Resting state EEG biomarkers in translational neuroscience

Cristina Gil Ávila

Graduate School of Systemic Neurosciences

LMU Munich

Dissertation at the

Graduate School of Systemic Neurosciences

Ludwig-Maximilians-Universität München

8 August 2023

Supervisor

Prof. Dr. Markus Ploner
Department of Neurology
TUM School of Medicine
Technical University of Munich

First Reviewer: Prof. Dr. Markus Ploner
Second Reviewer: Prof. Dr. Valentin Riedl
External Reviewer: Prof. Dr. Simone Schütz-Bosbach

Date of Submission: 8 August 2023
Date of Defense: 15 December 2023

ABSTRACT

Brain biomarkers are patterns that relate brain structure or function to cognitive or clinical characteristics. Their successful identification can help in the diagnosis, prevention, monitoring, and treatment of neurological and psychiatric disorders. Abnormalities in brain structure and function have been reported in a wide range of neurological and psychiatric disorders, and many features common to these disorders are encoded in dynamic functional large-scale networks. Thus, electroencephalography (EEG), a non-invasive, widely-used, and cost-effective method able to capture brain dynamics, is a promising tool for discovering brain biomarkers.

In Project 1, we explored a particular case of biomarker discovery in chronic pain. We investigated the temporal dynamics of brain activity using microstate analysis in a large resting state EEG dataset, including 101 patients with chronic pain and 89 healthy participants. We applied this novel method to chronic pain, a highly prevalent and severely disabling neurological condition. As chronic pain has repeatedly been associated with changes in brain function and incorrect processing of information, we expected to see differences between patients with chronic pain and healthy participants in microstates' topographies or temporal characteristics. Our results indicated a decreased presence of microstate D in patients with chronic pain during resting state with eyes closed. Subgroup analysis replicated this finding in patients with chronic back pain, but patients with chronic widespread pain presented no differences in microstates' characteristics. Thus, if future studies validate these findings, microstates' characteristics could turn into diagnostic or subtyping biomarkers of chronic pain.

In Project 2, we developed DISCOVER-EEG, an open, fully automated EEG pipeline to preprocess, analyze, and visualize resting state EEG data. In recent years, notable efforts have been made to increase the transparency of EEG research. They have yielded the creation of a standardized data structure for the efficient organization, sharing, and reuse of data (BIDS-EEG) and the development of automatic EEG preprocessing pipelines for specific populations, settings, and study designs. DISCOVER-EEG builds on and extends these advances by extracting and visualizing physiologically relevant EEG features (including oscillatory power, connectivity, and network characteristics) for

biomarker identification. It builds upon and combines two open-source and widely used Matlab toolboxes (EEGLAB and FieldTrip) and follows the most recent guidelines and standards for reproducible EEG research. We tested it in two large and openly available datasets: the LEMON dataset, including 213 healthy participants, and the TD-BRAIN dataset, including 1274 participants with different psychiatric conditions. We demonstrate the robustness of the pipeline across datasets with different characteristics and its reliability in capturing well-known EEG effects, such as the reduction of alpha power during eyes open. Finally, we provide an example analysis in the LEMON dataset that could inspire biomarkers of healthy aging.

During both projects, we followed open science practices and directed our research toward transparency and collaboration. Therefore, this thesis' publications, data, and code are openly available. In conclusion, this work adds to the understanding of the pathophysiology of chronic pain, facilitates and advances the analysis of large EEG datasets, and promotes open and reproducible research on brain function.

INDEX

1. INTRODUCTION

Brain disorders, including neurological and mental health disorders, are highly prevalent and the leading cause of disability worldwide, with depression, anxiety, and pain states being among the top causes of lasting disability (Vos et al., 2020). They impose a massive burden on patients, healthcare systems, and society, and their global cost is estimated in trillions (Patel et al., 2018). Still, the causes and mechanisms of these conditions are poorly understood, and treatments are often unsatisfactory, with small therapeutic responses (Leichsenring et al., 2022). Thus, further understanding of brain disorders to improve patient care is urgently needed.

Translational neuroscience, i.e., the field that aims to translate basic neuroscience knowledge into clinical applications, plays a vital role in understanding and treating brain disorders. Evidence from the last decades has shown that many features common to neurological and mental health disorders, such as pain, anhedonia, and negative emotions, are encoded in dynamic large-scale brain networks (de Lange et al., 2019; Fornito et al., 2015). Thus, characterizing these disorders as "brain network disorders" could help in their understanding and treatment (Scangos et al., 2023). In this regard, neuroimaging modalities such as functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET), and Electroencephalography (EEG) are essential to capture non-invasively alterations of brain networks.

An increasingly rising area of translational neuroscience is the development of *brain biomarkers* (Figure 1), i.e., patterns that relate structural or functional brain features to cognitive or clinical characteristics (FDA-NIH Biomarker Working Group, 2016; Woo et al., 2017). Correlating brain features to clinical outcomes, such as common symptoms across disorders or cognitive and affective components, has the potential to redefine diagnostic categories based on brain traits and even propose new treatments accordingly (Abi-Dargham and Horga, 2016; Woo et al., 2017). Consequently, biomarkers can contribute to the development of precision medicine by individualizing treatments for mental health disorders (Abi-Dargham and Horga, 2016) and provide new targets for new, non-invasive therapies, such as neuromodulation (Ploner et al., 2023).

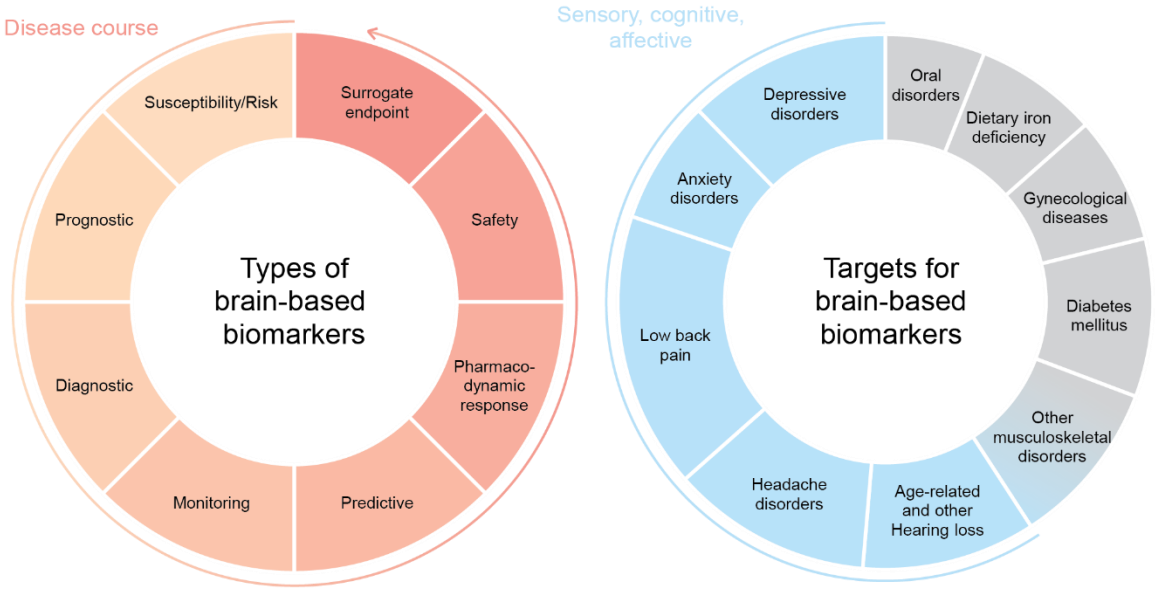

Figure 1. Illustrating the relevance of brain-based biomarkers across different disease stages and types of disorders. Left panel: types of biomarkers according to the BEST (FDA/NIH) categories (FDA-NIH Biomarker Working Group, 2016). The outer circular arrow and the graded colors of the circular sectors indicate the different stages of the disease course, including the pre-disease stage. Right panel: top 10 level-3 causes of disability worldwide, as accounted by the years lived with disability (YLDs) in the 2019 Global Burden of Disease (GBD) study (Vos et al., 2020). Blue circular sectors indicate the contribution of diseases with prominent sensory, cognitive, and/or affective symptoms. The outer blue circular line indicates the disorders that can be benefited from brain-based biomarkers. The panel was modified with permission from Ploner et al. (2023), © Springer Nature.

In the following sections, we will review the biomarker framework in translational neuroscience and its associated challenges (section 1.1.), we will examine possible targets for brain biomarkers, especially measures capturing brain dysfunction during the resting state (section 1.2.), we will explore the advantages and challenges of electroencephalography for biomarker discovery (section 1.3.), and we will introduce chronic pain as an example of a brain disorder to which the biomarker framework can be applied (section 1.4). Finally, the aims and the outline of this doctoral thesis will be detailed (section 1.5.).

1.1. BRAIN BIOMARKERS IN TRANSLATIONAL NEUROSCIENCE

Brain biomarkers can assess the risk of developing a disorder, predict disease progression or therapeutic responses, identify individuals that will benefit from treatment, and provide objective outcomes in clinical practice (Davis et al., 2020; FDA-NIH Biomarker Working Group, 2016). Despite the unquestionable utility of biomarkers, identifying, developing, and validating them is a long and costly process involving many stakeholders, such as patients, researchers, medical doctors, the life science industry, and governments. In the next paragraphs, we zoom in on the first stages of biomarker creation and its associated challenges: *generalizability*, *explainability*, and *scalability* of biomarkers (Woo et al., 2017).

First, *large and representative datasets* are needed to create biomarkers that generalize to different populations. Such datasets can only be compiled through institutional collaboration and the establishment of research consortia committed to data aggregation and sharing (Poldrack and Gorgolewski, 2014). An essential first step for collaboration was the creation of the Brain Imaging Data Structure (BIDS) standard for organizing, describing, and sharing neuroimaging data (Gorgolewski et al., 2016). Still, publicly sharing standardized data has not yet been widely adopted (Paret et al., 2022), and there is room for improvement in the creation of comprehensive datasets, including ethnicities underrepresented in science (Ricard et al., 2023).

Second, *reproducible and transparent methods* are required to generate robust and explainable biomarkers. Neuroimaging data is characterized by its high dimensionality, which allows multiple ways to analyze a dataset (Botvinik-Nezer et al., 2020). This 'analytical flexibility' (Parsons et al., 2022) can lead to a lack of reproducibility when only a written analysis description is provided. Thus, creating open and transparent workflows and publicly sharing the code and models used to generate biomarker candidates would help to mitigate this problem (Niso et al., 2022). Moreover, the trust in a biomarker is also influenced by the delivered measure's complexity and potential failure points. Therefore, biomarkers that are simple to understand, validate, and interpret are more likely to be successfully translated into clinical practice (Davis et al., 2020; Woo et al., 2017).

Third, *easily deployable technologies* are desired to render scalable biomarkers. Ideally, biomarkers would be developed and validated in real-world scenarios, where subtle pathological features of neurological and psychiatric disease might appear (Abi-Dargham and Horga, 2016). Therefore, safe, wearable, low-cost neurotechnologies are desirable for developing ecologically valid biomarkers (Stangl et al., 2023). The improvement of technologies widely established in clinical practice, such as EEG, now enables the fast and easy recording of neuronal data with mobile dry-electrode systems with comparable data quality to conventional wet-electrode systems (Kam et al., 2019). These technological developments promise improvements in developing scalable brain biomarkers.

From the challenges mentioned above, it is evident that transparent, collaborative efforts are needed to obtain clinically valid brain biomarkers. Hence, biomarker discovery, development, and validation in translational neuroscience should be addressed from an *open science* perspective (Allen and Mehler, 2019; Munafo et al., 2017; Niso et al., 2022).

1.2. TARGETS FOR BIOMARKER DISCOVERY: BRAIN DYSFUNCTION DURING RESTING STATE

Structural and functional changes in the brain have been reported across a wide range of neurological and psychiatric conditions (Fornito et al., 2015; Goodkind et al., 2015; Uhlhaas and Singer, 2012). Recent evidence has revealed that many features common across brain disorders are encoded in functional large-scale brain networks (de Lange et al., 2019). These functional networks are intrinsic to the brain and can therefore be captured during *resting state*, i.e., when the subject does not perform any explicit task (Fox and Raichle, 2007). In accordance, abnormalities of brain function during resting state have been observed in neurodegenerative disorders (Hohenfeld et al., 2018), schizophrenia (Baker et al., 2019), major depressive disorder (Drysdale et al., 2017), and chronic pain (Kucyi and Davis, 2015) among others. Therefore, measures capturing network disruptions during resting state could turn into successful biomarkers of brain disorders. In the next paragraphs, we provide recent examples of biomarker candidates

based on such features, specifically functional connectivity measures and network measures derived from graph theory.

Changes in *functional connectivity* and its dynamics have been repeatedly described in brain disorders during resting state (Fornito et al., 2015). For example, using fMRI, Drysdale et al. (2017) managed to stratify patients with major depressive disorder into four distinct biotypes based on their functional connectivity profiles. Biotypes reflected different symptomatology and predictive responsiveness to transcranial magnetic stimulation therapy, making them potential stratification biomarkers for treatment. In a different study, Lee et al. (2021) discovered a dynamic functional connectivity pattern with fMRI that could predict ongoing pain ratings in an experimental pain model. When the marker was translated to patients with chronic pain, it could also predict their ongoing pain ratings, showing potential as a pain monitoring biomarker. Functional connectivity, in consequence, can reveal patterns in brain networks able to subtype patients based on neuronal characteristics, predict treatment response, or monitor a condition.

Brain networks can be further characterized with *complex network measures* derived from *graph theory* by treating brain regions as nodes, and connections between nodes as edges in the graph (Rubinov and Sporns, 2010). Graph theory measures can reduce the information encoded in sparse connectivity matrices and help in their interpretation, e.g., by analyzing the segregation and integration of information in the network. For example, changes in functional network architecture have been described in schizophrenia and young adults at risk of psychosis, making changes in network structure promising risk biomarkers for schizophrenia (Bassett et al., 2018). Furthermore, common disruptions in network structure have been observed across twelve different psychiatric and neurological conditions (de Lange et al., 2019). These disruptions, related to global network communication and integration, could reveal common mechanisms underlying brain dysfunction and have the potential to redefine clinical categories based on brain activity.

1.3. ELECTROENCEPHALOGRAPHY FOR BIOMARKER DISCOVERY

Electroencephalography is a non-invasive neuroimaging method that captures the electrical activity generated primarily by pyramidal neurons in the cerebral cortex (Nunez and Srinivasan, 2006). Thus, oscillations are the primary subject of study of EEG. Compared to other neuroimaging modalities, such as PET and fMRI, EEG is a direct measure of neuronal activity and has a high temporal resolution, allowing the record of neuronal activity in the millisecond range at the cost of low spatial resolution. Magnetoencephalography combines some of the benefits of fMRI, PET, and EEG, as it can capture signals with high spatial and temporal resolution. However, compared to EEG, it is an expensive method, and its portability is still under development (Stangl et al., 2023). In brief, EEG is a well-established method that is safe, mobile, cost-efficient, and widely used in clinical and research contexts. All these characteristics, together with its potential to acquire ecologically valid signals, make EEG a particularly attractive tool for biomarker discovery in translational neuroscience.

Although EEG presents many advantages, it also has some pitfalls. For example, data acquisition is laborious, and traditional preprocessing and analysis of EEG data involve manual steps that are time-consuming, subjective, and dependent on expert knowledge. This results in EEG datasets often having small sample sizes. The replicability of EEG studies is aggravated by the multiple possibilities of correctly preprocessing and analyzing the data (Pavlov et al., 2021). To minimize such reproducibility issues, the Organization of Human Brain Mapping established general recommendations for data acquisition, analysis, reporting, and sharing of MEEG data and analysis (Pernet et al., 2020a). Additionally, much progress has been achieved in automatizing and speeding up EEG preprocessing (Klug et al., 2022; Pedroni et al., 2019; Pernet et al., 2020b; Pion-Tonachini et al., 2019).

Considering the advantages mentioned above and the new developments that aim to overcome some of the EEG challenges, studies proposing resting state EEG biomarkers have started to be published. One of the earliest and most controversial EEG biomarkers was the ratio between power at theta/beta frequencies as a diagnostic marker of Attention-Deficit/Hyperactivity Disorder (Arns et al., 2013; Kiiski et al., 2020). This approach, as well as the majority of EEG literature of the last 80 years, relies on the static

interpretation of the EEG power spectrum into broad frequency bands termed delta, theta, alpha, beta, and gamma (Newson and Thiagarajan, 2019). *Frequency band analysis* aggregates brain features across time, overseeing subtle temporal structures that may play a significant role in the coordination of functional networks (Baker et al., 2014). Small sample sizes, overlooking brain dynamics, and a high risk of bias are some of the reasons why EEG-based biomarkers have shown mild consensus so far (Newson and Thiagarajan, 2019). However, new studies using larger samples and based on dynamic functional measures of brain activity hold great promise for the discovery of EEG-based biomarkers. For example, Zhang et al. (2021) stratified patients with major depression and post-traumatic stress disorder based on their EEG connectivity profiles and predicted their clinical outcome to a variety of treatments.

Exploiting the high temporal precision of EEG to assess brain dynamics is, consequently, a promising approach for generating EEG biomarkers. New methods exploring the temporal dynamics of large-scale brain networks, such as EEG *microstate analysis* (Khanna et al., 2015; Michel and Koenig, 2018), have been popularized. Microstate analysis describes EEG resting state activity as a sequence of a limited number of voltage topographies, named microstates, that remain stable for hundreds of milliseconds. Microstate topographies can be reliably identified across participants, and their temporal characteristics (e.g., the average duration of a microstate or its occurrence) can be assessed in patients and healthy participants. Alterations in the temporal dynamics of microstates have been observed in schizophrenia (da Cruz et al., 2020), major depression disorder (Murphy et al., 2020), and Lewy body dementia (Schumacher et al., 2019). Therefore, they could represent predictive, monitoring, or diagnostic biomarkers of neurological and psychiatric diseases.

In conclusion, EEG can potentially unravel brain biomarkers in neuropsychiatric disorders due to its wide availability and its ability to capture functional brain networks at a high temporal resolution.

1.4. BRAIN BIOMARKERS IN CHRONIC PAIN

Chronic pain is a disease that affects around 30% of people worldwide and severely decreases the quality of life of those affected from it (Cohen et al., 2021). Despite its enormous societal and economic impact, its pathophysiology is not yet well understood, and its treatment is insufficient (Turk et al., 2011). Improvement of pain management via non-addictive pharmacological therapeutics and non-pharmacological interventions is urgently needed (Cohen et al., 2021). Hence, it is essential to develop translational tools such as validated biomarkers that can help to identify individuals at risk of developing chronic pain after injury, improve participant selection in clinical trials and identify objective outcomes to quantify the response of new therapies (Cohen et al., 2021; Davis et al., 2020; Tracey et al., 2019).

In contrast to acute pain, which has a protective function and is essential for healing and survival, chronic pain is a disease on its own characterized by persisting pain and accompanying sensory, cognitive, and affective abnormalities (Treede et al., 2019). As such, structural and functional alterations of the peripheral and central nervous systems have been observed in chronic pain (Baliki and Apkarian, 2015; Kuner and Flor, 2017). In the last years, converging lines of evidence have demonstrated structural and functional reorganization of the brain during chronic pain across spatial scales and species (Baliki and Apkarian, 2015; Kuner and Kuner, 2021). In particular, the prefrontal, sensory, motor, and cingulate cortices and subcortical areas, including the amygdala, hippocampus, and striatal areas, have been implicated in chronic pain (Kuner and Flor, 2017).

Importantly, brain areas and circuits involved in acute pain overlap but are not the same as those implicated in chronic pain (Baliki et al., 2006). Biomarkers of acute pain might have thus difficulties translating into clinical practice (Mouraux and Iannetti, 2018). As an example, a solid and specific signature of acute pain described with fMRI, the Neurologic Pain Signature (Wager et al., 2013), is thought to capture the nociceptive component of pain but not the emotional and cognitive aspects of chronic pain (Lopez-Sola et al., 2017). Biomarkers developed directly in patients with chronic pain, or at least in sustained pain models, might have better chances of being translated to clinical practice; see Lee et al. (2021) for a representative example.

As pain is an integrative phenomenon that emerges from dynamic functional networks (Kucyi and Davis, 2015; Ploner et al., 2017), using brain oscillations as a framework to discover pain biomarkers seems a promising approach. Diagnostic EEG-based biomarkers of chronic pain present, in general, higher theta and beta power in patients with chronic pain compared to healthy participants (Zebhauser et al., 2022). Nonetheless, the risk of bias is high in many studies and domains. As mentioned earlier in section 1.3., focusing on dynamic brain measures might hold more promise for developing chronic pain biomarkers than traditional frequency band analysis. For instance, Ta Dinh et al. (2019) found no differences in static frequency bands between chronic pain patients and healthy participants. In contrast, they observed changes in functional connectivity and network architecture in patients with chronic pain.

In summary, predictive, stratifying, and monitoring biomarkers are particularly needed to improve the understanding and management of chronic pain. The study of brain oscillations in clinical populations might unravel such a biomarker. However, high-quality studies with sufficient statistical power and exhaustive validation are needed before obtaining a clinical biomarker of chronic pain.

1.5. AIMS AND OUTLINE

The primary goal of this thesis was to investigate non-invasive biomarkers of chronic pain with electroencephalography. During the process of finding such a biomarker, we identified the need to develop reliable and robust tools for EEG biomarker discovery. Therefore, a second goal was set: developing an automatic and transparent tool that could facilitate the aggregation, preprocessing, and analysis of resting state EEG data in a wide range of neuropsychiatric disorders.

This thesis is organized in chronological order, which intends to reflect the challenges found during the process of finding non-invasive biomarkers in chronic pain and our consequent learning of new methods and perspectives. For that reason, we part from a particular case of biomarker discovery in chronic pain using EEG microstate analysis (project 1), to the creation of an automatic EEG pipeline to preprocess, visualize, and extract physiologically meaningful brain features that could turn into potential biomarkers (project 2).

In Project 1, we investigated the temporal dynamics of brain activity in a large cohort of patients with chronic pain and healthy participants using microstate analysis on resting state EEG. We applied this novel method to chronic pain, a highly prevalent and severely disabling neurological condition. As chronic pain has repeatedly been associated with changes in brain function and incorrect processing of information, we expected to see differences between microstate topographies or temporal characteristics between patients with chronic pain and healthy participants. Differences in microstates properties could turn into diagnostic or subtyping biomarkers of chronic pain.

In Project 2, we developed DISCOVER-EEG, an open, fully automated pipeline to enable fast and easy preprocessing, analysis, and visualization of resting state EEG data. Much progress has been made recently to speed up data collection and preprocessing. However, no tool could preprocess and extract physiologically meaningful EEG features (including oscillatory power, connectivity, and network characteristics) automatically following the most recent EEG guidelines and standards for use in biomarker identification. We approached this endeavor with an open science mindset to facilitate the aggregation, reuse, and analysis of large EEG datasets and promote transparent and reproducible research on brain function.

Finally, a common goal of both projects was to expand our knowledge of open science practices and direct our research toward transparency and collaboration. Therefore, all data and code used in this thesis are openly available. Anonymized raw resting state EEG data, behavioral and clinical outcomes, data derivatives, and related code of Project 1 are available at https://osf.io/srpbg/. The code of the DISCOVER-EEG pipeline, as well as script examples demonstrating its use and related data of project 2, are available at https://osf.io/mru42/. All raw EEG data are structured according to the BIDS-EEG standard (Pernet et al., 2019) to facilitate their reusability.

2. PROJECT 1. EEG MICROSTATE ANALYSIS IN PATIENTS WITH CHRONIC PAIN

This section includes the research article 'Dynamics of brain function in patients with chronic pain assessed by microstate analysis of resting-state electroencephalography', published under an open access license (CC BY-NC-ND) in the journal PAIN in December 2021 (May et al., 2021).

Authors

Elisabeth S. May*, Cristina Gil Ávila*, Son Ta Dinh, Henrik Heitmann, Vanessa D. Hohn, Moritz M. Nickel, Laura Tiemann, Thomas R. Tölle, Markus Ploner

Contributions

Conceptualization: CGA, ESM, and MP; Methodology: CGA, ESM, and MP; Software: CGA and ESM; Validation: CGA and ESM; Formal analysis: CGA and ESM; Investigation: CGA, ESM, STD, HH, VDH, MMN, and LT; Data curation: CGA and ESM; Visualization: CGA, ESM, and MP; Writing – original draft: CGA, ESM, and MP; Writing – review and editing: CGA, ESM, STD, HH, VDH, MMN, LT, TRT, and MP; Supervision: MP; Project administration: MP; Funding acquisition: MP

* Elisabeth S. May and Cristina Gil Ávila contributed equally to this work and shared the first authorship of the article.

Research Paper

PAIN®

OPEN

Dynamics of brain function in patients with chronic pain assessed by microstate analysis of resting-state electroencephalography

Elisabeth S. May[a,b], Cristina Gil Ávila[a,b], Son Ta Dinh[a,b], Henrik Heitmann[a,b,c], Vanessa D. Hohn[a,b], Moritz M. Nickel[a,b], Laura Tiemann[a,b], Thomas R. Tölle[a,c], Markus Ploner[a,b,c,*]

Abstract

Chronic pain is a highly prevalent and severely disabling disease that is associated with substantial changes of brain function. Such changes have mostly been observed when analyzing static measures of resting-state brain activity. However, brain activity varies over time, and it is increasingly recognized that the temporal dynamics of brain activity provide behaviorally relevant information in different neuropsychiatric disorders. Here, we therefore investigated whether the temporal dynamics of brain function are altered in chronic pain. To this end, we applied microstate analysis to eyes-open and eyes-closed resting-state electroencephalography data of 101 patients suffering from chronic pain and 88 age- and sex-matched healthy controls. Microstate analysis describes electroencephalography activity as a sequence of a limited number of topographies termed microstates that remain stable for tens of milliseconds. Our results revealed that sequences of 5 microstates, labelled with the letters A to E, consistently described resting-state brain activity in both groups in the eyes-closed condition. Bayesian analysis of the temporal characteristics of microstates revealed that microstate D has a less predominant role in patients than in controls. As microstate D has previously been related to attentional networks and functions, these abnormalities might relate to dysfunctional attentional processes in chronic pain. Subgroup analyses replicated microstate D changes in patients with chronic back pain, while patients with chronic widespread pain did not show microstates alterations. Together, these findings add to the understanding of the pathophysiology of chronic pain and point to changes of brain dynamics specific to certain types of chronic pain.

Keywords: Chronic pain, Dynamics, EEG, Microstate analysis, Resting-state

1. Introduction

Chronic pain is a highly disabling disease that affects 20% to 30% of the adult population.[7,24] Its pathophysiology is not fully understood, and treatment is often insufficient,[60] imposing a tremendous burden on patients, health care systems, and society.[46] Converging lines of evidence have shown that chronic pain is associated with extensive changes of brain structure and function.[2,29] Understanding these changes promises fundamental insights into the underlying pathophysiology and might eventually help to establish a much sought-after biomarker of chronic pain.[14,58]

Brain function in chronic pain has mostly been assessed using functional magnetic resonance imaging (fMRI)[2] and electroencephalography (EEG)/magnetoencephalography.[44] Most studies have analyzed static measures of brain activity during the resting state, usually by aggregating a certain feature of brain function across several minutes. However, brain activity varies over time, and it is increasingly recognized that these temporal dynamics provide behaviorally and clinically relevant information that complements static measures.[19,45] Correspondingly, it has been proposed that the dynamics of brain activity and connectivity critically shape the perception of pain.[28] By assessing brain activity and connectivity at ultra-low frequencies below 0.1 Hz, recent fMRI studies have provided support for this concept in chronic pain.[3,5,10,59] However, the temporal dynamics of chronic pain–related brain activity at frequencies higher than 1 Hz have not been consistently explored yet.

Electroencephalography and magnetoencephalography are well suited to study such dynamic changes of brain activity at higher frequencies. One of the best-established methods in this field is microstate analysis (see Refs. 25,34 for reviews) that has revealed that temporal changes of EEG activity do not occur randomly. Instead, EEG activity switches between a limited number of so-called microstates. During a microstate, the EEG topography remains stable for tens of milliseconds before abruptly transitioning to another microstate. Electroencephalography resting-state activity is usually well-described with 4 to 6 microstates, which are remarkably similar across participants. Thus, microstate analysis quantifies resting-state EEG recordings

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

E.S. May and C. Gil Ávila contributed to this work equally.

[a] Department of Neurology, School of Medicine, Technical University of Munich (TUM), Munich, Germany, [b] TUM-Neuroimaging Center, School of Medicine, TUM, Munich, Germany, [c] Center for Interdisciplinary Pain Medicine, School of Medicine, TUM, Munich, Germany

**Corresponding author. Address: Department of Neurology, Technical University of Munich (TUM), Ismaninger Str. 22, 81675 Munich, Germany. Tel.: +49-89-4140-4608. E-mail address: markus.ploner@tum.de (M. Ploner).*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painjournalonline.com).

PAIN 162 (2021) 2894–2908

http://dx.doi.org/10.1097/j.pain.0000000000002281

as sequences of a limited number of microstates. The temporal characteristics of these microstates carry important information about mental processes.[6,34] Moreover, abnormalities of temporal microstate characteristics have been observed in different neuropsychiatric disorders.[12,38,48] During the writing of this article, a first microstate study in patients suffering from chronic pain was published. The results showed lower occurrence and time coverage of microstate C in patients with chronic widespread pain.[20] However, these findings need to be replicated and extended to other chronic pain conditions.

Here, we investigated whether the temporal dynamics of brain activity are changed in a large cohort of patients suffering from chronic pain. To this end, we applied microstate analysis to EEG resting-state recordings of 101 patients suffering from different types of chronic pain and 88 matched healthy control participants. Thereby, the study aimed to further the understanding of the pathophysiology of chronic pain and to potentially contribute to the development of a brain-based biomarker of chronic pain.

2. Materials and methods

2.1. Participants

The current study represents a re-analysis of previously published data obtained at the Technical University of Munich for the large-scale study of brain dysfunction in chronic pain.[54] One hundred one patients (69 women; age = 58.1 ± 13.6 years [mean ± SD]) suffering from different types of chronic pain and 88 age- and sex-matched healthy controls (60 women, age = 57.5 ± 14.2 years) participated in the study. Inclusion criteria for patients were a clinical diagnosis of chronic pain, with pain lasting at least 6 months, and a minimum reported average pain intensity of at least 4 of 10 during the past 4 weeks (0 = no pain and 10 = worst imaginable pain). Exclusion criteria for patients were acute changes of the pain condition during the past 3 months (eg, due to recent injuries or surgeries), major neurological diseases (eg, epilepsy, stroke, or dementia), major psychiatric diseases aside from depression, and severe general diseases. Patients taking benzodiazepines were also excluded. Other medication was not restricted and was maintained. In total, 47 patients with chronic back pain, 30 patients with chronic widespread pain, 6 patients with joint pain, and 18 patients with neuropathic pain were included in the study. Exclusion criteria for healthy participants were a medical history of pain lasting more than 6 months, having any pain on the day of testing, surgery, or acute injury during the past 3 months, and any neurological or psychiatric diseases. All participants provided written informed consent. The study was approved by the ethics committee of the Medical Faculty of the Technical University of Munich and conducted according to the relevant guidelines and regulations.

Questionnaires were used to assess pain characteristics and comorbidities immediately before the EEG recording. All patients completed the following questionnaires: Pain characteristics were assessed by the short-form McGill Pain Questionnaire (SF-MPQ),[33] depression by the Beck Depression Inventory II (BDI-II),[4] and anxiety by the State-Trait Anxiety Inventory (STAI).[52] The medication was quantified for all patients using the Medication Quantification Scale (MQS),[21] which quantifies a patient's pain medication profile in a single numerical value. Eighty-one patients additionally completed the painDETECT questionnaire[18] to assess the neuropathic pain component, and 47 patients completed the Pain Disability Index (PDI)[16] and the Veteran's RAND 12-Item Health Survey (VR-12)[50] to assess pain disability and quality of life, respectively. All healthy control participants completed BDI-II and STAI questionnaires to assess potential comorbidities. Detailed characteristics of the participants can be found in **Table 1**.

2.2. Recordings

Brain activity was recorded using EEG during the resting state. Participants were instructed to stay in a wakeful and relaxed state without performing any particular task. For most participants, two 5-minute blocks of continuous resting-state data were recorded, one with eyes closed and the other with eyes open. During the eyes-open condition, participants were asked to rest their gaze on a centrally presented visual fixation cross. The temporal order of the blocks was counterbalanced. During the recording, participants were comfortably seated and listened to white noise played through headphones to mask any ambient noise. For 5 patients with chronic widespread pain and 7 healthy controls, only one 5-minute block with eyes closed was recorded. Thus, final sample sizes were 101 patients and 88 healthy controls for the eyes-closed condition and 96 patients and 81 healthy controls for the eyes-open condition.

Data were recorded with 64 electrodes and a BrainAmp MR plus amplifier (Brain Products, Munich, Germany). The electrodes included all electrodes from the International 10-20 system and the additional electrodes Fpz, CPz, POz, Oz, Iz, AF3/4, F5/6, FC1/2/3/4/5/6, FT7/8/9/10, C1/2/5/6, CP1/2/3/4/5/6, TP7/8/9/10, P1/2/5/6/7/8, and PO3/4/7/8/9/10 (Easycap, Herrsching, Germany). Two electrodes were placed below the outer canthus of each eye to monitor eye movements. All EEG electrodes were referenced to electrode FCz and grounded at electrode AFz. For 81 patients and 69 healthy controls, muscle activity was simultaneously recorded with 2 bipolar electromyography (EMG) electrode montages and a BrainAmp ExG MR amplifier (Brain Products, Munich, Germany). Electromyography electrodes were placed on the right masseter and neck (semispinalis capitis and splenius capitis) muscles.[13] The EMG ground electrode was placed at vertebra C2. Data were obtained at a sampling frequency of 1000 Hz, with 0.1-μV resolution, and were band-pass filtered online between 0.016 and 250 Hz. Impedances were kept below 20 kΩ.

2.3. Preprocessing

Preprocessing was performed with the Brain Vision Analyzer software (Brain Products, Munich, Germany) on the appended data from the eyes-open and the eyes-closed conditions. For artifact identification, a high pass filter at 1 Hz and a notch filter at 50 Hz were applied to remove low frequency drifts and electrical line noise, respectively. Independent component analysis was performed.[23] Components representing eye movements and muscle artifacts were identified based on their time courses and topographies and subtracted from the raw unfiltered EEG time series.[61] Signal jumps higher than ± 100 μV and their adjacent time intervals (200 ms before and after the jump) were marked for rejection. Subsequently, all data sets were visually inspected, and remaining bad intervals were marked for rejection. Finally, data were re-referenced to the average reference, and the reference electrode FCz was added to the electrode array.

2.4. Microstate analysis

Microstate analysis was performed using the free academic software Cartool version 3.8,[9] MATLAB (MathWorks, Natick, MA), and the MATLAB toolbox Fieldtrip.[41] Analyses were performed separately for the eyes-open and eyes-closed

Table 1

Demographic data and questionnaire results.

	Patients with chronic pain (mean ± SD)	Healthy controls (mean ± SD)
Number	101	88
Sex (m/f)	32/69	28/60
Age (y)	58.2 ± 13.5	57.5 ± 14.3
BDI	15.8 ± 8.9	3.5 ± 4.5
STAI—state	39.5 ± 10.6	30.6 ± 6.1
STAI—trait	44.0 ± 11.2	30.9 ± 7.1
SF-MPQ total pain score	27.1 ± 9.4	—
Current pain intensity (0-10)	5.2 ± 1.9	—
Avg. pain intensity in the past 4 wk (0-10)	5.6 ± 1.6	—
Pain duration (mo)	121.8 ± 114.4	—
PDQ	17.4 ± 6.5	—
PDI	27.4 ± 14.2	—
VR-12 PCS	31.8 ± 7.8	—
VR-12 MCS	46.4 ± 11.9	—
MQS	6.8 ± 8.1	—

Please note that data for avg. pain intensity in the past 4 weeks, pain duration, and PDQ were only available for a subset of 81 patients. Data from PDI, VR-12 PCS, and VR-12 MCS were only available for a subset of 47 patients. For most patients (n = 81), current pain intensity ratings were obtained from the painDETECT questionnaire, which uses a combination of numerical rating scale anchored at 0 (no pain) and 10 (max pain) with a color gradient. For n = 20 patients with chronic widespread pain, current pain intensity ratings were obtained from the SF-MPQ, which uses a visual analogue scale anchored at 0 (no pain) and 100 (worst imaginable pain). These ratings were divided by 10 to match rating scales across questionnaires.

Avg. pain intensity, average pain intensity in the past 4 weeks; BDI, Beck Depression Inventory; MQS, medication quantification scale; PDI, pain disability index; PDQ, painDETECT questionnaire; STAI, State-Trait Anxiety Inventory; SF-MPQ, Short-form McGill Pain Questionnaire; VAS, visual analogue scale; VR-12 PCS, Veteran's RAND 12-Item Physical Component Summary; VR-12 MCS, Veteran's RAND 12-Item Mental Component Summary.

conditions. Each 5-minute recording was first band-pass filtered between 1 and 40 Hz and downsampled to 125 Hz, in line with previous studies.[8,11,56] Subsequently, intervals marked as bad during preprocessing were rejected, and microstate analysis was performed using all remaining clean segments concatenated. An overview of the microstate analysis pipeline can be found in **Figure 1**.

2.4.1. Definition of microstates

We defined microstates through a well-established 2-step clustering procedure using a modified k-means algorithm.[42] In line with previous studies,[12,38,48,55] this was performed separately for each group and condition.

The first step consisted of a k-means clustering performed at the individual level. For each participant, EEG topographies at global field power (GFP) peaks were clustered, yielding a variable number of individual-level topographies. The GFP is a measure of the instantaneous strength of EEG activity measured over the whole scalp and mathematically defined as the SD of the signals of all electrodes.[39] EEG topographies were clustered at GFP maxima since they represent the time points of highest signal-to-noise ratio.[39,42]

The clustering algorithm requires an a priori definition of k, which is the number of clusters into which the data will be grouped. To select the optimal number of clusters, we performed the clustering with different numbers of k = 1 to 12 initial clusters, following Cartool default settings for resting-state data. First, an initial number of k topographies was randomly selected from all GFP-peak topographies of the individual EEG time series. Second, the selected topographies were spatially correlated with the remaining topographies at GFP peaks, ignoring polarity. The spatial correlation is a scalar value computed as the Pearson correlation coefficient between all matched electrodes of 2 different topographies.[39] Third, the topographies at GFP peaks were assigned to the cluster with the highest spatial correlation. If the highest correlation was smaller than 0.5 in absolute value (ie, in the range of −0.5 to 0.5), the topography was not assigned to any cluster. This threshold represents a trade-off between rejecting too many and too few topographies during labelling and was chosen in line with previous studies[6,11,63] and the Cartool default settings. Fourth, the center of each cluster was computed, resulting in k new "average" cluster topographies. The new cluster topographies were then again correlated with the topographies at GFP peaks, closing the loop. The algorithm stopped when the variance of the clusters converged to a limit. To overcome the random selection of the initial cluster topographies, the clustering was repeated 100 times per set of k clusters and the set explaining most variance of the data was selected. The optimal number of clusters was identified for each individual separately according to a meta-criterion with 7 independent optimization criteria (for more details refer to Ref. 6). This procedure resulted in 4 to 8 topographies for each participant and condition.

In the second step, a second k-means clustering was performed at group level, clustering the concatenated individual topographies obtained in the previous step. For the second clustering, an initial number of k = 4 to 15 clusters and 200 k-means initializations were set. Again, the polarity was ignored, and a maximum absolute Pearson correlation coefficient higher than 0.5 was needed for cluster assignment. The same meta-criterion as before was used to identify the optimal number of clusters on a group level.

This 2-step clustering is a nondeterministic algorithm and can thus yield varying results when repeated. To assess the reliability of our findings, we repeated the entire procedure for the definition of group microstates 5 times for both the eyes-closed and the eyes-open conditions. The identified optimal numbers of group microstates were then compared across reruns. For the eyes-closed condition, the optimal number of group microstates

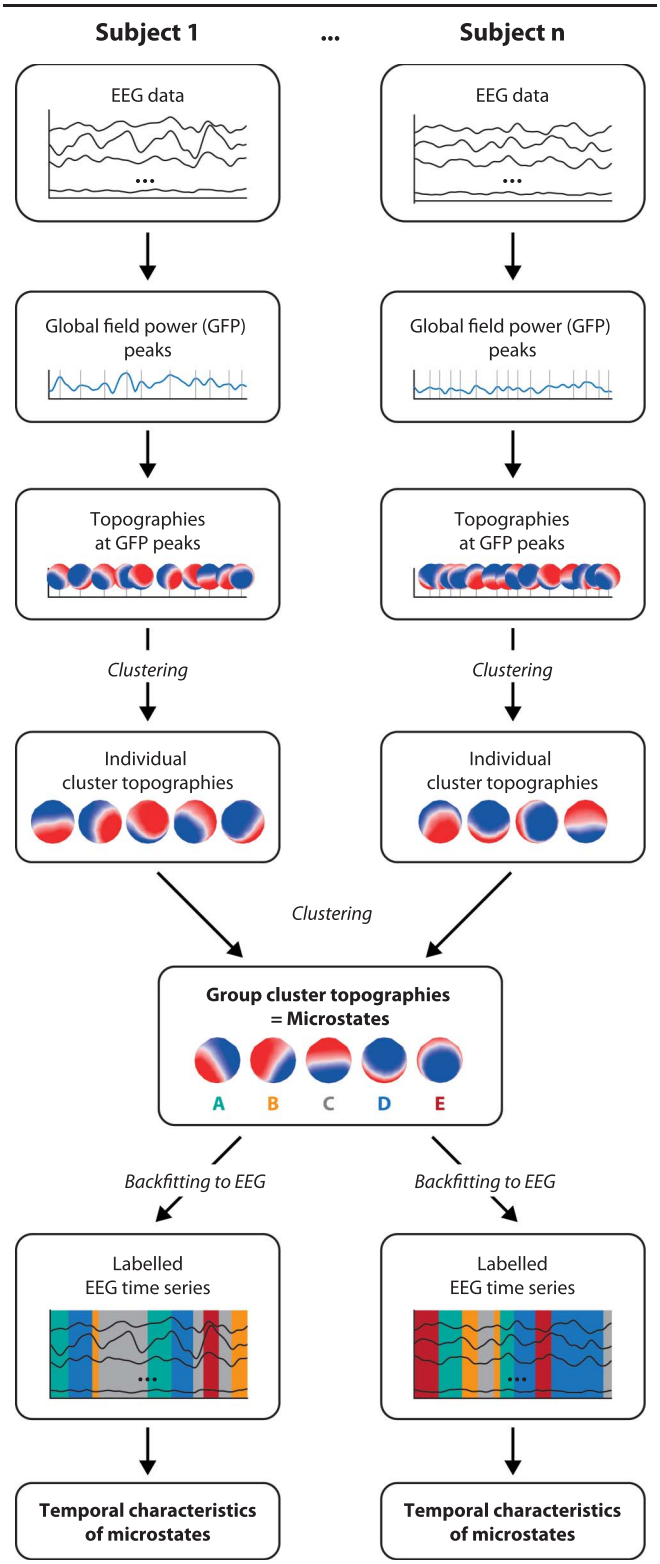

Figure 1. Microstate analysis. For each participant, the global field power (GFP) is calculated and topographies at GFP peaks are selected for individual clustering. Topographies at GFP peaks are clustered with a modified k-means clustering, leading to a variable number of individual cluster topographies per individual. Next, individual cluster topographies are concatenated and clustered on a group level. This consistently resulted in 5 different group cluster topographies for the eyes-closed condition, labelled as microstates A to E. Microstate topographies are then fitted back to the individual EEG data, resulting in a labelled EEG time series in which each time point is associated to a microstate. From the labelled EEG time series, the temporal characteristics of microstates are derived. This analysis was performed separately per group (patients with chronic pain and healthy controls). EEG, electroencephalography.

showed considerable stability (eyes-closed; number of microstates for 5 reruns for patients/controls: 5, 5, 5, 4, 5/5, 5, 5, 5, 5). For the eyes-open condition, by contrast, the optimal number of group microstates strongly varied across reruns, especially for patients, and could not reliably be estimated (eyes-open; number of microstates for 5 reruns for patients/controls: 5, 6, 4, 6, 5/5, 5, 5, 4, 5). See Supplementary Figure 1 for a depiction of group-level microstates for all reruns (available at http://links.lww.com/PAIN/B351). In light of this lack of reliability in the eyes-open condition, all further analyses were restricted to the eyes-closed condition (see below for a discussion of potential reasons for this discrepancy).

Group microstates of a representative rerun of the eyes-closed condition are shown in **Figure 2A**. Results of this rerun will be exemplarily shown throughout the article. To show the reliability of the findings, analyses of the other 4 eyes-closed reruns are also summarized in the article and their detailed results are shown in the supplementary material (available at http://links.lww.com/PAIN/B351). Topographies were visually inspected and compared with topographies reported in the literature. For both groups, the first 4 topographies closely resembled the 4 well-known "canonical" microstates A to D reported previously and were labeled accordingly.[25,27,34] The topography of the fifth microstate closely resembled a microstate that has been consistently reported in more recent studies,[6,11,63] since an increasing number of studies is now using a data-driven approach to define the optimal number of microstates. We labeled it with the letter E. Throughout the article, these 5 group-level topographies are referred to as microstates A to E. Similarities and differences of microstate topographies between groups were assessed by calculating spatial correlations and topographic analyses of variance (TANOVAs) for all microstates (A to E), respectively. Topographic analysis of variance is a nonparametric randomization test based on the global map dissimilarity of individual topographies.[39] The global map dissimilarity is a measure of the difference between 2 topographies directly related to the spatial correlation.[39] For each microstate, global map dissimilarity was computed between the microstate topographies of the patient and control groups using Cartool.[9] To obtain a *P*-value, this dissimilarity was compared with a distribution of dissimilarities, which was generated by randomly shuffling individual topographies between patient and controls and re-computing the dissimilarity between the center topographies of the randomized groups. The process was repeated 5000 times. This comparison resulted in a *P*-value per microstate, which was given by the proportion of permutations in which the dissimilarity was smaller than the dissimilarity originally observed in the data. Resulting *P*-values were corrected for multiple comparisons across the 5 microstate classes using the resampling-based false discovery rate (FDR).[62] Adjusted *P*-values are reported.

2.4.2. Temporal microstate characteristics

Next, we determined the temporal characteristics of the 5 microstates for both groups. To this end, individual EEG time series were construed as time series of microstates through a "fitting procedure," that is, a microstate was assigned to every time point. For each participant, the EEG topographies of all time points were spatially correlated to the microstate topographies of the participant's group (patients/controls) using absolute Pearson correlation coefficients. Next, each EEG time point was assigned to a microstate (A to E). To ensure a certain continuity in the microstate time series, the relabeling was performed based

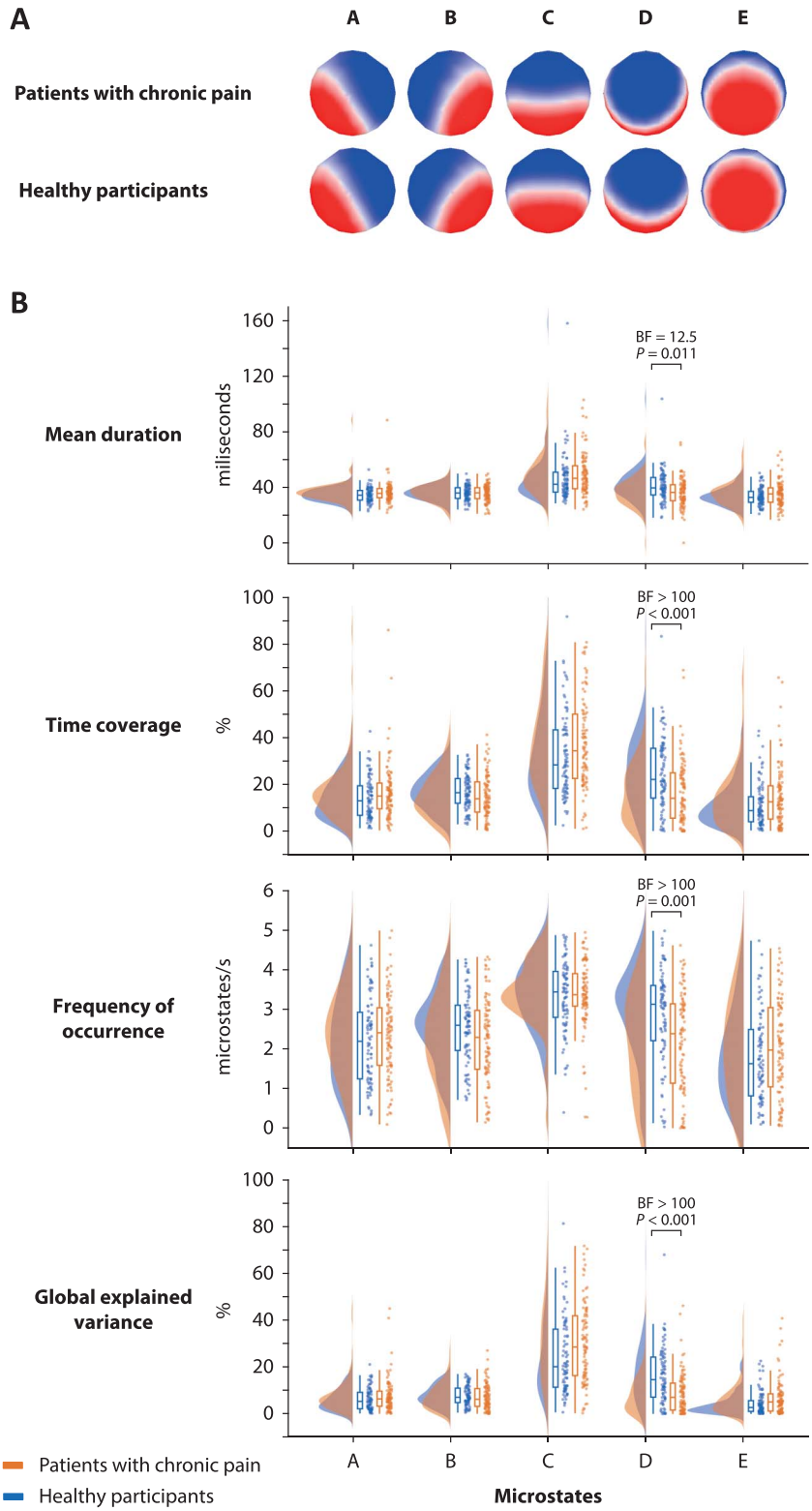

Figure 2. Microstate topographies and their temporal characteristics for all patients with chronic pain (n = 101) compared with healthy controls (n = 88) in the eyes-closed condition (representative rerun). (A) Microstate topographies were defined for the entire mixed chronic pain group and healthy controls separately. Microstates were labelled with the letters A to E according to previous literature.[34] (B) Temporal characteristics. Mean duration, time coverage, frequency of occurrence, and global explained variance of each microstate were calculated for each participant. Raincloud plots[1] show unmirrored violin plots displaying the probability density function of the data, boxplots, and individual data points. Boxplots depict the sample median as well as first (Q1) and third quartiles (Q3). Whiskers extend from Q1 to the smallest value within Q1 − 1.5 × interquartile range (IQR) and from Q3 to the largest values within Q3 + 1.5 × IQR. BF, Bayes factor in favor of the alternative hypothesis.

on 2 criteria: (1) the correlation should be high, and (2) most surrounding time points should belong to the same microstate.[9] To fulfill this compromise between goodness of fit and smoothness, standard temporal smoothing (window half size = 5 and strength (Besag factor) = 10) was applied.[42,55] No label was assigned if the highest (absolute) spatial correlation was smaller than 0.5. On average, the percentage of unlabeled time points was smaller than 0.1% and no differences existed between groups (representative rerun: $mean_{patients} = 0.080\%$, $mean_{controls} = 0.095\%$, $t = 0.630$, $P = 0.530$, $BF_{10} = 0.190$, median $\delta = 0.085$, 95% credible interval = [−0.190 to 0.362]; two-sided independent-samples *t* tests).

Based on the time series of microstates, 4 measures were calculated to quantify the temporal characteristics of each microstate: mean duration, time coverage, frequency of occurrence, and global explained variance. The mean duration is the average time (in milliseconds) for which a microstate persists before transitioning to a different microstate. The time coverage is the percentage of total time that a microstate is present. The frequency of occurrence is the number of times that a microstate recurs per second. The global explained variance is the percentage of global variance that is explained by every microstate.

Temporal characteristics of microstates were determined for all 5 reruns of the eyes-closed condition. As outlined above, the meta-criterion indicated a number of 5 microstates for both groups for all but 1 rerun, for which it indicated 4 optimal group-level microstates for patients. To enable a comparison of temporal characteristics between groups for this rerun, temporal characteristics were determined for the 5-microstate solution of the microstate analysis. Results of the representative rerun can be found in **Figure 2B**, results for all other reruns in Supplementary Table 1 (available at http://links.lww.com/PAIN/B351).

Finally, we investigated the microstate sequence by examining the transition probabilities from each microstate to the others for the representative rerun.[31,40,55] To this end, we computed the matrix of transition counts among all microstates for each participant and divided it by the overall count of transitions.

2.5. Statistical analysis

Group differences of temporal microstate measures (mean duration, time coverage, frequency of occurrence, and global explained variance) and transition probabilities were analyzed in JASP version 0.13.1[22] using 2-sided independent-samples *t* tests in both frequentist and Bayesian frameworks. For the frequentist approach, significance level was set to 0.05. For *P*-values of temporal measures, resampling-based FDR correction[62] was performed in MATLAB (MathWorks, Natick, MA) across the 5 microstates and the 4 different temporal measures, resulting in a correction for 20 statistical tests per rerun. Adjusted *P*-values[62] are reported throughout the article. For the transition matrix of the representative rerun, FDR correction was performed across all 20 transitions. For the Bayesian analysis, default priors (Cauchy distributions with a scale parameter $r = 0.707$) were used. In addition to *t*-values and FDR-adjusted *P*-values, results are reported using the two-tailed Bayes factor BF_{10}. Effect size estimates for the BF_{10} are reported as the median of the posterior δ distribution together with its 95% credibility interval.

Finally, we investigated relationships between temporal microstate measures and clinical parameters for the representative rerun using JASP version 0.13.1.[22] To this end, temporal microstate measures of microstate D (mean duration, time coverage, frequency of occurrence, and global explained variance) were selected for a correlation analysis, since they consistently showed significant differences between patients and controls across all reruns. Pearson correlations were calculated between the microstate measures and major clinical parameters that were available for all patients (current pain intensity, SF-MPQ total pain score, depression [BDI], and medication [MQS]). Please note that for most patients (n = 81), current pain intensity ratings were obtained from the painDETECT questionnaire, which uses a combination of a numerical rating scale anchored at 0 (no pain) and 10 (max pain) with a color gradient. Twenty patients with chronic widespread pain did not complete painDETECT questionnaires. For these patients, current pain intensity ratings were obtained from the SF-MPQ, which uses a visual analogue scale anchored at 0 (no pain) and 100 (worst imaginable). These ratings were divided by 10 to match rating scales across questionnaires. Correlations were again calculated in both frequentist and Bayesian frameworks. In the Bayesian analysis, default priors (stretched beta priors with width = 1) were used. Results are reported using the Pearson correlation coefficient, its FDR-adjusted *P*-value, its Bayes factor (BF_{10}), and the 95% credibility interval of the correlation coefficient. FDR correction of *P*-values[62] was performed in MATLAB (MathWorks, Natick, MA) across all 16 performed correlations.

2.6. Subgroup analyses

Finally, we investigated whether the results could be replicated within particular patient subgroups of our mixed sample. To this end, we repeated the definition of microstates and the investigation of their temporal characteristics in the eyes-closed condition for the 2 largest subgroups: patients with chronic back pain and patients with chronic widespread pain. These 2 groups were chosen based on the sample size. With a medium effect size Cohen's d of 0.5 and a type 1 error of 0.05, a post hoc power analysis revealed a power of 0.78 and 0.65 for 2-sided independent-samples *t* tests comparing our healthy control group (n = 88) with the chronic back pain (n = 47) and chronic widespread pain (n = 30) patient groups, respectively. For patients with neuropathic pain (n = 18) and joint pain (n = 6), power was even lower (0.48 and 0.29, respectively). Thus, microstate analysis was repeated for patients with chronic back pain and chronic widespread pain only. Individual topographies of each patient subgroup were selected for a new second clustering to obtain subgroup-specific microstate topographies. Subgroup-specific temporal characteristics were obtained through back-fitting and were statistically compared with the temporal characteristics of the healthy control group for each patient group as described above. Again, 5 reruns of these subgroup analyses were performed to assess the reliability of findings. For each subgroup, a representative rerun is shown in the article, and details of additional reruns are presented in the supplementary material (available at http://links.lww.com/PAIN/B351).

2.7. Data and code availability

Electroencephalography data in BIDS format[43] as well as scripts for statistical analyses are openly available at https://osf.io/srpbg/.

3. Results

The current study investigated whether the dynamics of resting-state brain activity are altered in patients suffering from chronic pain. We performed microstate analysis, which describes the

Table 2

Comparisons of temporal microstate measures for all patients with chronic pain (n = 101) compared with healthy controls (n = 88) in the eyes-closed condition (representative rerun).

Microstate	Measure	t	*P*	BF_{10}	Median effect size (δ)	95% CI
A	Mean dur.	−1.947	0.132	0.919	−0.265	−0.547 to 0.012
	Time cov.	−1.586	0.167	0.510	−0.216	−0.496 to 0.061
	Freq. of occ.	−0.809	0.471	*0.215*	−0.110	−0.387 to 0.166
	GEV	−1.635	0.167	0.548	−0.222	−0.503 to 0.055
B	Mean dur.	0.491	0.656	*0.177*	0.067	−0.209 to 0.343
	Time cov.	1.583	0.167	0.507	0.215	−0.062 to 0.495
	Freq. of occ.	2.316	0.072	1.892	0.317	0.037 to 0.600
	GEV	0.801	0.471	*0.214*	0.109	−0.167 to 0.386
C	Mean dur.	−1.356	0.230	0.373	−0.184	−0.463 to 0.092
	Time cov.	−1.777	0.154	0.686	−0.242	−0.523 to 0.035
	Freq. of occ.	−0.285	0.775	*0.164*	−0.039	−0.315 to 0.237
	GEV	−2.441	0.062	2.482	−0.334	−0.618 to 0.054
D	Mean dur.	3.087	**0.011**	**12.530**	0.425	0.142 to 0.711
	Time cov.	4.010	**<0.001**	**>100**	0.556	0.268 to 0.847
	Freq. of occ.	3.803	**0.001**	**>100**	0.526	0.240 to 0.816
	GEV	5.220	**<0.001**	**>100**	0.730	0.436 to 1.027
E	Mean dur.	−1.575	0.167	0.501	−0.214	−0.494 to 0.063
	Time cov.	−1.856	0.144	0.783	−0.253	−0.534 to 0.025
	Freq. of occ.	−1.333	0.230	0.362	−0.181	−0.460 to 0.095
	GEV	−2.239	0.075	1.611	−0.306	−0.589 to −0.027

Results of 2-sided independent-samples *t*-tests (frequentist and Bayesian approach) comparing the entire mixed chronic pain group with the healthy control group. *P*-values are FDR-adjusted. Mentioned in bold $P < 0.05$ and $BF_{10} > 3$, indicating at least moderate evidence for the alternative hypothesis, and in italics $BF_{10} < 1/3$, indicating at least moderate evidence for the null hypothesis. Median effect sizes (δ) and their respective 95% credible interval (CI) are reported.

BF_{10}, Bayes factor in favor of the alternative hypothesis; Freq. of occ., frequency of occurrence; GEV, global explained variance; Mean dur., mean duration; Time cov., time coverage.

time course of EEG activity as a sequence of a limited number of short stable topographies termed microstates. We applied microstate analysis to resting-state EEG activity and compared temporal characteristics of microstates between a large cohort of patients suffering from chronic pain and age- and sex-matched healthy control participants.

3.1. Definition of microstates A to E in patients and controls

We identified microstates using a standard two-step k-means clustering procedure.[42] Five repetitions of the entire analysis showed reliable microstate definitions for the eyes-closed condition but very variable microstates for the eyes-open condition (Supplementary Figure 1, available at http://links.lww.com/PAIN/B351). Thus, all further analyses were restricted to the eyes-closed condition. For this condition, the clustering procedure consistently revealed 5 different microstates in both groups in 4 of 5 reruns. In a single rerun, only 4 microstates were found to be optimal for patients. Throughout the article, results and further analyses of a representative rerun with 5 microstates for both groups are presented. Results of additional reruns are presented in the supplementary material (available at http://links.lww.com/PAIN/B351).

In accordance with previous studies,[6,8,11,26,27,34,35,63] the microstates were labeled as microstates A to E for both groups. Topographies for the representative rerun are depicted in **Figure 2A**. Together, these 5 microstates explained 81.05% and 81.08% of the variance across individuals for patients with chronic pain and healthy controls, respectively, which is in good accordance with previous studies.[12,34,49] The high similarity of topographies between groups was confirmed by high spatial correlations (microstate A: r = 0.99, B: r = 1.00, C: r = 0.96, D: 0.93, and E: r = 0.92). In addition, TANOVAs revealed subtle group differences between topographies hardly visible to the naked eye for microstates C to E (microstate A: $P = 0.534$, B: $P < 0.091$, C: $P < 0.001$, D: $P < 0.001$, and E: $P < 0.001$).

Taken together, the clustering procedures for both groups resulted in 5 microstate topographies, which were largely similar between groups.

3.2. Temporal characteristics of microstates in patients and controls

To investigate the dynamics of brain activity, we next analyzed whether the temporal characteristics of microstates differed between patients and healthy controls. To this end, 5 microstates were backfitted to the individual EEG time series by correlating microstate topographies with the EEG topographies at every time point for all reruns of the eyes-closed condition. This allowed to assign each time point to a microstate and, thus, to construe the EEG time series as time series of microstates.

We specifically calculated the mean duration, time coverage, frequency of occurrence, and global explained variance of each microstate. This was performed for each patient and each healthy control participant. We next compared these temporal microstate characteristics between groups for each rerun. Results are depicted in **Figure 2B** and **Table 2** for the representative rerun. Results for all other reruns can be found in Supplementary Table 1 (available at http://links.lww.com/PAIN/B351). Across reruns, microstate analysis consistently revealed strong evidence for changes in microstate D characteristics in patients compared with healthy participants. We found strong to very strong evidence for a lower time coverage, a lower frequency of occurrence, and lower global explained variance of microstate D in patients compared with controls in all 5 reruns (**Table 2**, Supplementary Table 1; all $BF_{10} > 10$, all FDR-adjusted *P*-values < 0.011). In addition, results showed moderate to very strong

Table 3

Comparisons of transition probabilities between microstates for all patients with chronic pain (n = 101) compared with healthy controls (n = 88) in the eyes-closed condition (representative rerun).

	Mean trans. prob. patients with chronic pain	Mean trans. prob. HC	t	*P*	BF_{10}	Median effect size (δ)	95% CI
From A to B	0.229	0.251	1.335	0.305	0.363	0.181	−0.095 to 0.460
From A to C	0.394	0.340	−2.273	0.071	1.727	−0.311	−0.594 to 0.031
From A to D	0.190	0.258	3.501	**0.004**	**42.514**	0.483	0.198 to 0.772
From A to E	0.183	0.147	−2.269	0.071	1.713	−0.310	−0.593 to −0.031
From B to A	0.244	0.218	−1.524	0.238	0.466	−0.207	−0.487 to 0.069
From B to C	0.385	0.360	−1.080	0.414	*0.273*	−0.146	−0.425 to 0.129
From B to D	0.187	0.273	4.548	**<0.001**	**>100**	0.633	0.342 to 0.927
From B to E	0.181	0.146	−2.046	0.105	1.102	−0.279	−0.561 to −0.001
From C to A	0.265	0.216	−2.470	0.071	2.647	−0.338	−0.622 to −0.058
From C to B	0.244	0.264	1.178	0.375	*0.302*	0.160	−0.116 to 0.438
From C to D	0.277	0.339	2.391	0.071	2.223	0.327	0.047 to 0.611
From C to E	0.211	0.178	−1.857	0.147	0.785	−0.253	−0.534 to 0.025
From D to A	0.192	0.191	−0.052	1	0.159	−0.007	−0.283 to 0.268
From D to B	0.177	0.230	3.673	**0.003**	**73.594**	0.508	0.222 to 0.797
From D to C	0.412	0.400	−0.498	0.860	*0.178*	−0.067	−0.344 to 0.208
From D to E	0.205	0.175	−1.513	0.238	0.459	−0.206	−0.485 to 0.071
From E to A	0.206	0.177	−2.250	0.071	1.647	−0.307	−0.590 to −0.028
From E to B	0.195	0.198	0.250	1	*0.163*	0.034	−0.242 to 0.310
From E to C	0.374	0.342	−1.508	0.238	0.456	−0.205	−0.485 to 0.072
From E to D	0.222	0.280	2.709	**0.046**	4.648	0.371	0.090 to 0.656

Results of 2-sided independent-samples *t* tests (frequentist and Bayesian approach) comparing the entire mixed chronic pain group with the healthy control group. *P* values are FDR-adjusted. Mentioned in bold $P < 0.05$ and $BF_{10} > 3$, indicating at least moderate evidence for the alternative hypothesis, and in italics $BF_{10} < 1/3$, indicating at least moderate evidence for the null hypothesis. Median effect sizes (δ) and their respective 95% credible interval (CI) are reported.

BF_{10}, Bayes factor in favor of the alternative hypothesis; HC, healthy controls; Mean trans. prob., mean transition probability.

evidence for a shorter mean duration of microstate D in patients in 4 of 5 reruns (**Table 2**; Supplementary Table 1, available at http://links.lww.com/PAIN/B351; $BF_{10} > 7$, FDR-adjusted *P*-values < 0.016). In the fifth rerun, evidence was inconclusive (Supplementary Table 1, available at http://links.lww.com/PAIN/B351; rerun 4: $BF_{10} = 2.211$, FDR-adjusted *P*-value = 0.089). Regarding the other microstates, a single rerun showed evidence for changes in the global explained variance of microstate C, while 2 reruns showed evidence for alterations in microstate E characteristics (Supplementary Table 1 for details, available at http://links.lww.com/PAIN/B351). However, these changes could not be replicated consistently.

For the representative rerun, we further investigated whether the sequences of microstates differed between groups. To this end, transition probabilities from each microstate to all other microstates were calculated. Mean transition probabilities for both groups as well as statistical results are presented in **Table 3**. In line with the analysis of temporal characteristics, we found moderate to very strong evidence for a lower transition probability from microstates A, B, and E to microstate D in patients compared with controls (**Table 3**; A to D: $P = 0.004$, $BF_{10} = 43$; B to D: $P < 0.001$, $BF_{10} > 100$; and E to D: $P = 0.046$, $BF_{10} = 5$; FDR-adjusted *P*-values). In addition, healthy controls were more likely to transition from microstate D to microstate B than patients (**Table 3**; $P = 0.003$, $BF_{10} = 74$; FDR-adjusted *P*-value). Evidence for differences of all other transition probabilities was either inconclusive or against a group difference (see **Table 3** for details).

In summary, the analysis of the temporal dynamics of microstates revealed consistent evidence for a less predominant role of microstate D in eyes-closed resting-state brain activity of patients with chronic pain. Evidence for changes of the temporal characteristics of microstates other than D was inconsistent and could not be replicated across reruns of the analysis.

3.3. Relationships between temporal microstate characteristics and clinical characteristics

Having observed consistent evidence for changes of microstate D temporal characteristics in patients with chronic pain, we explored whether microstate D temporal measures were significantly related to clinical characteristics. To this end, we performed correlation analyses between the temporal measures of microstate D and clinical parameters of the patients. We specifically related microstate D characteristics obtained in the representative rerun to the current pain intensity, the SF-MPQ total pain score, as well as measures of depression (BDI) and medication (MQS). Frequentist statistics did not reveal significant correlations (**Table 4**; all $P > 0.05$, FDR-adjusted). This was confirmed using Bayesian statistics, which consistently provided moderate evidence for an absence of relations (**Table 4**; $1/10 < BF_{10} < 1/3$). Thus, the results did not provide evidence for relationships between microstate D temporal measures and clinical characteristics.

Table 4

Relationships between microstate D temporal measures and clinical parameters for all patients with chronic pain (n = 101) in the eyes-closed condition (representative rerun).

	Current pain	SF-MPQ	BDI	MQS
Mean dur.				
Pearson's r	0.031	0.047	−0.020	0.127
P	0.927	0.927	0.956	0.927
BF_{10}	*0.131*	*0.139*	*0.128*	*0.276*
95% CI	−0.163 to 0.224	−0.148 to 0.238	−0.213 to 0.174	−0.069 to 0.311
Time cov.				
Pearson's r	0.036	−0.035	−0.002	0.058
P	0.927	0.927	0.984	0.927
BF_{10}	*0.133*	*0.133*	*0.125*	*0.147*
95% CI	−0.159 to 0.228	−0.227 to 0.160	−0.196 to 0.192	−0.136 to 0.248
Freq. of occ.				
Pearson's r	0.067	−0.072	0.084	−0.073
P	0.927	0.927	0.927	0.927
BF_{10}	*0.156*	*0.161*	*0.176*	*0.162*
95% CI	−0.129 to 0.257	−0.261 to 0.124	−0.112 to 0.273	−0.261 to 0.122
GEV				
Pearson's r	0.046	−0.066	−0.011	0.059
P	0.927	0.927	0.927	0.927
BF_{10}	*0.139*	*0.155*	*0.126*	*0.147*
95% CI	−0.149 to 0.237	−0.256 to 0.130	−0.204 to 0.183	−0.136 to 0.248

Pearson correlations (frequentist and Bayesian approach) were performed for microstate D temporal measures that had consistently shown evidence for differences between patients and controls across reruns of previous analyses. *P*-values are FDR-adjusted. Mentioned in italics $BF_{10} < 1/3$, indicating at least moderate evidence for the null hypothesis.
95% CI, 95% credible interval; BF_{10}, Bayes factor in favor of the alternative hypothesis; BDI, Beck Depression Inventory; Freq. of occ., frequency of occurrence; GEV, global explained variance; Mean dur., mean duration; MQS, medication quantification scale; SF-MPQ, short-form McGill Pain Questionnaire; Time cov., time coverage.

3.4. Subgroup analyses

Since the patient group comprised patients with different types of chronic pain, we finally investigated whether our main finding of a less predominant role of microstate D in chronic pain could be replicated in different subgroups of patients. To this end, we repeated the definition of microstates and the investigation of their temporal characteristics in the eyes-closed condition for the 2 largest subgroups of patients, that is, patients suffering from chronic back pain and chronic widespread pain. For patients with chronic back pain, 5 reruns consistently revealed 5 microstates. For patients with chronic widespread pain, 4 of 5 reruns also revealed 5 microstates, while a single rerun revealed 6 microstates. Results of a representative rerun for each subgroup can be found in **Figures 3 and 4** and **Tables 5 and 6**. Results of additional reruns are provided in Supplementary Tables 2 and 3 (available at http://links.lww.com/PAIN/B351). To enable comparisons of temporal characteristics, further analyses were performed on the 5-microstate solutions of all reruns.

For patients with chronic back pain (representative rerun shown in **Figure 3**), all reruns consistently showed moderate to very strong evidence for a lower time coverage, frequency of occurrence, and global explained variance of microstate D in patients compared with controls (**Table 5**, Supplementary Table 2, available at http://links.lww.com/PAIN/B351; all $BF_{10} > 6$, all FDR-adjusted *P*-values < 0.037). In addition, all but 1 reruns showed at least moderate evidence for a shorter mean duration of microstate D in patients (**Table 5**, Supplementary Table 2, available at http://links.lww.com/PAIN/B351; rerun 1: $BF_{10} = 0.308$, FDR-adjusted *P*-value = 0.517; all other reruns: $BF_{10} > 3$, all FDR-adjusted *P*-values < 0.038). With respect to the other microstates, some reruns additionally showed alterations in isolated measures of other microstates (see Supplementary Table 2 for details, available at http://links.lww.com/PAIN/B351), which were, however, not consistently replicated.

For patients with chronic widespread pain (representative rerun shown in **Figure 4**), Bayesian statistics showed moderate evidence against a group difference for most microstates and temporal characteristics (**Table 6**, Supplementary Table 3, available at http://links.lww.com/PAIN/B351). The only significant finding was an increased global explained variance of microstate B in patients compared with controls in a single rerun (Supplementary Table 3, rerun 4, available at http://links.lww.com/PAIN/B351), which was, however, not replicated across reruns.

Thus, although the small size of the subgroups has to be considered, our main finding of a less predominant role of microstate D could convincingly be replicated in patients with chronic back pain. By contrast, patients with chronic widespread pain did not show reliable microstate alterations.

4. Discussion

In this study, we investigated the dynamics of brain function in patients suffering from different types of chronic pain. To this end, we performed microstate analysis of resting-state EEG recordings in a large cohort of patients and age- and sex-matched healthy control participants. In both groups, resting-state brain activity could consistently be described as sequences of 5 microstates labeled A to E in the eyes-closed condition. However, a varying number of microstates were obtained in the eyes-open condition. Analyses of the temporal characteristics of these microstates in the eyes-closed condition revealed a decreased presence of microstate D in patients as compared to healthy participants. No consistent evidence for differences in other microstates was found. When investigating specific chronic pain pathologies, these findings were replicated for patients with chronic back pain. By contrast, patients with chronic widespread pain did not present microstate alterations. Thus, the present

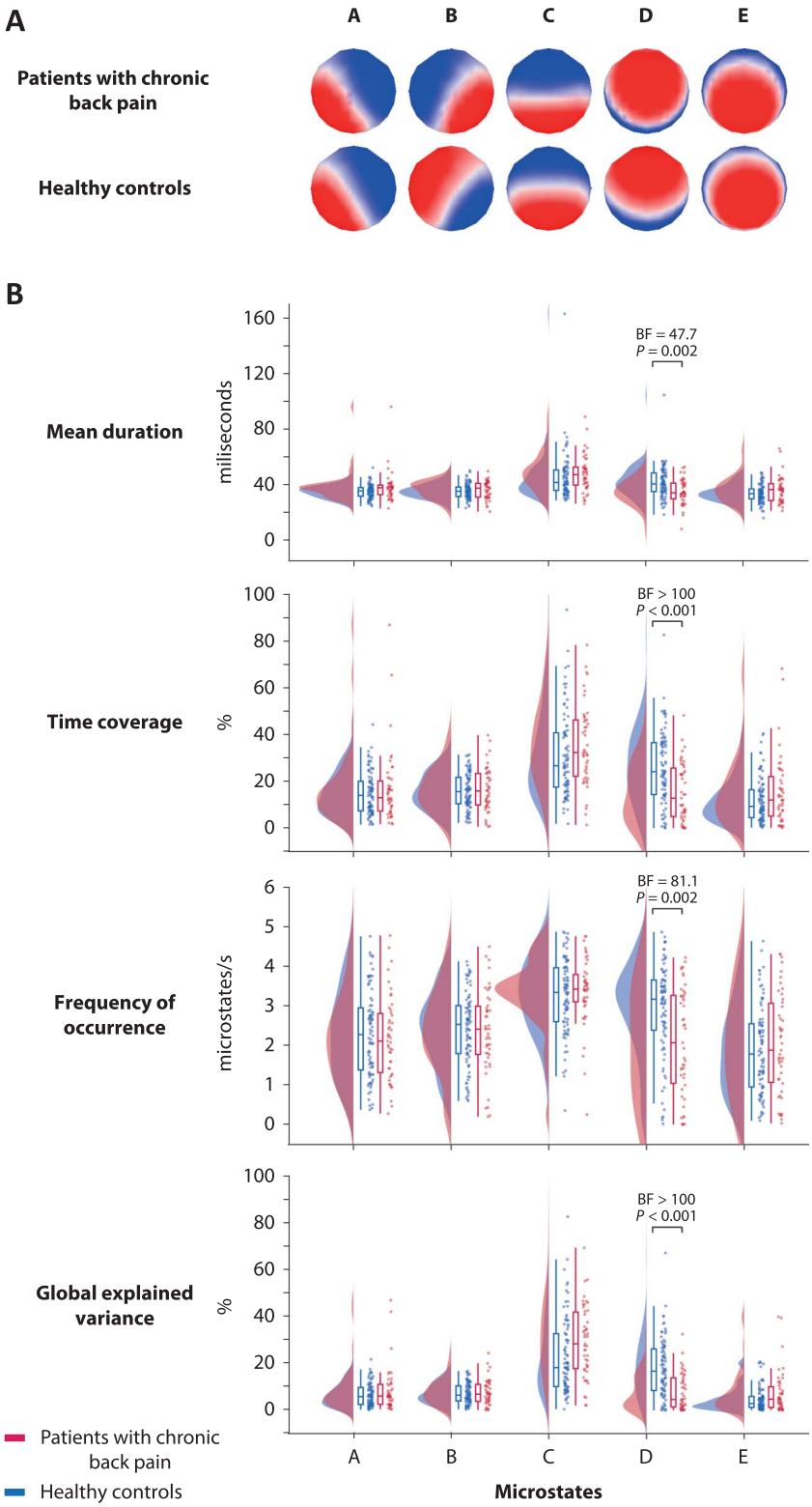

Figure 3. Microstate topographies and their temporal characteristics for patients with chronic back pain (n = 47) compared with healthy controls (n = 88) in the eyes-closed condition (representative rerun). (A) Microstate topographies were defined for the patients with chronic back pain and healthy controls separately. Microstates were labelled with the letters A to E according to previous literature.[34] (B) Temporal characteristics. Mean duration, time coverage, frequency of occurrence, and global explained variance of each microstate were calculated for each participant. Raincloud plots[1] show unmirrored violin plots displaying the probability density function of the data, boxplots, and individual data points. Boxplots depict the sample median as well as first (Q1) and third quartiles (Q3). Whiskers extend from Q1 to the smallest value within Q1 − 1.5 × interquartile range (IQR) and from Q3 to the largest values within Q3 + 1.5 × IQR. BF, Bayes factor in favor of the alternative hypothesis.

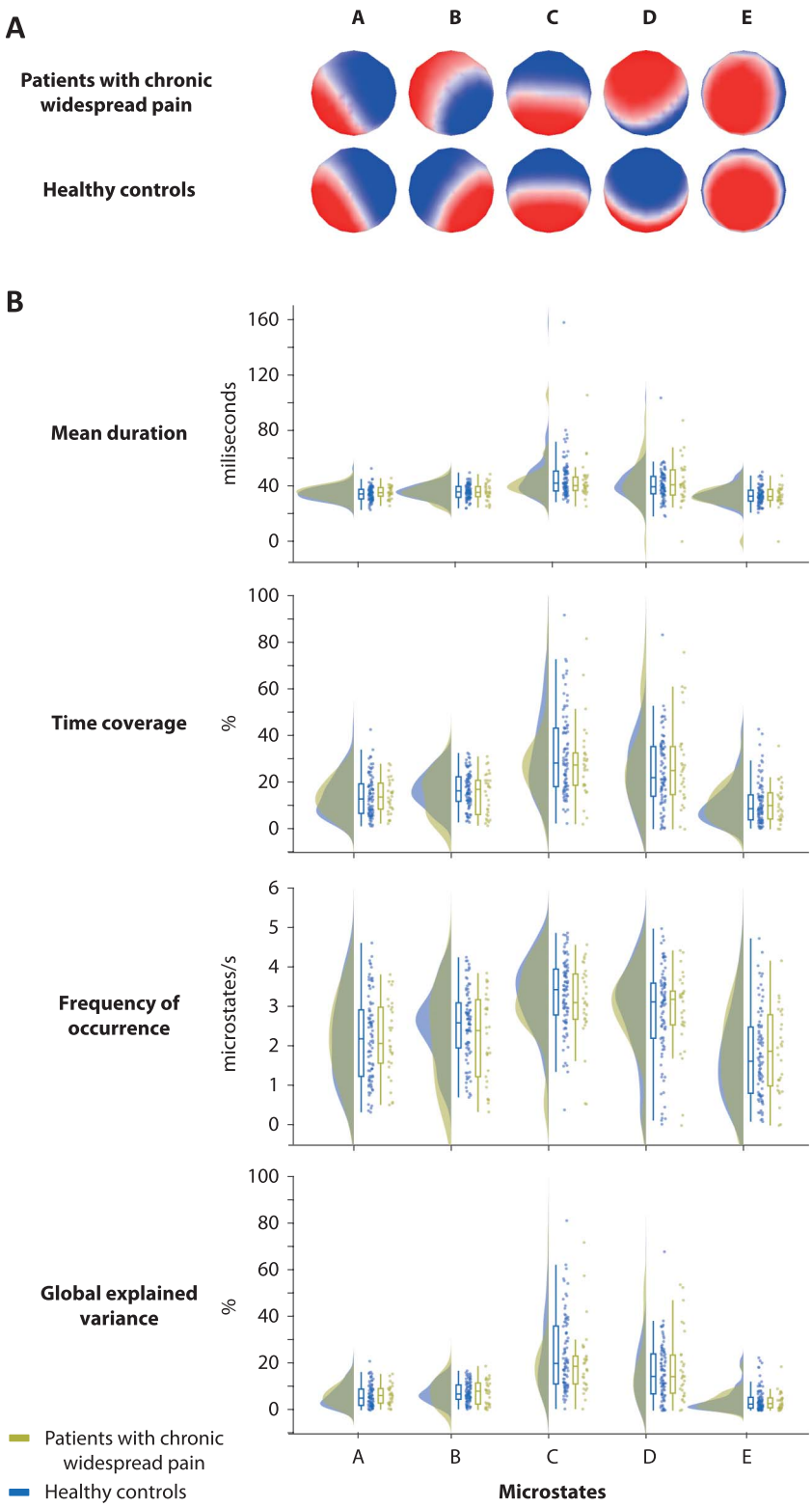

Figure 4. Microstate topographies and their temporal characteristics for patients with chronic widespread pain (n = 30) compared with healthy controls (n = 88) in the eyes-closed condition (representative rerun). (A) Microstate topographies were defined for the patients with chronic widespread pain and healthy controls separately. Microstates were labelled with the letters A to E according to previous literature.[34] (B) Temporal characteristics. Mean duration, time coverage, frequency of occurrence, and global explained variance of each microstate were calculated for each participant. Raincloud plots[1] show unmirrored violin plots displaying the probability density function of the data, boxplots, and individual data points. Boxplots depict the sample median as well as first (Q1) and third quartiles (Q3). Whiskers extend from Q1 to the smallest value within Q1 − 1.5 × interquartile range (IQR) and from Q3 to the largest values within Q3 + 1.5 × IQR.

Table 5

Comparisons of temporal microstate measures for patients with chronic back pain compared (n = 47) with healthy controls (n = 88) in the eyes-closed condition (representative rerun).

Microstate	Measure	t	*P*	BF_{10}	Median effect size (δ)	95% CI
A	Mean dur.	2.234	0.093	1.802	0.368	0.026 to 0.719
	Time cov.	0.957	0.523	*0.291*	0.156	−0.180 to 0.496
	Freq. of occ.	−0.617	0.631	*0.229*	−0.100	−0.439 to 0.235
	GEV	1.524	0.259	0.549	0.249	−0.089 to 0.594
B	Mean dur.	0.868	0.552	*0.271*	0.141	−0.195 to 0.481
	Time cov.	0.431	0.702	*0.209*	0.070	−0.265 to 0.408
	Freq. of occ.	−0.274	0.784	*0.199*	−0.044	−0.382 to 0.291
	GEV	0.572	0.631	*0.223*	0.093	−0.242 to 0.431
C	Mean dur.	0.990	0.523	*0.300*	0.161	−0.175 to 0.502
	Time cov.	1.430	0.282	0.484	0.233	−0.104 to 0.578
	Freq. of occ.	0.664	0.631	*0.235*	0.108	−0.288 to 0.477
	GEV	2.219	0.093	1.749	0.366	0.023 to 0.717
D	Mean dur.	−3.540	**0.002**	**47.788**	−0.595	−0.956 to −0.240
	Time cov.	−4.114	**<0.001**	**>100**	−0.697	−1.063 to −0.336
	Freq. of occ.	−3.711	**0.002**	**81.160**	−0.625	−0.988 to −0.268
	GEV	−4.962	**<0.001**	**>100**	−0.849	−1.22 to −0.480
E	Mean dur.	2.000	0.118	1.160	0.329	−0.012 to 0.678
	Time cov.	1.672	0.215	0.678	0.274	−0.065 to 0.620
	Freq. of occ.	0.789	0.575	*0.255*	0.128	−0.207 to 0.468
	GEV	2.154	0.094	1.542	0.355	0.013 to 0.705

Results of 2-sided independent-samples *t* tests (frequentist and Bayesian approach) comparing patients with chronic back pain with the healthy control group. *P* values are FDR-adjusted. Mentioned in bold $P < 0.05$ and $BF_{10} > 3$, indicating at least moderate evidence for the alternative hypothesis, and in italics $BF_{10} < 1/3$, indicating at least moderate evidence for the null hypothesis. Median effect sizes (δ) and their respective 95% credible interval (CI) are reported.

BF_{10}, Bayes factor in favor of the alternative hypothesis; Freq. of occ., frequency of occurrence; GEV, global explained variance; Mean dur., mean duration; Time cov., time coverage.

findings describe microstate D–specific changes of the dynamics of brain function in eyes-closed resting-state EEG recordings of patients suffering from chronic pain. Beyond, they indicate that alterations of brain dynamics as measured by microstate analysis might be specific for certain types of chronic pain.

Our analyses reveal shorter mean duration, lower time coverage, fewer occurrences, and less explained variance of microstate D in patients compared with controls but no consistent alterations in other microstates. This pattern of results was found both when investigating the entire mixed chronic pain group or when specifically focusing on

Table 6

Comparisons of temporal microstate measures for patients with chronic widespread pain (n = 30) compared to healthy controls (n = 88) in the eyes-closed condition (representative rerun).

Microstate	Measure	t	*P*	BF_{10}	Median effect size (δ)	95% CI
A	Mean dur.	0.871	0.954	*0.309*	0.160	−0.226 to 0.555
	Time cov.	−0.057	0.954	*0.222*	−0.010	−0.398 to 0.377
	Freq. of occ.	−0.086	0.954	*0.222*	−0.016	−0.404 to 0.371
	GEV	0.284	0.954	*0.229*	0.052	−0.334 to 0.441
B	Mean dur.	−0.293	0.954	*0.230*	−0.054	−0.443 to 0.333
	Time cov.	−1.173	0.954	0.405	−0.217	−0.615 to 0.171
	Freq. of occ.	−1.575	0.954	0.655	−0.293	-0.696 to 0.097
	GEV	−0.123	0.954	*0.223*	−0.022	−0.411 to 0.364
C	Mean dur.	−0.654	0.954	*0.267*	−0.120	−0.513 to 0.266
	Time cov.	−0.958	0.954	0.331	−0.177	−0.572 to 0.210
	Freq. of occ.	−1.066	0.954	0.365	−0.197	−0.594 to 0.190
	GEV	−0.982	0.954	0.338	−0.181	−0.577 to 0.206
D	Mean dur.	0.654	0.954	*0.267*	0.120	−0.266 to 0.513
	Time cov.	0.839	0.954	*0.302*	0.154	−0.232 to 0.549
	Freq. of occ.	0.114	0.954	*0.223*	0.021	−0.366 to 0.409
	GEV	0.717	0.954	*0.277*	0.132	−0.254 to 0.525
E	Mean dur.	−0.637	0.954	*0.264*	−0.117	−0.509 to 0.269
	Time cov.	−0.065	0.954	*0.222*	−0.012	−0.400 to 0.375
	Freq. of occ.	0.211	0.954	*0.226*	0.039	−0.348 to 0.427
	GEV	−0.454	0.954	*0.242*	−0.083	−0.474 to 0.303

Results of 2-sided independent-samples *t* tests (frequentist and Bayesian approach) comparing patients with chronic widespread pain with the healthy control group. *P* values are FDR-adjusted. Mentioned in italics $BF_{10} < 1/3$, indicating at least moderate evidence for the null hypothesis. Median effect sizes (δ) and their respective 95% credible interval (CI) are reported.

BF_{10}, Bayes factor in favor of the alternative hypothesis; Freq. of occ., frequency of occurrence; GEV, global explained variance; Mean dur., mean duration; Time cov., time coverage.

the chronic back pain group. By contrast, no microstate alterations were found in patients suffering from chronic widespread pain. These findings are in contrast to the only study that applied microstate analysis to resting-state EEG recordings of patients suffering from chronic pain so far.[20] That study found a lower occurrence and time coverage of microstate C in eyes-open resting-state recordings of 43 patients suffering from chronic widespread pain. This difference between studies might at least in part be due to methodological differences. For example, the current study defined microstates separately for patients and controls, while the previous study defined microstates for both groups together. In addition, our sample size was smaller (n = 30 vs n = 47), and therefore, small effects might have been missed. Together, the 2 studies prompt further studies in larger groups of patients, ideally from different recording sites, to resolve these differences and to further clarify changes of microstates common to and different in distinct chronic pain populations.

Our observations further complement recent fMRI studies that have shown changes of the dynamics of brain function in chronic pain at ultra-low frequencies below 0.1 Hz.[3,5,10] They extend this evidence by showing alterations of the dynamics of brain function at frequencies higher than 1 Hz, in line with the dynamic pain connectome concept.[28,45]

Microstate analysis is an emerging tool for investigating the dynamics of brain activity. Although the functional interpretation of microstates is not fully clear yet, microstate analysis has been increasingly used to identify changes of brain dynamics in various neuropsychiatric diseases[25,34] that have furthered the understanding of the pathology of these disorders. Beyond, alterations of microstates characteristics might be useful as clinical biomarkers. For instance, a recent study has identified the dynamics of microstates C and D as a promising candidate endophenotype for schizophrenia.[12] However, our data did not provide evidence for a correlation between alterations of microstate D and clinical characteristics. As the brain processes discriminating patients with chronic pain from healthy people differ from those encoding momentary pain intensity,[3,32,54,59,64] the observed changes might reflect the abnormal disease state per se rather than its specific characteristics. Beyond, our analyses showed that results of our standard microstate analysis varied remarkably across repeated runs. This was in particular the case for patient data from the eyes-open condition, for which no stable optimal number of microstates could be obtained. This instability is likely due to higher variance and a stronger contamination of data by artifacts in eyes-open compared with eyes-closed resting-state recordings. Thus, future studies using microstate analyses should explicitly confirm reliability of findings.

Our most consistent and replicable finding was a reduced presence of microstate D in chronic pain. Microstate D has been related to attentional brain networks and functions (for reviews, see Refs. 34,51). In particular, microstate D has been associated with brain activity in frontoparietal regions,[8,11] the dorsal attentional control network,[49] and focus-switching and attentional reorientation.[36] Interestingly, deficits of cognitive function and particularly of attentional switching have been extensively reported in patients suffering from chronic pain.[37] A common hypothesis is that pain competes with other stimuli for limited cognitive resources, thereby "demanding attention" and potentially impairing higher-order attentional control mechanisms.[17,30,57] Thus, a decreased presence of microstate D might represent a neurophysiological correlate of altered attentional functioning in chronic pain. However, as we have not obtained direct measures of attentional functioning, we cannot directly test this hypothesis in this study. Future microstate studies on chronic pain might therefore include tasks and/or questionnaires assessing attentional functions.

Several limitations of the current study need to be discussed. First, the specificity of the decreased presence of microstate D for chronic pain is unclear. In particular, studies in patients suffering from schizophrenia[12,47] and major depressive disorder[38] also showed a decreased presence of microstate D. However, investigating symptom- and disease-specificity of these findings is challenging. Substantial progress in this endeavor requires large samples of patients suffering from different neuropsychiatric symptoms and diseases, standardized assessments, and, ideally, sharing of data acquired at different sites. As a first step in that direction, we share data and code of this study in a standardized format with the research community. Second, comparisons of microstate topographies showed slight but statistically significant differences between patients and healthy controls. However, considering these subtle differences together with the overwhelming similarity of microstate topographies, the microstates of both groups likely capture the same underlying neural networks. Third, the causal relationship between altered microstate dynamics and chronic pain is unclear. First studies have shown that the dynamics of microstates can be changed by neurofeedback[15] and noninvasive brain stimulation.[53] These approaches might thus be useful to prove the causal link between changes in microstate dynamics and neuropsychiatric disorders including chronic pain. Moreover, they highlight the potential utility of microstate dynamics as targets for neurofeedback- and/or brain stimulation–based treatments of chronic pain.

In conclusion, our findings provide evidence for altered and potentially pathology-specific dynamics of brain function in a large cohort of patients with chronic pain using EEG microstate analysis. We particularly observed alterations of microstate D. As this microstate has been associated with attentional brain networks and functions, changes of microstate D might relate to dysfunctional attentional processes in chronic pain. These results add to the understanding of the pathophysiology of chronic pain and indicate the need for future large-scale studies including patients suffering from chronic pain of different types.

Conflict of interest statement

The authors have no conflicts of interest to declare.

Acknowledgments

Supported by the Deutsche Forschungsgemeinschaft (PL 321/10-2, PL321/11-2, and PL321/13-1).
E.S. May and C.Gil Ávila contributed to this work equally.

Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/B351.

Article history:
Received 2 October 2020
Received in revised form 10 March 2021
Accepted 17 March 2021
Available online 8 April 2021

References

[1] Allen M, Poggiali D, Whitaker K, Marshall TR, Kievit RA. Raincloud plots: a multi-platform tool for robust data visualization. Wellcome Open Res 2019;4:63.
[2] Baliki MN, Apkarian AV. Nociception, pain, negative moods, and behavior selection. Neuron 2015;87:474–91.

[3] Baliki MN, Geha PY, Apkarian AV, Chialvo DR. Beyond feeling: chronic pain hurts the brain, disrupting the default-mode network dynamics. J Neurosci 2008;28:1398–403.
[4] Beck AT, Steer RA, Brown G. Manual for the Beck Depression Inventory-II. San Antonio: Psychological Corporation, 1996.
[5] Bosma RL, Kim JA, Cheng JC, Rogachov A, Hemington KS, Osborne NR, Oh J, Davis KD. Dynamic pain connectome functional connectivity and oscillations reflect multiple sclerosis pain. PAIN 2018;159:2267–76.
[6] Bréchet L, Brunet D, Birot G, Gruetter R, Michel CM, Jorge J. Capturing the spatiotemporal dynamics of self-generated, task-initiated thoughts with EEG and fMRI. NeuroImage 2019;194:82–92.
[7] Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain 2006;10:287–333.
[8] Britz J, Van De Ville D, Michel CM. BOLD correlates of EEG topography reveal rapid resting-state network dynamics. NeuroImage 2010;52:1162–70.
[9] Brunet D, Murray MM, Michel CM. Spatiotemporal analysis of multichannel EEG: cartool. Comput intelligence Neurosci 2011;2011: 813870.
[10] Cheng JC, Rogachov A, Hemington KS, Kucyi A, Bosma RL, Lindquist MA, Inman RD, Davis KD. Multivariate machine learning distinguishes cross-network dynamic functional connectivity patterns in state and trait neuropathic pain. PAIN 2018;159:1764–76.
[11] Custo A, De Ville DV, Wells WM, Tomescu MI, Brunet D, Michel CM. Electroencephalographic resting-state networks: source localization of microstates. Brain Connectivity 2017;7:671–82.
[12] da Cruz JR, Favrod O, Roinishvili M, Chkonia E, Brand A, Mohr C, Figueiredo P, Herzog MH. EEG microstates are a candidate endophenotype for schizophrenia. Nat Commun 2020;11:3089.
[13] Davis JF. Manual of surface electromyography. In: WADC Technical Report (59-184). Montreal: Aerospace Medical Laboratory, United States Air Force, 1959.
[14] Davis KD, Aghaeepour N, Ahn AH, Angst MS, Borsook D, Brenton A, Burczynski ME, Crean C, Edwards R, Gaudilliere B, Hergenroeder GW, Iadarola MJ, Iyengar S, Jiang Y, Kong JT, Mackey S, Saab CY, Sang CN, Scholz J, Segerdahl M, Tracey I, Veasley C, Wang J, Wager TD, Wasan AD, Pelleymounter MA. Discovery and validation of biomarkers to aid the development of safe and effective pain therapeutics: challenges and opportunities. Nat Rev Neurol 2020;16:381–400.
[15] Diaz Hernandez L, Rieger K, Baenninger A, Brandeis D, Koenig T. Towards using microstate-neurofeedback for the treatment of psychotic symptoms in schizophrenia. A feasibility study in healthy participants. Brain Topogr 2016;29:308–21.
[16] Dillmann U, Nilges P, Saile H, Gerbershagen HU. [Assessing disability in chronic pain patients. Schmerz (Berlin, Germany) 1994;8:100–10.
[17] Eccleston C, Crombez G. Pain demands attention: a cognitive-affective model of the interruptive function of pain. Psychol Bull 1999;125:356–66.
[18] Freynhagen R, Baron R, Gockel U, Tolle TR. painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. Curr Med Res Opin 2006;22:1911–20.
[19] Garrett DD, Samanez-Larkin GR, MacDonald SW, Lindenberger U, McIntosh AR, Grady CL. Moment-to-moment brain signal variability: a next frontier in human brain mapping? Neurosci Biobehav Rev 2013;37: 610–24.
[20] Gonzalez-Villar AJ, Trinanes Y, Gomez-Perretta C, Carrillo-de-la-Pena MT. Patients with fibromyalgia show increased beta connectivity across distant networks and microstates alterations in resting-state electroencephalogram. Neuroimage 2020;223:117266.
[21] Harden RN, Weinland SR, Remble TA, Houle TT, Colio S, Steedman S, Kee WG. Medication Quantification Scale Version III: update in medication classes and revised detriment weights by survey of American Pain Society Physicians. J Pain 2005;6:364–71.
[22] JASP Team (Version 0.14.1) [Computer software], 2020.
[23] Jung TP, Makeig S, Humphries C, Lee TW, McKeown MJ, Iragui V, Sejnowski TJ. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 2000;37:163–78.
[24] Kennedy J, Roll JM, Schraudner T, Murphy S, McPherson S. Prevalence of persistent pain in the U.S. adult population: new data from the 2010 national health interview survey. J Pain 2014;15:979–84.
[25] Khanna A, Pascual-Leone A, Michel CM, Farzan F. Microstates in resting-state EEG: current status and future directions. Neurosci Biobehav Rev 2015;49:105–13.
[26] Koenig T, Lehmann D, Merlo MC, Kochi K, Hell D, Koukkou M. A deviant EEG brain microstate in acute, neuroleptic-naive schizophrenics at rest. Eur Arch Psychiatry Clin Neurosci 1999;249:205–11.
[27] Koenig T, Prichep L, Lehmann D, Sosa PV, Braeker E, Kleinlogel H, Isenhart R, John ER. Millisecond by millisecond, year by year: normative EEG microstates and developmental stages. NeuroImage 2002;16:41–8.
[28] Kucyi A, Davis KD. The dynamic pain connectome. Trends Neurosci 2015;38:86–95.
[29] Kuner R, Flor H. Structural plasticity and reorganisation in chronic pain. Nat Rev Neurosci 2017;18:113.
[30] Legrain V, Damme SV, Eccleston C, Davis KD, Seminowicz DA, Crombez G. A neurocognitive model of attention to pain: behavioral and neuroimaging evidence. PAIN 2009;144:230–2.
[31] Lehmann D, Faber PL, Galderisi S, Herrmann WM, Kinoshita T, Koukkou M, Mucci A, Pascual-Marqui RD, Saito N, Wackermann J, Winterer G, Koenig T. EEG microstate duration and syntax in acute, medication-naïve, first-episode schizophrenia: a multi-center study. Psychiatry Res Neuroimaging 2005;138:141–56.
[32] May ES, Nickel MM, Ta Dinh S, Tiemann L, Heitmann H, Voth I, Tölle TR, Gross J, Ploner M. Prefrontal gamma oscillations reflect ongoing pain intensity in chronic back pain patients. Hum Brain Mapp 2019;40: 293–305.
[33] Melzack R. The short-form McGill pain questionnaire. PAIN 1987;30: 191–7.
[34] Michel CM, Koenig T. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. Neuroimage 2018; 180:577–93.
[35] Michel CM, Murray MM. Towards the utilization of EEG as a brain imaging tool. NeuroImage 2012;61:371–85.
[36] Milz P, Faber PL, Lehmann D, Koenig T, Kochi K, Pascual-Marqui RD. The functional significance of EEG microstates-Associations with modalities of thinking. Neuroimage 2016;125:643–56.
[37] Moriarty O, McGuire BE, Finn DP. The effect of pain on cognitive function: a review of clinical and preclinical research. Prog Neurobiol 2011;93: 385–404.
[38] Murphy M, Whitton AE, Deccy S, Ironside ML, Rutherford A, Beltzer M, Sacchet M, Pizzagalli DA. Abnormalities in electroencephalographic microstates are state and trait markers of major depressive disorder. Neuropsychopharmacology 2020;45:2030–2037.
[39] Murray MM, Brunet D, Michel CM. Topographic ERP analyses: a step-by-step tutorial review. Brain Topography 2008;20:249–64.
[40] Nishida K, Morishima Y, Yoshimura M, Isotani T, Irisawa S, Jann K, Dierks T, Strik W, Kinoshita T, Koenig T. EEG microstates associated with salience and frontoparietal networks in frontotemporal dementia, schizophrenia and Alzheimer's disease. Clin Neurophysiol 2013;124: 1106–14.
[41] Oostenveld R, Fries P, Maris E, Schoffelen JM. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput intelligence Neurosci 2011;2011: 156869.
[42] Pascual-Marqui RD, Michel CM, Lehmann D. Segmentation of brain electrical activity into microstates: model estimation and validation. Ieee T Bio-med Eng 1995;42:658–65.
[43] Pernet CR, Appelhoff S, Gorgolewski KJ, Flandin G, Phillips C, Delorme A, Oostenveld R. EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. Sci Data 2019;6:103.
[44] Ploner M, May ES. Electroencephalography and magnetoencephalography in pain research - current state and future perspectives. PAIN 2018;159:206–11.
[45] Preti MG, Bolton TA, Van De Ville D. The dynamic functional connectome: state-of-the-art and perspectives. Neuroimage 2017;160:41–54.
[46] Rice AS, Smith BH, Blyth FM. Pain and the global burden of disease. PAIN 2016;157:791–6.
[47] Rieger K, Diaz Hernandez L, Baenninger A, Koenig T. 15 Years of microstate research in schizophrenia - where are we? A meta-analysis. Front Psychiatry 2016;7:22.
[48] Schumacher J, Peraza LR, Firbank M, Thomas AJ, Kaiser M, Gallagher P, O'Brien JT, Blamire AM, Taylor JP. Dysfunctional brain dynamics and their origin in Lewy body dementia. Brain 2019;142:1767–82.
[49] Seitzman BA, Abell M, Bartley SC, Erickson MA, Bolbecker AR, Hetrick WP. Cognitive manipulation of brain electric microstates. Neuroimage 2017;146:533–43.
[50] Selim AJ, Rogers W, Fleishman JA, Qian SX, Fincke BG, Rothendler JA, Kazis LE. Updated U.S. Population standard for the veterans RAND 12-item health survey (VR-12). Qual Life Res 2009;18:43–52.
[51] Shaw SB, Dhindsa K, Reilly JP, Becker S. Capturing the forest but missing the trees: microstates inadequate for characterizing shorter-scale EEG dynamics. Neural Comput 2019;31:2177–211.
[52] Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. Manual for the State-Trait Anxiety Inventory. Palo Alto: Consulting Psychologists Press, 1983.
[53] Sverak T, Albrechtova L, Lamos M, Rektorova I, Ustohal L. Intensive repetitive transcranial magnetic stimulation changes EEG microstates in schizophrenia: a pilot study. Schizophr Res 2018;193:451–2.

[54] Ta Dinh S, Nickel MM, Tiemann L, May ES, Heitmann H, Hohn VD, Edenharter G, Utpadel-Fischler D, Tolle TR, Sauseng P, Gross J, Ploner M. Brain dysfunction in chronic pain patients assessed by resting-state electroencephalography. PAIN 2019;160:2751–65.
[55] Tomescu MI, Rihs TA, Becker R, Britz J, Custo A, Grouiller F, Schneider M, Debbané M, Eliez S, Michel CM. Deviant dynamics of EEG resting state pattern in 22q11.2 deletion syndrome adolescents: a vulnerability marker of schizophrenia? Schizophrenia Res 2014;157:175–81.
[56] Tomescu MI, Rihs TA, Rochas V, Hardmeier M, Britz J, Allali G, Fuhr P, Eliez S, Michel CM. From swing to cane: sex differences of EEG resting-state temporal patterns during maturation and aging. Develop Cogn Neurosci 2018;31:58–66.
[57] Torta DM, Legrain V, Mouraux A, Valentini E. Attention to pain! A neurocognitive perspective on attentional modulation of pain in neuroimaging studies. Cortex 2017;89:120–34.
[58] Tracey I, Woolf CJ, Andrews NA. Composite pain biomarker signatures for objective assessment and effective treatment. Neuron 2019;101:783–800.
[59] Tu Y, Fu Z, Mao C, Falahpour M, Gollub RL, Park J, Wilson G, Napadow V, Gerber J, Chan ST, Edwards RR, Kaptchuk TJ, Liu T, Calhoun V, Rosen B, Kong J. Distinct thalamocortical network dynamics are associated with the pathophysiology of chronic low back pain. Nat Commun 2020;11:3948.
[60] Turk DC, Wilson HD, Cahana A. Treatment of chronic non-cancer pain. Lancet 2011;377:2226–35.
[61] Winkler I, Debener S, Muller KR, Tangermann M. On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. Conf Proc 2015;2015:4101–5.
[62] Yekutieli D, Benjamini Y. Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. J Stat Plann Inference 1999;82:171–96.
[63] Zanesco AP, King BG, Skwara AC, Saron CD. Within and between-person correlates of the temporal dynamics of resting EEG microstates. Neuroimage 2020;211:116631.
[64] Zhang B, Jung M, Tu Y, Gollub R, Lang C, Ortiz A, Park J, Wilson G, Gerber J, Mawla I, Chan ST, Wasan A, Edwards R, Lee J, Napadow V, Kaptchuk T, Rosen B, Kong J. Identifying brain regions associated with the neuropathology of chronic low back pain: a resting-state amplitude of low-frequency fluctuation study. Br J Anaesth 2019;123:e303–11.

3. PROJECT 2. DISCOVER-EEG: AN AUTOMATIC EEG PIPELINE FOR BIOMARKER DISCOVERY

This section includes the research article 'DISCOVER-EEG: an open, fully automated EEG pipeline for biomarker discovery in clinical neuroscience', published under an open access license (CC-BY 4.0) in Scientific Data (Gil Ávila et al., 2023b).

Authors

Cristina Gil Ávila, Felix S. Bott, Laura Tiemann, Vanessa D. Hohn, Elisabeth S. May, Moritz M. Nickel, Paul Theo Zebhauser, Joachim Gross, Markus Ploner

Contributions

Conceptualization: CGA and MP; Methodology: CGA and MP; Software: CGA, FSB, and JG; Validation: CGA, FSB, and JG; Formal analysis: CGA; Investigation: CGA and PTZ; Data curation: CGA; Visualization: CGA and MP; Writing – original draft: CGA and MP; Writing – review and editing: CGA, FSB, LT, VDH, ESM, MMN, PTZ, JG, and MP; Supervision: MP; Project administration: MP; Funding acquisition: MP

scientific data

Check for updates

OPEN

ANALYSIS

DISCOVER-EEG: an open, fully automated EEG pipeline for biomarker discovery in clinical neuroscience

Cristina Gil Ávila[1,2,3], Felix S. Bott[1,2], Laura Tiemann[1,2], Vanessa D. Hohn[1,2], Elisabeth S. May[1,2], Moritz M. Nickel[1,2], Paul Theo Zebhauser[1,2,4], Joachim Gross[5] & Markus Ploner[1,2,4] ✉

Biomarker discovery in neurological and psychiatric disorders critically depends on reproducible and transparent methods applied to large-scale datasets. Electroencephalography (EEG) is a promising tool for identifying biomarkers. However, recording, preprocessing, and analysis of EEG data is time-consuming and researcher-dependent. Therefore, we developed DISCOVER-EEG, an open and fully automated pipeline that enables easy and fast preprocessing, analysis, and visualization of resting state EEG data. Data in the Brain Imaging Data Structure (BIDS) standard are automatically preprocessed, and physiologically meaningful features of brain function (including oscillatory power, connectivity, and network characteristics) are extracted and visualized using two open-source and widely used Matlab toolboxes (EEGLAB and FieldTrip). We tested the pipeline in two large, openly available datasets containing EEG recordings of healthy participants and patients with a psychiatric condition. Additionally, we performed an exploratory analysis that could inspire the development of biomarkers for healthy aging. Thus, the DISCOVER-EEG pipeline facilitates the aggregation, reuse, and analysis of large EEG datasets, promoting open and reproducible research on brain function.

Introduction

Biomarkers that relate brain function to cognitive and clinical phenotypes can help in the prediction, treatment, monitoring, and diagnosis of neurological and psychiatric disorders[1,2]. The successful identification of biomarkers crucially depends on the application of reproducible and transparent methods[3] to large-scale datasets[4]. Furthermore, to translate biomarkers into clinical practice, they need to be generalizable, interpretable, and easy to deploy in clinical settings.

Electroencephalography (EEG) is a promising tool for biomarker discovery, as it is non-invasive, safe, widely used in clinical and research contexts, portable, and cost-efficient. Consequently, EEG biomarker candidates have been described in depression[5–7], post-traumatic stress disorder[7,8], and chronic pain[9]. Most of these biomarker candidates have been discovered in resting state data, during which spontaneous neural activity is captured. Still, their translation into clinical practice has not yet been successful[1]. This is partly due to small sample sizes and the low availability of objective, transparent, and reproducible EEG preprocessing and analysis methods.

In recent years, notable efforts have been made to automatize, speed up, and increase the transparency of EEG research. A standardized EEG Brain Imaging Data Structure (EEG-BIDS) has been created, which allows for the efficient organization, sharing, and reuse of EEG data[10]. Moreover, automatic preprocessing pipelines have been developed, some for specific populations, settings, and study designs, including pediatric populations[11,12],

[1]Department of Neurology, TUM School of Medicine, Technical University of Munich, München, Germany. [2]TUM-Neuroimaging Center, TUM School of Medicine, Technical University of Munich, München, Germany. [3]Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München, München, Germany. [4]Center for Interdisciplinary Pain Medicine, TUM School of Medicine, Technical University of Munich, München, Germany. [5]Institute for Biomagnetism and Biosignalanalysis, University of Münster, Münster, Germany. ✉e-mail: markus.ploner@tum.de

mobile brain-body imaging[13], and event-related potentials (ERPs)[14,15], and others for more generic purposes[15–18]. Beyond, guidelines for the standardized reporting of EEG studies have been established[19]. A logical next step is to integrate these solutions into an automatic workflow that can preprocess and extract physiologically informative features in large EEG datasets.

Here, we present DISCOVER-EEG, a comprehensive EEG pipeline for resting state data that extends current preprocessing pipelines by extracting and visualizing physiologically relevant EEG features for biomarker identification. As translation of biomarkers benefits from being neuroscientifically plausible and interpretable[1], DISCOVER-EEG extracts EEG features, such as oscillatory power, connectivity, and network characteristics that have previously been related to brain dysfunction in neurological and psychiatric disorders[20–23]. It builds upon and combines two open-source and widely-used Matlab toolboxes (EEGLAB[24] and FieldTrip[25]) and adheres to COBIDAS-MEEG guidelines for reproducible MEEG research[19]. It facilitates the aggregation and analysis of large-scale datasets, as it applies to a wide range of EEG setups, and fosters sharing and reusability of the data by handling EEG-BIDS standardized data[10].

We tested DISCOVER-EEG in two large and openly-available datasets, the LEMON dataset[26], which includes resting state EEG recordings of 213 young and old healthy participants, and the TDBRAIN dataset[27], which includes resting state EEG recordings of 1274 participants mainly with psychiatric conditions. In both datasets, we demonstrate the capability of the pipeline to capture a well-known EEG effect: the reduction of alpha power during eyes open compared to eyes closed[28]. Finally, using the LEMON dataset, we present an example analysis investigating differences in EEG features between old and young healthy populations that could inspire the development of biomarkers of healthy aging. Thus, the DISCOVER-EEG pipeline facilitates the preprocessing and analysis of large EEG datasets, promoting open and reproducible research on brain function.

Methods

Design principles. *Open-source and FAIR code.* We developed an automated workflow for fast preprocessing, analysis, and visualization of resting state EEG data (Fig. 1) following open science and FAIR principles (Findability, Accessibility, Interoperability, and Reusability)[29]. The code of the DISCOVER-EEG pipeline is published on GitHub and co-deposited at Zenodo, where it is uniquely referenced by a DOI[30] (*Findability*). The code can be easily downloaded (*Accessibility*) and receive contributions (please refer to the section *Code availability*). To ensure its *Interoperability* and *Reusability*, the pipeline is based on two open-source Matlab toolboxes, EEGLAB[24] and FieldTrip[25], which are widely used, maintained, and supported by the developers and the neuroimaging community (i.e., through forums and mailing lists). Basing the pipeline on validated and established software ensures its compatibility with future software updates. Moreover, it facilitates interaction with experts in EEG analysis, who also gave advice and supported the pipeline during its development. The code of the current pipeline is intended to represent a basis that will integrate and benefit from feedback from the neuroimaging community.

Data reusability and large-scale data handling. As biomarker discovery needs large datasets, we designed the pipeline in view of data reusability and large-scale data handling. To this end, we followed the FAIR principles of scientific data management[31] and incorporated the EEG-BIDS standardized data structure[10] as a mandatory input of the pipeline. Most EEG setups and electrode configurations are compatible with the pipeline thanks to the EEGLAB plugin *bids-matlab-tools*. We refer the reader to the *Results* section for a demonstration of performance across two published datasets with different characteristics, such as participant sample, number of channels, recording length, and sampling rate.

The pipeline was designed for resting state EEG data, during which spontaneous neural activity is recorded. Resting state data can be recorded easily in healthy and patient populations, in different study designs (e.g., cross-sectional, longitudinal), and in different types of neuropsychiatric disorders. Therefore, the use of resting state data facilitates the application to different settings and research questions. EEG recordings might also be accompanied by standardized patient-reported outcomes, such as the PROMIS questionnaires[32], which can assess symptoms (e.g., pain, fatigue, anxiety, depression) across different neuropsychiatric disorders and, thus, enable cross-disorder analyses. Together, these considerations contribute to the scalability and generalizability of the workflow.

Ease of use, transparency, and interpretability. The pipeline consists of a main function, *main_pipeline.m*, in which the preprocessing, feature extraction, and visualization of the data are carried out, and a *params.json* file, in which parameters for preprocessing and feature extraction are defined. This latter file is the only one that needs to be configured by the user and can be easily adapted to dataset and/or user-specific demands. When executing the pipeline, parameters are saved to a separate json file to ensure reproducibility. We recommend DISCOVER-EEG users reporting all parameter configurations as well as software versions when reporting their findings.

Additionally, the pipeline focuses on transparency and interpretability of results. For that reason, single images and an optional PDF report can be generated for each recording to visualize the intermediate steps of the preprocessing (Fig. 2) and the extracted EEG features (Fig. 3). These visualizations can serve as quality control checkpoints and help to detect shorter or corrupted files, misalignment of electrodes, or missing data through fast visual inspection[33]. In that way, we do not exclude any recordings during preprocessing based on quality criteria. Instead, we provide visualization tools that let the users decide how conservative to be in their analysis. Along with their visualization, the preprocessed data and the extracted features of each recording are saved to separate files that can be later used for statistical group analysis and biomarker discovery. These output files can be easily imported to other statistical packages or working environments, such as R[34] or Python[35], e.g., for applying machine learning or deep learning models.

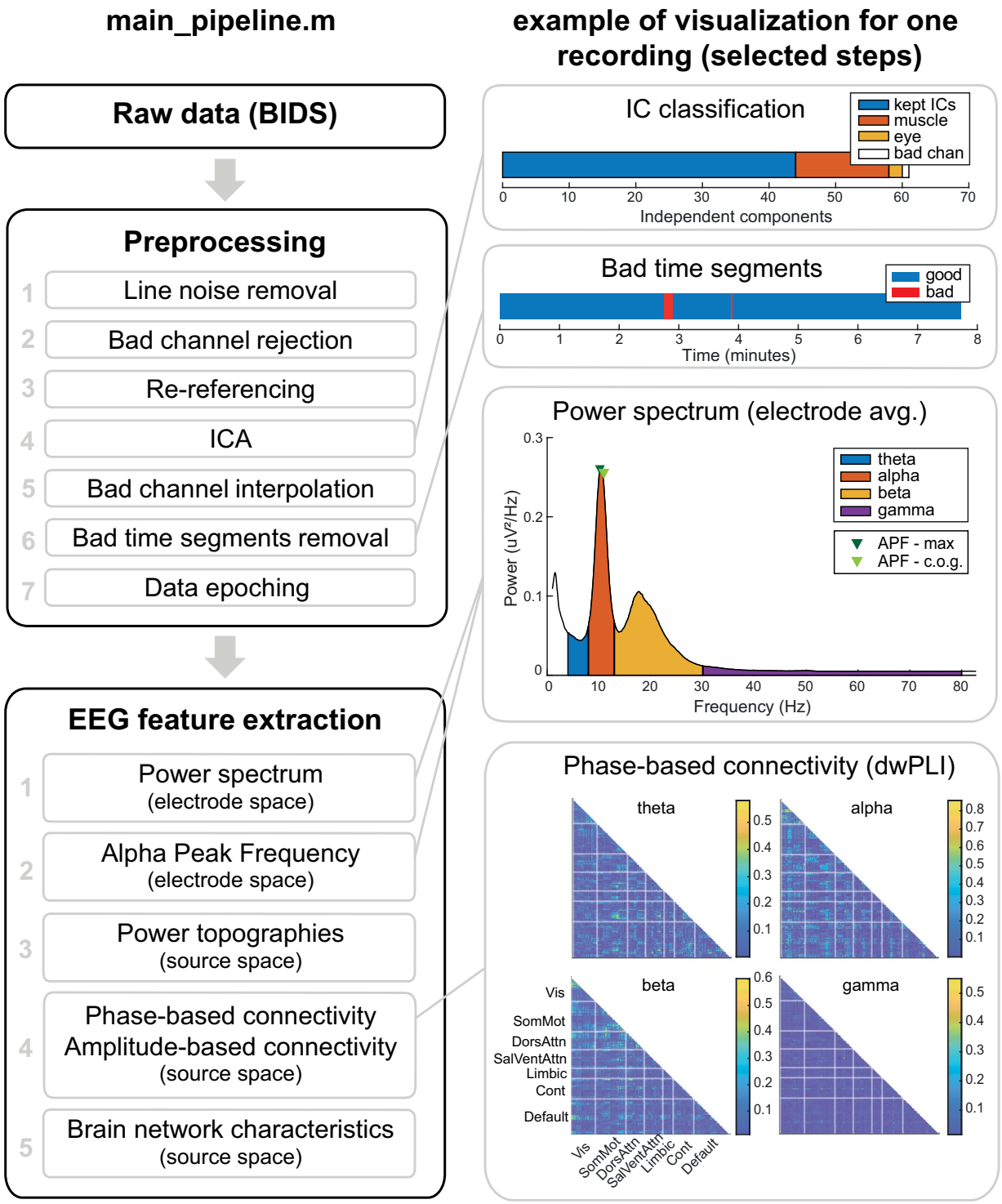

Fig. 1 Outline of the DISCOVER pipeline. The left column shows all preprocessing steps and extracted EEG features. The right column shows visualizations for selected steps and features for one EEG recording of the LEMON dataset. APF = Alpha Peak Frequency; BIDS = Brain Image Data Structure; c.o.g. = center of gravity; Cont = Control; DorsAttn = Dorsal Attention; dwPLI = debiased weighted Phase Lag Index; ICA = Independent Component Analysis; SalVentAttn = Salience-Ventral Attention; SomMot = Somato-Motor; Vis = Visual.

Preprocessing. We developed a Matlab-based EEG processing pipeline that complements previous approaches in other programming languages, such as the MNE-BIDS-pipeline (https://mne.tools/mne-bids-pipeline/1.3/index.html) in Python. Matlab is widely used by the EEG community and enabled us to use well-established Matlab-based EEG toolboxes (please refer to the *Design principles* section) that provide

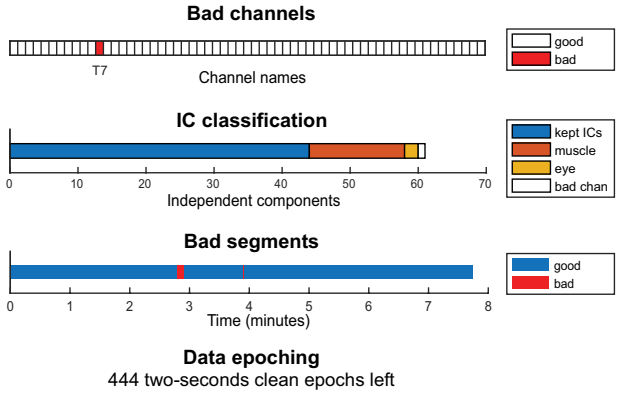

Fig. 2 Visualization of the outcome of the preprocessing part of the DISCOVER pipeline. Example of one EEG recording of the LEMON dataset. In the independent component classification (second row), bad channels correspond to the channels that were removed in the bad channel removal step (first row). Bad channels were not included in the independent component analysis. IC = independent component.

robust functions for computing functional connectivity measures. Thus, we followed a pragmatic approach towards preprocessing and adopted a simple, established, and automatic workflow in EEGLAB proposed by Pernet *et al.*[14] and originally developed for ERP data. We adapted this pipeline to resting-state data and detail the seven preprocessing steps below.

Loading the data. The pipeline input must be EEG data in BIDS format[10], including all mandatory sidecar files. We recommend checking the BIDS compliance with the BIDS validator (https://bids-standard.github.io/bids-validator). By default, only EEG channels with standard electrode positions in the 10-5 system[36] are preprocessed. However, it is possible to preprocess non-standard channels if all electrode positions and their coordinate system are specified in the BIDS sidecar file. Optionally, after loading, the data can be downsampled to a frequency specified by the user after applying an anti-aliasing filter.

1. Line noise removal. Line noise is removed with the EEGLAB function *pop_cleanline()*. The CleanLine plugin adaptively estimates and removes sinusoidal artifacts using a frequency-domain (multi-taper) regression technique. CleanLine, compared to band-stop filters, does not introduce gaps in the power spectrum and avoids the frequency distortions filters create. This step was added to the Pernet *et al.*[14] pipeline to explore brain activity at gamma frequencies (>30 Hz). The line noise frequency (e.g., 50 Hz or 60 Hz) must be specified in the BIDS file *sub-<label>_eeg.json* to be appropriately removed.

2. High pass filtering and bad channel rejection. Artifactual channels are detected and removed with the function *pop_clean_rawdata()*. The first step of this function is the application of a high pass filter with the function *clean_drifts()* with a default transition band of 0.25 to 0.75 Hz. A channel is considered artifactual if it meets any of the following criteria: 1) If it is flat for more than 5 seconds, 2) If the z-scored noise-to-signal ratio of the channel is higher than a threshold set to 4 by default, 3) If the channel's time course cannot be predicted from a randomly selected subset of remaining channels at least 80% of the recorded time. Channels marked as artifactual are removed from the data. The mentioned parameters are the defaults proposed by Pernet *et al.*[14].

3. Re-referencing. Data is re-referenced to the average reference with the function *pop_reref()*. Optionally, the time course of the original reference channel can be reconstrued and added back to the data if the user specifies it in the file *params.json*. The name of the reference electrode name must be specified in the BIDS file *sub-<label>_eeg.json*.

4. Independent Component Analysis and automatic IC rejection. Independent Component Analysis (ICA) is performed with the function *pop_runica()* using the algorithm *runica*. Artifactual components are automatically classified into seven distinct categories ('Brain', 'Muscle', 'Eye', 'Heart', 'Line Noise', 'Channel Noise', and 'Other') by the ICLabel classifier[37]. Note that ICA is performed only on clean channels, as bad channels detected in step 2 were removed from the data. Therefore, the category 'Channel Noise' in the IC classification refers to channel noise remaining after bad channel removal in step 2. By default, only components whose probability of being 'Muscle' or 'Eye' is higher than 80% are subtracted from the data[14]. Due to the non-deterministic nature of the ICA algorithm, its results vary across repetitions. That is, every repetition of the ICA algorithm leads to small differences in the reconstructed time series after removing artifactual components. While these deviations are small, we noticed that they affect the removal of bad time segments (step 6). For that reason, the pipeline performs 10 times steps 4, 5, and 6 by default and selects the bad time segment mask most similar to the average bad

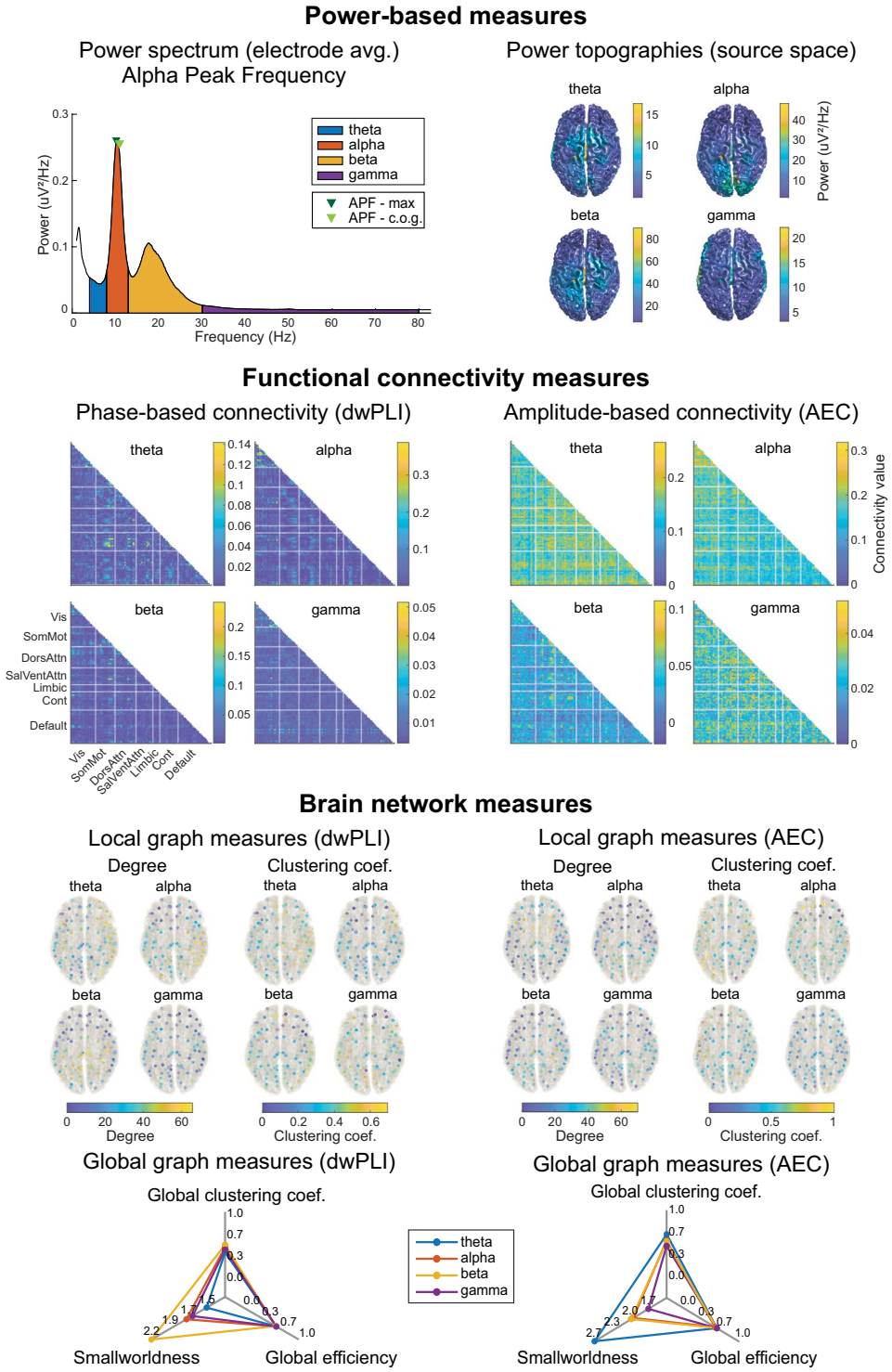

Fig. 3 Visualization of the outcome of the feature extraction part of the pipeline. Example for one recording of the LEMON dataset (same recording as Fig. 2). **Power-based measures** are extracted in electrode (power spectrum averaged across channels, APF) and source space (power topographies). **Functional connectivity measures** are estimated in source space for 100 pairs of brain regions organized in 7 different functional networks (Visual, Somato-Motor, Salience, Ventral-Attention, Limbic, Control, and Default). Connectivity matrices are symmetric, and thus only lower triangular matrices are shown. **Brain network measures** are characterized by local (degree and clustering coefficient) and global (global clustering coefficient, global efficiency, and smallworldness) graph measures computed on the thresholded connectivity matrices. AEC = Amplitude Envelope Correlation; APF = Alpha Peak Frequency; c.o.g. = center of gravity; Cont = Control; DorsAttn = Dorsal Attention; dwPLI = debiased weighted Phase Lag Index; SalVentAttn = Salience-Ventral Attention; SomMot = Somato-Motor; Vis = Visual.

time segment mask across all repetitions. This reduces the variability of rejected time segments. However, for computational economy, the number of iterations may be set to one in the *params.json* file.

5. Interpolation of removed channels. Channels that were removed in step 2 are interpolated with the function *pop_interp()* using spherical splines[38]. This step was added to the Pernet *et al.*[14] pipeline to hold the number of channels constant across participants.

6. Bad time segment removal. Time segments containing artifacts are removed with the Artifact Subspace Reconstruction (ASR) method[39] implemented in the function *pop_clean_rawdata()*. This method automatically removes segments in which power is abnormally strong. First, a clean data segment is identified according to the default ASR settings and used for calibration. Calibration data contains all data points in which less than 7.5% of channels are noisy. Here, a channel is defined as noisy if the standard deviation of its RMS is higher than 5.5. Therefore, the length of the calibration data depends on the specific recording. Then, in a sliding window fashion, the whole EEG signal is decomposed via PCA, and the principal subspaces of the window segment are compared with those of the calibration data. Segments with principal subspaces deviating from the calibration data (20 times higher variance) are removed. Again, default parameters are in line with Pernet *et al.*[14].

7. Data segmentation into epochs. Lastly, the continuous data are segmented into epochs with the function *pop_epoch()*. By default, data are segmented into 2-second epochs with a 50% overlap. Although longer epochs might be desirable for the Alpha Peak Frequency estimation to increase frequency resolution, short epochs favor the reliability of functional connectivity measures[40,41]. Thus, we propose 2-second epochs to establish a balance between frequency resolution, stationarity of the signal, and reliability of the later extracted features. Fifty percent overlap was chosen to provide a smooth estimation of the power spectra and mitigate the loss of signal due to tapering[42]. Epochs containing a discontinuity (e.g., because a segment containing an artifact was discarded) are rejected automatically. Data segmentation was adapted from Pernet *et al.*[14], which focused on event-related data.

EEG feature extraction. DISCOVER-EEG extracts EEG features that have previously been related to different neurological and psychiatric disorders and have the potential to be translated into neuroscientifically plausible and interpretable biomarkers.

In electrode space, power spectra and the Alpha Peak Frequency, i.e., the frequency at which a peak in the power spectrum in the alpha range occurs, are extracted. Power changes in different frequency bands have been found in a broad spectrum of neuropsychiatric disorders[20], and APF has been correlated with behavioral and cognitive characteristics[43] in aging and disease[44].

In source space, two measures of functional connectivity in the theta, alpha, beta, and gamma bands, are extracted. Additionally, brain networks derived from these connectivity matrices are further characterized with common graph theory measures[45]. Functional connectivity is a promising biomarker candidate, as it has been successfully used for the classification, tracking, and stratification of patients with several neurological and psychiatric disorders, such as Alzheimer's disease[46], Post Traumatic Stress Disorder[8], and Major Depressive Disorder[7,47] (for a recent review see[48]). In EEG, functional connectivity measures are commonly classified as phase-based or amplitude-based, each type capturing different and complementary communication processes in the brain[49]. This pipeline thus includes a phase-based measure, the debiased weighted Phase Lag Index (dwPLI)[50], and an amplitude-based measure, the orthogonalized Amplitude Envelope Correlation (AEC)[51]. Both functional connectivity measures are undirected and have low susceptibility to volume conduction. The dwPLI is also more sensitive and capable of capturing non-linear relationships compared to other phase-based measures[42]. For each connectivity matrix, two local graph measures (the degree and the clustering coefficient) are calculated at each source location, and three global graph measures (the global clustering coefficient, the global efficiency, and the smallworldness) summarize the whole network in one value.

Power and connectivity features are computed using the preprocessed and segmented data in FieldTrip. Graph theory measures are computed with the Brain Connectivity Toolbox[45]. Specific parameters on feature extraction are defined in the file *params.json* and detailed below.

1. Power spectrum. Power spectra are computed with the FieldTrip function *ft_freqanalysis* between 1 and 100 Hz using Slepian multitapers with +/−1 Hz frequency smoothing. For 2-second epochs, the maximum frequency resolution of the power spectrum is, by definition, the inverse of the epoch length, i.e., 0.5 Hz. As frequency band limits are determined by the spectral resolution, the pipeline zero-pads the epochs to 5 seconds, which yields a resolution of 0.1 Hz, to better capture the frequency range of the bands. Thus, the frequency band limits for theta, alpha, beta, and gamma are 4 to 7.9 Hz, 8 to 12.9 Hz, 13 to 30 Hz, and 30.1 to 80 Hz, respectively. These limits are defined by the COBIDAS-MEEG[19] guidelines and maintained throughout all features. Power spectra are computed for every channel and then averaged across epochs and channels to obtain a single global power spectrum. This global power spectrum is saved to *sub-<label>_power.mat* files and visualized with the different frequency bands highlighted in different colors (Fig. 3).

2. Alpha Peak Frequency. APF is computed based on the global power spectrum in the alpha range (8 to 12.9 Hz). There are two widely accepted strategies to assess the APF in the literature[52]. The first one reports the frequency at which the highest peak in the alpha range occurs (peak maximum). This comes with the problem that not always a peak is present in the power spectrum. Therefore, the second strategy, calculating the center of gravity (c.o.g.) of the power spectra in the alpha band is also frequently used[53]. DISCOVER-EEG reports both measures for estimating APF. The peak maximum is computed with the Matlab function *findpeaks()*. If there is

no peak in the alpha range, no value is returned. The center of gravity is computed as the weighted average of frequencies in the alpha band, each frequency being weighted by their power[53]. Both measures are visualized in the same figure as the global power spectrum (Fig. 3). Individual APF values are also saved to the *sub-<label>_peakfrequency.mat* files.

3. Source reconstruction. To mitigate the volume conduction problem when computing functional connectivity measures[54], we perform a source reconstruction of the preprocessed data with an atlas-based beamforming approach[55]. For each frequency band, the band-pass filtered data from sensor space is projected into source space using an array-gain Linear Constrained Minimum Variance (LCMV) beamformer[56]. As source model, we selected the centroids of 100 regions of interest (ROIs) of the 7-network version of the Schaefer atlas[57]. This atlas is a refined version of the Yeo atlas and follows a data-driven approach in which 100 parcellations are clustered and assigned to 7 brain networks (Visual, Somato-Motor, Dorsal attention, Salience-Ventral attention, Limbic, Control, and Default networks). This atlas can be easily changed according to the user's preferences in the *params.json* file. The lead field is built using a realistically shaped volume conduction model based on the Montreal Neurological Institute (MNI) template available in FieldTrip (*standard_bem.mat*) and the source model. Spatial filters are finally constructed with the covariance matrices of the band-passed filtered data and the described lead fields. A 5% regularization parameter is set to account for rank deficiencies in the covariance matrix, and the dipole orientation is fixed to the direction of the maximum variance following the most recent recommendations[58]. The power for each source location is estimated using the spatial filter and band-passed data. A visualization of the source power in each frequency band (Fig. 3) is provided by projecting the band-specific source power to a cortical surface model provided as a template in FieldTrip (*surface_white_both.mat*).

4. Functional connectivity. After creating spatial filters in the four frequency bands, virtual time series in the 100 source locations are reconstructed for each frequency band by applying the respective band-specific spatial filter to the band-pass filtered sensor data. Then, the two functional connectivity measures (dwPLI and AEC) are computed for each frequency band and combination of the 100 reconstructed virtual time series. Average connectivity matrices for each band are visualized (Fig. 3) and saved to separate files (*sub-<label>_<conmeasure>_<band>.mat*).

The phase-based connectivity measure dwPLI is computed using the FieldTrip function *ft_connectivityanalysis* with the method *wpli_debiased*, which requires a frequency structure as input. Therefore, Fourier decompositions of the virtual time series are calculated in each frequency band with a frequency resolution of 0.5 Hz. Thereby, a connectivity matrix is obtained for each frequency of interest in the current frequency band. Connectivity matrices are then averaged across each frequency band resulting in one 100×100 connectivity matrix for each frequency band.

The amplitude-based connectivity measure AEC is computed according to the original equations with a custom function *compute_aec*, as the original implementation was not available in FieldTrip. For each epoch, the analytical signal of the virtual time series is extracted at each source location with the Hilbert transform. For each source pair, the analytical signal at source A is orthogonalized with respect to the analytical signal at source B, yielding the signal $A_{\perp B}$. Then, the Pearson correlation is computed between the amplitude envelope of signals B and $A_{\perp B}$. To obtain the average connectivity between sources A and B, the Pearson correlation between the amplitude envelopes of analytical signal A and signal $B_{\perp A}$ is also computed, and the two correlation coefficients are averaged. In this way, we obtain a 100×100 connectivity matrix for each epoch. We finally average the connectivity matrices across epochs, resulting in one 100×100 connectivity matrix for each frequency band.

5. Brain network characteristics. Graph measures were computed on thresholded and binarized connectivity matrices[45]. Matrices were binarized by keeping the 20% strongest connections, as this threshold delivers fairly reproducible graph measures based on dwPLI and AEC connectivity[59]. Nevertheless, it is good practice to test the reliability of final results with different binarizing thresholds[60]. This threshold can be easily changed in the file *params.json*.

The computed local network measures are the degree and the clustering coefficient. The degree is the number of connections of a node in the network. The clustering coefficient is the percentage of triangle connections surrounding a node. The measures, thus, assess the global and local connectedness of a node, respectively.

The computed global network measures are the global clustering coefficient, the global efficiency, and the smallworldness. The global clustering coefficient is a measure of functional segregation in the network and is defined as the average clustering coefficient of all nodes. Global efficiency is a measure of functional integration in the network and is defined as the average of the inverse shortest path length between all pairs of nodes. High global efficiency indicates that information can travel efficiently between regions that are far away. Smallworldness compares the ratio between functional integration and segregation in the network against a random network of the same size and degree. Smallworld networks are highly clustered and have short characteristic path length (average shortest path length between all nodes) compared to random networks[61]. Local and global measures are visualized per recording and saved to files named *sub-<label>_graph_<conmeasure>_<band>.mat* (Fig. 3).

Results

Testing DISCOVER-EEG in two large, public datasets. We tested the DISCOVER-EEG pipeline in two well-documented and openly-available resting state EEG datasets, the LEMON dataset[26,62], including 213 healthy participants, and the TDBRAIN dataset[27,63], including 1274 participants, mainly with a psychiatric condition. Below we detail the characteristics of both datasets and the effects of preprocessing using the DISCOVER-EEG

	Rejected channels	Rejected ICs	Bad segments
LEMON closed	1 ± 2 (2%)	6 ± 3 (11%)	32 ± 42 s (7%)
LEMON open	2 ± 1 (3%)	9 ± 5 (15%)	29 ± 38 s (6%)
TDBRAIN closed	1 ± 1 (5%)	1 ± 1 (7%)	5 ± 8 s (4%)
TDBRAIN open	2 ± 2 (7%)	2 ± 1 (10%)	6 ± 2 s (5%)

Table 1. DISCOVER-EEG preprocessing summary of the LEMON and TDBRAIN datasets for eyes open and closed conditions. Average ± standard deviation across participants is stated for the number of rejected channels, number of rejected ICs, and length of rejected bad time segments in seconds. The average of each variable in percentage is stated between brackets. The maximum number of ICs varied between participants as it depended on the number of channels that were rejected for each participant. ICs = Independent components.

pipeline for resting state recordings under eyes open and eyes closed conditions. These results demonstrate that the pipeline works successfully across different EEG systems, electrode numbers and layouts, sampling rates, and recording lengths.

We additionally show that the pipeline can capture a well-known neurophysiological phenomenon in both datasets: the difference in oscillatory alpha power between eyes open and eyes closed conditions[28]. A comparison of power spectra between eyes open and eyes closed has previously been reported for the TDBRAIN dataset to prove the neurophysiological validity of the data[27]. Reproducing this effect with a different preprocessing strategy for the TDBRAIN dataset and confirming it for the first time for the LEMON dataset provides evidence for the capability of the DISCOVER-EEG pipeline to capture well-known features of EEG data.

LEMON dataset. The LEMON dataset is a publicly available resting state EEG dataset of 213 young and old healthy participants acquired in Leipzig (Germany) to study mind-body-emotion interactions[26]. This dataset is divided into two groups based on the age of the participants: a 'young' group (20 to 35 years old; N = 143; 43 females) and an 'old' group (55 to 80 years old; N = 70; 35 females). One participant (male) had an intermediate age and was not included in the posterior statistical analyses (next Results section). Resting state EEG was recorded with a BrainAmp MR plus amplifier using 62 active ActiCAP electrodes (61 scalp electrodes in the 10-10 system positions and 1 VEOG below the right eye) provided by Brain Products Gmbh, Gilching, Germany. The ground electrode was located at the sternum, and the reference electrode was FCz. The recording sampling rate was 2500 Hz. Recordings contained 16 blocks of 1-minute duration, 8 with eyes closed and 8 with eyes open in an interleaved fashion. Before the execution of the pipeline, all blocks corresponding to eyes closed and eyes open conditions were extracted and concatenated by condition, including boundaries between blocks. In both conditions, two recordings were discarded because the data were truncated. In the eyes closed condition, an additional recording was discarded because there were no eyes closed marker events. Therefore, the final sample sizes were 210 recordings for eyes closed and 211 for eyes open. For both conditions, recordings of 8 minutes duration were entered into the pipeline.

TDBRAIN dataset. The TDBRAIN dataset is a publicly available, heterogenous resting state EEG dataset aggregated to obtain neurophysiological insights into psychiatric disorders[27]. It includes 1274 participants, most of them suffering from a psychiatric disorder. The primary diagnoses are Major Depressive Disorder (N = 426), Attention Deficit Hyperactivity Disorder (N = 271), Subjective Memory Complains (N = 119), and Obsessive-Compulsive Disorder (N = 75). The dataset also includes healthy participants (N = 47) and participants with unknown diagnoses (N = 255). Resting state EEG was recorded with a Compumedics Quickcap or ANT-Neuro Wavegurard Cap using 26 EEG Ag/AgCl electrodes positioned according to the 10-10 system. Additionally, five electrodes recorded vertical and horizontal eye movements, one electrode recorded muscle activity at the masseter, and one electrode captured the electrocardiogram at the clavicle bone. The ground electrode was placed at AFz, and recordings were referenced offline to averaged mastoids (A1 and A2). The recording sampling frequency was 500 Hz. Resting state data were acquired with eyes closed and eyes open in two respective blocks of 2 minutes duration. Some participants had more than one recording session, but only one recording per participant was entered into the pipeline. In total, resting state recordings for both eyes closed and eyes open conditions were available for 1274 individuals.

DISCOVER-EEG preprocessing. Table 1 and Fig. 4 show an overview of the DISCOVER-EEG preprocessing results, comparing datasets with different numbers of EEG electrodes (61 for the LEMON and 26 for the TDBRAIN dataset) and recording lengths (8 minutes for the LEMON and 2 minutes for the TDBRAIN datasets). These results point to a fair amount of data remaining after preprocessing and are in the same range as other automatic preprocessing pipelines[11,16]. They could be useful in future studies interested in benchmarking different EEG configurations or EEG preprocessing strategies.

DISCOVER-EEG validation. To validate the pipeline's outcomes, we replicated the well-known physiological effect of alpha power attenuation in eyes open compared to eyes closed conditions in both datasets. To this end, we estimated the power spectrum for each recording and condition as described in the feature extraction section. For visualization, we performed a grand average across participants and channels for each dataset and condition (Fig. 4). To test for differences between conditions, we performed a dependent samples cluster-based permutation test[64] across frequencies in the range of 1 to 100 Hz. The tests were two-tailed and based on 500 randomizations. Cluster-level statistics were calculated by taking the sum of the t-values within each cluster.

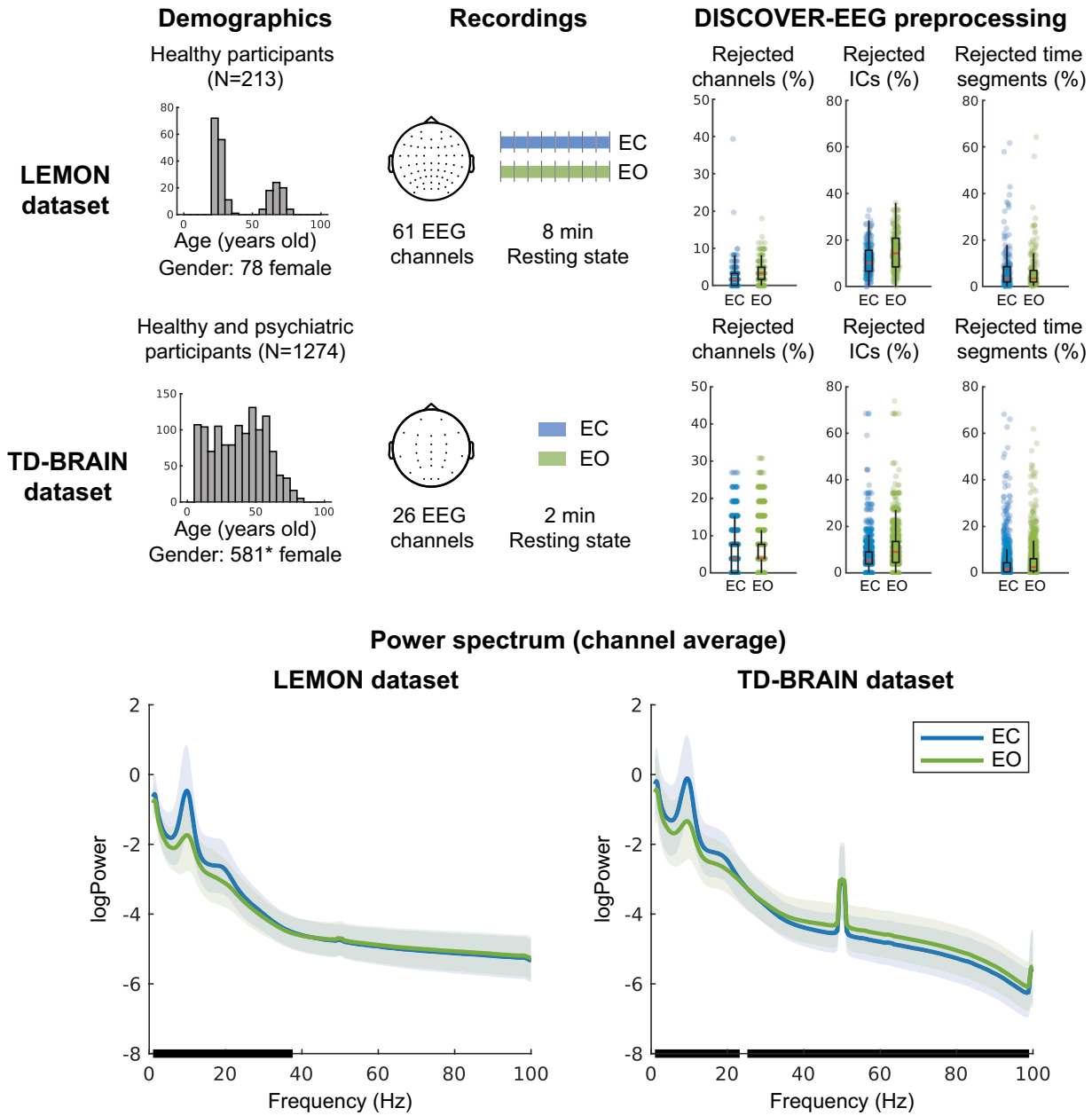

Fig. 4 Overview of the LEMON and TDBRAIN datasets, results of their preprocessing for resting state eyes closed (blue) and eyes open (green), and average power spectra across participants and channels. **Demographics** includes a histogram depicting the age of the participants in bins of 5 years. **Recordings** depicts the EEG setup layout and the recordings' average duration. Vertical bars in the recording duration of the LEMON dataset represent the boundaries between the concatenated 1-minute blocks in which the data was acquired. **DISCOVER-EEG preprocessing** presents the fractions of rejected channels, rejected ICs, and rejected time segments. Boxplots visualize the distribution of these variables. The median is indicated by a red horizontal line, and the first and third quartiles are indicated with black boxes. The whiskers extend to 1.5 times the interquartile range. Blue dots overlaid to the boxplots represent individual recordings. **Power spectrum** depicts the grand average power spectra across participants and channels for eyes open and eyes closed conditions. Shaded areas indicate the standard deviation of the power spectra across participants. Black bars over the x axis indicate frequency intervals of significant clusters between conditions (dependent samples cluster-based permutation test). Note that in the TDBRAIN dataset, line noise could not be completely removed due to high levels of noise in some recordings. EC = Eyes Closed. EO = Eyes Open. ICs = Independent Components. *The gender of 18 participants was missing in the TDBRAIN dataset.

The cluster threshold and alpha level were set to 0.05. In the LEMON dataset, a significant positive cluster indicated higher power values in the eyes closed condition in the 1 to 37.2 Hz range ($p = 0.002$, cluster statistic = 1310). In the TDBRAIN dataset, a positive cluster was found in the 1 to 23.2 Hz range ($p = 0.002$, cluster

statistic = 2037) along with a negative cluster in the 25.2 to 99 Hz range ($p = 0.002$, cluster statistic = −3254). These results indicate lower power in the eyes open condition in the alpha range (8 to 12.9 Hz) for both datasets, as shown with previous analyses of the TDBRAIN dataset[27].

Example analysis with DISCOVER-EEG. To exemplify how DISCOVER-EEG features could be aggregated and analyzed, we conducted an exploratory analysis of the LEMON dataset to investigate age-related differences in resting state EEG. Recently, the measure of brain age, i.e., the expected level of cognitive function of a person with the same chronological age, has been proven a valuable marker of cognitive decline in healthy and clinical populations[65,66]. Brain age is usually estimated by machine and deep learning models trained to predict the chronological age of a participant based on neuroimaging data. One criticism of these methods is their limited neuroscientific interpretability, as it is not straightforward which features the models used to predict brain age[65]. It is, therefore, highly useful to explore which physiologically meaningful brain features change with age.

To this end, we statistically compared the features automatically extracted by the pipeline in the 'old' and 'young' groups using Bayesian statistics performed in Matlab with the bayesFactor package[67] (Fig. 5). We also provide the scripts used to perform this group analysis and the visualization of group-level data together with the pipeline code.

We tested whether old participants had different APF values than young participants with two-sided independent samples Bayesian t-tests. Two tests were performed, one for each APF measure (the local maximum and the center of gravity). Results show very strong evidence in favor of the alternative hypothesis, i.e., a lower APF in the old compared to the young group, when the APF is computed as local maximum peak ($BF_{10} = 325.5$), but inconclusive evidence when the APF is computed as the center of gravity ($BF_{10} = 1.1$) (Fig. 5, first row).

We further tested whether there were differences in the connectivity matrices between young and old participants. For each connectivity measure (dwPLI and AEC) and frequency band (theta, alpha, beta, and gamma), we compared the connectivity values of each undirected source pair between the young and the old groups. Thus, we performed 9900 two-sided independent samples Bayesian t-tests per connectivity matrix. In Fig. 5, second row, we depict t-values color-coded to show the direction of effects. Statistical tests showing strong evidence in favor ($BF_{10} > 30$) or against ($BF_{10} < 1/30$) of the alternative hypothesis are not faded out. We observed strong evidence in favor of a reduction of phase-based connectivity in old participants, predominantly in the alpha band as well as a reduction of amplitude-based connectivity (Fig. 5, second row, non-masked blue values).

We finally tested whether there were any differences between old and young participants in the graph measures. For the local graph measures, we performed one test per source location, i.e., 100 independent sample Bayesian t-tests for each graph measure, connectivity measure, and frequency band. For the global measures, we performed a two-sided independent sample Bayesian t-test per graph measure, connectivity measure, and frequency band. With regard to global measures, the most prevalent differences appeared at low frequencies (theta and alpha) using the AEC (Fig. 5, third row, blue and red dots in brain sketches have $BF_{10} > 30$). The strongest effects concerning global measures are a reduction of global efficiency and smallworldness of the older group in the beta band for the dwPLI and the alpha band for the AEC (Fig. 5, third row, raincloud plots with inset BF_{10} indicating strong evidence). Together, these results show a reduction of local connectivity at theta and alpha frequencies and an increase in network integration at alpha and beta frequencies in older participants.

Overall, the current findings align with previous EEG literature, which has reported a slowing of APF and a general decrease in functional connectivity and network integrity in older individuals[44,68].

This example analysis could inspire future studies aimed at discovering explainable biomarkers of healthy aging or risk biomarkers of cognitive decline. However, it does not intend to present a ready-to-use biomarker for aging, which would require validation beyond the scope of this manuscript.

Discussion

Here, we present DISCOVER-EEG, an open and fully automated pipeline that enables fast and easy aggregation, preprocessing, analysis, and visualization of resting state EEG data. The current pipeline builds upon state-of-the-art automated preprocessing elements and extends them by including the computation and visualization of physiologically relevant EEG features. These EEG features follow recent COBIDAS guidelines for MEEG research[19], are implemented in widely used EEG toolboxes, and have been repeatedly associated with cognitive and behavioral measures in healthy as well in neuropsychiatric populations. Therefore, they could represent promising biomarker candidates for neurological and psychiatric disorders.

Importantly, this pipeline presents one of many ways of preprocessing and analyzing resting state data and is not intended to be the ultimate solution for EEG analysis. Instead, it aims to represent a reasonable and pragmatic realization for accelerating the acquisition, preprocessing, and analysis of large-scale datasets with the potential to discover physiologically plausible and interpretable biomarkers. To this end, we adopted a previously published, simple, and robust preprocessing strategy and selected a well-defined set of EEG features that could be useful for biomarker discovery. However, this pipeline might not suit all populations, settings, or research paradigms, such as children, real-life ecological EEG assessments, or paradigms assessing event-related potentials. Additionally, specific EEG layouts, such as those with few electrodes and limited scalp coverage, might not be adequate for source localization or average referencing. However, it is compatible with different types of EEG systems, and its focus on resting state data makes it suitable for patients and healthy populations. The features extracted by DISCOVER-EEG represent a basis for developing neurophysiologically plausible and interpretable biomarkers of neurological and psychiatric disorders. They can be easily extended and adapted to specific study designs, as the modular configuration of the code allows for substituting, removing, or adding specific steps of the preprocessing and feature extraction.

By default, the pipeline uses a realistically shaped volume conduction model based on the Montreal Neurological Institute (MNI) template for source localization. For optimal accuracy of source reconstruction,

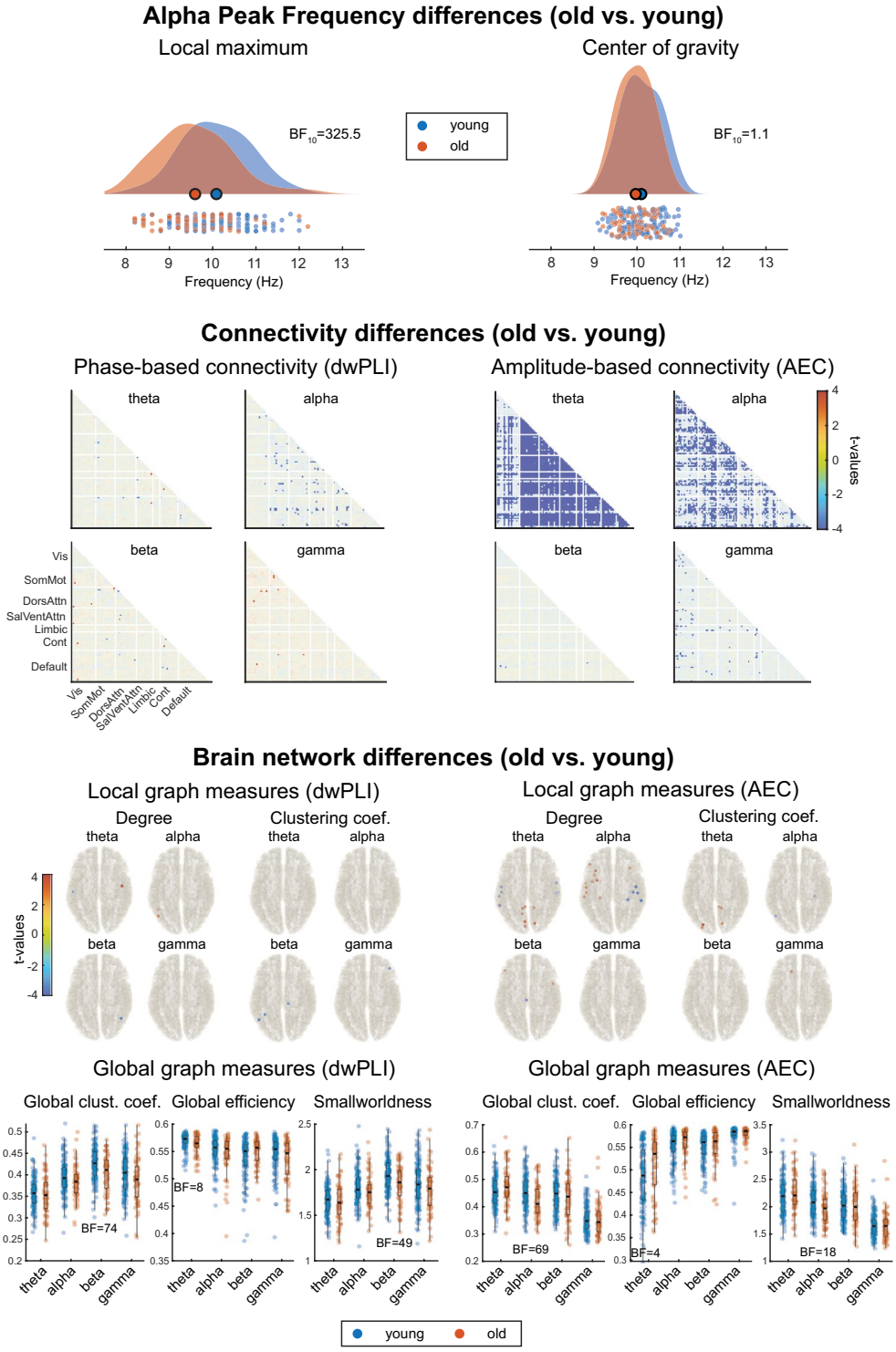

Fig. 5 Age-related differences in resting state EEG between old and young participants of the LEMON dataset. **Alpha Peak Frequency differences** between the old and the young group are visualized using raincloud plots[70]. **Connectivity differences** between the old and young groups. Blue values indicate lower connectivity in the old group. All connections not showing strong evidence for or against a connectivity difference ($1/30 < BF_{10} < 30$) are faded out. **Brain network differences** between the old and young group. With regard to local graph measures, only locations with strong evidence for ($BF_{10} > 30$) or against ($BF_{10} < 1/30$) a difference in the graph measures are depicted. Only results in favor of a difference between groups were found (blue dots indicate lower local graph measures in the old group, and red dots indicate higher local graph measures in the old group). With regard to global graph measures, only BF_{10} showing substantial evidence for ($BF_{10} > 3$) or against ($BF_{10} < 1/3$) the alternative hypothesis were included as insets to facilitate the reading. AEC = Amplitude Envelope Correlation; BF = Bayes Factor; Cont = Control; DorsAttn = Dorsal Attention; dwPLI = debiased weighted Phase Lag Index; SalVentAttn = Salience-Ventral Attention; SomMot = Somato-Motor; Vis = Visual.

individual MRIs or at least individual electrode positions could be recorded along with the EEG data. However, this would substantially increase the effort and time needed to acquire data and would hinder the fast generation of large new datasets. For that reason, the pipeline uses a generic template for source localization. However, it might be desirable to have different templates that better reflect the variability of head shapes in the future.

On a broader scope, it should always be considered whether datasets used for biomarker discovery are representative of the population of interest or whether they are biased towards young, Caucasian, highly educated populations, as is often the case[69]. The participants of the studies used here to test the pipeline were recruited based on convenience sampling and, therefore, might not cover the entire population. The creation and adoption of data standards such as BIDS will help to mitigate this fact by promoting collaboration and data sharing around the world.

Our intention with this pipeline was to push the field of EEG biomarker discovery forward to acquiring and analyzing large datasets, as needed in neuroimaging and artificial intelligence. Moreover, the provided example analysis can serve as a starting point for researchers who want to use complex measures of brain function. However, we do not claim to present a ready-to-use biomarker for aging, nor that the approach we propose is the best or only way to develop such a biomarker. Instead, it is intended as an offer to the community to assess physiologically plausible and interpretable EEG features in an efficient, transparent, and reproducible manner. Therefore, it can help the discovery of EEG-based biomarkers in neuropsychiatric disorders and promotes and facilitates open and reproducible assessments of brain function in EEG communities and beyond.

Data availability

The LEMON dataset[62] was accessed via http from the Max Plax Institute Leipzig webpage, and is described in an accompanying publication[26].

The TDBRAIN dataset[63] was accessed from the webpage of Brainclinics Foundation, and is described in an accompanying publication[27].

Code availability

The EEG pipeline code is available at GitHub under the CC-BY 4.0 license, and it is co-deposited in Zenodo, and referenced with a unique DOI[30].

The pipeline was created and tested in Matlab 2020b (The Mathworks, Inc.) on Ubuntu 18.04.5 LTS with the Signal Processing and Statistical and Machine Learning Toolboxes installed. EEGLab (v2022.0)[24] with the plugins bids-matlab-tools (v6.1), bva-io (v1.7), firfilt, (v2.4), cleanLine (v2.0), ICLabel (v1.3), clean_rawdata (v2.6) and dipfilt (v4.3) were installed and used for preprocessing. FieldTrip (revision ee916f5e5)[25] was used for source reconstruction and EEG feature extraction, and the Brain Connectivity Toolbox (version 03 2019)[45] was used for network analysis.

Received: 19 May 2023; Accepted: 31 August 2023;
Published online: 11 September 2023

References

1. Woo, C. W., Chang, L. J., Lindquist, M. A. & Wager, T. D. Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci* **20**, 365–377, https://doi.org/10.1038/nn.4478 (2017).
2. FDA-NIH Biomarker Working Group. BEST *(Biomarkers, EndpointS, and other Tools) Resource*, https://www.ncbi.nlm.nih.gov/books/NBK326791/ (2016).
3. Botvinik-Nezer, R. *et al.* Variability in the analysis of a single neuroimaging dataset by many teams. *Nature* **582**, 84–88, https://doi.org/10.1038/s41586-020-2314-9 (2020).
4. Marek, S. *et al.* Reproducible brain-wide association studies require thousands of individuals. *Nature* **603**, 654–660, https://doi.org/10.1038/s41586-022-04492-9 (2022).
5. Drysdale, A. T. *et al.* Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med* **23**, 28–38, https://doi.org/10.1038/nm.4246 (2017).
6. Wu, W. *et al.* An electroencephalographic signature predicts antidepressant response in major depression. *Nat Biotechnol* **38**, 439–447, https://doi.org/10.1038/s41587-019-0397-3 (2020).
7. Zhang, Y. *et al.* Identification of psychiatric disorder subtypes from functional connectivity patterns in resting-state electroencephalography. *Nat Biomed Eng* **5**, 309–323, https://doi.org/10.1038/s41551-020-00614-8 (2021).
8. Toll, R. T. *et al.* An Electroencephalography Connectomic Profile of Posttraumatic Stress Disorder. *Am J Psychiatry* **177**, 233–243, https://doi.org/10.1176/appi.ajp.2019.18080911 (2020).
9. Ta Dinh, S. *et al.* Brain dysfunction in chronic pain patients assessed by resting-state electroencephalography. *Pain* **160**, 2751–2765, https://doi.org/10.1097/j.pain.0000000000001666 (2019).
10. Pernet, C. R. *et al.* EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Sci Data* **6**, 103, https://doi.org/10.1038/s41597-019-0104-8 (2019).
11. Gabard-Durnam, L. J., Mendez Leal, A. S., Wilkinson, C. L. & Levin, A. R. The Harvard Automated Processing Pipeline for Electroencephalography (HAPPE): Standardized Processing Software for Developmental and High-Artifact Data. *Front Neurosci* **12**, 97, https://doi.org/10.3389/fnins.2018.00097 (2018).
12. Debnath, R. *et al.* The Maryland analysis of developmental EEG (MADE) pipeline. *Psychophysiology* **57**, e13580, https://doi.org/10.1111/psyp.13580 (2020).
13. Klug, M. *et al.* The BeMoBIL Pipeline for automated analyses of multimodal mobile brain and body imaging data. *bioRxiv* https://doi.org/10.1101/2022.09.29.510051 (2022).
14. Pernet, C. R., Martinez-Cancino, R., Truong, D., Makeig, S. & Delorme, A. From BIDS-Formatted EEG Data to Sensor-Space Group Results: A Fully Reproducible Workflow With EEGLAB and LIMO EEG. *Front Neurosci* **14**, 610388, https://doi.org/10.3389/fnins.2020.610388 (2020).
15. Pedroni, A., Bahreini, A. & Langer, N. Automagic: Standardized preprocessing of big EEG data. *NeuroImage* **200**, 460–473, https://doi.org/10.1016/j.neuroimage.2019.06.046 (2019).
16. Rodrigues, J., Weiss, M., Hewig, J. & Allen, J. J. B. EPOS: EEG Processing Open-Source Scripts. *Frontiers in Neuroscience* **15**, https://doi.org/10.3389/fnins.2021.660449 (2021).

17. Bailey, N. W. *et al*. Introducing RELAX: An automated pre-processing pipeline for cleaning EEG data- Part 1: Algorithm and application to oscillations. *Clinical Neurophysiology* **149**, 178–201, https://doi.org/10.1016/j.clinph.2023.01.017 (2023).
18. Appelhoff, S. *et al*. MNE-BIDS: Organizing electrophysiological data into the BIDS format and facilitating their analysis. *J Open Source Softw* **4**, https://doi.org/10.21105/joss.01896 (2019).
19. Pernet, C. *et al*. Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research. *Nat Neurosci* **23**, 1473–1483, https://doi.org/10.1038/s41593-020-00709-0 (2020).
20. Newson, J. J. & Thiagarajan, T. C. EEG Frequency Bands in Psychiatric Disorders: A Review of Resting State Studies. *Frontiers in human neuroscience* **12**, https://doi.org/10.3389/fnhum.2018.00521 (2019).
21. Ploner, M. & Tiemann, L. Exploring Dynamic Connectivity Biomarkers of Neuropsychiatric Disorders. *Trends in cognitive sciences* **25**, 336–338, https://doi.org/10.1016/j.tics.2021.03.005 (2021).
22. Bassett, D. S., Xia, C. H. & Satterthwaite, T. D. Understanding the Emergence of Neuropsychiatric Disorders With Network Neuroscience. *Biol Psychiatry Cogn Neurosci Neuroimaging* **3**, 742–753, https://doi.org/10.1016/j.bpsc.2018.03.015 (2018).
23. de Lange, S. C. *et al*. Shared vulnerability for connectome alterations across psychiatric and neurological brain disorders. *Nat Hum Behav* **3**, 988–998, https://doi.org/10.1038/s41562-019-0659-6 (2019).
24. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods* **134**, 9–21, https://doi.org/10.1016/j.jneumeth.2003.10.009 (2004).
25. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J. M. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience* **2011**, 156869, https://doi.org/10.1155/2011/156869 (2011).
26. Babayan, A. *et al*. A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults. *Sci Data* **6**, 180308, https://doi.org/10.1038/sdata.2018.308 (2019).
27. van Dijk, H. *et al*. The two decades brainclinics research archive for insights in neurophysiology (TDBRAIN) database. *Sci Data* **9**, 333, https://doi.org/10.1038/s41597-022-01409-z (2022).
28. Adrian, E. D. & Matthews, B. H. C. The Berger rhythm potential changes from the occipital lobes in man. *Brain* **57**, 355–385, https://doi.org/10.1093/brain/57.4.355 (1934).
29. Barker, M. *et al*. Introducing the FAIR Principles for research software. *Sci Data* **9**, 622, https://doi.org/10.1038/s41597-022-01710-x (2022).
30. Gil Ávila, C., Bott, F. S., Gross, J. & Ploner, M. crisglav/discover-eeg: 1.0.0 v. 1.0.0. *Zenodo* https://doi.org/10.5281/zenodo.8207523 (2023).
31. Wilkinson, M. D. *et al*. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**, 160018, https://doi.org/10.1038/sdata.2016.18 (2019).
32. Cella, D. *et al*. The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap Cooperative Group During its First Two Years. *Medical Care* **45**, S3–S11, https://doi.org/10.1097/01.mlr.0000258615.42478.55 (2007).
33. Niso, G. *et al*. Open and reproducible neuroimaging: From study inception to publication. *NeuroImage* **263**, 119623, https://doi.org/10.1016/j.neuroimage.2022.119623 (2022).
34. R Core Team. R: A Language and Environement for Statistical Computing *R Foundation for Statistical Computing*, Vienna, Austria (2021).
35. Van Rossum, G. a. D., Fred L. *Python 3 Reference Manual*. (CreateSpace, 2009*).
36. Oostenveld, R. & Praamstra, P. The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology* **112**, 713–719, https://doi.org/10.1016/s1388-2457(00)00527-7 (2001).
37. Pion-Tonachini, L., Kreutz-Delgado, K. & Makeig, S. ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage* **198**, 181–197, https://doi.org/10.1016/j.neuroimage.2019.05.026 (2019).
38. Perrin, F., Pernier, J., Bertrand, O. & Echallier, J. F. Spherical splines for scalp potential and current density mapping. *Electroencephalogr Clin Neurophysiol* **72**, 184–187, https://doi.org/10.1016/0013-4694(89)90180-6 (1989).
39. Mullen, T. R. *et al*. Real-Time Neuroimaging and Cognitive Monitoring Using Wearable Dry EEG. *Ieee Transactions on Biomedical Engineering* **62**, 2553–2567, https://doi.org/10.1109/Tbme.2015.2481482 (2015).
40. Fraschini, M. *et al*. The effect of epoch length on estimated EEG functional connectivity and brain network organisation. *Journal of neural engineering* **13**, 036015, https://doi.org/10.1088/1741-2560/13/3/036015 (2016).
41. van Diessen, E. *et al*. Opportunities and methodological challenges in EEG and MEG resting state functional brain network research. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology* **126**, 1468–1481, https://doi.org/10.1016/j.clinph.2014.11.018 (2015).
42. Cohen, M. X. *Analyzing Neural Time Series Data: Theory and Practice*. (MIT Press, 2014).
43. Haegens, S., Cousijn, H., Wallis, G., Harrison, P. J. & Nobre, A. C. Inter- and intra-individual variability in alpha peak frequency. *NeuroImage* **92**, 46–55, https://doi.org/10.1016/j.neuroimage.2014.01.049 (2014).
44. Scally, B., Burke, M. R., Bunce, D. & Delvenne, J. F. Resting-state EEG power and connectivity are associated with alpha peak frequency slowing in healthy aging. *Neurobiol Aging* **71**, 149–155, https://doi.org/10.1016/j.neurobiolaging.2018.07.004 (2018).
45. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* **52**, 1059–1069, https://doi.org/10.1016/j.neuroimage.2009.10.003 (2010).
46. Badhwar, A. *et al*. Resting-state network dysfunction in Alzheimer's disease: A systematic review and meta-analysis. *Alzheimers Dement (Amst)* **8**, 73–85, https://doi.org/10.1016/j.dadm.2017.03.007 (2017).
47. Gallo, S. *et al*. Functional connectivity signatures of major depressive disorder: machine learning analysis of two multicenter neuroimaging studies. *Mol Psychiatry* https://doi.org/10.1038/s41380-023-01977-5 (2023).
48. Cao, J. *et al*. Brain functional and effective connectivity based on electroencephalography recordings: A review. *Hum Brain Mapp* **43**, 860–879, https://doi.org/10.1002/hbm.25683 (2022).
49. Engel, A. K., Gerloff, C., Hilgetag, C. C. & Nolte, G. Intrinsic coupling modes: multiscale interactions in ongoing brain activity. *Neuron* **80**, 867–886, https://doi.org/10.1016/j.neuron.2013.09.038 (2013).
50. Vinck, M., Oostenveld, R., van Wingerden, M., Battaglia, F. & Pennartz, C. M. An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *NeuroImage* **55**, 1548–1565, https://doi.org/10.1016/j.neuroimage.2011.01.055 (2011).
51. Hipp, J. F., Hawellek, D. J., Corbetta, M., Siegel, M. & Engel, A. K. Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nat Neurosci* **15**, 884–890, https://doi.org/10.1038/nn.3101 (2012).
52. Corcoran, A. W., Alday, P. M., Schlesewsky, M. & Bornkessel-Schlesewsky, I. Toward a reliable, automated method of individual alpha frequency (IAF) quantification. *Psychophysiology* **55**, e13064, https://doi.org/10.1111/psyp.13064 (2018).
53. Klimesch, W., Schimke, H. & Pfurtscheller, G. Alpha frequency, cognitive load and memory performance. *Brain Topogr* **5**, 241–251, https://doi.org/10.1007/BF01128991 (1993).
54. Bastos, A. M. & Schoffelen, J. M. A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Frontiers in systems neuroscience* **9**, 175, https://doi.org/10.3389/fnsys.2015.00175 (2015).
55. Pellegrini, F., Delorme, A., Nikulin, V. & Haufe, S. Identifying good practices for detecting inter-regional linear functional connectivity from EEG. *NeuroImage* **277**, 120218, https://doi.org/10.1016/j.neuroimage.2023.120218 (2023).

56. Van Veen, B. D., van Drongelen, W., Yuchtman, M. & Suzuki, A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE transactions on bio-medical engineering* **44**, 867–880, https://doi.org/10.1109/10.623056 (1997).
57. Schaefer, A. *et al*. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex* **28**, 3095–3114, https://doi.org/10.1093/cercor/bhx179 (2018).
58. Westner, B. U. *et al*. A unified view on beamformers for M/EEG source reconstruction. *NeuroImage* **246**, 118789, https://doi.org/10.1016/j.neuroimage.2021.118789 (2022).
59. Pourmotabbed, H., de Jongh Curry, A. L., Clarke, D. F., Tyler-Kabara, E. C. & Babajani-Feremi, A. Reproducibility of graph measures derived from resting-state MEG functional connectivity metrics in sensor and source spaces. *Hum Brain Mapp* **43**, 1342–1357, https://doi.org/10.1002/hbm.25726 (2022).
60. Adamovich, T., Zakharov, I., Tabueva, A. & Malykh, S. The thresholding problem and variability in the EEG graph network parameters. *Scientific Reports* **12**, 18659, https://doi.org/10.1038/s41598-022-22079-2 (2022).
61. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442, https://doi.org/10.1038/30918 (1998).
62. Babayan, A. *et al. Functional Connectomes Project International Neuroimaging Data-Sharing Initiative* https://doi.org/10.15387/fcp_indi.mpi_lemon (2018).
63. van Dijk, H. *et al*. Two Decades - Brainclinics Research Archive for Insights in Neuroscience (TD-BRAIN). *Synapse* https://doi.org/10.7303/syn25671079 (2021).
64. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *Journal of neuroscience methods* **164**, 177–190 (2007).
65. Cole, J. H. & Franke, K. Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers. *Trends Neurosci* **40**, 681–690, https://doi.org/10.1016/j.tins.2017.10.001 (2017).
66. Engemann, D. A. *et al*. A reusable benchmark of brain-age prediction from M/EEG resting-state signals. *NeuroImage* **262**, 119521, https://doi.org/10.1016/j.neuroimage.2022.119521 (2022).
67. Krekelberg, B. BayesFactor: Release 2022 v. 2.3.0. *Zenodo* https://doi.org/10.5281/zenodo.7006300 (2022).
68. Onoda, K., Ishihara, M. & Yamaguchi, S. Decreased Functional Connectivity by Aging Is Associated with Cognitive Decline. *J Cognitive Neurosci* **24**, 2186–2198, https://doi.org/10.1162/jocn_a_00269 (2012).
69. Henrich, J., Heine, S. J. & Norenzayan, A. The weirdest people in the world? *Behav Brain Sci* **33**, 61–83; discussion 83–135, https://doi.org/10.1017/S0140525X0999152X (2010).
70. Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R. & Kievit, R. A. Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Res* **4**, 63, https://doi.org/10.12688/wellcomeopenres.15191.1 (2019).

Acknowledgements

The study has been supported by the TUM Innovation Network Neurotechnology for Mental Health (NEUROTECH), the Deutsche Forschungsgemeinschaft (PL321/14-1) and the TUM School of Medicine (KKF).

Author contributions

Conceptualization: C.G.A. and M.P., Methodology: C.G.A. and M.P., Software: C.G.A., F.S.B., and J.G., Validation: C.G.A., F.S.B., and J.G., Formal analysis: C.G.A., Investigation: C.G.A. and P.T.Z., Data curation: C.G.A., Visualization: C.G.A. and M.P., Writing – Original draft: C.G.A. and M.P., Writing – Review and editing: C.G.A., F.S.B., L.T., V.D.H., E.S.M., M.M.N., P.T.Z., J.G., and M.P., Supervision: M.P., Project administration: M.P., Funding acquisition: M.P.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

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

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.

4. DISCUSSION

In this work, we used electroencephalography to investigate the development of brain biomarkers in translational neuroscience. In the first project, we explored the brain's temporal dynamics with microstate analysis in a diverse group of patients with chronic pain. In the second project, we developed DISCOVER-EEG, an open EEG pipeline to automatically preprocess, analyze, and visualize resting state EEG data. Following the most recent EEG methods, guidelines, and standards, we explored a biomarker candidate in chronic pain and proposed a new tool for fast biomarker identification across neurological and psychiatric disorders. With an open science mindset, we aimed to promote transparent and reproducible research on brain function.

4.1. PROJECT 1. EEG MICROSTATE ANALYSIS IN PATIENTS WITH CHRONIC PAIN

In this study, we investigated brain dynamics in patients suffering from different types of chronic pain. We, therefore, applied microstate analysis to a large, cross-sectional resting state EEG dataset, including patients with chronic pain and healthy participants. We found that EEG data was consistently described in both groups with five microstates previously described in the literature and labeled A to E. Bayesian statistics indicated that the presence of microstate D was decreased in patients with chronic pain, while there was no consistent evidence for changes in other microstates. However, temporal characteristics of microstate D did not correlate with patients' clinical scores. Subgroup analysis based on specific chronic pain pathologies replicated a decreased presence of microstate D in patients with chronic back pain but not in patients with chronic widespread pain. These results indicate that alterations of brain dynamics measured with microstate analysis might be specific for certain types of chronic pain. However, due to the limited physiological interpretability of microstates and the technical limitations of this method, further studies are needed to consider microstates' characteristics a biomarker of chronic pain.

Limitations and future work

Microstates are thought to emerge from coordinated neural activity across the brain, with transitions between microstates usually interpreted as "sequential activation of different neuronal networks" (Khanna et al., 2015). Yet, the functional interpretation of microstates is still uncertain. Microstates have been associated across the literature with fMRI resting state networks, cognitive processes, and disease states (Khanna et al., 2015; Michel and Koenig, 2018). Specifically, microstate D, which we found to be less present in chronic pain, has been associated with attentional brain networks and functions (Milz et al., 2016). Thus, our observed changes in microstate D might reflect attentional deficits in patients with chronic pain, which have been previously reported (Moriarty et al., 2011). However, as we have not obtained direct measures of attentional functioning in this study, we cannot test this hypothesis. Follow-up microstate studies in chronic pain might include tasks and questionnaires specifically assessing attentional functions.

Changes in microstate D have also been observed in other brain disorders, such as schizophrenia (da Cruz et al., 2020) and major depressive disorder (Murphy et al., 2020). Therefore, microstate D alterations might not be specific to chronic pain. This, nevertheless, does not diminish the potential clinical utility of microstates as diagnostic or subtyping biomarkers of chronic pain. Symptoms such as pain and depression are commonly shared across neurological and psychiatric disorders. Similarly, altered brain functions and dynamics could be shared across neuropsychiatric disorders as well (Scangos et al., 2023). Further studies using cross-disorder datasets could test the specificity of microstate D to chronic pain or specific brain networks.

In contrast to our results, which did not show any microstate changes in chronic widespread pain, a recent study published lower occurrence and time coverage of microstate C in resting state eyes-open recordings of patients suffering from chronic widespread pain (Gonzalez-Villar et al., 2020). This discrepancy might be due to our smaller sample size —30 patients with chronic widespread pain with respect to 47 of Gonzalez-Villar et al. (2020)— but also due to methodological differences between studies and the limitations of microstate analysis itself.

In microstate analysis, the definition of microstates is based on a non-deterministic clustering procedure in which an a priori number of microstates has to be defined by the researcher. The non-deterministic nature of the algorithm leads to variability across results when repeating the analysis with the same data and parameters. For this reason, we reported the results of 5 repetitions of the microstate analysis, performed separately on resting state eyes-closed and eyes-open data. Our most consistent and replicable finding was a reduced presence of microstate D in chronic pain during resting state with eyes closed. This effect, however, could not be replicated in resting state with eyes open, where the number of microstates varied notably across repetitions.

Additionally, the software used to define the microstates and their temporal characteristics was provided as a compiled executable, precluding any audit of the source code to quantify the sources of variability in the implementation of the algorithm. Based on this experience, we recommend that future studies use an open-source implementation of microstate analysis, such as Ragu (Habermann et al., 2018). Despite having fewer functionalities, Ragu provides more transparent access to microstate analysis. In our view, further efforts to increase the transparency and robustness of microstate analysis should be pushed forward.

Conclusion

This study contributed to the understanding of the brain dynamics of patients suffering from different types of chronic pain. Our results presented differences in the brain dynamics between patients with chronic pain and healthy participants, specifically a decreased presence of microstate D in patients, which putatively relates to attention deficits during chronic pain. Additionally, our results point towards different brain dynamics in patients with chronic back pain and chronic widespread pain. If future studies demonstrate that these results are generalizable, they could serve to establish new treatments based on neural features of chronic pain (Baron et al., 2023). Overall, we showed the utility of electroencephalography for analyzing brain dynamics in chronic pain. This work can inspire future studies aiming to find EEG-based biomarkers of chronic pain.

4.2. PROJECT 2. DISCOVER-EEG: AN AUTOMATIC EEG PIPELINE FOR BIOMARKER DISCOVERY

In this study, we developed DISCOVER-EEG, an open, fully automated pipeline to enable fast and easy aggregation, preprocessing, analysis, and visualization of resting state EEG data. We extended state-of-the-art EEG preprocessing pipelines by integrating them with the computation and visualization of physiologically meaningful EEG features, such as oscillatory power, connectivity, and graph theory network measures. These features are based on the most recent EEG guidelines for MEEG research (Pernet et al., 2020a) and have been repeatedly associated with alterations in brain disorders (Cao et al., 2022; Newson and Thiagarajan, 2019). Therefore, they represent promising EEG-based biomarker candidates for brain disorders. The pipeline was tested in two publicly available resting state EEG datasets, the LEMON dataset, containing 213 EEG recordings of healthy participants (Babayan et al., 2019), and the TD-BRAIN dataset, including 1274 participants suffering from different psychiatric conditions (van Dijk et al., 2022). We also performed an exploratory analysis of the LEMON dataset that could inspire the identification of biomarkers of healthy aging. Approaching this endeavor with an open science mindset, we followed the FAIR principles of scientific software management (Wilkinson et al., 2019). Thus, the pipeline code is publicly available and open for contribution (Gil Ávila et al., 2023a). With DISCOVER-EEG, we aim to facilitate the aggregation, reuse, and analysis of large EEG datasets and promote transparent and reproducible research on brain function.

Limitations and future work

DISCOVER-EEG presents one out of many possible ways of preprocessing and analyzing resting state EEG data. Therefore, we do not intend to imply that it is the unique or best solution for EEG analysis, as this endeavor is neither possible nor necessary (Pernet et al., 2020a). Instead, DISCOVER-EEG represents a reasonable and pragmatic way of automatically processing resting state data to extract physiologically meaningful features that have the potential to turn into explainable biomarkers. To achieve this, we based the preprocessing on a previously published, simple, and robust pipeline for analyzing ERPs (Pernet et al., 2020b) and optimized it for resting state recordings. Consequently, specific populations, settings, and study designs might not benefit from the

proposed preprocessing strategy and feature extraction. We recommend applying DISCOVER-EEG cautiously to very contaminated data, such as that recorded in children or during motion, and EEG settings in which the number of electrodes is low and do not cover the entire scalp. In this latter case, average referencing and source reconstruction are not indicated.

Despite these specific cases, DISCOVER-EEG is robust and compatible with many settings and study designs. We tested it on two datasets with different characteristics (EEG system, electrode layout, number of electrodes, sampling rate, recording length, and population type), and we found that the preprocessing outcome was in the same range as other automatic preprocessing pipelines (Gabard-Durnam et al., 2018; Rodrigues et al., 2021). We could also replicate a well-known EEG effect in both datasets: the reduction of alpha power in resting state eyes open compared to eyes closed. This proves the reliability of the pipeline to capture a well-known effect in EEG.

DISCOVER-EEG features represent only a basis for developing neurophysiologically plausible and interpretable biomarkers. Other approaches that do not require the intermediate extraction of hand-crafted features, such as machine or deep learning, might as well complement this pipeline and contribute to the generation and validation of clinical biomarkers. The adaptation and extension of DISCOVER-EEG to specific study designs is relatively simple due to the modular structure of the code. In this way, specific steps of the preprocessing and feature extraction can be substituted, removed, or added. Future extensions of DISCOVER-EEG might include its adaptation to analyze ERPs, which are widely investigated in the EEG community.

We also included an example analysis to inspire biomarkers of healthy aging. Importantly, it is not our intention to present a ready-to-use biomarker, as this task would naturally require extensive validation. Conversely, this analysis and visualization are an example of how to aggregate and analyze the features extracted by the pipeline. They should be taken as a starting point for researchers who want to use advanced EEG measures of brain function.

Finally, as large datasets are needed for biomarker discovery, the computation time of preprocessing and feature extraction can increase with large datasets. Therefore, future work could facilitate the execution of DISCOVER-EEG on high-performance

clusters, which would help to reduce computation time. Using software containers, such as Docker or Singularity (Kurtzer et al., 2017), would be helpful for this purpose and ensure cross-platform compatibility. Another possibility would be integrating DISCOVER-EEG with platforms such as Brainlife (brainlife.io) or DataLad (Halchenko et al., 2021). Brainlife allows the execution of neuroimaging pipelines on the cloud, using data stored in publicly available repositories. DataLad complements this platform by allowing the version control of neuroimaging data and metadata while being integrated with the biggest neuroimaging data-sharing repositories. In this way, raw data and derivatives (outputs from processed data) could be more efficiently managed and tracked, further easing data reuse and providing a basis for rigor and reproducibility (Niso et al., 2022).

Conclusion

We developed an open and automatic workflow to facilitate fast preprocessing, analysis, and visualization of resting EEG data across neurological and psychiatric disorders. It produces physiologically plausible and interpretable EEG features in an efficient, transparent, and reproducible manner. Thus, DISCOVER-EEG promotes transparency and reproducibility in neuroimaging, which is essential for developing clinically useful biomarkers.

4.3. IMPLICATIONS ACROSS PROJECTS

In this work, we have reviewed and addressed three of the current challenges that the biomarker framework poses. We have supported the creation and reuse of large datasets by making available a large dataset of chronic pain and healthy participants in the BIDS-EEG structure (project 1). We have created a reproducible and transparent workflow for generating explainable EEG biomarkers (project 2). In both projects, we have used electroencephalography, a common, safe, non-invasive neuroimaging modality that is widely used and has the potential to be easily deployed in research, clinical and ecological environments. However, our contribution covers only a small part of this huge nascent field. Therefore, we will discuss the limitations of this thesis and future work.

Limitations and future work

The work presented here includes only a subset of potential biomarker targets, i.e., microstates' characteristics derived from EEG microstate analysis, power-based EEG measures, and functional connectivity and network measures derived from resting state EEG. However, exciting new measures of brain function, informative about behavioral and clinical characteristics, are being unraveled. For example, the *excitation/inhibition ratio* has been recently assessed non-invasively at the whole brain level with several EEG measures (Ahmad et al., 2022). The aperiodic (1/f) component of the power spectrum is a particularly promising measure that would help to better characterize the power spectrum (Donoghue et al., 2020; Gao et al., 2017). Additionally, intermediate brain phenotypes, that do not characterize a disease state per se but indicate a fragility or risk of developing a brain disease, are an interesting concept worth exploring. *Brain age*, i.e., the expected level of cognitive function of a person with the same chronological age, is one of such intermediate phenotypes (Cole and Franke, 2017; Engemann et al., 2022). Initial evidence indicated that brain age is predictive of the risk of developing cardiovascular and neurological disorders (Cole and Franke, 2017). Thus, these measures deserve further investigation.

A general limitation of the biomarker framework presented here is that biomarkers rely mostly on correlations between brain and clinical features. However, correlation does not imply causation. The combination of causal information with neuroimaging techniques can offer new insights into the brain in health and disease (Siddiqi et al., 2022). Also, the integration of features from different imaging modalities and fields into *composite biomarkers* could help to mitigate this "causality gap" (Tracey et al., 2019). For instance, two modalities pointing to convergent results would enhance the credibility of the biomarker and enable stronger causal inference. New tools such as NeuroMaps (Markello et al., 2022) are working in this direction. Alternatively, *machine learning* and deep learning have shown considerable success in the integration of modalities in bioinformatics, genetics, and cancer research and hold, therefore, great promise in neuroscience (Abi-Dargham and Horga, 2016; Tracey et al., 2019). Future work assessing potential biomarkers might therefore use these algorithms as well.

Both projects presented in this thesis included datasets that were acquired through convenience sampling. This raises the question of the representativity, and generalizability of the results acquired in this work, but also of brain biomarkers in general. It is known that minority groups are underrepresented in neuroscientific studies, which affects the generalizability of biomarkers to different populations (Webb et al., 2022). Defining what is a representative population and how to gather sensible data regarding ethnicity is an enormous question that requires expert knowledge (Muller et al., 2023). Nevertheless, efforts towards the gathering of *diverse and representative datasets* that reflect the variability of human phenotypes should be encouraged. For that reason, integrating minority groups into research and technology development, as well as potentiating *patient involvement,* is crucial when developing new neurotechnologies, biomarkers, and treatments. Stakeholders' involvement is, in consequence, essential for developing generalizable and inclusive biomarkers.

4.4. CONTRIBUTION TO OPEN SCIENCE

Open science practices are paramount for progressing in the neuroscience field. Collaboration between research sites, wide adoption of standardized data structures, creation of inclusive and representative datasets, and patient and society engagement, especially including minority groups, are essential for the development of accurate and inclusive biomarkers.

This thesis aimed to conduct and promote transparent and open research on brain function. Our alignment with open science principles is, consequently, visible in several ways. First, we contributed to the sharing and reuse of neuroimaging data by making openly available a well-documented EEG resting state dataset of patients with chronic pain and healthy participants in the standard BIDS-EEG structure. Second, we promoted reproducibility and transparency in EEG research by developing an automatic pipeline for preprocessing and analysis resting state EEG data. This could facilitate the accurate reporting of future studies and mitigate flexible analysis and p-hacking (Munafo et al., 2017; Parsons et al., 2022). Third, the use of Bayesian statistics in both studies allowed quantifying the credibility of both null and alternative hypotheses, contributing to reducing the problem of p-hacking and fighting the publication bias in the literature, i.e.,

publishing only significant results (Munafo et al., 2017). Fourth, we allowed rapid scientific dissemination by publishing both articles as preprints in bioRxiv. Preprints are open-access versions of the scientific manuscript that are published before undergoing peer review. Open-access versions accelerate scientific discovery and collaboration and make knowledge accessible to both humans and machines (Niso et al., 2022). Finally, and in relation to the last point, the first publication was published as open access, eliminating the paywall barrier for this study.

As a result, adhering to these open science practices facilitated communication with leaders in the field, who gave advice on the implementation of the DISCOVER-EEG pipeline. Also, this communication worked bi-directionally and allowed to increase the robustness of existing EEG methods with the reporting of bugs and suggestions to existing GitHub repositories. Scientific dissemination of the results via participation at international conferences and social media led to a rapid awareness of our research by the neuroimaging and pain communities and the general public. As an example, the publication on bioRxiv on the 20th of January 2023 of the DISCOVER-EEG preprint allowed thousands of researchers to discover, read, test, and use the DISCOVER-EEG pipeline during the review process of this manuscript (to date, 31 July 2023, bioRxiv: >4200 abstract reads, >1200 pdf downloads; Twitter: >34.000 views, >360 likes, >110 tweets).

Thus, open science accelerates research and creates a lively debate, facilitating the integration of comments from people all around the world. It additionally promotes equity and reduces socio-economic barriers. For example, sharing data and code can provide tools and resources to those researchers who lack the resources to acquire the data themselves. In conclusion, with this work, we have promoted and contributed to the dissemination of open science practices and encouraged the creation of high-quality studies and fair practices on human brain research.

REFERENCES

Abi-Dargham, A. & Horga, G. The search for imaging biomarkers in psychiatric disorders. *Nat Med* 22, 1248-1255, doi:10.1038/nm.4190 (2016).

Ahmad, J. *et al.* From mechanisms to markers: novel noninvasive EEG proxy markers of the neural excitation and inhibition system in humans. *Transl Psychiatry* 12, 467, doi:10.1038/s41398-022-02218-z (2022).

Allen, C. & Mehler, D. M. A. Open science challenges, benefits and tips in early career and beyond. *PLoS Biol* 17, e3000246, doi:10.1371/journal.pbio.3000246 (2019).

Arns, M., Conners, C. K. & Kraemer, H. C. A decade of EEG Theta/Beta Ratio Research in ADHD: a meta-analysis. *J Atten Disord* 17, 374-383, doi:10.1177/1087054712460087 (2013).

Babayan, A. *et al.* A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults. *Sci Data* 6, 180308, doi:10.1038/sdata.2018.308 (2019).

Baker, A. P. *et al.* Fast transient networks in spontaneous human brain activity. *eLife* 3, e01867, doi:10.7554/eLife.01867 (2014).

Baker, J. T. *et al.* Functional connectomics of affective and psychotic pathology. *PNAS* 116, 9050-9059, doi:10.1073/pnas.1820780116 (2019).

Baliki, M. N. & Apkarian, A. V. Nociception, Pain, Negative Moods, and Behavior Selection. *Neuron* 87, 474-491, doi:10.1016/j.neuron.2015.06.005 (2015).

Baliki, M. N. *et al.* Chronic pain and the emotional brain: specific brain activity associated with spontaneous fluctuations of intensity of chronic back pain. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 26, 12165-12173, doi:10.1523/JNEUROSCI.3576-06.2006 (2006).

Baron, R., Dickenson, A. H., Calvo, M., Dib-Hajj, S. D. & Bennett, D. L. Maximizing treatment efficacy through patient stratification in neuropathic pain trials. *Nat Rev Neurol* 19, 53-64, doi:10.1038/s41582-022-00741-7 (2023).

Bassett, D. S., Xia, C. H. & Satterthwaite, T. D. Understanding the Emergence of Neuropsychiatric Disorders With Network Neuroscience. *Biol Psychiatry Cogn Neurosci Neuroimaging* 3, 742-753, doi:10.1016/j.bpsc.2018.03.015 (2018).

Botvinik-Nezer, R. *et al.* Variability in the analysis of a single neuroimaging dataset by many teams. *Nature* 582, 84-88, doi:10.1038/s41586-020-2314-9 (2020).

Cao, J. *et al.* Brain functional and effective connectivity based on electroencephalography recordings: A review. *Hum Brain Mapp* 43, 860-879, doi:10.1002/hbm.25683 (2022).

Cohen, S. P., Vase, L. & Hooten, W. M. Chronic pain: an update on burden, best practices, and new advances. *Lancet* 397, 2082-2097, doi:10.1016/S0140-6736(21)00393-7 (2021).

Cole, J. H. & Franke, K. Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers. *Trends Neurosci* 40, 681-690, doi:10.1016/j.tins.2017.10.001 (2017).

da Cruz, J. R. *et al.* EEG microstates are a candidate endophenotype for schizophrenia. *Nature communications* 11, 3089, doi:10.1038/s41467-020-16914-1 (2020).

Davis, K. D. *et al.* Discovery and validation of biomarkers to aid the development of safe and effective pain therapeutics: challenges and opportunities. *Nat Rev Neurol* 16, 381-400, doi:10.1038/s41582-020-0362-2 (2020).

de Lange, S. C. *et al.* Shared vulnerability for connectome alterations across psychiatric and neurological brain disorders. *Nat Hum Behav* 3, 988-998, doi:10.1038/s41562-019-0659-6 (2019).

Donoghue, T. *et al.* Parameterizing neural power spectra into periodic and aperiodic components. *Nat Neurosci* 23, 1655-1665, doi:10.1038/s41593-020-00744-x (2020).

Drysdale, A. T. *et al.* Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med* 23, 28-38, doi:10.1038/nm.4246 (2017).

Engemann, D. A. *et al.* A reusable benchmark of brain-age prediction from M/EEG resting-state signals. *NeuroImage* 262, 119521, doi:10.1016/j.neuroimage.2022.119521 (2022).

FDA-NIH Biomarker Working Group. *BEST (Biomarkers, EndpointS, and other Tools) Resource*, <https://www.ncbi.nlm.nih.gov/books/NBK326791/> (2016).

Fornito, A., Zalesky, A. & Breakspear, M. The connectomics of brain disorders. *Nat Rev Neurosci* 16, 159-172, doi:10.1038/nrn3901 (2015).

Fox, M. D. & Raichle, M. E. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 8, 700-711, doi:10.1038/nrn2201 (2007).

Gabard-Durnam, L. J., Mendez Leal, A. S., Wilkinson, C. L. & Levin, A. R. The Harvard Automated Processing Pipeline for Electroencephalography (HAPPE): Standardized Processing Software for Developmental and High-Artifact Data. *Front Neurosci* 12, 97, doi:10.3389/fnins.2018.00097 (2018).

Gao, R., Peterson, E. J. & Voytek, B. Inferring synaptic excitation/inhibition balance from field potentials. *NeuroImage* 158, 70-78, doi:10.1016/j.neuroimage.2017.06.078 (2017).

Gil Ávila, C., Bott, F. S., Gross, J. & Ploner, M. crisglav/discover-eeg: 1.0.0 v. 1.0.0 *Zenodo*, doi:10.5281/zenodo.8207523 (2023a).

Gil Ávila, C. *et al.* DISCOVER-EEG: an open, fully automated EEG pipeline for biomarker discovery in clinical neuroscience. *Sci Data* 10, 613, doi:10.1038/s41597-023-02525-0 (2023b).

Gonzalez-Villar, A. J., Trinanes, Y., Gomez-Perretta, C. & Carrillo-de-la-Pena, M. T. Patients with fibromyalgia show increased beta connectivity across distant networks and microstates alterations in resting-state electroencephalogram. *NeuroImage* 223, 117266, doi:10.1016/j.neuroimage.2020.117266 (2020).

Goodkind, M. *et al.* Identification of a common neurobiological substrate for mental illness. *JAMA psychiatry* 72, 305-315, doi:10.1001/jamapsychiatry.2014.2206 (2015).

Gorgolewski, K. J. *et al.* The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci Data* 3, 160044, doi:10.1038/sdata.2016.44 (2016).

Habermann, M., Weusmann, D., Stein, M. & Koenig, T. A Student's Guide to Randomization Statistics for Multichannel Event-Related Potentials Using Ragu. *Frontiers in Neuroscience* 12, 355, doi:10.3389/fnins.2018.00355 (2018).

Halchenko, Y. O. *et al.* DataLad: distributed system for joint management of code, data, and their relationship. *Journal of Open Source Software* 6, 3262, doi:10.21105/joss.03262 (2021).

Hohenfeld, C., Werner, C. J. & Reetz, K. Resting-state connectivity in neurodegenerative disorders: Is there potential for an imaging biomarker? *NeuroImage. Clinical* 18, 849-870, doi:10.1016/j.nicl.2018.03.013 (2018).

Kam, J. W. Y. *et al.* Systematic comparison between a wireless EEG system with dry electrodes and a wired EEG system with wet electrodes. *NeuroImage* 184, 119-129, doi:10.1016/j.neuroimage.2018.09.012 (2019).

Khanna, A., Pascual-Leone, A., Michel, C. M. & Farzan, F. Microstates in resting-state EEG: current status and future directions. *Neuroscience and biobehavioral reviews* 49, 105-113, doi:10.1016/j.neubiorev.2014.12.010 (2015).

Kiiski, H. *et al.* EEG spectral power, but not theta/beta ratio, is a neuromarker for adult ADHD. *European Journal of Neuroscience* 51, 2095-2109, doi:10.1111/ejn.14645 (2020).

Klug, M. *et al.* The BeMoBIL Pipeline for automated analyses of multimodal mobile brain and body imaging data. *bioRxiv*, doi:10.1101/2022.09.29.510051 (2022).

Kucyi, A. & Davis, K. D. The dynamic pain connectome. *Trends Neurosci* 38, 86-95, doi:10.1016/j.tins.2014.11.006 (2015).

Kuner, R. & Flor, H. Structural plasticity and reorganisation in chronic pain. *Nat Rev Neurosci* 18, 113, doi:10.1038/nrn.2017.5 (2017).

Kuner, R. & Kuner, T. Cellular Circuits in the Brain and Their Modulation in Acute and Chronic Pain. *Physiol Rev* 101, 213-258, doi:10.1152/physrev.00040.2019 (2021).

Kurtzer, G. M., Sochat, V. & Bauer, M. W. Singularity: Scientific containers for mobility of compute. *PloS one* 12, e0177459, doi:10.1371/journal.pone.0177459 (2017).

Lee, J. J. *et al.* A neuroimaging biomarker for sustained experimental and clinical pain. *Nat Med* 27, 174-182, doi:10.1038/s41591-020-1142-7 (2021).

Leichsenring, F., Steinert, C., Rabung, S. & Ioannidis, J. P. A. The efficacy of psychotherapies and pharmacotherapies for mental disorders in adults: an umbrella review and meta-analytic evaluation of recent meta-analyses. *World Psychiatry* 21, 133-145, doi:10.1002/wps.20941 (2022).

Lopez-Sola, M. *et al.* Towards a neurophysiological signature for fibromyalgia. *Pain* 158, 34-47, doi:10.1097/j.pain.0000000000000707 (2017).

Markello, R. D. *et al.* neuromaps: structural and functional interpretation of brain maps. *Nature methods* 19, 1472-1479, doi:10.1038/s41592-022-01625-w (2022).

May, E. S. *et al.* Dynamics of brain function in patients with chronic pain assessed by microstate analysis of resting-state electroencephalography. *PAIN* 162, 2894-2908, doi:10.1097/j.pain.0000000000002281 (2021).

Michel, C. M. & Koenig, T. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review. *NeuroImage* 180, 577-593, doi:10.1016/j.neuroimage.2017.11.062 (2018).

Milz, P. *et al.* The functional significance of EEG microstates-Associations with modalities of thinking. *NeuroImage* 125, 643-656, doi:10.1016/j.neuroimage.2015.08.023 (2016).

Moriarty, O., McGuire, B. E. & Finn, D. P. The effect of pain on cognitive function: a review of clinical and preclinical research. *Prog Neurobiol* 93, 385-404, doi:10.1016/j.pneurobio.2011.01.002 (2011).

Mouraux, A. & Iannetti, G. D. The search for pain biomarkers in the human brain. *Brain* 141, 3290-3307, doi:10.1093/brain/awy281 (2018).

Muller, R. *et al.* Next steps for global collaboration to minimize racial and ethnic bias in neuroscience. *Nat Neurosci* 26, 1132-1133, doi:10.1038/s41593-023-01369-6 (2023).

Munafo, M. R. *et al.* A manifesto for reproducible science. *Nat Hum Behav* 1, 0021, doi:10.1038/s41562-016-0021 (2017).

Murphy, M. *et al.* Abnormalities in electroencephalographic microstates are state and trait markers of major depressive disorder. *Neuropsychopharmacology*, doi:10.1038/s41386-020-0749-1 (2020).

Newson, J. J. & Thiagarajan, T. C. EEG Frequency Bands in Psychiatric Disorders: A Review of Resting State Studies. *Frontiers in human neuroscience* 12, doi:10.3389/fnhum.2018.00521 (2019).

Niso, G. *et al.* Open and reproducible neuroimaging: From study inception to publication. *NeuroImage* 263, 119623, doi:10.1016/j.neuroimage.2022.119623 (2022).

Nunez, P. L. & Srinivasan, R. *Eletric Fields of the Brain: The Neurophysics of EEG*. 2nd edition edn, (Oxford University Press, 2006).

Paret, C. *et al.* Survey on Open Science Practices in Functional Neuroimaging. *NeuroImage* 257, 119306, doi:10.1016/j.neuroimage.2022.119306 (2022).

Parsons, S. *et al.* A community-sourced glossary of open scholarship terms. *Nat Hum Behav* 6, 312-318, doi:10.1038/s41562-021-01269-4 (2022).

Patel, V. *et al.* The Lancet Commission on global mental health and sustainable development. *Lancet* 392, 1553-1598, doi:10.1016/S0140-6736(18)31612-X (2018).

Pavlov, Y. G. *et al.* #EEGManyLabs: Investigating the replicability of influential EEG experiments. *Cortex* 144, 213-229, doi:10.1016/j.cortex.2021.03.013 (2021).

Pedroni, A., Bahreini, A. & Langer, N. Automagic: Standardized preprocessing of big EEG data. *NeuroImage* 200, 460-473, doi:10.1016/j.neuroimage.2019.06.046 (2019).

Pernet, C. *et al.* Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research. *Nat Neurosci* 23, 1473-1483, doi:10.1038/s41593-020-00709-0 (2020a).

Pernet, C. R. *et al.* EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Sci Data* 6, 103, doi:10.1038/s41597-019-0104-8 (2019).

Pernet, C. R., Martinez-Cancino, R., Truong, D., Makeig, S. & Delorme, A. From BIDS-Formatted EEG Data to Sensor-Space Group Results: A Fully Reproducible Workflow With EEGLAB and LIMO EEG. *Front Neurosci* 14, 610388, doi:10.3389/fnins.2020.610388 (2020b).

Pion-Tonachini, L., Kreutz-Delgado, K. & Makeig, S. ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage* 198, 181-197, doi:10.1016/j.neuroimage.2019.05.026 (2019).

Ploner, M. *et al.* Reengineering neurotechnology: placing patients first. *Nat Mental Health* 1, 5-7, doi:10.1038/s44220-022-00011-x (2023).

Ploner, M., Sorg, C. & Gross, J. Brain Rhythms of Pain. *Trends in cognitive sciences* 21, 100-110, doi:10.1016/j.tics.2016.12.001 (2017).

Poldrack, R. A. & Gorgolewski, K. J. Making big data open: data sharing in neuroimaging. *Nat Neurosci* 17, 1510-1517, doi:10.1038/nn.3818 (2014).

Ricard, J. A. *et al.* Confronting racially exclusionary practices in the acquisition and analyses of neuroimaging data. *Nat Neurosci* 26, 4-11, doi:10.1038/s41593-022-01218-y (2023).

Rodrigues, J., Weiss, M., Hewig, J. & Allen, J. J. B. EPOS: EEG Processing Open-Source Scripts. *Frontiers in Neuroscience* 15, doi:10.3389/fnins.2021.660449 (2021).

Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* 52, 1059-1069, doi:10.1016/j.neuroimage.2009.10.003 (2010).

Scangos, K. W., State, M. W., Miller, A. H., Baker, J. T. & Williams, L. M. New and emerging approaches to treat psychiatric disorders. *Nat Med* 29, 317-333, doi:10.1038/s41591-022-02197-0 (2023).

Schumacher, J. *et al.* Dysfunctional brain dynamics and their origin in Lewy body dementia. *Brain* 142, 1767-1782, doi:10.1093/brain/awz069 (2019).

Siddiqi, S. H., Kording, K. P., Parvizi, J. & Fox, M. D. Causal mapping of human brain function. *Nat Rev Neurosci* 23, 361-375, doi:10.1038/s41583-022-00583-8 (2022).

Stangl, M., Maoz, S. L. & Suthana, N. Mobile cognition: imaging the human brain in the 'real world'. *Nat Rev Neurosci* 24, 347-362, doi:10.1038/s41583-023-00692-y (2023).

Ta Dinh, S. *et al.* Brain dysfunction in chronic pain patients assessed by resting-state electroencephalography. *Pain* 160, 2751-2765, doi:10.1097/j.pain.0000000000001666 (2019).

Tracey, I., Woolf, C. J. & Andrews, N. A. Composite Pain Biomarker Signatures for Objective Assessment and Effective Treatment. *Neuron* 101, 783-800, doi:10.1016/j.neuron.2019.02.019 (2019).

Treede, R.-D. *et al.* Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). *Pain* 160, 19-27, doi:10.1097/j.pain.0000000000001384 (2019).

Turk, D. C., Wilson, H. D. & Cahana, A. Treatment of chronic non-cancer pain. *Lancet* 377, 2226-2235, doi:10.1016/S0140-6736(11)60402-9 (2011).

Uhlhaas, P. J. & Singer, W. Neuronal dynamics and neuropsychiatric disorders: toward a translational paradigm for dysfunctional large-scale networks. *Neuron* 75, 963-980, doi:10.1016/j.neuron.2012.09.004 (2012).

van Dijk, H. *et al.* The two decades brainclinics research archive for insights in neurophysiology (TDBRAIN) database. *Sci Data* 9, 333, doi:10.1038/s41597-022-01409-z (2022).

Vos, T. *et al.* Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 396, 1204-1222, doi:10.1016/S0140-6736(20)30925-9 (2020).

Wager, T. D. *et al.* An fMRI-based neurologic signature of physical pain. *N Engl J Med* 368, 1388-1397, doi:10.1056/NEJMoa1204471 (2013).

Webb, E. K., Etter, J. A. & Kwasa, J. A. Addressing racial and phenotypic bias in human neuroscience methods. *Nat Neurosci* 25, 410-414, doi:10.1038/s41593-022-01046-0 (2022).

Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3, 160018, doi:10.1038/sdata.2016.18 (2019).

Woo, C. W., Chang, L. J., Lindquist, M. A. & Wager, T. D. Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci* 20, 365-377, doi:10.1038/nn.4478 (2017).

Zebhauser, P. T., Hohn, V. D. & Ploner, M. Resting state EEG and MEG as biomarkers of chronic pain: a systematic review. *Pain* 164, 1200-1221, doi:10.1097/j.pain.0000000000002825 (2022).

Zhang, Y. *et al.* Identification of psychiatric disorder subtypes from functional connectivity patterns in resting-state electroencephalography. *Nat Biomed Eng* 5, 309-323, doi:10.1038/s41551-020-00614-8 (2021).

ABBREVIATIONS

BIDS	Brain Imaging Data Structure
EEG	Electroencephalography
ERP	Event Related Potentials
FAIR	Findability, Accessibility, Interoperability, Reuse
fMRI	Functional Magnetic Resonance Imaging
MEEG	Magneto and Electroencephalography
PET	Positron Emission Tomography

ACKNOWLEDGMENTS

I would like to thank in first place my supervisor, Markus Ploner, for believing in me and giving me the opportunity to do this PhD. His outstanding scientific vision, guidance, support, and optimism throughout these years have made me grow personally and as a researcher. I would also like to thank the members of my thesis advisory committee, Valentin Riedl and Paul Sauseng, and additionally Joachim Gross, for their helpful comments and approachability. I would also like to thank all past and present members of the PainLab Munich for creating such a fantastic work atmosphere and teaching me what real teamwork is. Special thanks to Elisabeth for going all the way with me through 'microstates' of joy and despair, her empathy and invaluable support; to Laura, for always having my back, especially with bureaucracy, and making everything smooth, fun, and salamander-full; to Vanessa, for her encouragement, funny stories, and becoming such a good friend, to Felix, for all his methodological knowledge and lending a hand every time was needed, to Henrik and Paul, for their kindness, help with data, and medical knowledge, and to Clara, for her bringing always a big smile to the office and, occasionally, delicious crepes as well. I would also like to thank the TUMNIC members, especially the AG Mühlau, for sharing their servers, that saved me more once, and so many fun lunches, hikes, and celebrations. I would also like to thank the GSN for always caring about their students and offering great hard and soft skills courses. I can't think of a better environment for carrying out my PhD. Finally, I would also like to thank my family and friends in Spain, for their support and understanding, even in the distance, and David, for his love, patience, and thirst for knowledge that inspires me every day.

LIST OF PUBLICATIONS

Müller R, Ruess AK, Schönweitz FB, Buyx A, **Gil Ávila C**, Ploner M. Next steps for global collaboration to minimize racial and ethnic bias in neuroscience. *Nat Neurosci.* 2023. doi: 10.1038/s41593-023-01369-6

Bott FS, Nickel MM, Hohn VD, May ES, **Gil Ávila C**, Tiemann L, Gross J, Ploner M. Local brain oscillations and inter-regional connectivity differentially serve sensory and expectation effects on pain*. Sci. Adv.* 2023. 9:eadd7572. doi:10.1126/sciadv.add7572

Gil Ávila C, Bott FS, Tiemann L, Hohn VD, May ES, Nickel MM, Zebhauser PT, Gross J, Ploner M. DISCOVER-EEG: an open, fully automated EEG pipeline for biomarker discovery in clinical neuroscience. *Sci. Data.* 2023. 10, 613. doi:10.1038/s41597-023-02525-0

Hohn VD, Bott FS, May ES, Tiemann L, Fritzen C, Nickel MM, **Gil Ávila C**, Ploner M. How do alpha oscillations shape the perception of pain? - An EEG-based neurofeedback study. *PLoS Biol,* Registered Report, in-principle accepted. 2022. doi: 10.17605/OSF.IO/QBKJ2

Nickel MM, Tiemann L, Hohn VD, May ES, **Gil Ávila C**, Eippert F, Ploner M. Temporal-spectral signaling of sensory information and expectations in the cerebral processing of pain. *PNAS*. 2022. 119:e2116616119. doi: 10.1073/pnas.2116616119.

Heitmann H, **Gil Ávila C**, Nickel MM, Ta Dinh S, May ES, Tiemann L, Hohn VD, Tölle TR, Ploner M. Longitudinal resting-state electroencephalography in patients with chronic pain undergoing interdisciplinary multimodal pain therapy. *Pain*. 2022. 163:e997-e1005. doi: 10.1097/j.pain.0000000000002565.

May ES*, **Gil Ávila C***, Ta Dinh S, Heitmann H, Hohn VD, Nickel MM, Tiemann L, Tölle TR, Ploner M. Dynamics of brain function in chronic pain patients assessed by microstate analysis of resting-state electroencephalography. *Pain*. 2021. 162:2894-2908. doi: 10.1097/j.pain.0000000000002281.

May ES*, Hohn VD*, Nickel MM, Tiemann L, **Gil Ávila C**, Heitmann H, Sauseng P, Ploner M. Modulating brain rhythms of pain using transcranial alternating current stimulation (tACS) - A sham-controlled study in healthy human participants. *J Pain.* 2021. 22, 1256-1272. doi: 10.1016/j.jpain.2021.03.150.

* Equal contribution

AFFIDAVIT

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation "Resting state EEG biomarkers in translational neuroscience" selbstständig angefertigt habe, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

I hereby confirm that the dissertation "Resting state EEG biomarkers in translational neuroscience" is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

Munich, 8 August 2023 Cristina Gil Ávila

DECLARATION OF AUTHOR CONTRIBUTIONS

Project 1. Dynamics of brain function in patients with chronic pain assessed by microstate analysis of resting-state electroencephalography

Authors: Elisabeth S. May*, Cristina Gil Ávila*, Son Ta Dinh, Henrik Heitmann, Vanessa D. Hohn, Moritz M. Nickel, Laura Tiemann, Thomas R. Tölle, Markus Ploner

Contributions: Conceptualization: CGA, ESM, and MP; Methodology: CGA, ESM, and MP; Software: CGA and ESM; Validation: CGA and ESM; Formal analysis: CGA and ESM; Investigation: CGA, ESM, STD, HH, VDH, MMN, and LT; Data curation: CGA and ESM; Visualization: CGA, ESM, and MP; Writing – original draft: CGA, ESM, and MP; Writing – review and editing: CGA, ESM, STD, HH, VDH, MMN, LT, TRT, and MP; Supervision: MP; Project administration: MP; Funding acquisition: MP.

* Elisabeth S. May and Cristina Gil Ávila contributed equally to this work and shared the first authorship of the article.

Project 2. DISCOVER-EEG: an open, fully automated EEG pipeline for biomarker discovery in clinical neuroscience

Authors: Cristina Gil Ávila, Felix S. Bott, Laura Tiemann, Vanessa D. Hohn, Elisabeth S. May, Moritz M. Nickel, Paul Theo Zebhauser, Joachim Gross, Markus Ploner

Contributions: Conceptualization: CGA and MP; Methodology: CGA and MP; Software: CGA, FSB, and JG; Validation: CGA, FSB, and JG; Formal analysis: CGA; Investigation: CGA and PTZ; Data curation: CGA; Visualization: CGA and MP; Writing – original draft: CGA and MP; Writing – review and editing: CGA, FSB, LT, VDH, ESM, MMN, PTZ, JG, and MP; Supervision: MP; Project administration: MP; Funding acquisition: MP.

Munich, 8 August 2023 Cristina Gil Ávila