of health-related physical and mental functioning statuses. It assesses functioning on eight subscales including domains of physical functioning, physical role, bodily pain, general health, social functioning, vitality, emotional role, and mental health, and two general subscales covering the physical and mental health domains rated on a scale from 0–100. Lower scores indicate a more negative impact on health and daily functioning [41].

2.7. Sample Size Estimation and Power Analysis

This trial was the first exploratory, population-based proof-of-concept study in people with ME/CFS. Sixty participants were enrolled with 30 patients being randomly allocated to each arm.

2.8. Compliance Monitoring and Adverse Events

All participants were asked to return any remaining study products after the intervention. Adherence was measured by calculating all remaining capsules for all patients, including withdrawals. Participants who did not take the supplement for more than two days (either consecutive or non-consecutive) were considered non-compliant (n = 0). All adverse events following administration and intake of the study product were monitored until the end of the study.

2.9. Statistical Analysis

A total of 51 patients were analyzed, divided into two groups: active (n = 29) and placebo (n = 22). The variables were evaluated (1) at different times (questionnaire domains) and (2) at the time of inclusion (demographic and clinical characteristics). For the purpose of questionnaire analysis, a total of 23 numerical values corresponding to subscales of the FIS-40, PSQI, HADS, and SF-36 outcome measures were used. For each one, two samples collected at different times were analyzed, as shown in the CONSORT chart. For variables describing demographic and clinical characteristics, such as age, BMI, heart rate, and medication use, a total of four numerical and five dichotomous variables obtained at baseline were analyzed. The normality of the data was evaluated for all numerical variables using the Shapiro–Wilk test, and the statistical value W was calculated using the appropriate formula. For all cases, results were considered significant for an alpha value = 0.05, and the *p*-value was reported with four digits or as indicated if the value was less than 0.001. All statistical analyses were performed using Python package Pingouin (version 0.5.2) for Windows (University of California, Berkeley, CA, USA).

$$W = \frac{\left(\sum_{i=1}^{n} a_i\, x_i\right)^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2}$$

2.9.1. Independent Sample Analysis

The normality of independent samples was analyzed using the Welch–Satterthwaite equation, an approximation of the adjusted degrees of freedom, since the sizes were assumed to be unequal. The formula used was as follows (Delacre, Lakens, and Leys, n.d.):

$$v = \frac{\left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}\right)^2}{\frac{\left(\frac{s_x^2}{n_x}\right)^2}{(n_x - 1)} + \frac{\left(\frac{s_y^2}{n_y}\right)^2}{(n_y - 1)}}$$

The Mann–Whitney *U* test, a nonparametric test of the null hypothesis in which a value chosen at random from a sample is equally likely to be less than or greater than a random value, was analyzed. A brute force version of the formula from Vargha and Delaney, 2000, was used. For all cases, results were considered significant for an alpha value = 0.05, and the *p*-value was reported with four digits or as indicated if the value was less than 0.001.

2.9.2. Paired Data Analysis

To analyze paired samples of questionnaires whose values were collected at different times, the paired sample t-test was used to assess the data normality. If data were non-normal, the Wilcoxon signed-rank test was applied.

2.9.3. Categorical Data Analysis

To evaluate the balance of the samples of the two groups, the Chi-squared test of independence was used consisting of a 2 × 2 contingency table, since the variables were dichotomous and the degrees of freedom were 1 in all cases. For all cases, results were considered significant for an alpha value = 0.05, and the *p*-value was reported with four digits or as indicated if the value was less than 0.001.

3. Results

3.1. Participants' Characteristics

The baseline demographic and clinical features of the participants are displayed in Table 1. In this study, there was no statistically significant differences in participants' demographic and clinical data during the intervention (Table 1). The mean (SD) age of patients was 52.90 (6.47) in the intervention arm and 52.5 (7.48) in the placebo. All participants had a normal BMI with means of 23.58 (3.15) and 22.95 (2.83), for intervention and placebo, respectively. All received medications comprised anticonvulsants, antidepressants, anxiolytics, analgesics, and NSAIDs.

Table 1. Baseline demographic and clinical features for each arm of the study participants.

Variables	Active (n = 29)	Placebo (n = 22)	*p*-Values
Age (years)	52.90 ± 6.47	52.50 ± 7.48	0.90
BMI (kg/m^2)	23.58 ± 3.15	22.95 ± 2.83	0.39
Heart rate (bpm)	75.27 ± 11.69	74.50 ± 7.71	1.00
Concomitant drugs			
Anticonvulsants	9	6	0.98
Antidepressants	24	16	0.60
Anxiolytics	19	14	0.87
Analgesics	21	13	0.48
NSAIDs	6	5	0.86

Data are expressed as means ± standard deviation (SD) for continuous variables. In both cases, the non-parametric Mann–Whitney U test was used because the samples were not normally distributed. No significant differences were observed in either case, confirming that both samples were derived from the same population. The Chi-squared test of independence was performed with one degree of freedom, and no significant differences in the proportions were observed, so no variable was considered to exert a different influence depending on the group. Abbreviations: BMI, body mass index; NSAIDs, non-steroidal anti-inflammatory drugs.

3.2. Changes in Fatigue Perception

Cognitive fatigue (Table 2), as indicated by the change in the cognitive domain, improved significantly from the baseline at the 36-week visit in the intervention group ($p = 0.0338$). The FIS-40 domain scores evolved in parallel between groups over the course of the study. The actual percentage difference was 5.7%, so the improvement effect could be due to the patient-self-reported interindividual variability and/or synergic effects of beta-glucans plus multivitamins among other potential effects (Figure 3).

Table 2. Changes in the fatigue severity assessed with the FIS-40 score from the baseline to the final assessment (at week 36) in the study participants.

FIS-40 Domains	Baseline	36-Weeks	*p*-Values [1]
Active arm (n = 29)			
Cognitive	33.31 ± 7.03	31.45 ± 9.76	0.0338 *
Psychosocial	61.79 ± 14.59	60.45 ± 16.92	0.3681
Physical functioning	34.52 ± 5.20	33.28 ± 6.78	0.0725
Total FIS-40 score	129.62 ± 25.37	125.17 ± 31.43	0.1070
Placebo (n = 22)			
Cognitive	33.82 ± 5.42	32.73 ± 5.76	0.3725
Psychosocial	64.14 ± 11.79	59.45 ± 12.92	0.1398
Physical functioning	33.73 ± 5.90	32.32 ± 6.95	0.1313
Total FIS-40 score	131.68 ± 21.16	124.50 ± 23.83	0.1375

Data are expressed as means ± SD. [1] The statistical significance of the paired data is analyzed. Statistical significance was set at * $p < 0.05$. The cognitive domain in the active arm presented a slight difference (5.7% from the baseline). Depending on the normality of the data, the test of independence of parametric (*t*-test) or non-parametric (Wilcoxon) paired data was used. There were no significant differences between the groups at baseline. Intervention group was compared with the baseline. Abbreviations: FIS-40, 40-item fatigue impact scale. Lower scores indicate an improvement in fatigue perception. [1] *p*-value for intergroup analysis.

3.3. Changes in Sleep Quality, Anxiety/Depression, and Health-Related Quality of Life

3.3.1. Sleep Quality Assessment

The changes observed in the sleep quality between study groups from the baseline to the final assessment are displayed in Table 3. The intervention with dietary supplements significantly improved daytime dysfunction at the 36-week visit in the active group ($p = 0.0131$ with respect to the baseline) compared to the placebo group. The other PSQI domain scores of the two groups evolved in parallel over the course of the study.

Table 3. Changes in sleep quality assessed using the PSQI questionnaire from the baseline to the final assessment (week 36).

PSQI Domains	Baseline	36 Weeks	*p*-Values [1]
Active arm (n = 29)			
Subjective sleep quality	2.14 ± 0.88	2.17 ± 0.89	0.7389
Sleep latency	2.07 ± 1.03	2.14 ± 1.16	0.5271
Sleep duration	1.69 ± 1.00	1.62 ± 0.98	0.4795
Habitual sleep efficiency	1.93 ± 1.33	1.83 ± 1.20	0.5728
Sleep disturbances	2.10 ± 0.72	2.21 ± 0.73	0.1797
Use of sleeping medication	2.07 ± 1.31	2.00 ± 1.39	0.6256
Daytime dysfunction *	2.31 ± 0.89	2.28 ± 0.75	0.7963
Overall PSQI score	14.31 ± 4.91	14.24 ± 4.93	0.8845

Table 3. *Cont.*

PSQI Domains	Baseline	36 Weeks	*p*-Values [1]
Placebo (n = 22)			
Subjective sleep quality	2.05 ± 0.84	1.68 ± 0.99	0.0588
Sleep latency	2.32 ± 0.78	2.18 ± 0.85	0.3173
Sleep duration	1.41 ± 1.05	1.36 ± 1.09	0.7630
Habitual sleep efficiency	1.77 ± 1.31	1.59 ± 1.40	0.2575
Sleep disturbances	2.09 ± 0.53	1.95 ± 0.58	0.2568
Use of sleeping medication	1.95 ± 1.36	2.14 ± 1.17	0.3795
Daytime dysfunction *	1.59 ± 0.73	1.73 ± 0.83	0.4386
Overall PSQI score	13.18 ± 3.85	12.64 ± 4.70	0.4567

Data are expressed as means ± SD. [1] The statistical significance of the paired data is analyzed. Depending on the normality of the samples, parametric (*t*-test) or non-parametric (Wilcoxon) tests for paired data were used. Baseline data showed no significant differences between the groups, indicating that there was no bias at the baseline between the two groups, (*) except in daytime dysfunction, in which significant differences were found using the Mann–Whitney U test for independence with a *p*-value of 0.0015; therefore, this variable should not be considered.

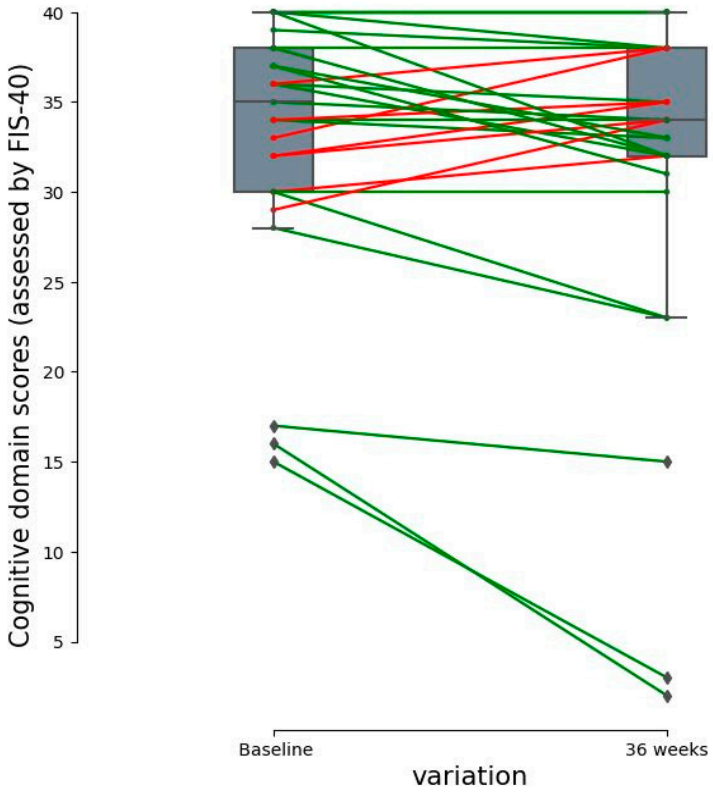


Figure 3. Dot plot of individual paired data of the cognitive fatigue domain (assessed using FIS-40) with significant differences. The differences between the values of the cognitive subscale domain are shown. Each point represents the score of each patient. Bars depict group medians. Lines indicate that cognitive scores belong to the same patient. Green lines depict a decrease in the perceived cognitive fatigue score (better cognitive function), and the red lines indicate an increase in the perceived cognitive fatigue score (poor cognitive function) from the baseline as quantified on the *Y* axis.

3.3.2. Anxiety and Depression

As shown in Table 4, the anxiety and depression symptoms did not present statistical differences between groups over the course of the study.

Table 4. Changes in the anxiety/depression symptoms assessed through the HADS questionnaire from the baseline to the final assessment (at week 36).

FIS-40 Domains	Baseline	36-Weeks	*p*-Values [1]
Active arm (n = 29)			
Anxiety	12.86 ± 5.05	12.76 ± 4.70	0.7940
Depression	11.86 ± 5.11	11.86 ± 5.59	0.5354
Total HADS	24.72 ± 9.60	24.62 ± 9.67	0.6716
Placebo (n = 22)			
Anxiety	11.82 ± 4.55	11.00 ± 4.65	0.1821
Depression	11.59 ± 4.07	10.68 ± 3.98	0.1218
Total HADS	23.41 ± 8.19	21.68 ± 8.20	0.1890

Data are expressed as means ± SD. [1] The statistical significance of the paired data is analyzed. Depending on the normality of the samples, parametric (*t*-test) or non-parametric (Wilcoxon) tests for paired data were used. Baseline data showed no significant differences between the groups, indicating that there was no bias at the baseline between the two groups.

3.3.3. Health-Related Quality of Life

As display in Table 5, social role functioning improved significantly at the 36-week visit compared to the baseline in the placebo group ($p = 0.0131$). The SF-36 domain scores evolved in parallel between groups over the course of the study.

Table 5. Changes in the health-related quality of life evaluated using the SF-36 questionnaire from the baseline to the final assessment (at week 36).

SF-36 Domains	Baseline	36-Weeks	*p*-Values [1]
Active arm (n = 29)			
Physical functioning	38.45 ± 20.49	40.17 ± 22.50	0.8178
Physical role functioning	6.03 ± 16.66	10.34 ± 20.61	0.2020
Bodily pain	15.93 ± 15.58	20.90 ± 18.68	0.1468
General health perception	22.76 ± 14.18	21.90 ± 13.19	0.5683
Vitality	12.24 ± 13.53	16.55 ± 15.93	0.0526
Social role functioning	32.33 ± 26.63	36.64 ± 25.65	0.3344
Emotional role functioning	34.48 ± 46.70	28.74 ± 43.39	0.1695
Mental health status	38.21 ± 21.18	42.07 ± 22.69	0.1393
Placebo (n = 22)			
Physical functioning	39.32 ± 21.17	38.18 ± 20.21	0.9243
Physical role functioning	1.14 ± 5.33	4.55 ± 16.61	0.3287
Bodily pain	20.14 ± 19.22	23.73 ± 15.33	0.1089
General health perception	25.32 ± 14.49	27.36 ± 17.57	0.3225
Vitality	17.27 ± 16.24	15.23 ± 12.77	0.5847
Social role functioning	30.68 ± 20.68	42.05 ± 22.34	0.0131 *
Emotional role functioning	33.33 ± 44.84	37.88 ± 45.19	0.4797
Mental health status	38.73 ± 23.11	44.73 ± 20.01	0.1918

Data are expressed as means ± SD. [1] The statistical significance of the paired data is analyzed. (*) Social role functioning in the placebo group presented a significant difference. Depending on the normality of the samples, parametric (*t*-test) or non-parametric (Wilcoxon) tests for paired data were used. Baseline data were analyzed between the two groups, indicating that there was no bias at the baseline between the two groups.

3.4. Clinical Safety and Intervention Tolerability

No relevant treatment-related adverse events were recorded in the study population. The intervention with dietary supplement containing yeast-derived beta-glucans + vitamin D3 + vitamin B6 + zinc over the period of 9 months was safe and well tolerated by the patients.

4. Discussion

Emerging evidence linking alterations in the gut microbiota composition with the clinical symptoms reported by ME/CFS patients suggests a potential therapeutic role

for microbiome-targeted interventions, such as prebiotic supplementation. By restoring intestinal homeostasis, this approach may help alleviate sufferers' health complaints and improve their quality of life. Therefore, to investigate the advantages of the proposed intervention, this clinical trial was designed to evaluate the effect of a food supplement ImmunoVita®, containing yeast-derived beta-glucan, combined with vitamin D3, vitamin B6, and zinc, on fatigue, sleep problems, anxiety/depression, and overall quality of life in ME/CFS patients.

In this study, beta-glucan supplements were administered to individuals with ME/CFS over a prolonged time period (36 weeks). A slight reduction in cognitive fatigue symptoms was reported along with an improvement in self-reported HR-QoL in the active arm of the study participants. Another controlled study assessing the effects of beta-glucan supplementation in ME/CFS patients over a 12-week period showed a significant reduction in patients' self-perceived fatigue and an improvement in their immune function compared to the placebo group [42].

The main findings of this study provide early evidence that administration of beta-glucan may influence certain neurocognitive outcomes related to perceived fatigue in ME/CFS patients after 36 months' treatment. The outcomes are consistent with the growing body of evidence supporting the presence of a two-way hormonal and neural signaling pathway between the gut microbiome–brain axis and its ability to influence metabolism, behavior, and neurocognitive functions through the production of microbial by-products exerting various local and systemic effects on gut hormones, oxidative stress, and inflammation in ME/CFS [21,43–45]. The gut–brain axis is considered to be a key regulator of cognitive function. In the present study, we demonstrated beneficial effects of beta-glucan supplementation on cognitive fatigue in ME/CFS and presented the first evidence of 36 weeks of beta-glucan supplementation improving cognitive impairment, assessed using FIS-40 measures.

Cognitive fatigue became less severe after the 36-week treatment with beta-glucan, as did exhaustion, but the differences with respect to placebo were not significant. The lack of cognitive energy was less severe in the active group compared to the baseline. These results are in line with pre-clinical studies showing that beta-glucan attenuates cognitive impairment via the gut–brain axis in diet-induced obese mice [46–48]. Experimental studies using murine models have shown that beta-glucan treatment can exert significant anti-fatigue effects, as demonstrated by the increased exhaustive swimming time in mice during the forced swimming test [30]. These benefits were attributed to the potential beneficial effects of beta-glucan on energy metabolism and oxidative stress, demonstrated in the form of reduced levels of exercise fatigue and injury-related blood biomarkers, including lactate, blood urea nitrogen (BUN), creatinine kinase (CK), alanine transaminase (ALT), and aspartate transaminase (AST), reduced serum glucose, and improvements in the response to exercise-induced oxidative stress [30].

Evidence from human trials has shown that increasing prebiotic fiber intake can influence perceived fatigue. For example, consuming cereals high in wheat fiber for two weeks reduced fatigue compared to the control in healthy adults [49], whereas a 13-week long supplementation with a prebiotic containing inulin and fru-oligosaccharides reduced exhaustion in elderly subjects when compared to the control [50]. In healthy adults, supplementation with oats (1 g), providing 1.9 g of dietary fiber as beta-glucan, significantly decreased the occurrence of exhaustion and fatigue compared with the baseline ($p < 0.05$). Interestingly, individuals who received beta-glucan had less severe headaches and lower perception of cold than controls. In addition, changes in inflammatory markers, C-reactive-protein (CRP), and oxidized-LDL with gastrointestinal symptom severity were associated with the occurrence and severity of several non-GI symptoms [51]. Furthermore, the results of a phase I/II clinical trial conducted in patients with advanced malignancies receiving chemotherapy indicated that beta-(1,3)-(1,6)-D-glucan administered as an adjunctive therapy may improve white blood cells and platelet counts, and may also raise levels of

hemoglobin, which in up to 40% of patients reduces the feeling of fatigue compared with the period before the supplementation [52].

Fatigue is associated with a lack of sleep, an association that may be mediated by increased inflammation [53]. To date, the therapeutic effects of beta-glucans in ME/CFS have not been assessed. Beta-glucans are naturally occurring polysaccharides found in the cell walls of fungi, algae, bacteria, and some cereals. It has been suggested that they may have immunomodulatory and anti-inflammatory effects, which could be beneficial for ME/CFS. Some preliminary studies have shown improvements in ME/CFS patients' perception of fatigue after beta-glucan administration [54].

In relation to sleep disturbances and anxiety/depression, there is limited evidence on the effects of beta-glucan on these symptoms in ME/CFS. Some studies have suggested that beta-glucan may have anxiolytic and antidepressant properties, but further research is required to support these claims and to determine its efficacy in this situation [55]. In a recent study assessing the effects of beta-glucan administration on mental health and HPA axis, reactivity was assessed in ME/CFS, and this approach improved anxiety/depression and HPA axis function [56]. In another study, apparently healthy adults who regularly consume an oats-based supplement containing 1.9 g of dietary fiber like beta-glucan had significantly fewer feelings of anxiety/depression [30].

Another preliminary study found that beta-glucan administration improved mood and vitality among HR-QoL measures in ME/CFS [51]. Overall, while there are some preliminary research studies suggesting possible potential benefits of beta-glucan administration on fatigue, mental health, and quality of life in ME/CFS, the current evidence is limited, and more research is needed to fully evaluate its efficacy and understand the pathophysiological mechanisms after beta-glucan administration in ME/CFS and other chronic post-infectious conditions.

These previous studies suggest that beta-glucans have potential beneficial effects in ME/CFS, including improving quality of life, reducing fatigue, decreasing anxiety/depression, and modulating the immune response. However, research on this topic is still ongoing, and more studies are needed to confirm and extend these findings, and also to determine the different sources, functionality (viscosity, fermentability, solubility, etc.), optimal doses, and duration of beta-glucan treatment that may affect physiological outcomes in ME/CFS.

The current results provide consistent evidence linking increased beta-glucan intake to enhanced cognitive function in ME/CFS. The beneficial effect suggests that the relationship between gut microbiota alteration and cognitive impairment may be causal. In addition to highlighting the adverse impact of western diets on the gut–brain axis, the findings of this study suggest that enhanced consumption of beta-glucan-rich foods is an easily implementable nutritional strategy for attenuating the diet-induced cognitive decline in ME/CFS patients [57].

Strengths and Limitations

This study has several strengths and some limitations that must be considered. To the best of our knowledge, this trial is the first clinical evaluation of the effects of beta-glucan supplementation with multivitamins (including vitamin D3, vitamin B6, and zinc) involved in energy metabolism and immune function enhancement on ME/CFS patients' perceptions of their neurocognitive symptoms, fatigue, sleep quality, and anxiety/depression. It should be noted that a cognition improvement found through percentage differences of 5.7% in active arm is more likely to be due to a real treatment effect using beta-glucans plus multivitamins among the participants. In addition, study participants were assessed according to the ME/CFS case criteria based on the 1994 CDC/Fukuda definition, and a potential gender bias was avoided by restricting the study to females.

The main limitation of this study was the relatively small sample size. In addition, these participants were recruited from a single tertiary referral center, so the results cannot be generalized to other ME/CFS populations. It should also be noted that the primary endpoint, the perception of fatigue assessed with the FIS-40 questionnaire, was a self-

reported parameter. Finally, the lack of statistical differences between groups may suggest a placebo effect for microcrystalline cellulose and plant capsule.

In future studies, the doses and follow-up timing of the intervention should be pre-established, so as to ensure appropriate monitoring beyond the intervention. In addition, objective physiological and biological parameters should be used to evaluate beneficial effects of beta-glucan in ME/CFS and other post-viral fatigue syndromes.

5. Conclusions and Future Directions

These findings provide hypothesis-generating evidence that beta-glucan may be beneficial for several affective and physical states in ME/CFS. For example, its positive effect on cognitive function and gut microbiome could be extended and applied in a disease model of cognitive decline. Since information is lacking on the effects on non-GI symptoms in ME/CFS of dietary fibers in general, and of beta-glucan in particular, these results provide potentially useful data for the design of controlled studies required to confirm these observations.

This study provides the first evidence that beta-glucan administration improves indices of cognition function with major beneficial effects all along the gut microbiota–immune–brain axis. Our data suggest that the consumption of 1 g of beta-glucan is an easily implementable nutritional strategy for alleviating the detrimental features of gut–brain dysregulation and for preventing cognitive decline in ME/CFS.

Author Contributions: Conceptualization, J.A.-M., R.S.S., L.V-S. and J.C.-M.; methodology, M.L., J.A.-M. and J.J.; software, M.L.; validation, J.A.-M., L.V.-S. and J.C.-M.; formal analysis, M.L.; investigation, J.C.-M.; resources, L.V.-S.; data curation, M.L. and J.A.-M.; writing—original draft preparation, M.L., J.J. and J.C.-M.; writing—review and editing, M.L., L.V.-S., J.J. and J.C.-M.; supervision, L.V.-S.; project administration, J.A.-M. and L.V.-S.; funding acquisition, J.A.-M. and L.V.-S. All authors have read and agreed to the published version of the manuscript.

Funding: J.C.-M. received financial support and honoraria from Vitae Health Innovation S.L. (Montmeló, Barcelona, Spain). This study was supported by Vitae Health Innovation S.L. (Montmeló, Spain) which supplied both treatments (funding number: 2020-0445).

Institutional Review Board Statement: All study procedures were reviewed and approved by the local Clinical Research Ethics Committee of the Vall d'Hebron University Hospital in Barcelona, Spain (protocol code: InmunoVitaME-PR(AG)-447-2019; approved on 28 February 2019). The study was conducted in accordance with the guidelines of the Declaration of Helsinki, the current Spanish regulations on clinical research, and the EU's standards of Good Clinical Practice. It also followed the Consolidated Standards of Reporting Trials (CONSORT) guidelines.

Informed Consent Statement: Informed consent was obtained from all participants in the study.

Data Availability Statement: All relevant data analyzed during the current trial are included in the article. Access to raw datasets may be shared upon reasonable request to the corresponding author.

Acknowledgments: The authors thank the patients who participated. The authors are grateful to Maria Cleofé Zaragoza for her help in data curation and for her contribution to the paper. We also thank Michael Maudsley for editing and his useful linguistic advice on the updated manuscript.

Conflicts of Interest: L.V. is an employee of VITAE Health Innovation S.L. (Montmeló, Barcelona, Spain). As stated in the author contributions, the funders declare that they had no role in the design, execution, analysis and interpretation, and presentation of data. J.C.-M. received financial support from VITAE Health Innovation S.L. to conduct this intervention study. The rest of the authors declare no conflicts of interest.

References

1. Estevez-Lopez, F.; Mudie, K.; Wang-Steverding, X.; Bakken, I.J.; Ivanovs, A.; Castro-Marrero, J.; Nacul, L.; Alegre, J.; Zalewski, P.; Słomko, J.; et al. Systematic Review of the Epidemiological Burden of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome Across Europe: Current Evidence and EUROMENE Research Recommendations for Epidemiology. *J. Clin. Med.* **2020**, *9*, 1557. [CrossRef] [PubMed]
2. Faro, M.; Saez-Francas, N.; Castro-Marrero, J.; Aliste, L.; Fernandez de Sevilla, T.; Alegre, J. Gender differences in chronic fatigue syndrome. *Reumatol. Clin.* **2016**, *12*, 72–77. [CrossRef]
3. Nacul, L.; Authier, F.J.; Scheibenbogen, C.; Lorusso, L.; Helland, I.B.; Martin, J.A.; Sirbu, C.A.; Mengshoel, A.M.; Polo, O.; Behrends, U.; et al. European Network on Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (EUROMENE): Expert Consensus on the Diagnosis, Service Provision, and Care of People with ME/CFS in Europe. *Medicina* **2021**, *57*, 510. [CrossRef] [PubMed]
4. Cortes Rivera, M.; Mastronardi, C.; Silva-Aldana, C.T.; Arcos-Burgos, M.; Lidbury, B.A. Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Comprehensive Review. *Diagnostics* **2019**, *9*, 91. [CrossRef] [PubMed]
5. Castro-Marrero, J.; Saez-Francas, N.; Santillo, D.; Alegre, J. Treatment and management of chronic fatigue syndrome/myalgic encephalomyelitis: All roads lead to Rome. *Br. J. Pharmacol.* **2017**, *174*, 345–369. [CrossRef]
6. Castro-Marrero, J.; Faro, M.; Zaragoza, M.C.; Aliste, L.; de Sevilla, T.F.; Alegre, J. Unemployment and work disability in individuals with chronic fatigue syndrome/myalgic encephalomyelitis: A community-based cross-sectional study from Spain. *BMC Public Health* **2019**, *19*, 840. [CrossRef]
7. Roberts, D. Chronic fatigue syndrome and quality of life. *Patient Relat. Outcome Meas.* **2018**, *9*, 253–262. [CrossRef]
8. Fernandez-Quiros, J.; Lacasa-Cazcarra, M.; Alegre-Martin, J.; Sanmartin-Sentanes, R.; Almirall, M.; Launois-Obregon, P.; Castro-Marrero, J.; Rodríguez-Urrutia, A.; Navarro-Sanchis, J.A.; Ramos-Quiroga, J.A. The Conners Continuous Performance Test CPT3^{TM}: Is it a reliable marker to predict neurocognitive dysfunction in Myalgic encephalomyelitis/chronic fatigue syndrome? *Front. Psychol.* **2023**, *14*, 1127193. [CrossRef]
9. Jain, V.; Arunkumar, A.; Kingdon, C.; Lacerda, E.; Nacul, L. Prevalence of and risk factors for severe cognitive and sleep symptoms in ME/CFS and MS. *BMC Neurol.* **2017**, *17*, 117. [CrossRef]
10. Wirth, K.J.; Scheibenbogen, C.; Paul, F. An attempt to explain the neurological symptoms of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *J. Transl. Med.* **2021**, *19*, 471. [CrossRef]
11. Carruthers, B.M.; van de Sande, M.I.; De Meirleir, K.L.; Klimas, N.G.; Broderick, G.; Mitchell, T.; Staines, D.; Powles, A.P.; Speight, N.; Vallings, R.; et al. Myalgic encephalomyelitis: International Consensus Criteria. *J. Intern. Med.* **2011**, *270*, 327–338. [CrossRef]
12. What's in a name? Systemic exertion intolerance disease. *Lancet* **2015**, *385*, 663. [CrossRef] [PubMed]
13. Cvejic, E.; Birch, R.C.; Vollmer-Conna, U. Cognitive Dysfunction in Chronic Fatigue Syndrome: A Review of Recent Evidence. *Curr. Rheumatol. Rep.* **2016**, *18*, 24. [CrossRef] [PubMed]
14. Castro-Marrero, J.; Saez-Francas, N.; Segundo, M.J.; Calvo, N.; Faro, M.; Aliste, L.; de Sevilla, T.F.; Alegre, J. Effect of coenzyme Q10 plus nicotinamide adenine dinucleotide supplementation on maximum heart rate after exercise testing in chronic fatigue syndrome—A randomized, controlled, double-blind trial. *Clin. Nutr.* **2016**, *35*, 826–834. [CrossRef] [PubMed]
15. Maes, M.; Twisk, F.N. Why myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) may kill you: Disorders in the inflammatory and oxidative and nitrosative stress (IO&NS) pathways may explain cardiovascular disorders in ME/CFS. *Neuroendocrinol. Lett.* **2009**, *30*, 677–693.
16. Castro-Marrero, J.; Zaragoza, M.C.; Gonzalez-Garcia, S.; Aliste, L.; Saez-Francas, N.; Romero, O.; Ferré, A.; Fernandez de Sevilla, T.; Alegre, J. Poor self-reported sleep quality and health-related quality of life in patients with chronic fatigue syndrome/myalgic encephalomyelitis. *J. Sleep Res.* **2018**, *27*, e12703. [CrossRef]
17. Lakhan, S.E.; Kirchgessner, A. Gut inflammation in chronic fatigue syndrome. *Nutr. Metab.* **2010**, *7*, 79. [CrossRef]
18. Rao, A.V.; Bested, A.C.; Beaulne, T.M.; Katzman, M.A.; Iorio, C.; Berardi, J.M.; Logan, A.C. A randomized, double-blind, placebo-controlled pilot study of a probiotic in emotional symptoms of chronic fatigue syndrome. *Gut Pathog.* **2009**, *1*, 6. [CrossRef]
19. Wallis, A.; Ball, M.; Butt, H.; Lewis, D.P.; McKechnie, S.; Paull, P.; Jaa-Kwee, A.; Bruck, D. Open-label pilot for treatment targeting gut dysbiosis in myalgic encephalomyelitis/chronic fatigue syndrome: Neuropsychological symptoms and sex comparisons. *J. Transl. Med.* **2018**, *16*, 24. [CrossRef]
20. Martin, F.; Blanco-Suarez, M.; Zambrano, P.; Caceres, O.; Almirall, M.; Alegre-Martin, J.; Lobo, B.; González-Castro, A.M.; Santos, J.; Domingo, J.C.; et al. Increased gut permeability and bacterial translocation are associated with fibromyalgia and myalgic encephalomyelitis/chronic fatigue syndrome: Implications for disease-related biomarker discovery. *Front. Immunol.* **2023**, *14*, 1253121. [CrossRef]
21. Guo, C.; Che, X.; Briese, T.; Ranjan, A.; Allicock, O.; Yates, R.A.; Cheng, A.; March, D.; Hornig, M.; Komaroff, A.L.; et al. Deficient butyrate-producing capacity in the gut microbiome is associated with bacterial network disturbances and fatigue symptoms in ME/CFS. *Cell Host Microbe* **2023**, *31*, 288–304.e288. [CrossRef] [PubMed]
22. Borren, N.Z.; Plichta, D.; Joshi, A.D.; Bonilla, G.; Peng, V.; Colizzo, F.P.; Luther, J.; Khalili, H.; Garber, J.J.; van der Woude, C.J.; et al. Alterations in Fecal Microbiomes and Serum Metabolomes of Fatigued Patients With Quiescent Inflammatory Bowel Diseases. *Clin. Gastroenterol. Hepatol.* **2021**, *19*, 519–527.e5. [CrossRef] [PubMed]

23. Xiao, C.; Fedirko, V.; Beitler, J.; Bai, J.; Peng, G.; Zhou, C.; Gu, J.; Zhao, H.; Lin, I.H.; Chico, C.E.; et al. The role of the gut microbiome in cancer-related fatigue: Pilot study on epigenetic mechanisms. *Support. Care Cancer* **2021**, *29*, 3173–3182. [CrossRef] [PubMed]
24. Miyake, S.; Kim, S.; Suda, W.; Oshima, K.; Nakamura, M.; Matsuoka, T.; Chihara, N.; Tomita, A.; Sato, W.; Kim, S.W.; et al. Dysbiosis in the Gut Microbiota of Patients with Multiple Sclerosis, with a Striking Depletion of Species Belonging to Clostridia XIVa and IV Clusters. *PLoS ONE* **2015**, *10*, e0137429. [CrossRef]
25. Leiva-Gea, I.; Sanchez-Alcoholado, L.; Martin-Tejedor, B.; Castellano-Castillo, D.; Moreno-Indias, I.; Urda-Cardona, A.; Tinahones, F.J.; Fernández-García, J.C.; Queipo-Ortuño, M.I. Gut Microbiota Differs in Composition and Functionality Between Children With Type 1 Diabetes and MODY2 and Healthy Control Subjects: A Case-Control Study. *Diabetes Care* **2018**, *41*, 2385–2395. [CrossRef]
26. Singh, R.P.; Bhardwaj, A. beta-glucans: A potential source for maintaining gut microbiota and the immune system. *Front. Nutr.* **2023**, *10*, 1143682. [CrossRef]
27. Hughes, S.A.; Shewry, P.R.; Gibson, G.R.; McCleary, B.V.; Rastall, R.A. In vitro fermentation of oat and barley derived beta-glucans by human faecal microbiota. *FEMS Microbiol. Ecol.* **2008**, *64*, 482–493. [CrossRef]
28. Varesi, A.; Deumer, U.S.; Ananth, S.; Ricevuti, G. The Emerging Role of Gut Microbiota in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS): Current Evidence and Potential Therapeutic Applications. *J. Clin. Med.* **2021**, *10*, 5077. [CrossRef]
29. Konig, R.S.; Albrich, W.C.; Kahlert, C.R.; Bahr, L.S.; Lober, U.; Vernazza, P.; Scheibenbogen, C.; Forslund, S.K. The Gut Microbiome in Myalgic Encephalomyelitis (ME)/Chronic Fatigue Syndrome (CFS). *Front. Immunol.* **2021**, *12*, 628741. [CrossRef]
30. Xu, X.; Ding, Y.; Yang, Y.; Gao, Y.; Sun, Q.; Liu, J.; Yang, X.; Wang, J.; Zhang, J. beta-glucan Salecan Improves Exercise Performance and Displays Anti-Fatigue Effects through Regulating Energy Metabolism and Oxidative Stress in Mice. *Nutrients* **2018**, *10*, 858. [CrossRef]
31. Dharsono, T.; Rudnicka, K.; Wilhelm, M.; Schoen, C. Effects of Yeast (1,3)-(1,6)-Beta-Glucan on Severity of Upper Respiratory Tract Infections: A Double-Blind, Randomized, Placebo-Controlled Study in Healthy Subjects. *J. Am. Coll. Nutr.* **2019**, *38*, 40–50. [CrossRef] [PubMed]
32. Ganda Mall, J.P.; Casado-Bedmar, M.; Winberg, M.E.; Brummer, R.J.; Schoultz, I.; Keita, A.V. A beta-Glucan-Based Dietary Fiber Reduces Mast Cell-Induced Hyperpermeability in Ileum From Patients With Crohn's Disease and Control Subjects. *Inflamm. Bowel Dis.* **2017**, *24*, 166–178. [CrossRef] [PubMed]
33. Asano, T.; Tanaka, K.; Suemasu, S.; Ishihara, T.; Tahara, K.; Suzuki, T.; Suzuki, H.; Fukudo, S.; Mizushima, T. Effects of beta-(1,3-1,6)-D-glucan on irritable bowel syndrome-related colonic hypersensitivity. *Biochem. Biophys. Res. Commun.* **2012**, *420*, 444–449. [CrossRef]
34. Ciacci, C.; Franceschi, F.; Purchiaroni, F.; Capone, P.; Buccelletti, F.; Iacomini, P.; Ranaudo, A.; Andreozzi, P.; Tondi, P.; Gentiloni Siveri, N.; et al. Effect of beta-glucan, inositol and digestive enzymes in GI symptoms of patients with IBS. *Eur. Rev. Med. Pharmacol. Sci.* **2011**, *15*, 637–643.
35. Glynn, L.M.; Sandman, C.A. Evaluation of the association between placental corticotrophin-releasing hormone and postpartum depressive symptoms. *Psychosom. Med.* **2014**, *76*, 355–362. [CrossRef] [PubMed]
36. Kano, M.; Muratsubaki, T.; Van Oudenhove, L.; Morishita, J.; Yoshizawa, M.; Kohno, K.; Yagihashi, M.; Tanaka, Y.; Mugikura, S.; Dupont, P.; et al. Altered brain and gut responses to corticotropin-releasing hormone (CRH) in patients with irritable bowel syndrome. *Sci. Rep.* **2017**, *7*, 12425. [CrossRef]
37. Fukuda, K.; Straus, S.E.; Hickie, I.; Sharpe, M.C.; Dobbins, J.G.; Komaroff, A. The chronic fatigue syndrome: A comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Ann. Intern. Med.* **1994**, *121*, 953–959. [CrossRef]
38. Fisk, J.D.; Ritvo, P.G.; Ross, L.; Haase, D.A.; Marrie, T.J.; Schlech, W.F. Measuring the functional impact of fatigue: Initial validation of the fatigue impact scale. *Clin. Infect. Dis.* **1994**, *18* (Suppl. S1), S79–S83. [CrossRef]
39. Buysse, D.J.; Reynolds, C.F., 3rd; Monk, T.H.; Berman, S.R.; Kupfer, D.J. The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Res.* **1989**, *28*, 193–213. [CrossRef]
40. Herrero, M.J.; Blanch, J.; Peri, J.M.; De Pablo, J.; Pintor, L.; Bulbena, A. A validation study of the hospital anxiety and depression scale (HADS) in a Spanish population. *Gen. Hosp. Psychiatry* **2003**, *25*, 277–283. [CrossRef]
41. Alonso, J.; Prieto, L.; Anto, J.M. The Spanish version of the SF-36 Health Survey (the SF-36 health questionnaire): An instrument for measuring clinical results. *Med. Clin.* **1995**, *104*, 771–776.
42. Davis, J.M.; Murphy, E.A.; Brown, A.S.; Carmichael, M.D.; Ghaffar, A.; Mayer, E.P. Effects of oat beta-glucan on innate immunity and infection after exercise stress. *Med. Sci. Sport. Exerc.* **2004**, *36*, 1321–1327. [CrossRef] [PubMed]
43. Wachsmuth, H.R.; Weninger, S.N.; Duca, F.A. Role of the gut-brain axis in energy and glucose metabolism. *Exp. Mol. Med.* **2022**, *54*, 377–392. [CrossRef] [PubMed]
44. Marano, G.; Mazza, M.; Lisci, F.M.; Ciliberto, M.; Traversi, G.; Kotzalidis, G.D.; De Berardis, D.; Laterza, L.; Sani, G.; Gasbarrini, A.; et al. The Microbiota-Gut-Brain Axis: Psychoneuroimmunological Insights. *Nutrients* **2023**, *15*, 1496. [CrossRef] [PubMed]
45. Xiong, R.; Gunter, C.; Fleming, E.; Vernon, S.D.; Bateman, L.; Unutmaz, D.; Oh, J. Multi-'omics of gut microbiome-host interactions in short- and long-term myalgic encephalomyelitis/chronic fatigue syndrome patients. *Cell Host Microbe* **2023**, *31*, 273–287.e275. [CrossRef]
46. Tornero-Aguilera, J.F.; Jimenez-Morcillo, J.; Rubio-Zarapuz, A.; Clemente-Suarez, V.J. Central and Peripheral Fatigue in Physical Exercise Explained: A Narrative Review. *Int. J. Environ. Res. Public. Health* **2022**, *19*, 3909. [CrossRef]

47. Germain, A.; Giloteaux, L.; Moore, G.E.; Levine, S.M.; Chia, J.K.; Keller, B.A.; Stevens, J.; Franconi, C.J.; Mao, X.; Shungu, D.C.; et al. Plasma metabolomics reveals disrupted response and recovery following maximal exercise in myalgic encephalomyelitis/chronic fatigue syndrome. *JCI Insight* **2022**, *7*, e157621. [CrossRef]
48. Joseph, P.; Singh, I.; Oliveira, R.; Capone, C.A.; Mullen, M.P.; Cook, D.B.; Stovall, M.C.; Squires, J.; Madsen, K.; Waxman, A.B.; et al. Exercise Pathophysiology in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome and Postacute Sequelae of SARS-CoV-2: More in Common Than Not? *Chest* **2023**, *164*, 717–726. [CrossRef]
49. Smith, A.; Bazzoni, C.; Beale, J.; Elliott-Smith, J.; Tiley, M. High fibre breakfast cereals reduce fatigue. *Appetite* **2001**, *37*, 249–250. [CrossRef]
50. Buigues, C.; Fernandez-Garrido, J.; Pruimboom, L.; Hoogland, A.J.; Navarro-Martinez, R.; Martinez-Martinez, M.; Verdejo, Y.; Mascarós, M.C.; Peris, C.; Cauli, O. Effect of a Prebiotic Formulation on Frailty Syndrome: A Randomized, Double-Blind Clinical Trial. *Int. J. Mol. Sci.* **2016**, *17*, 932. [CrossRef]
51. Wolever, T.M.S.; Rahn, M.; Dioum, E.H.; Jenkins, A.L.; Ezatagha, A.; Campbell, J.E.; Chu, Y. Effect of Oat beta-Glucan on Affective and Physical Feeling States in Healthy Adults: Evidence for Reduced Headache, Fatigue, Anxiety and Limb/Joint Pains. *Nutrients* **2021**, *13*, 1534. [CrossRef] [PubMed]
52. Weitberg, A.B. A phase I/II trial of beta-(1,3)/(1,6) D-glucan in the treatment of patients with advanced malignancies receiving chemotherapy. *J. Exp. Clin. Cancer Res.* **2008**, *27*, 40. [CrossRef] [PubMed]
53. Thomas, K.S.; Motivala, S.; Olmstead, R.; Irwin, M.R. Sleep depth and fatigue: Role of cellular inflammatory activation. *Brain Behav. Immun.* **2011**, *25*, 53–58. [CrossRef] [PubMed]
54. Vetvicka, V.; Vannucci, L.; Sima, P.; Richter, J. Beta Glucan: Supplement or Drug? From Laboratory to Clinical Trials. *Molecules* **2019**, *24*, 1251. [CrossRef] [PubMed]
55. Murphy, E.J.; Rezoagli, E.; Major, I.; Rowan, N.J.; Laffey, J.G. beta-Glucan Metabolic and Immunomodulatory Properties and Potential for Clinical Application. *J. Fungi* **2020**, *6*, 356. [CrossRef]
56. Cryan, J.F.; O'Riordan, K.J.; Cowan, C.S.M.; Sandhu, K.V.; Bastiaanssen, T.F.S.; Boehme, M.; Codagnone, M.G.; Cussotto, S.; Fulling, C.; Golubeva, A.V.; et al. The Microbiota-Gut-Brain Axis. *Physiol. Rev.* **2019**, *99*, 1877–2013. [CrossRef]
57. Hu, M.; Zhang, P.; Wang, R.; Zhou, M.; Pang, N.; Cui, X.; Ge, X.; Liu, X.; Huang, X.F.; Yu, Y. Three Different Types of beta-Glucans Enhance Cognition: The Role of the Gut-Brain Axis. *Front. Nutr.* **2022**, *9*, 848930. [CrossRef]

Chapter 8

Discussion

8.1 The Complexity of ME/CFS

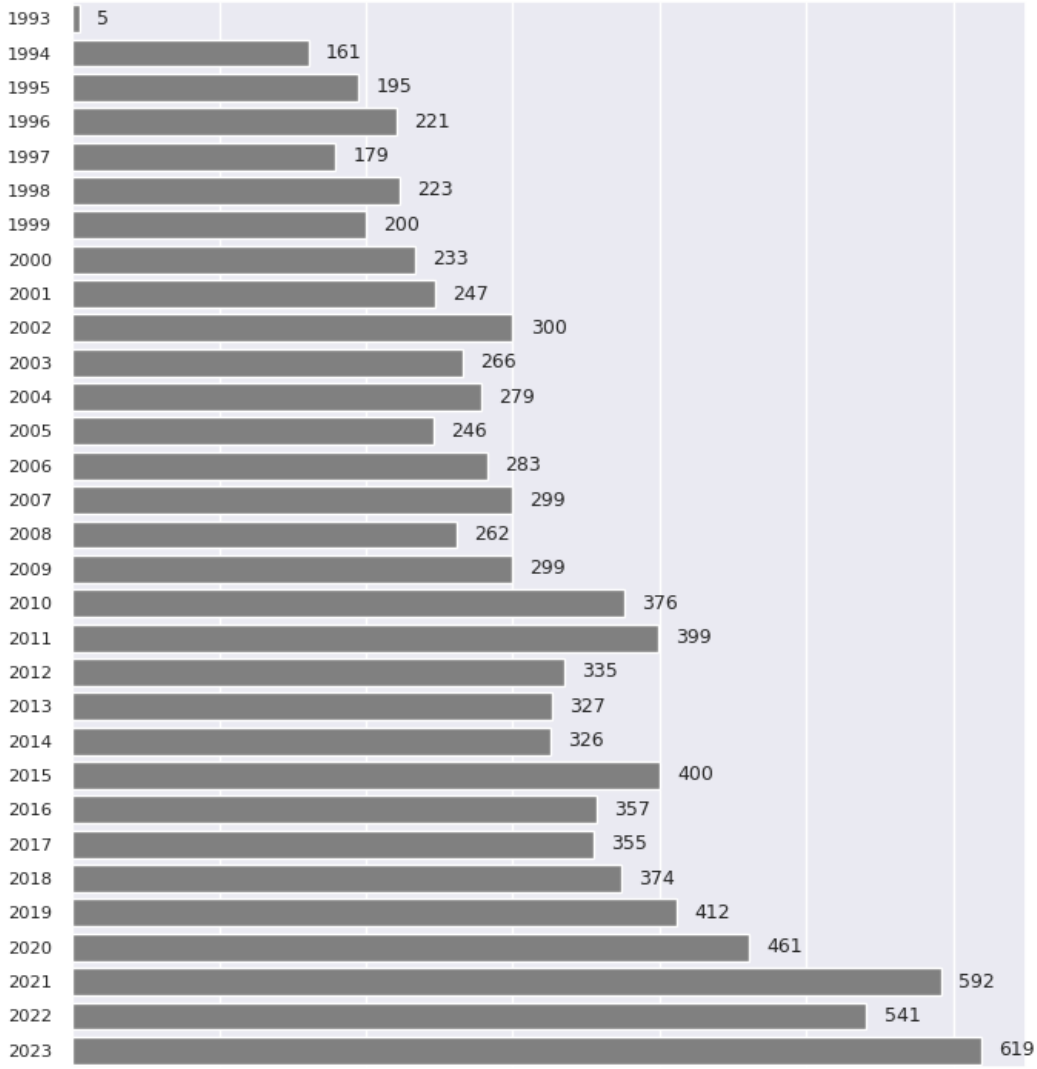
Academic interest in ME/CFS can be measured by analyzing indexed scientific publications. This is shown in Figure 8.1. An analysis of the data obtained is shown to examine both the number of publications and the scientific journals that publish them[1]. An application is provided that is configured using the API (acronym for Application Programming Interface) to generate databases through statements that define search configurations. In this case, I use the library that implements the API in Python. Following the instructions, all publications from 1993 to 2023 with the keywords "Chronic Fatigue Syndrome" are displayed.

The first conclusion is that there is a growing interest in MS/CFS research, especially in the last 10 years. However, we want to know which areas of publications would indicate a greater effort. For this, a count by journals is performed, but in this case, the search is adjusted from 2015 to see recent trends. It can be observed that journals with a general scope are the ones that group the most publications, but also journals with a very different scope, such as the *Journal of Psychosomatic Research*, which is a multidisciplinary research journal covering all aspects of psychology. But we also find the *International Journal of Molecular Sciences*, which focuses on biochemistry, molecular and cell biology, molecular biophysics, molecular medicine, and all aspects of molecular research in chemistry. We can also find journals focused on the fields of rheumatology, pediatrics, and neurology in the list.

A review of the status of biomarkers has been shown in section 2.5.8, showing the difficulty of defining any of them for diagnosing ME/CFS. The review concludes that further studies with much larger numbers of samples are required to identify markers based on protein or genetic sequencing. The rest of the biochemical analyses yield information on the presence of diseases compatible with ME/CFS, but no **discriminating marker** has been found that can be used for the diagnosis of ME/CFS.

[1] *Fuente para la obtención de datos: hppt://pubmed.ncbi.nlm.nih.gov*

Figure 8.1: Number of indexed publications on ME/CFS



The number of articles published in journals indexed in the National Library of Medicine is shown. The image shows a steady increase in the number of publications, especially in the last 10 years.

Table 8.1: List of scientific journals according to number of publications on ME/CFS

Journals	*n*
Journal of translational medicine	106
PloS one	87
Journal of psychosomatic research	62
Journal of clinical medicine	60
Cureus	50
International journal of molecular sciences	48
Journal of health psychology	46
Scientific reports	43
Healthcare (Basel, Switzerland)	43
Frontiers in immunology	42
International journal of environmental research and public health	41
BMJ open	39
Frontiers in medicine	37
Frontiers in neurology	34
Medicina (Kaunas, Lithuania)	31
Clinical and experimental rheumatology	31
Fatigue : biomedicine, health & behavior	31
Diagnostics (Basel, Switzerland)	28
Medicine	28
Frontiers in pediatrics	27

From the above selection, the 20 scientific journals are shown in order in number of publications whose articles deal with ME/CFS from 2015 to 2023. The characteristic of this list is the variability on the fields of interest. As a journal specializing in ME/CFS, one could name Fatigue: Biomedicine, Health, and Behavior, an international interdisciplinary journal addressing fatigue symptoms in specific medical diseases, behavioral disorders, and environmental conditions. The journal also refers to articles on chronic fatigue syndrome, fibromyalgia, and related diseases.

Diagnosis

Medical diagnosis is the process of evaluating medical conditions by analyzing symptoms, medical history, and accepted test results. The goal is to determine the cause and effective treatment. Imaging tests such as X-rays or MRIs, blood tests, or biopsy procedures for analysis in anatomic pathology are some of the most common. The importance of data science as part of the diagnostic process has grown in recent years. In particular, there has been a breakthrough in diagnostic imaging. The evolution of server power, computer language specialization, and data availability have enabled continuous improvement of artificial intelligence models [169].

There have also been major advances in the use of data science to define biomarkers. Metabolomics is the study of metabolites in the blood. It is less expensive than proteomics or genomics. Machine learning facilitates clinical metabolomics because it analyzes large amounts of data efficiently. It is a key tool for the classification and risk prediction of diseases such as cancer [170].

Genetic analysis of thousands of genes simultaneously is possible by using different algorithms for gene expression analysis using RNA sequencing and DNA microarray data for cancer detection. Advances in machines and computer languages have made it possible to develop in various fields such as cancer. This allows a disease to be classified by genetic analysis of a patient and effective therapy to be administered based not only on the disease itself but also on the patient's genetic profile [171].

In the case of ME/CFS, the patient will experience a variety of complaints and will receive an initial visit to a primary care physician where treatments to alleviate the symptoms will be evaluated. A longer-term follow-up visit may lead to consideration of other illnesses or a referral for a hospital consultation if the symptoms do not go away. Diagnosing ME/CFS is neither quick nor intuitive and depends largely on the sensitivity and knowledge of the physicians consulted. In addition, it can take many years of identification and consensus among the consulted medical team before ME/CFS is defined, which is itself an anomaly. Thus, the patient with ME/CFS faces a first complex difficulty: being diagnosed within a reasonable period of time under the criteria accepted by the specialized scientific community.

Treatment

There are no specific and effective treatments to cure ME/CFS. The lack of knowledge about the mechanism of action that causes the syndrome makes the development of a drug very complicated. The physician may perceive a lack of urgency in diagnosis and follow-up because there is no accepted treatment with a specific indication for ME/CFS.

Follow-up

It requires a specialized unit with different specialists covering the ME/CFS symptomatology. The lack of such a unit makes monitoring and treating symptoms difficult. Report evaluation must be carried out by multidisciplinary teams. These teams should include specialists in different fields, such as internal medicine, psychology, rehabilitation, biochemical analysis, and data science.

8.2 Relating Questionnaires to Objective Variables

In a specialized unit, evaluating the patient's medical history is essential in the multidisciplinary assessment of the patient's current condition. However, it is the specific questionnaires that determine the degree of symptomatology. The answers to these questionnaires are given by the patients themselves. In this sense, using objective tests is crucial in assessing the patient's state of health. The key to contex-

tualizing the information in the medical history is the interpretation of the results of the scales of these questionnaires.

In this research, the answers to the questionnaires are precisely related to objective tests. Exercise tests are performed to assess the severity of the physical condition. Ergometric tests to determine exercise intolerance must be performed on these patients by a highly specialized professional, and the physical functional capacity and ability to perform a programmed and individualized physical exercise program must be determined by an experienced specialist.

The publication model defined in 6.1.1 allows for screening patients with ME/CFS and identifying those at risk of a critical outcome in hospital stress tests [172].

8.3 Contextualizing ME/SFC Research

Present and future perspectives on biomarker research.

In section 2.5.8 It has analyzed the open lines of research and the contributions made so far. The strategy that has mainly been followed is the comparison of biological samples from patients diagnosed with ME/CFS and healthy patients using different **machine-learning** techniques. These bioinformatics techniques resolve that ME/CFS patients show differences in certain symptoms.

A review of 101 scientific articles searched using the keywords "biomarker" and "ME/CFS" [70] showed that half of the publications focused on genetic, epigenetic, and immunological studies. The rest were mostly distributed in metabolomic and endo-vascular studies. Eighty percent were based on blood tests and the biomarkers did not have primary selectivity (which is the ability of the biomarker to identify a disease- causing agent).

It is concluded that MS/CFS research needs to reduce heterogeneity, require multidisciplinary research and unify biomarker research protocols.

Treatment perspectives now and in the future

A recent review [173] There has been an analysis of randomized clinical trials on the efficacy of treatments for ME/CFS. Again, we find heterogeneity in the designs. Argentina is the only country to have approval for a treatment for MS/CFS under the trade name Ampligen (rintatolimod), which was denied by the FDA in 2009 and 2013[2].

Clinical trials focus on immune modulators, redox modulators and mitochondrial dysfunction, as well as circadian rhythm and intestinal microbiota modulators for mitochondrial dysfunction, antioxidants, probiotics, and neurological or neuroendocrine diseases. Self-administered and validated questionnaires are commonly used to measure primary outcomes. The SF-36 was found to be the most commonly used questionnaire in an analysis of 540 publications [174].

[2]Source: https://www.fda.gov/media/131114/download, – *consulted on 8/6/2024* –

Objective tests, such as maximal oxygen consumption in ergometry, have been used in clinical trials to demonstrate the efficacy of rintatolimod. [175].

8.4 Limitations

As indicated in the publications provided in this thesis by compendium, the studies' limitations are centered on the uniqueness of the research center. The studies' database comes from the specialized unit of the Hospital Vall d'Hebron in Barcelona.

8.4.1 Results in another research unit

It is appropriate to point out that, at the time of writing, the study:

"Confirmatory clinical trial on unsupervised cluster analysis reveals distinct subtypes of ME/CFS patients based on peak oxygen consumption and SF-36 score"

It is under review in a scientific journal. A validation study in another research center shows that the classification algorithm produces results similar to those of the publication referred to in point 6.1.1. These results are considered relevant to the use of the application referred to in Section 1.5.

8.5 Conclusions

Myalgic encephalomyelitis or chronic fatigue syndrome (ME/CFS) is a chronic, debilitating and multifaceted clinical entity. The heterogeneous onset and clinical presentation with additional comorbidities makes diagnosis difficult due to the lack of available tests and biomarkers. Current treatment guidelines focus on symptom management without addressing the causal mechanism. There is an urgent need for robust clinical trials to identify effective treatments. A limiting element is the lack of objective measures of efficacy, which are currently based on validated, self-administered questionnaires answered by the patient.

One of the most important conclusions is the need to use objective measures. This research provides 2 biomarkers that can be used for this purpose: the peak oxygen consumption in the exercise test and the result of the CPT3 test for the measurement of cognitive impairment.

Oxygen consumption in the stress test

Publication 6.1.1 shows that the measurement of peak oxygen consumption in the exercise test determines the patient's physical status with ME/CFS. In addition, screening can be performed using an artificial intelligence-based model that predicts which patients require exercise testing due to a limiting physical condition.

The CPT3 test for the detection of cognitive impairment.

It is demonstrated in publication 6.1.4 that the use of the CPT3 test measures the cognitive impairment of the MS/CFS patient and may be a measure of treatment efficacy in future clinical trials.

Knowledge transfer

An application that offers a multidisciplinary analysis and predicts the physical risk of a patient with ME/CFS is designed by requiring the answers to the SF-36 questionnaire as input data. This application not only favors the early detection of physical deterioration and the suggestion of referral to a specialized unit but would also favor, on the one hand, the early stratification of the severity of fatigue and, on the other hand, the individualized design of an individualized physical exercise program by a specialist.

Bibliography

[1] A. Simmons and S. Chappell, "Artificial intelligence-definition and practice," *IEEE Journal of Oceanic Engineering*, vol. 13, no. 2, pp. 14–42, 1988.

[2] D. Monett and C. W. P. Lewis, "Getting clarity by defining artificial intelligence—a survey," in *Philosophy and Theory of Artificial Intelligence 2017* (V. C. Müller, ed.), (Cham), pp. 212–214, Springer International Publishing, 2018.

[3] W. G. de Sousa, E. R. P. de Melo, P. H. D. S. Bermejo, R. A. S. Farias, and A. O. Gomes, "How and where is artificial intelligence in the public sector going? a literature review and research agenda," *Government Information Quarterly*, vol. 36, no. 4, p. 101392, 2019.

[4] J. M. Helm, A. M. Swiergosz, H. S. Haeberle, J. M. Karnuta, J. L. Schaffer, V. E. Krebs, A. I. Spitzer, and P. N. Ramkumar, "Machine learning and artificial intelligence: Definitions, applications, and future directions," *Curr. Rev. Musculoskelet. Med.*, vol. 13, pp. 69–76, Feb. 2020.

[5] K. P. Murphy, "Machine learning: A probabilistic perspective (adaptive computation and machine learning series)," *The MIT Press: London, UK*, 2018.

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[7] A. Bosch, J. C. Roma, and A. L. Bagén, *Deep learning: principios y fundamentos*. Editorial UOC, 2019.

[8] W. E. Tapprich, L. Reichart, D. M. Simon, G. Duncan, W. McClung, N. Grandgenett, and M. A. Pauley, "An instructional definition and assessment rubric for bioinformatics instruction," *Biochemistry and Molecular Biology Education*, vol. 49, no. 1, pp. 38–45, 2021.

[9] K. Strimbu and J. A. Tavel, "What are biomarkers?," *Curr. Opin. HIV AIDS*, vol. 5, pp. 463–466, Nov. 2010.

[10] N. I. Al-Dewik, S. N. Younes, M. M. Essa, S. Pathak, and M. W. Qoronfleh, "Making biomarkers relevant to healthcare innovation and precision medicine," *Processes*, vol. 10, no. 6, 2022.

[11] N. Rifai, M. A. Gillette, and S. A. Carr, "Protein biomarker discovery and validation: the long and uncertain path to clinical utility," *Nat. Biotechnol.*, vol. 24, pp. 971–983, Aug. 2006.

[12] S. Bonassi, M. Neri, and R. Puntoni, "Validation of biomarkers as early predictors of disease," *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, vol. 480-481, pp. 349–358, 2001. Molecular Mechanisms of Anticarcinogenesis and Antimutagenesis.

[13] O. F. Laterza, R. C. Hendrickson, and J. A. Wagner, "Molecular biomarkers," *Drug Information Journal*, vol. 41, no. 5, pp. 573–585, 2007.

[14] R. P. Horgan and L. C. Kenny, "'omic'technologies: genomics, transcriptomics, proteomics and metabolomics," *The Obstetrician & Gynaecologist*, vol. 13, no. 3, pp. 189–195, 2011.

[15] S. H.-D. F. S. E. E. B. S. H. M. Ali Khodadadian, Somaye Darzi and M. Nazari, "Genomics and transcriptomics: The powerful technologies in precision medicine," *International Journal of General Medicine*, vol. 13, pp. 627–640, 2020. PMID: 32982380.

[16] S. M. Foltz, C. S. Greene, and J. N. Taroni, "Cross-platform normalization enables machine learning model training on microarray and RNA-seq data simultaneously," *Commun Biol*, vol. 6, p. 222, Feb. 2023.

[17] V. Trevino, F. Falciani, and H. A. Barrera-Saldaña, "DNA microarrays: a powerful genomic tool for biomedical and clinical research," *Mol. Med.*, vol. 13, no. 9-10, pp. 527–541, 2007.

[18] R. Goodacre, S. Vaidyanathan, W. B. Dunn, G. G. Harrigan, and D. B. Kell, "Metabolomics by numbers: acquiring and understanding global metabolite data," *Trends in Biotechnology*, vol. 22, no. 5, pp. 245–252, 2004.

[19] C. Baumgartner, M. Osl, M. Netzer, and D. Baumgartner, "Bioinformatic-driven search for metabolic biomarkers in disease," *J. Clin. Bioinforma.*, vol. 1, p. 2, Jan. 2011.

[20] M. Dobbins, K. DeCorby, and T. Twiddy, "A knowledge transfer strategy for public health decision makers," *Worldviews Evid. Based. Nurs.*, vol. 1, no. 2, pp. 120–128, 2004.

[21] G. N. Thompson, C. A. Estabrooks, and L. F. Degner, "Clarifying the concepts in knowledge transfer: a literature review," *J. Adv. Nurs.*, vol. 53, pp. 691–701, Mar. 2006.

[22] D. W. Bates, W. Schmitt, D. Buchwald, N. C. Ware, J. Lee, E. Thoyer, R. J. Kornish, and A. L. Komaroff, "Prevalence of fatigue and chronic fatigue syndrome in a primary care practice," *Arch. Intern. Med.*, vol. 153, pp. 2759–2765, Dec. 1993.

[23] G. Ranjith, "Epidemiology of chronic fatigue syndrome," 2005.

[24] L. Nacul, F. J. Authier, C. Scheibenbogen, L. Lorusso, I. B. Helland, J. A. Martin, C. A. Sirbu, A. M. Mengshoel, O. Polo, U. Behrends, H. Nielsen, P. Grabowski, S. Sekulic, N. Sepulveda, F. Estévez-López, P. Zalewski, D. F. H. Pheby, J. Castro-Marrero, G. K. Sakkas, E. Capelli, I. Brundsdlund, J. Cullinan, A. Krumina, J. Bergquist, M. Murovska, R. C. W. Vermuelen, and E. M. Lacerda, "European network on myalgic Encephalomyelitis/Chronic fatigue syndrome (EUROMENE): Expert consensus on the diagnosis, service provision, and care of people with ME/CFS in europe," *Medicina*, vol. 57, May 2021.

[25] C. D. Reciprocated, "The lancet review: Chronic fatigue syndrome prins JB, van der meer JW, bleijenberg g. invited review: Chronic fatigue syndrome. the Lancet-Vol. 367 . . . ," *Lancet*.

[26] A. C. Bested and L. M. Marshall, "Review of myalgic Encephalomyelitis/Chronic fatigue syndrome: an evidence-based approach to diagnosis and management by clinicians," *Rev. Environ. Health*, vol. 30, no. 4, pp. 223–249, 2015.

[27] K. G. Brurberg, M. S. Fønhus, L. Larun, S. Flottorp, and K. Malterud, "Case definitions for chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME): a systematic review," *BMJ Open*, vol. 4, p. e003973, Feb. 2014.

[28] B. M. Carruthers, A. K. Jain, K. L. De Meirleir, D. L. Peterson, N. G. Klimas, A. M. Lerner, A. C. Bested, P. Flor-Henry, P. Joshi, A. C. P. Powles, J. A. Sherkey, and M. I. van de Sande, "Myalgic encephalomyelitis/chronic fatigue syndrome," *J. Chronic Fatigue Syndr.*, vol. 11, pp. 7–115, Jan. 2003.

[29] K. Fukuda, S. E. Straus, I. Hickie, M. C. Sharpe, J. G. Dobbins, and A. Komaroff, "The chronic fatigue syndrome: A comprehensive approach to its definition and study," *Ann. Intern. Med.*, vol. 121, pp. 953–959, Dec. 1994.

[30] G. P. Holmes, J. E. Kaplan, N. M. Gantz, A. L. Komaroff, L. B. Schonberger, S. E. Straus, J. F. Jones, R. E. Dubois, C. Cunningham-Rundles, and S. Pahwa, "Chronic fatigue syndrome: a working case definition," *Ann. Intern. Med.*, vol. 108, pp. 387–389, Mar. 1988.

[31] B. M. Carruthers, M. I. van de Sande, K. L. De Meirleir, N. G. Klimas, G. Broderick, T. Mitchell, D. Staines, A. C. P. Powles, N. Speight, R. Vallings, L. Bateman, B. Baumgarten-Austrheim, D. S. Bell, N. Carlo-Stella, J. Chia, A. Darragh, D. Jo, D. Lewis, A. R. Light, S. Marshall-Gradisnik, I. Mena, J. A. Mikovits, K. Miwa, M. Murovska, M. L. Pall, and S. Stevens, "Myalgic encephalomyelitis: International consensus criteria," *J. Intern. Med.*, vol. 270, pp. 327–338, Oct. 2011.

[32] M. C. Sharpe, L. C. Archard, J. E. Banatvala, L. K. Borysiewicz, A. W. Clare, A. David, R. H. Edwards, K. E. Hawton, H. P. Lambert, and R. J. Lane, "A report–chronic fatigue syndrome: guidelines for research," *J. R. Soc. Med.*, vol. 84, pp. 118–121, Feb. 1991.

[33] Institute of Medicine, Board on the Health of Select Populations, and Committee on the Diagnostic Criteria for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome, *Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness*. National Academies Press, Mar. 2015.

[34] L. A. Jason, M. L. Zinn, and M. A. Zinn, "Myalgic encephalomyelitis: Symptoms and biomarkers," *Curr. Neuropharmacol.*, vol. 13, no. 5, pp. 701–734, 2015.

[35] B. M. Carruthers, A. K. Jain, K. L. De Meirleir, D. L. Peterson, N. G. Klimas, A. Martin Lerner, A. C. Bested, P. Flor-Henry, P. Joshi, A. C. Peter Powles, J. A. Sherkey, and M. I. van de Sande, "Myalgic Encephalomyelitis/Chronic fatigue syndrome," 2003.

[36] P. D. White, "Beyond myalgic Encephalomyelitis/Chronic fatigue syndrome: Redefining an illness. institute of medicine. (pp. 282; ISBN 978-0-309-31689-7; $£29.99; paperback.) the national academies press: Washington. 2015," 2016.

[37] J. Castro-Marrero, M. Faro, L. Aliste, N. Sáez-Francàs, N. Calvo, A. Martínez-Martínez, T. F. de Sevilla, and J. Alegre, "Comorbidity in chronic fatigue Syndrome/Myalgic encephalomyelitis: A nationwide Population-Based cohort study," *Psychosomatics*, vol. 58, pp. 533–543, Apr. 2017.

[38] E.-J. Lim, Y.-C. Ahn, E.-S. Jang, S.-W. Lee, S.-H. Lee, and C.-G. Son, "Systematic review and meta-analysis of the prevalence of chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME)," *J. Transl. Med.*, vol. 18, p. 100, Feb. 2020.

[39] T. Norris, S. M. Collin, K. Tilling, R. Nuevo, S. A. Stansfeld, J. A. Sterne, J. Heron, and E. Crawley, "Natural course of chronic fatigue syndrome/myalgic encephalomyelitis in adolescents," *Arch. Dis. Child.*, vol. 102, pp. 522–528, June 2017.

[40] Q. Wu, J. Gao, D. Bai, Y. Zhong, Z. Yang, and X. Jiang, "Prevalence of chronic fatigue syndrome in china: a meta-analysis," *Youjiang Med J.*

[41] A. R. Lloyd, I. Hickie, C. R. Boughton, O. Spencer, and D. Wakefield, "Prevalence of chronic fatigue syndrome in an australian population," *Med. J. Aust.*, vol. 153, pp. 522–528, Nov. 1990.

[42] C. Bierl, R. Nisenbaum, D. C. Hoaglin, B. Randall, A.-B. Jones, E. R. Unger, and W. C. Reeves, "Regional distribution of fatiguing illnesses in the united states: a pilot study," *Popul. Health Metr.*, vol. 2, p. 1, Feb. 2004.

[43] P. H. Levine, "Epidemic neuromyasthenia and chronic fatigue syndrome: epidemiological importance of a cluster definition," *Clin. Infect. Dis.*, vol. 18 Suppl 1, pp. S16–20, Jan. 1994.

[44] F. Estévez-López, K. Mudie, X. Wang-Steverding, I. J. Bakken, A. Ivanovs, J. Castro-Marrero, L. Nacul, J. Alegre, P. Zalewski, J. Słomko, E. B. Strand, D. Pheby, E. Shikova, L. Lorusso, E. Capelli, S. Sekulic, C. Scheibenbogen, N. Sepúlveda, M. Murovska, and E. Lacerda, "Systematic review of the epidemiological burden of myalgic Encephalomyelitis/Chronic fatigue syndrome across europe: Current evidence and EUROMENE research recommendations for epidemiology," *J. Clin. Med. Res.*, vol. 9, May 2020.

[45] M. Cortes Rivera, C. Mastronardi, C. T. Silva-Aldana, M. Arcos-Burgos, and B. A. Lidbury, "Myalgic Encephalomyelitis/Chronic fatigue syndrome: A comprehensive review," *Diagnostics (Basel)*, vol. 9, Aug. 2019.

[46] L. Lorusso, S. V. Mikhaylova, E. Capelli, D. Ferrari, G. K. Ngonga, and G. Ricevuti, "Immunological aspects of chronic fatigue syndrome," *Autoimmun. Rev.*, vol. 8, pp. 287–291, Feb. 2009.

[47] E. W. Brenu, M. L. van Driel, D. R. Staines, K. J. Ashton, S. L. Hardcastle, J. Keane, L. Tajouri, D. Peterson, S. B. Ramos, and S. M. Marshall-Gradisnik, "Longitudinal investigation of natural killer cells and cytokines in chronic fatigue syndrome/myalgic encephalomyelitis," *J. Transl. Med.*, vol. 10, p. 88, May 2012.

[48] M. Maes, F. N. M. Twisk, M. Kubera, and K. Ringel, "Evidence for inflammation and activation of cell-mediated immunity in myalgic Encephalomyelitis/Chronic fatigue syndrome (ME/CFS): increased interleukin-1, tumor necrosis factor-α, PMN-elastase, lysozyme and neopterin," *J. Affect. Disord.*, vol. 136, pp. 933–939, Feb. 2012.

[49] J. R. Kerr, F. Barah, D. L. Mattey, I. Laing, S. J. Hopkins, I. V. Hutchinson, and D. A. J. Tyrrell, "Circulating tumour necrosis factor-alpha and interferon-gamma are detectable during acute and convalescent parvovirus B19 infection and are associated with prolonged and chronic fatigue," *J. Gen. Virol.*, vol. 82, pp. 3011–3019, Dec. 2001.

[50] J. A. G. Glassford, "The neuroinflammatory etiopathology of myalgic Encephalomyelitis/Chronic fatigue syndrome (ME/CFS)," *Front. Physiol.*, vol. 8, p. 88, Feb. 2017.

[51] M. Maes, J.-C. Leunis, M. Geffard, and M. Berk, "Evidence for the existence of myalgic Encephalomyelitis/Chronic fatigue syndrome (ME/CFS) with and without abdominal discomfort (irritable bowel) syndrome," *Neuro Endocrinol. Lett.*, vol. 35, no. 6, pp. 445–453, 2014.

[52] J. Nijs, A. Nees, L. Paul, M. De Kooning, K. Ickmans, M. Meeus, and J. Van Oosterwijck, "Altered immune response to exercise in patients with chronic fatigue syndrome/myalgic encephalomyelitis: a systematic literature review," *Exerc. Immunol. Rev.*, vol. 20, pp. 94–116, 2014.

[53] J. Blomberg, C.-G. Gottfries, A. Elfaitouri, M. Rizwan, and A. Rosén, "Infection elicited autoimmunity and myalgic Encephalomyelitis/Chronic fatigue syndrome: An explanatory model," *Front. Immunol.*, vol. 9, p. 229, Feb. 2018.

[54] J. J. Dibble, S. J. McGrath, and C. P. Ponting, "Genetic risk factors of ME/CFS: a critical review," *Hum. Mol. Genet.*, vol. 29, pp. R117–R124, Sept. 2020.

[55] J. P. Griffith and F. A. Zarrouf, "A systematic review of chronic fatigue syndrome: don't assume it's depression," *Prim. Care Companion J. Clin. Psychiatry*, vol. 10, no. 2, pp. 120–128, 2008.

[56] J. Fernández-Quirós, M. Lacasa-Cazcarra, J. Alegre-Martín, R. Sanmartín-Sentañes, M. Almirall, P. Launois-Obregón, J. Castro-Marrero, A. Rodríguez-Urrutia, J. A. Navarro-Sanchis, and J. A. Ramos-Quiroga, "The conners continuous performance test CPT3™: Is it a reliable marker to predict neurocognitive dysfunction in myalgic encephalomyelitis/chronic fatigue syndrome?," *Front. Psychol.*, vol. 14, p. 1127193, Feb. 2023.

[57] C. Tomas, J. Newton, and S. Watson, "A review of hypothalamic-pituitary-adrenal axis function in chronic fatigue syndrome," *ISRN Neurosci*, vol. 2013, p. 784520, Sept. 2013.

[58] M. A. Arroll, "Allostatic overload in myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS)," *Med. Hypotheses*, vol. 81, pp. 506–508, Sept. 2013.

[59] S. Du Preez, M. Corbitt, H. Cabanas, N. Eaton, D. Staines, and S. Marshall-Gradisnik, "A systematic review of enteric dysbiosis in chronic fatigue syndrome/myalgic encephalomyelitis," *Syst. Rev.*, vol. 7, p. 241, Dec. 2018.

[60] A. Proal and T. Marshall, "Myalgic Encephalomyelitis/Chronic fatigue syndrome in the era of the human microbiome: Persistent pathogens drive chronic symptoms by interfering with host metabolism, gene expression, and immunity," *Front Pediatr*, vol. 6, p. 373, Dec. 2018.

[61] G. F. D. Lupo, G. Rocchetti, L. Lucini, L. Lorusso, E. Manara, M. Bertelli, E. Puglisi, and E. Capelli, "Potential role of microbiome in chronic fatigue Syndrome/Myalgic encephalomyelits (CFS/ME)," *Sci. Rep.*, vol. 11, p. 7043, Mar. 2021.

[62] D. Missailidis, S. J. Annesley, and P. R. Fisher, "Pathological mechanisms underlying myalgic Encephalomyelitis/Chronic fatigue syndrome," *Diagnostics (Basel)*, vol. 9, July 2019.

[63] P. Luczynski, K.-A. McVey Neufeld, C. S. Oriach, G. Clarke, T. G. Dinan, and J. F. Cryan, "Growing up in a bubble: Using Germ-Free animals to assess the influence of the gut microbiota on brain and behavior," *Int. J. Neuropsychopharmacol.*, vol. 19, Aug. 2016.

[64] M. Lacasa, F. Prados, J. Alegre, and J. Casas-Roma, "A synthetic data generation system for myalgic encephalomyelitis/chronic fatigue syndrome questionnaires," *Sci. Rep.*, vol. 13, p. 14256, Aug. 2023.

[65] E. W. Brenu, K. J. Ashton, J. Batovska, D. R. Staines, and S. M. Marshall-Gradisnik, "High-throughput sequencing of plasma microRNA in chronic fatigue syndrome/myalgic encephalomyelitis," *PLoS One*, vol. 9, p. e102783, Sept. 2014.

[66] A. K. Cheema, L. Sarria, M. Bekheit, F. Collado, E. Almenar-Pérez, E. Martín-Martínez, J. Alegre, J. Castro-Marrero, M. A. Fletcher, N. G. Klimas, E. Oltra, and L. Nathanson, "Unravelling myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS): Gender-specific changes in the microRNA expression profiling in ME/CFS," *J. Cell. Mol. Med.*, vol. 24, pp. 5865–5877, May 2020.

[67] E. Nepotchatykh, W. Elremaly, I. Caraus, C. Godbout, C. Leveau, L. Chalder, C. Beaudin, E. Kanamaru, R. Kosovskaia, S. Lauzon, Y. Maillet, A. Franco, V. Lascau-Coman, S. Bouhanik, Y. P. Gaitan, D. Li, and A. Moreau, "Profile of circulating microRNAs in myalgic encephalomyelitis and their relation to symptom severity, and disease pathophysiology," *Sci. Rep.*, vol. 10, p. 19620, Nov. 2020.

[68] C. Del Rio, L. F. Collins, and P. Malani, "Long-term health consequences of COVID-19," *JAMA*, vol. 324, pp. 1723–1724, Nov. 2020.

[69] L. Townsend, A. H. Dyer, K. Jones, J. Dunne, A. Mooney, F. Gaffney, L. O'Connor, D. Leavy, K. O'Brien, J. Dowds, J. A. Sugrue, D. Hopkins, I. Martin-Loeches, C. Ni Cheallaigh, P. Nadarajan, A. M. McLaughlin, N. M. Bourke, C. Bergin, C. O'Farrelly, C. Bannan, and N. Conlon, "Persistent fatigue following SARS-CoV-2 infection is common and independent of severity of initial infection," *PLoS One*, vol. 15, p. e0240784, Nov. 2020.

[70] R. Maksoud, C. Magawa, N. Eaton-Fitch, K. Thapaliya, and S. Marshall-Gradisnik, "Biomarkers for myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS): a systematic review," *BMC Med.*, vol. 21, p. 189, May 2023.

[71] J. Blauensteiner, R. Bertinat, L. E. León, M. Riederer, N. Sepúlveda, and F. Westermeier, "Altered endothelial dysfunction-related mirs in plasma from ME/CFS patients," *Sci. Rep.*, vol. 11, p. 10604, May 2021.

[72] E. W. Brenu, K. J. Ashton, M. van Driel, D. R. Staines, D. Peterson, G. M. Atkinson, and S. M. Marshall-Gradisnik, "Cytotoxic lymphocyte microRNAs as prospective biomarkers for chronic fatigue Syndrome/Myalgic encephalomyelitis," *J. Affect. Disord.*, vol. 141, pp. 261–269, Dec. 2012.

[73] M. A. J. Kadhum, "The possible role of interleukin-17a elevation in the development of chronic fatigue syndrome," *Indian J. Public Health Res. Dev.*, vol. 9, no. 11, p. 1911, 2018.

[74] A. R. Light, L. Bateman, D. Jo, R. W. Hughen, T. A. Vanhaitsma, A. T. White, and K. C. Light, "Gene expression alterations at baseline and following moderate exercise in patients with chronic fatigue syndrome and fibromyalgia syndrome," *J. Intern. Med.*, vol. 271, pp. 64–81, Jan. 2012.

[75] S. Marshall-Gradisnik, S. Johnston, A. Chacko, T. Nguyen, P. Smith, and D. Staines, "Single nucleotide polymorphisms and genotypes of transient receptor potential ion channel and acetylcholine receptor genes from isolated B lymphocytes in myalgic encephalomyelitis/chronic fatigue syndrome patients," *J. Int. Med. Res.*, vol. 44, pp. 1381–1394, Dec. 2016.

[76] T. Saiki, T. Kawai, K. Morita, M. Ohta, T. Saito, K. Rokutan, and N. Ban, "Identification of marker genes for differential diagnosis of chronic fatigue syndrome," *Mol. Med.*, vol. 14, no. 9-10, pp. 599–607, 2008.

[77] A. T. White, A. R. Light, R. W. Hughen, T. A. Vanhaitsma, and K. C. Light, "Differences in metabolite-detecting, adrenergic, and immune gene expression after moderate exercise in patients with chronic fatigue syndrome, patients with multiple sclerosis, and healthy controls," *Psychosom. Med.*, vol. 74, pp. 46–54, Jan. 2012.

[78] E. W. Brenu, M. L. van Driel, D. R. Staines, K. J. Ashton, S. B. Ramos, J. Keane, N. G. Klimas, and S. M. Marshall-Gradisnik, "Immunological abnormalities as potential biomarkers in chronic fatigue Syndrome/Myalgic encephalomyelitis," *J. Transl. Med.*, vol. 9, p. 81, May 2011.

[79] P. Espinosa and J. M. Urra, "Decreased expression of the CD57 molecule in T lymphocytes of patients with chronic fatigue syndrome," *Mol. Neurobiol.*, vol. 56, pp. 6581–6585, Sept. 2019.

[80] M. A. Fletcher, M. Rosenthal, M. Antoni, G. Ironson, X. R. Zeng, Z. Barnes, J. M. Harvey, B. Hurwitz, S. Levis, G. Broderick, and N. G. Klimas, "Plasma neuropeptide y: a biomarker for symptom severity in chronic fatigue syndrome," *Behav. Brain Funct.*, vol. 6, p. 76, Dec. 2010.

[81] M. A. Fletcher, X. R. Zeng, K. Maher, S. Levis, B. Hurwitz, M. Antoni, G. Broderick, and N. G. Klimas, "Biomarkers in chronic fatigue syndrome: evaluation of natural killer cell function and dipeptidyl peptidase IV/CD26," *PLoS One*, vol. 5, p. e10817, May 2010.

[82] N. Groven, E. A. Fors, A. K. Stunes, and S. K. Reitan, "MCP-1 is increased in patients with CFS and FM, whilst several other immune markers are significantly lower than healthy controls," *Brain Behav Immun Health*, vol. 4, p. 100067, Apr. 2020.

[83] P. Halpin, M. V. Williams, N. G. Klimas, M. A. Fletcher, Z. Barnes, and M. E. Ariza, "Myalgic encephalomyelitis/chronic fatigue syndrome and gulf war illness patients exhibit increased humoral responses to the herpesviruses-encoded dUTPase: Implications in disease pathophysiology," *J. Med. Virol.*, vol. 89, pp. 1636–1645, Sept. 2017.

[84] K. Hanevik, E. K. Kristoffersen, S. Sørnes, K. Mørch, H. Næss, A. C. Rivenes, J. E. Bødtker, T. Hausken, and N. Langeland, "Immunophenotyping in post-giardiasis functional gastrointestinal disease and chronic fatigue syndrome," *BMC Infect. Dis.*, vol. 12, p. 258, Oct. 2012.

[85] S. L. Hardcastle, E. W. Brenu, S. Johnston, T. Nguyen, T. Huth, N. Wong, S. Ramos, D. Staines, and S. Marshall-Gradisnik, "Characterisation of cell functions and receptors in chronic fatigue Syndrome/Myalgic encephalomyelitis (CFS/ME)," *BMC Immunol.*, vol. 16, p. 35, June 2015.

[86] M. Maes, I. Mihaylova, M. Kubera, J.-C. Leunis, F. N. M. Twisk, and M. Geffard, "IgM-mediated autoimmune responses directed against anchorage epitopes are greater in myalgic Encephalomyelitis/Chronic fatigue syndrome (ME/CFS) than in major depression," *Metab. Brain Dis.*, vol. 27, pp. 415–423, Dec. 2012.

[87] S. F. Khaiboullina, K. L. DeMeirleir, S. Rawat, G. S. Berk, R. S. Gaynor-Berk, T. Mijatovic, N. Blatt, A. A. Rizvanov, S. G. Young, and V. C. Lombardi, "Cytokine expression provides clues to the pathophysiology of gulf war illness and myalgic encephalomyelitis," *Cytokine*, vol. 72, pp. 1–8, Mar. 2015.

[88] M. Milivojevic, X. Che, L. Bateman, A. Cheng, B. A. Garcia, M. Hornig, M. Huber, N. G. Klimas, B. Lee, H. Lee, S. Levine, J. G. Montoya, D. L. Peterson, A. L. Komaroff, and W. I. Lipkin, "Plasma proteomic profiling suggests an association between antigen driven clonal B cell expansion and ME/CFS," *PLoS One*, vol. 15, p. e0236148, July 2020.

[89] J. L. Rivas, T. Palencia, G. Fernández, and M. García, "Association of T and NK cell phenotype with the diagnosis of myalgic Encephalomyelitis/Chronic fatigue syndrome (ME/CFS)," *Front. Immunol.*, vol. 9, p. 1028, May 2018.

[90] A. P. Sung, J. J.-J. Tang, M. J. Guglielmo, J. Smith-Gagen, L. Bateman, L. Navarrete-Galvan, D. D. Redelman, and D. Hudig, "Antibody-Dependent cell-mediated cytotoxicity (ADCC) in familial myalgic Encephalomyelitis/Chronic fatigue syndrome (ME/CFS)," *Fatigue*, vol. 8, no. 4, pp. 226–244, 2020.

[91] M. Szklarski, H. Freitag, S. Lorenz, S. C. Becker, F. Sotzny, S. Bauer, J. Hartwig, H. Heidecke, K. Wittke, C. Kedor, L. G. Hanitsch, P. Grabowski, N. Sepúlveda, and C. Scheibenbogen, "Delineating the association between soluble CD26 and autoantibodies against G-Protein coupled receptors, immunological and cardiovascular parameters identifies distinct patterns in Post-Infectious vs. Non-Infection-Triggered myalgic Encephalomyelitis/Chronic fatigue syndrome," *Front. Immunol.*, vol. 12, p. 644548, Apr. 2021.

[92] S. D. Vernon and W. C. Reeves, "Evaluation of autoantibodies to common and neuronal cell antigens in chronic fatigue syndrome," *J. Autoimmune Dis.*, vol. 2, p. 5, May 2005.

[93] C. W. Armstrong, N. R. McGregor, J. R. Sheedy, I. Buttfield, H. L. Butt, and P. R. Gooley, "NMR metabolic profiling of serum identifies amino acid disturbances in chronic fatigue syndrome," *Clin. Chim. Acta*, vol. 413, pp. 1525–1531, Oct. 2012.

[94] F. Ciregia, L. Kollipara, L. Giusti, R. P. Zahedi, C. Giacomelli, M. R. Mazzoni, G. Giannaccini, P. Scarpellini, A. Urbani, A. Sickmann, A. Lucacchini, and L. Bazzichi, "Bottom-up proteomics suggests an association between differential expression of mitochondrial proteins and chronic fatigue syndrome," *Transl. Psychiatry*, vol. 6, p. e904, Sept. 2016.

[95] J. C. Domingo, B. Cordobilla, R. Ferrer, M. Giralt, J. Alegre-Martín, and J. Castro-Marrero, "Are circulating fibroblast growth factor 21 and N-Terminal prohormone of brain natriuretic peptide promising novel biomarkers in myalgic Encephalomyelitis/Chronic fatigue syndrome?," *Antioxid. Redox Signal.*, vol. 34, pp. 1420–1427, June 2021.

[96] D. Missailidis, O. Sanislav, C. Y. Allan, S. J. Annesley, and P. R. Fisher, "Cell-Based blood biomarkers for myalgic Encephalomyelitis/Chronic fatigue syndrome," *Int. J. Mol. Sci.*, vol. 21, Feb. 2020.

[97] D. Nagy-Szakal, B. L. Williams, N. Mishra, X. Che, B. Lee, L. Bateman, N. G. Klimas, A. L. Komaroff, S. Levine, J. G. Montoya, D. L. Peterson, D. Ramanan, K. Jain, M. L. Eddy, M. Hornig, and W. I. Lipkin, "Fecal metagenomic profiles in subgroups of patients with myalgic encephalomyelitis/chronic fatigue syndrome," *Microbiome*, vol. 5, p. 44, Apr. 2017.

[98] D. Nagy-Szakal, D. K. Barupal, B. Lee, X. Che, B. L. Williams, E. J. R. Kahn, J. E. Ukaigwe, L. Bateman, N. G. Klimas, A. L. Komaroff, S. Levine, J. G. Montoya, D. L. Peterson, B. Levin, M. Hornig, O. Fiehn, and W. I. Lipkin, "Insights into myalgic encephalomyelitis/chronic fatigue syndrome phenotypes through comprehensive metabolomics," *Sci. Rep.*, vol. 8, p. 10056, July 2018.

[99] M. Simonato, S. Dall'Acqua, C. Zilli, S. Sut, R. Tenconi, N. Gallo, P. Sfriso, L. Sartori, F. Cavallin, U. Fiocco, P. Cogo, P. Agostinis, A. Aldovini, D. Bruttomesso, R. Marcolongo, S. Comai, and A. Baritussio, "Tryptophan metabolites, cytokines, and fatty acid binding protein 2 in myalgic Encephalomyelitis/Chronic fatigue syndrome," *Biomedicines*, vol. 9, Nov. 2021.

[100] E. Almenar-Pérez, L. Sarría, L. Nathanson, and E. Oltra, "Assessing diagnostic value of microRNAs from peripheral blood mononuclear cells and extracellular vesicles in myalgic Encephalomyelitis/Chronic fatigue syndrome," *Sci. Rep.*, vol. 10, p. 2064, Feb. 2020.

[101] H. Bonilla, D. Hampton, E. G. Marques de Menezes, X. Deng, J. G. Montoya, J. Anderson, and P. J. Norris, "Comparative analysis of extracellular vesicles in patients with severe and mild myalgic Encephalomyelitis/Chronic fatigue syndrome," *Front. Immunol.*, vol. 13, p. 841910, Mar. 2022.

[102] A. Eguchi, S. Fukuda, H. Kuratsune, J. Nojima, Y. Nakatomi, Y. Watanabe, and A. E. Feldstein, "Identification of actin network proteins, talin-1 and filamin-a, in circulating extracellular vesicles as blood biomarkers for human myalgic encephalomyelitis/chronic fatigue syndrome," *Brain Behav. Immun.*, vol. 84, pp. 106–114, Feb. 2020.

[103] E. Fenouillet, A. Vigouroux, J. G. Steinberg, A. Chagvardieff, F. Retornaz, R. Guieu, and Y. Jammes, "Association of biomarkers with health-related quality of life and history of stressors in myalgic encephalomyelitis/chronic fatigue syndrome patients," *J. Transl. Med.*, vol. 14, p. 251, Aug. 2016.

[104] B. A. Lidbury, B. Kita, A. M. Richardson, D. P. Lewis, E. Privitera, S. Hayward, D. de Kretser, and M. Hedger, "Rethinking ME/CFS diagnostic reference intervals via machine learning, and the utility of activin B for defining symptom severity," *Diagnostics (Basel)*, vol. 9, July 2019.

[105] M. Maes, I. Mihaylova, and M. De Ruyter, "Decreased dehydroepiandrosterone sulfate but normal insulin-like growth factor in chronic fatigue syndrome (CFS): relevance for the inflammatory response in CFS," *Neuro Endocrinol. Lett.*, vol. 26, pp. 487–492, Oct. 2005.

[106] A. Melvin, E. Lacerda, H. M. Dockrell, S. O'Rahilly, and L. Nacul, "Circulating levels of GDF15 in patients with myalgic encephalomyelitis/chronic fatigue syndrome," *J. Transl. Med.*, vol. 17, p. 409, Dec. 2019.

[107] L. Nacul, B. de Barros, C. C. Kingdon, J. M. Cliff, T. G. Clark, K. Mudie, H. M. Dockrell, and E. M. Lacerda, "Evidence of clinical pathology abnormalities in people with myalgic Encephalomyelitis/Chronic fatigue syndrome (ME/CFS) from an analytic Cross-Sectional study," *Diagnostics (Basel)*, vol. 9, Apr. 2019.

[108] A. Nkiliza, M. Parks, A. Cseresznye, S. Oberlin, J. E. Evans, T. Darcey, K. Aenlle, D. Niedospial, M. Mullan, F. Crawford, N. Klimas, and L. Abdullah, "Sex-specific plasma lipid profiles of ME/CFS patients and their association with pain, fatigue, and cognitive symptoms," *J. Transl. Med.*, vol. 19, p. 370, Aug. 2021.

[109] N. Shishioh-Ikejima, T. Ogawa, K. Yamaguti, Y. Watanabe, H. Kuratsune, and H. Kiyama, "The increase of alpha-melanocyte-stimulating hormone in the plasma of chronic fatigue syndrome patients," *BMC Neurol.*, vol. 10, p. 73, Aug. 2010.

[110] K. Sørland, M. K. Sandvik, I. G. Rekeland, L. Ribu, M. C. Småstuen, O. Mella, and Ø. Fluge, "Reduced endothelial function in myalgic Encephalomyelitis/Chronic fatigue Syndrome-Results from Open-Label cyclophosphamide intervention study," *Front. Med.*, vol. 8, p. 642710, Mar. 2021.

[111] A. E. Jenike and M. K. Halushka, "mir-21: a non-specific biomarker of all maladies," *Biomark Res*, vol. 9, p. 18, Mar. 2021.

[112] P. Piscopo, M. Grasso, M. Puopolo, E. D'Acunto, G. Talarico, A. Crestini, M. Gasparini, R. Campopiano, S. Gambardella, A. E. Castellano, *et al.*, "Circulating mir-127-3p as a potential biomarker for differential diagnosis in frontotemporal dementia," *Journal of Alzheimer's Disease*, vol. 65, no. 2, pp. 455–464, 2018.

[113] M. H. Sahami-Fard, E. F. Yazd, Z. Jamalpoor, and A. Nezhadi, "Expression levels of mir-127-3p and mir-144-3p in gastric cancer and their relationships with clinicopathological parameters.," *Clinical Laboratory*, no. 6, 2020.

[114] R. Nisenbaum, A. Jones, J. Jones, and W. Reeves, "Longitudinal analysis of symptoms reported by patients with chronic fatigue syndrome," *Ann. Epidemiol.*, vol. 10, p. 458, Oct. 2000.

[115] K. J. Reynolds, S. D. Vernon, E. Bouchery, and W. C. Reeves, "The economic impact of chronic fatigue syndrome," *Cost Eff. Resour. Alloc.*, vol. 2, p. 4, June 2004.

[116] R. Cairns and M. Hotopf, "A systematic review describing the prognosis of chronic fatigue syndrome," *Occup. Med.*, vol. 55, pp. 20–31, Jan. 2005.

[117] M. M. Brown, D. S. Bell, L. A. Jason, C. Christos, and D. E. Bell, "Understanding long-term outcomes of chronic fatigue syndrome," *J. Clin. Psychol.*, vol. 68, pp. 1028–1035, Sept. 2012.

[118] J. Joyce, M. Hotopf, and S. Wessely, "The prognosis of chronic fatigue and chronic fatigue syndrome: a systematic review," *QJM*, vol. 90, pp. 223–233, Mar. 1997.

[119] D. S. Ciccone, H. K. Chandler, and B. H. Natelson, "Illness trajectories in the chronic fatigue syndrome: a longitudinal study of improvers versus non-improvers," *J. Nerv. Ment. Dis.*, vol. 198, pp. 486–493, July 2010.

[120] C. Rusu, M. E. Gee, C. Lagacé, and M. Parlor, "Chronic fatigue syndrome and fibromyalgia in canada: prevalence and associations with six health status indicators," *Health Promot Chronic Dis Prev Can*, vol. 35, pp. 3–11, Mar. 2015.

[121] L. A. Jason, K. Corradi, S. Gress, S. Williams, and S. Torres-Harding, "Causes of death among patients with chronic fatigue syndrome," *Health Care Women Int.*, vol. 27, pp. 615–626, Aug. 2006.

[122] W. R. Smith, C. Noonan, and D. Buchwald, "Mortality in a cohort of chronically fatigued patients," *Psychol. Med.*, vol. 36, pp. 1301–1306, Sept. 2006.

[123] L. A. Jason, M. C. Benton, L. Valentine, A. Johnson, and S. Torres-Harding, "The economic impact of ME/CFS: individual and societal costs," *Dyn. Med.*, vol. 7, p. 6, Apr. 2008.

[124] E. R. Unger, J.-M. S. Lin, H. Tian, B. H. Natelson, G. Lange, D. Vu, M. Blate, N. G. Klimas, E. G. Balbin, L. Bateman, A. Allen, C. W. Lapp, W. Springs, A. M. Kogelnik, C. C. Phan, J. Danver, R. N. Podell, T. Fitzpatrick, D. L. Peterson, C. G. Gottschalk, M. S. Rajeevan, and for the MCAM Study Group, "Multi-Site Clinical Assessment of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (MCAM): Design and Implementation of a Prospective/Retrospective Rolling Cohort Study," *American Journal of Epidemiology*, vol. 185, pp. 617–626, 03 2017.

[125] T. Pendergrast, A. Brown, M. Sunnquist, R. Jantke, J. L. Newton, E. B. Strand, and L. A. Jason, "Housebound versus nonhousebound patients with myalgic encephalomyelitis and chronic fatigue syndrome," *Chronic Illness*, vol. 12, no. 4, pp. 292–307, 2016. PMID: 27127189.

[126] G. Andersson, A. Rozental, R. Shafran, and P. Carlbring, "Long-term effects of internet-supported cognitive behaviour therapy," *Expert Rev. Neurother.*, vol. 18, pp. 21–28, Jan. 2018.

[127] K. J. Geraghty and C. Blease, "Cognitive behavioural therapy in the treatment of chronic fatigue syndrome: A narrative review on efficacy and informed consent," *J. Health Psychol.*, vol. 23, pp. 127–138, Jan. 2018.

[128] M. Vink and A. Vink-Niese, "The draft updated NICE guidance for ME/CFS highlights the unreliability of subjective outcome measures in non-blinded trials," *J. Health Psychol.*, vol. 27, pp. 9–12, Jan. 2022.

[129] U. S. Me and C. C. Coalition, "ME/CFS treatment recommendations. february 20, 2021," 2021.

[130] K. Fukuda, S. E. Straus, I. Hickie, M. C. Sharpe, J. G. Dobbins, and A. Komaroff, "The chronic fatigue syndrome: a comprehensive approach to its definition and study. international chronic fatigue syndrome study group," *Ann. Intern. Med.*, vol. 121, pp. 953–959, Dec. 1994.

[131] S. Adam, M. S. Y. Thong, E. Martin-Diener, B. Camey, C. Egger Hayoz, I. Konzelmann, S. M. Mousavi, C. Herrmann, S. Rohrmann, M. Wanner, K. Staehelin, R. T. Strebel, M. Randazzo, H. John, H.-P. Schmid, A. Feller, and V. Arndt, "Identifying classes of the pain, fatigue, and

depression symptom cluster in long-term prostate cancer survivors-results from the multi-regional prostate cancer survivorship study in switzerland (PROCAS)," *Support. Care Cancer*, vol. 29, pp. 6259–6269, Nov. 2021.

[132] S. Poenaru, S. J. Abdallah, V. Corrales-Medina, and J. Cowan, "COVID-19 and post-infectious myalgic encephalomyelitis/chronic fatigue syndrome: a narrative review," *Ther Adv Infect Dis*, vol. 8, p. 20499361211009385, Jan. 2021.

[133] C. Arpino, M. P. Carrieri, G. Valesini, E. Pizzigallo, P. Rovere, U. Tirelli, F. Conti, P. Dialmi, A. Barberio, N. Rusconi, O. Bosco, A. Lazzarin, A. Saracco, M. L. Moro, and D. Vlahov, "Idiopathic chronic fatigue and chronic fatigue syndrome: a comparison of two case-definitions," *Ann. Ist. Super. Sanita*, vol. 35, no. 3, pp. 435–441, 1999.

[134] J. Castro-Marrero, N. Sáez-Francàs, D. Santillo, and J. Alegre, "Treatment and management of chronic fatigue syndrome/myalgic encephalomyelitis: all roads lead to rome," *Br. J. Pharmacol.*, vol. 174, pp. 345–369, Mar. 2017.

[135] J. D. Fisk, P. G. Ritvo, L. Ross, D. A. Haase, T. J. Marrie, and W. F. Schlech, "Measuring the functional impact of fatigue: initial validation of the fatigue impact scale," *Clin. Infect. Dis.*, vol. 18 Suppl 1, pp. S79–83, Jan. 1994.

[136] J. D. Fisk and S. E. Doble, "Construction and validation of a fatigue impact scale for daily administration (D-FIS)," *Qual. Life Res.*, vol. 11, pp. 263–272, May 2002.

[137] D. J. Buysse, C. F. Reynolds, 3rd, T. H. Monk, S. R. Berman, and D. J. Kupfer, "The pittsburgh sleep quality index: a new instrument for psychiatric practice and research," *Psychiatry Res.*, vol. 28, pp. 193–213, May 1989.

[138] J. Alonso, L. Prieto, and J. M. Antó, "[the spanish version of the SF-36 health survey (the SF-36 health questionnaire): an instrument for measuring clinical results]," *Med. Clin.*, vol. 104, pp. 771–776, May 1995.

[139] N. R. McGregor, R. H. Dunstan, H. L. Butt, T. K. Roberts, I. J. Klineberg, and M. Zerbes, "A preliminary assessment of the association of SCL-90-R psychological inventory responses with changes in urinary metabolites in patients with chronic fatigue syndrome," *J. Chronic Fatigue Syndr.*, vol. 3, pp. 17–37, Jan. 1997.

[140] C. Castresana, A. G.-E. Perez, and J. L. G. de Rivera, "" hospital anxiety and depression scale" y psicopatología afectiva," in *Anales de psiquiatría*, vol. 11, pp. 126–130, 1995.

[141] M. Lacasa, J. Alegre-Martin, R. S. Sentañes, L. Varela-Sende, J. Jurek, and J. Castro-Marrero, "Yeast Beta-Glucan supplementation with multivitamins attenuates cognitive impairments in individuals with myalgic Encephalomyelitis/Chronic fatigue syndrome: A randomized, Double-Blind, Placebo-Controlled trial," *Nutrients*, vol. 15, Oct. 2023.

[142] D. G. Contopoulos-Ioannidis, A. Karvouni, I. Kouri, and J. P. A. Ioannidis, "Reporting and interpretation of SF-36 outcomes in randomised trials: systematic review," *BMJ*, vol. 338, p. a3006, Jan. 2009.

[143] A. Garratt, L. Schmidt, A. Mackintosh, and R. Fitzpatrick, "Quality of life measurement: bibliographic study of patient assessed health outcome measures," *BMJ*, vol. 324, p. 1417, June 2002.

[144] E. Ruiz, J. Alegre, A. García Quintana, L. Aliste, A. Blázquez, and T. Fernández de Sevilla, "Síndrome de fatiga crónica: estudio de una serie consecutiva de 824 casos evaluados en dos unidades especializadas," *Revista Clínica Española*, vol. 211, no. 8, pp. 385–390, 2011.

[145] S. Raschka, "Mlxtend: Providing machine learning and data science utilities and extensions to python's scientific computing stack," *The Journal of Open Source Software*, vol. 3, Apr. 2018.

[146] D. C. W. M. Vos-Vromans, R. J. E. M. Smeets, I. P. J. Huijnen, A. J. A. Köke, W. M. G. C. Hitters, L. J. M. Rijnders, M. Pont, B. Winkens, and J. A. Knottnerus, "Multidisciplinary rehabilitation treatment versus cognitive behavioural therapy for patients with chronic fatigue syndrome: a randomized controlled trial," *J. Intern. Med.*, vol. 279, pp. 268–282, Mar. 2016.

[147] M. Núñez, J. Fernández-Solà, E. Nuñez, J.-M. Fernández-Huerta, T. Godás-Sieso, and E. Gomez-Gil, "Health-related quality of life in patients with chronic fatigue syndrome: group cognitive behavioural therapy and graded exercise versus usual treatment. a randomised controlled trial with 1 year of follow-up," *Clin. Rheumatol.*, vol. 30, pp. 381–389, Mar. 2011.

[148] H. O'Dowd, P. Gladwell, C. A. Rogers, S. Hollinghurst, and A. Gregory, "Cognitive behavioural therapy in chronic fatigue syndrome: a randomised controlled trial of an outpatient group programme," *Health Technol. Assess.*, vol. 10, pp. iii–iv, ix–x, 1–121, Oct. 2006.

[149] P. K. Peterson, A. Pheley, J. Schroeppel, C. Schenck, P. Marshall, A. Kind, J. M. Haugland, L. J. Lambrecht, S. Swan, and S. Goldsmith, "A preliminary placebo-controlled crossover trial of fludrocortisone for chronic fatigue syndrome," *Arch. Intern. Med.*, vol. 158, pp. 908–914, Apr. 1998.

[150] P. D. White, K. A. Goldsmith, A. L. Johnson, L. Potts, R. Walwyn, J. C. DeCesare, H. L. Baber, M. Burgess, L. V. Clark, D. L. Cox, J. Bavinton, B. J. Angus, G. Murphy, M. Murphy, H. O'Dowd, D. Wilks, P. McCrone, T. Chalder, M. Sharpe, and PACE trial management group, "Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a randomised trial," *Lancet*, vol. 377, pp. 823–836, Mar. 2011.

[151] L. V. Clark, F. Pesola, J. M. Thomas, M. Vergara-Williamson, M. Beynon, and P. D. White, "Guided graded exercise self-help plus specialist medical care versus specialist medical care alone for chronic fatigue syndrome (GETSET): a pragmatic randomised controlled trial," *Lancet*, vol. 390, pp. 363–373, July 2017.

[152] I. Pinxsterhuis, L. Sandvik, E. B. Strand, E. Bautz-Holter, and U. Sveen, "Effectiveness of a group-based self-management program for people with chronic fatigue syndrome: a randomized controlled trial," *Clin. Rehabil.*, vol. 31, pp. 93–103, Jan. 2017.

[153] M. Burgess, M. Andiappan, and T. Chalder, "Cognitive behaviour therapy for chronic fatigue syndrome in adults: face to face versus telephone treatment: a randomized controlled trial," *Behav. Cogn. Psychother.*, vol. 40, pp. 175–191, Mar. 2012.

[154] M. Tummers, H. Knoop, A. van Dam, and G. Bleijenberg, "Implementing a minimal intervention for chronic fatigue syndrome in a mental health centre: a randomized controlled trial," *Psychol. Med.*, vol. 42, pp. 2205–2215, Oct. 2012.

[155] M. E. Roerink, S. J. H. Bredie, M. Heijnen, C. A. Dinarello, H. Knoop, and J. W. M. Van der Meer, "Cytokine inhibition in patients with chronic fatigue syndrome: A randomized trial," *Ann. Intern. Med.*, vol. 166, pp. 557–564, Apr. 2017.

[156] A. J. Wearden, C. Dowrick, C. Chew-Graham, R. P. Bentall, R. K. Morriss, S. Peters, L. Riste, G. Richardson, K. Lovell, G. Dunn, and Fatigue Intervention by Nurses Evaluation (FINE) trial writing group and the FINE trial group, "Nurse led, home based self help treatment for patients in primary care with chronic fatigue syndrome: randomised controlled trial," *BMJ*, vol. 340, p. c1777, Apr. 2010.

[157] R. A. Hobday, S. Thomas, A. O'Donovan, M. Murphy, and A. J. Pinching, "Dietary intervention in chronic fatigue syndrome," *J. Hum. Nutr. Diet.*, vol. 21, pp. 141–149, Apr. 2008.

[158] P. Powell, R. P. Bentall, F. J. Nye, and R. H. Edwards, "Randomised controlled trial of patient education to encourage graded exercise in chronic fatigue syndrome," *BMJ*, vol. 322, pp. 387–390, Feb. 2001.

[159] H. Walach, H. Bosch, G. Lewith, J. Naumann, B. Schwarzer, S. Falk, N. Kohls, E. Haraldsson, H. Wiesendanger, A. Nordmann, H. Tomasson, P. Prescott, and H. C. Bucher, "Effectiveness of distant healing for patients with chronic fatigue syndrome: a randomised controlled partially blinded trial (EUHEALS)," *Psychother. Psychosom.*, vol. 77, pp. 158–166, Feb. 2008.

[160] J. Castro-Marrero, M.-C. Zaragozá, I. López-Vílchez, J. L. Galmés, B. Cordobilla, S. Maurel, J. C. Domingo, and J. Alegre-Martín, "Effect of melatonin plus zinc supplementation on fatigue perception in myalgic Encephalomyelitis/Chronic fatigue syndrome: A randomized, Double-Blind, Placebo-Controlled trial," *Antioxidants (Basel)*, vol. 10, June 2021.

[161] C. M. C. van Campen, P. C. Rowe, F. W. A. Verheugt, and F. C. Visser, "Physical activity measures in patients with myalgic encephalomyelitis/chronic fatigue syndrome: correlations between peak oxygen consumption, the physical functioning scale of the SF-36 questionnaire, and the number of steps from an activity meter," *J. Transl. Med.*, vol. 18, p. 228, June 2020.

[162] J. Castro-Marrero, M. D. Cordero, M. J. Segundo, N. Sáez-Francàs, N. Calvo, L. Román-Malo, L. Aliste, T. Fernández de Sevilla, and J. Alegre, "Does oral coenzyme Q10 plus NADH supplementation improve fatigue and biochemical parameters in chronic fatigue syndrome?," *Antioxid. Redox Signal.*, vol. 22, pp. 679–685, Mar. 2015.

[163] D. R. Strayer, W. A. Carter, B. C. Stouch, S. R. Stevens, L. Bateman, P. J. Cimoch, C. W. Lapp, D. L. Peterson, the Chronic Fatigue Syndrome AMP-516 Study Group, and W. M. Mitchell, "A double-blind, placebo-controlled, randomized, clinical trial of the tlr-3 agonist rintatolimod in severe cases of chronic fatigue syndrome," *PLOS ONE*, vol. 7, pp. 1–9, 03 2012.

[164] S. Ryali, K. Supekar, D. A. Abrams, and V. Menon, "Sparse logistic regression for whole-brain classification of fMRI data," *Neuroimage*, vol. 51, pp. 752–764, June 2010.

[165] F. Pereira, T. Mitchell, and M. Botvinick, "Machine learning classifiers and fMRI: a tutorial overview," *Neuroimage*, vol. 45, pp. S199–209, Mar. 2009.

[166] D. Provenzano, S. D. Washington, and J. N. Baraniuk, "A machine learning approach to the differentiation of functional magnetic resonance imaging data of chronic fatigue syndrome (CFS) from a sedentary control," *Front. Comput. Neurosci.*, vol. 14, p. 2, Jan. 2020.

[167] S. Nikolin, Y. Y. Tan, A. Schwaab, A. Moffa, C. K. Loo, and D. Martin, "An investigation of working memory deficits in depression using the n-back task: A systematic review and meta-analysis," *Journal of Affective Disorders*, vol. 284, pp. 1–8, 2021.

[168] C. Jarrold and J. N. Towse, "Individual differences in working memory," *Neuroscience*, vol. 139, pp. 39–50, Apr. 2006.

[169] T. Boeken, J. Feydy, A. Lecler, P. Soyer, A. Feydy, M. Barat, and L. Duron, "Artificial intelligence in diagnostic and interventional radiology: Where are we now?," *Diagnostic and Interventional Imaging*, vol. 104, no. 1, pp. 1–5, 2023.

[170] H.-L. Ngan, K.-Y. Lam, Z. Li, J. Zhang, and Z. Cai, "Machine learning facilitates the application of mass spectrometry-based metabolomics to clinical analysis: A review of early diagnosis of high mortality rate cancers," *TrAC Trends in Analytical Chemistry*, vol. 168, p. 117333, 2023.

[171] F. Alharbi and A. Vakanski, "Machine learning methods for cancer classification using gene expression data: A review," *Bioengineering*, vol. 10, no. 2, 2023.

[172] U. Tirelli, M. Franzini, L. Valdenassi, S. Pandolfi, M. Berretta, G. Ricevuti, and S. Chirumbolo, "Patients with myalgic Encephalomyelitis/Chronic fatigue syndrome (ME/CFS) greatly improved fatigue symptoms when treated with Oxygen-Ozone autohemotherapy," *J. Clin. Med. Res.*, vol. 11, Dec. 2021.

[173] K. A. Seton, J. A. Espejo-Oltra, K. Giménez-Orenga, R. Haagmans, D. J. Ramadan, and J. Mehlsen, "Advancing research and treatment: An overview of clinical trials in myalgic encephalomyelitis/chronic fatigue syndrome (me/cfs) and future perspectives," *Journal of Clinical Medicine*, vol. 13, no. 2, 2024.

[174] D.-Y. Kim, J.-S. Lee, and C.-G. Son, "Systematic review of primary outcome measurements for chronic fatigue syndrome/myalgic encephalomyelitis (cfs/me) in randomized controlled trials," *Journal of Clinical Medicine*, vol. 9, no. 11, 2020.

[175] W. M. Mitchell, "Efficacy of rintatolimod in the treatment of chronic fatigue syndrome/myalgic encephalomyelitis (cfs/me)," *Expert Review of Clinical Pharmacology*, vol. 9, no. 6, pp. 755–770, 2016. PMID: 27045557.

Appendices

Support documentation:

Activities and publications related to the thesis.

INTERUNIVERSITY DOCTORATE PROGRAM IN BIOINFORMATICS (UAB, UPC, UdG ,UdL, UOC, UVic-UCC, URV, UB)

TRAINING PLAN FOR DOCTORAL STUDENT MARCOS LACASA

DIRECTORS: Dr. Jordi Casas Roma and Dr. José Alegre-Martín

MANDATORY ACTIVITIES

1. Development of research communication and dissemination skills.

Basic Level (MANDATORY)
No. of hours: 36

Each year, each Program Selection and Monitoring Committee will organize a follow-up conference that will bring together the doctoral students of the same promotion (same year of commencement) by sessions. These sessions will also be used to develop the students' skills in research communication, an aspect to which specific attention will be paid during the presentations.

The **conferences will take place in one of the participating universities**, with rotation, and will contemplate the mobility of the doctoral students. In justified cases (for example, stays abroad or residence abroad), the doctoral student will have the option to carry out the follow-up by videoconference. At **least one face-to-face follow-up** during the thesis is mandatory.

EVIDENCE:

- **Attendance June 2022 at UOC Barcelona.**
- **Attendance: June 2023 at UOC Barcelona.**

2. Assistance a working doctoral group

(REQUIRED)
No. of hours: 48

Annually, the doctoral program will organize a one-day scientific workshop. The doctoral student **must participate in at least two of these workshops**, presenting a **poster**. The workshops will aim to present, raise, share and discuss the status of the research process of the thesis project in its initial, central and final stages. The sessions will always be organized on three levels:

INITIAL: conceptual approaches and complete research plan,
CENTRAL: applied methodologies and methodological problems to be solved,
FINAL: expected results and results already achieved.

The workshops will follow a typical format, with oral presentations and two poster sessions (morning and afternoon). The language of the workshop will be the language of the program (English). Doctoral students must previously submit an abstract to the Academic Committee, from which twenty-minute talks (plus ten minutes for discussion) will be selected for each of the three levels mentioned above.

This activity will allow the sharing of experiences among students at three crucial moments in the development of their training process and will provide the PhD Coordination with a clear idea of the trajectory, level and pace of PhD research.

The workshops will be held at one of the participating universities, on a rotational basis, and provide for the mobility of doctoral students, with the exception of international students in online mode, who will be able to participate in the workshops by telematic means.

EVIDENCE:

Oral presentation: 3 FEBRUARY 2023 (Vic).

Oral presentation: 2 FEBRUARY 2024 (Girona).

3. Attendance at seminars or conferences given by experts in the field of knowledge.

(REQUIRED).
No. of hours: 9

The doctoral program will organize a cycle of scientific seminars with a periodicity of approximately two months (**four seminars per course**), in order to improve the doctoral student's knowledge on the topics related to the research of the doctoral thesis. **The doctoral student must attend at least three of these seminars per year for three years**. The seminars, which will deal with the topics covered by the program, will be given by researchers of international prestige.

The seminars will be held at one of the participating universities, on a rotational basis, and provide for the mobility of doctoral students, with the exception of international students in online mode, who will be able to participate in the seminars by telematic means.

EVIDENCE:
21st European Congress of Internal Medicine, ECIM joint with the 12th International Congress of Internal Medicine 2023.

25 credits

4. Presentation of a communication (poster or oral) at a national or international congress.

(REQUIRED)
No. of hours: 100

The presentation of one's own research, as well as its discussion before an audience with training in the subject matter presented, is one of the most useful ways of learning and improving research activity. In this sense, the framework of an international congress, specifically dedicated to the subject of the doctoral thesis, with the participation of important researchers in their field of study, is a fundamental test in the definitive and complete training of the future researcher.

Therefore, the doctoral program proposes as an activity the participation of the student, with discussions and reflections on the main issues of his thematic area, in a scientific meeting such as a workshop or international congress, where he can present the results of his research and receive the comments and criticisms of specialist and renowned researchers in his field.

Due to its highly specific nature, it will require the student to be at an advanced stage of his or her thesis work. This activity will be supervised and guided by the thesis director or tutor.

EVIDENCE:

CONFERENCE PRESENTATIONS

21st European Congress of Internal Medicine, ECIM joint with the 12th International Congress of Internal Medicine 2023.

Oral Presentation: *"A synthetic data generation system for myalgic encephalomyelitis / chronic fatigue syndrome questionnaires".*

Poster Discussion: *"The use of oxygen consumption in the CPET test as a biomarker in chronic fatigue syndrome* patients".

st21 September 2023. SRMI (Societé Romania Medicine Interne) Bridging the Gap: Research and Innovation in Internal Medicine congress.
Oral Presentation: "*Artificial Intelligence Healthcare Models: How to use it in Internal Medicine Research*".

ACCEPTANCE

2023 IACFS/ME Conference: "*CPT3™ test detects early cognitive dysfunction in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome and could prevent further deterioration due to fatigue development time*" has been accepted for presentation as an **Oral Presentation**

2023 IACFS/ME Conference *"Unsupervised cluster analysis reveals different subtypes of ME/CFS patients based on peak oxygen consumption and SF-36 score in two studies"*. has been accepted for presentation as a **Poster.**

Summer University Teruel 2024: Advances in clinical and basic research in Central Sensitization Syndromes (CSS). *Oral Presentation: Artificial Intelligence in Central Sensitization Syndromes.*

5. Elaboration of a research article, submitted to an impact scientific journal.

(REQUIRED).
No. of hours: 500

The most representative results of the student will be presented in the form of a scientific article.

Within the elaboration of the doctoral thesis, and as a fundamental element of development and training of future researchers, it is relevant to work intensively on the formal and content capacity of the students of the doctoral program for the optimal development of their scientific activity within the framework of the international research community. To this end, it is essential that they know the mode, mechanisms, tools, expository and formal strategies, and the usual requirements for the publication of a research article in indexed journals.

Considering that articles are very often the main tool of current research, it is considered essential that, before obtaining a doctoral degree, doctoral students acquire experience in writing at least one scientific article for publication in an indexed journal.

This activity will be carried out as the student completes his experimental work and has results that can lead to a scientific publication.

EVIDENCE:

The report on the quality of the publications submitted to the compendium is received:

Publication 1

Lacasa-Cazcarra, M., Launois-Obregón, P., Prados, F., Alegre-Martín, J. & Casas-Roma, J. (2023). Unsupervised cluster analysis reveals distinct subtypes of ME/CFS patients based on peak oxygen consumption and SF-36 scores. Clinical Therapeutics, 45(12), 1228-1235. doi: 10.1016/j.clinthera.2023.09.007.
ISSN: 0149-2918 / 1879114X

Tipus de publicació	Indicis de qualitat	Punts
Article indexat	**JCR** (2022): 3.2 Q3 Pharmacology & Pharmacy	
	SJR (2023): 0.875 Q1 Pharmacology (medical) / Q2 Pharmacology	
	Nombre de **citacions WoS:** 1	
	Nombre de **citacions Scopus:** 1	
	Nombre de **citacions Google Scholar:** 1	
	Article publicat en **accés obert** a "Scientia. Dipòsit d'Informació Digital del Departament de Salut": https://hdl.handle.net/11351/10723	

Publication 2

Lacasa-Cazcarra, M., Alegre-Martín, J., Sanmartín-Sentañes, R., Varela Sende, Luisa, Jurek, Joanna & Castro-Marrero, J. (2023). Yeast Beta-Glucan Supplementation with Multivitamins Attenuates Cognitive Impairments in Individuals with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Randomized, Double-Blind, Placebo-Controlled Trial. Nutrients, 15(21), 1-16. doi: 10.3390/nu15214504
ISSN: 2072-6643

Tipus de publicació	Indicis de qualitat	Punts
Article indexat	**JCR** (2022): 5.9 Q1 Nutrition & Dietetics - 17/88 rank SCIE	
	SJR (2023): 1.301 Q1 Food science / Q1 Nutrition and Dietetics	
	Article publicat en **accés obert** a "Scientia. Dipòsit d'Informació Digital del Departament de Salut": https://hdl.handle.net/11351/10770	

Publication 3

Lacasa-Cazcarra, M., Prados, F., Alegre-Martín, J. & Casas-Roma, J. (2023). A synthetic data generation system for myalgic encephalomyelitis/chronic fatigue syndrome questionnaires. Scientific Reports, 13(1), 1-10. doi: 10.1038/s41598-023-40364-6
ISSN: 2045-2322

Tipus de publicació	Indicis de qualitat	Punts
Article indexat	**JCR** (2022)**:** 4.6 Q2 Multidisciplinary Sciences - rank 22/73 SCIE	
	SJR (2023)**:** 0.9 Q1 Multidisciplinary	
	Article publicat en **accés obert**	

Publication 4

Fernández-Quirós, J., Lacasa-Cazcarra, M., Alegre-Martín, J., Sanmartín-Sentañes, R., Almirall, M., Launois-Obregón, P., Castro-Marrero, J., Rodríguez-Urrutia, A., Navarro-Sanchis, J.A. & Ramos Quiroga, J.A. (2023). The Conners Continuous Performance Test CPT3™: Is it a reliable marker to predict neurocognitive dysfunction in Myalgic encephalomyelitis/chronic fatigue syndrome?. Frontiers in Psychology, 14, 1-13. doi: 10.3389/fpsyg.2023.1127193.
ISSN: 1664-1078

Tipus de publicació	Indicis de qualitat	Punts
Article indexat	**JCR** (2022)**:** 3.8 Q1 Psychology, Multidisciplinary- rank 34/147 SSCI	
	SJR (2023)**:** 0.8 Q2 Psychology, miscellaneous	
	Nombre de **citacions WoS: 4**	
	Nombre de **citacions Scopus**: 4	
	Nombre de **citacions Google Scholar:** 7	
	Article publicat en **accés obert** al DDD dipòsit digital de documents de la UAB: https://ddd.uab.cat/record/280993	

SENT - PENDING JOURNAL DECISION

Publication 5

Confirmatory clinical trial on unsupervised cluster analysis reveals distinct subtypes of ME/CFS patients based on peak oxygen consumption and SF-36 scores.

Marcos Lacasa1, Anaís Cuerva-Muñoz3,4, Patricia Launois2 , José Alegre2, Célia Ruth Nandja3,4, Jonathan Galán Carracedo3,4 , Andrea Suárez-Segade4 ,Myriam Guerra-Balic3, Jordi Casas-Roma1,5,64- Biomedical Network Research Centre on Mental Health (CIBERSAM), Barcelona, Catalonia, Spain.

AWARDS: MUTUA MADRILEÑA SCHOLARSHIP

Dr. José Alegre, coordinator of the Chronic Fatigue Syndrome Unit of the Vall d'Hebron University Hospital and principal investigator of the Rheumatology group of the Vall d'Hebron Research Institute (VHIR), which aims to deepen the knowledge of persistent Covid will be funded by the Mutua Madrileña Foundation. The research has been one of the 21 projects selected in the XIX call for Health Research Grants of the Foundation, which allocates two million euros to this call.

KNOWLEDGE TRANSMISSION

Working meetings for the design of a PPP. As part of the tasks of this research, we have collaborated with companies specialized in the development of applications to evaluate a computer application, which develops the main conclusions of the studies presented in this compendium in order to substantiate the generation of knowledge, precisely for the benefit of people.