ACTA SCIENTIFIC NEUROLOGY (ISSN: 2582-1121)

Volume 7 Issue 6 June 2024

Research Article

Construction and Validation of a New BrainView qEEG Discriminant Database

Annie TL Young[1], Slav Danev[2] and Jonathan RT Lakey[1]*

[1]*Department of Surgery, University of California Irvine, California, USA*

[2]*Medeia Inc, Santa Barbara, CA, USA*

***Corresponding Author:** Jonathan RT Lakey, Department of Surgery, University of California Irvine, California, USA.

DOI: 10.31080/ASNE.2024.07.0742

Received: March 11, 2024

Published: May 24, 2024

Abstract

A normative quantitative electroencephalogram (qEEG) database is vital for assessing brain disorders. However, constructing qEEG normative databases for research and clinical applications has posed challenges over the past 61 years, due to defining the 'normal' population and lack of standardized procedures for EEG data. This study aims to build a new BrainView qEEG discriminant database that meets strict normative data criteria derived from the field's challenges and milestones, using a method similar to that used to construct a normative database. It follows key procedures: data collection and preprocessing, feature extraction and selection, as well as classification and validation. BrainView comprises data for 28,283 subjects (7,798 healthy subjects) for eyes-open and eyes-closed conditions, spanning ages 4 to 85 years. Developed using patient data, BrainView's discriminant function identifies a patient's likelihood of belonging to a specific clinical group, aiding in precise diagnosis. The goal is to establish BrainView as a gold standard for diagnosis and prognosis of various brain disorders, enabling standardized use in clinical practice.

Keywords: Quantitative Electroencephalogram; BrainView; Database, qEEG; Brain Disorders

Abbreviations

qEEG: Quantitative Electroencephalogram; EEG: Electroencephalogram; IFCN: International Federation of Clinical Neurophysiology; mTBI: Mild Traumatic Brain Injury; BSS: Blind Source Separation; CSD: Current Source Density; FFT: Fast Fourier Transform; ERP: Event Related Potential; CT: Computed Tomography; MRI: Magnetic Resonance Imaging; MCI: Mild Cognitive Impairment; AD: Alzheimer's Disease; MS: Multiple Sclerosis; ADHD: Attention Deficit Hyperactivity Disorder; ASD: Autism Spectrum Disorder; PTSD: Post-Traumatic Stress Disorder; SVMs: Support Vector Machines; LDA: Linear Discriminant Analysis; KNNs: k-Nearest Neighbors; PSD: Power Spectral Density; MMN: Mismatch Negativity; MKL: Multiple Kernel Learning; RF: Random Forest; GC: Granger Causality; PECs: Power Envelope Correlations; FgMDM: Fisher Geodesic Minimum Distance to the Mean; CT: Classification Trees; ANN: Artificial Neural Networks; HFD: Higuchi's Fractal Dimension; MLP: Multilayer Back-Propagation Network; SampEn: Sample Entropy; PCA: Principal Component Analysis; LMT: Logistic Model Trees; FSL: Fuzzy Synchronization Likelihood; KFD: Katz's Fractal Dimension; m-ACO: Modified Ant Colony Optimization Method; AUC: Lower Area Under the Curve; EMCI: Early Mild Cognitive Impairment

Introduction

The history of normative database traces back to 1929 when Hans Berger conducted the inaugural quantitative electroencephalogram (qEEG) study, marking the commencement of human EEG measurements [1]. Utilizing the Fourier transform, he spectrally analyzed EEG data and compared various measures to a normative database. During the 1950s, Ross Adey, associated

with the UCLA Brain Research Institute, developed the first qEEG reference normative database between 1961 and 1974 [2]. Initially intended for astronaut selection in NASA space travel, this database relied on statistical tests such as means and standard deviations, checks for normal/Gaussian distribution, complex demodulation, Fourier spectral analysis, and other essential statistical parameters. The establishment of statistical standards for normative databases and the first peer-reviewed publication occurred in 1973, led by Swedish neurologists Dr. Milos Matousek and Dr. Ingemar Petersen [2]. Surveying 401 subjects (Female: 54.4%) aged 2 months to 22 years, the Swedish pair set stringent criteria, including age-specific sample sizes and standardized inclusion/exclusion criteria. Their work laid the foundation for parametric statistical tests and peer-reviewed publications. E. Roy John and colleagues, in 1975, validated the Swedish database's reliability by independently cross-validating it with EEG data from Harlem black children aged 9 to 11, who performed at grade level and lacked neurological disorders [2]. Recognizing the need for standardization, E. Roy John and colleagues formed a consortium of universities between 1982 and 1988 [2]. In 1994, the American EEG Association adopted statistical standards to ensure replicability, cross-validation, reliability, and Gaussian approximation for any normative qEEG database [2]. Subsequently, between 1993 and 2001, the establishment of the four Daubert factors set scientific standards for the admissibility of EEG findings in federal courts [2,3]. These standards paved the way for the evolution of qEEG and EEG norms, as currently endorsed by the International Federation of Clinical Neurophysiology (IFCN) [2].

Normative reference databases play a crucial role in contemporary clinical science and patient assessment. These databases adhere to common methodological, statistical, and scientific standards. A qEEG normative database comprises metrics derived from EEG data collected from a sufficiently large and diverse population, ensuring representation of the general populace. This facilitates the comparison of an individual's metrics to the qEEG database, aiming to identify any non-typical electrophysiological markers relative to the population, with potential clinical relevance in specific disorders when professionally interpreted. The term "normative" in this context pertains to the analytical and statistical procedures used in database creation to ensure valid comparisons. Thatcher., *et al*. outlines the crucial steps, including careful inclusion criteria, representative sampling, and balancing participant recruitment based on demographic variables like gender, age, and socioeconomic status [4]. Additionally, amplifier matching corrects individual qEEG metrics based on EEG amplifier frequency characteristics, making them comparable to the database. To achieve a Gaussian distribution characterized by mean and standard deviation, normative databases perform analytical transformations on qEEG metrics. This step ensures high sensitivity and test-retest reliability.

The construction of a qEEG database involves gathering EEG data during active tasks or resting states. Active tasks involve recording brain activity during perceptual, motor, or cognitive tasks, while resting state recordings capture brain activity in an awake, relaxed state with closed (EC) or open (EO) eyes. These recordings offer simplicity and replicability across laboratories worldwide. Several normative databases, such as the Neurometrics, Sterman-Kaiser Imaging Laboratory (SKIL), and NeuroGuide Lifespan databases, have been developed. These are 'normality' databases constituting data from healthy or normal individuals [2,4]. For example, the Neurometrics database, FDA-cleared in 1998, includes metrics measured in 782 normal individuals, with 356 aged between 6-16 years and 426 aged from 16 to 90 [2]. Similarly, the SKIL normative database, developed by Sterman and Kaiser, comprises healthy participants aged 18-55, made up of students, laboratory personnel, community volunteers, and United States Air Force personnel [2]. The commercially available NeuroGuide Lifespan database, renowned for its extensive sample size, covers 727 healthy individuals aged two months to 82.6 years, representing multi-ethnics, with 71.4% white, 24.2% black, and 3.2% oriental individuals [4,5]. Various countries, including South Korea, Taiwan, the Netherlands, and Cuba, have developed their own normative databases to assess a broad spectrum of clinical disorders [5]. It's crucial to note that while EEG normative databases are valuable tools, they are not standalone diagnostic tools. Accurate interpretation by experienced professionals of EEG results requires additional patient information, considering symptoms, medications, and age-related changes for a comprehensive diagnosis, prognosis, or treatment.

The evaluation of brain function continues to be a persistent challenge in the healthcare field, especially considering that brain disorders impact about one in three individuals [6]. This emphasizes the critical necessity for precise tools capable of

distinguishing between healthy and impaired brain function. It is a pressing, unmet medical need for a swift, objective, and physiological measure of brain function within the clinical arena [6]. Essentially, what is required is a vital sign for brain function. Existing potential measures for brain vital signs are predominantly utilized in research settings [7]. One such measure is Event-Related Potentials (ERPs), which are assessed through EEG [8]. ERPs, a subset of cognitive evoked potentials (EPs), are well-established in the research literature as a physiological evaluation of brain function. However, their integration into clinical practice is still limited, highlighting the need for further exploration and development in this area.

ERPs, minute voltages generated in the brain in response to specific stimulus events like images, auditory tones, or spoken words, have been extensively studied as indicators of brain function [8]. Non-invasive recording of these signals is made possible through EEG technology and scalp electrodes. Dating back to the 1930s, ERPs have undergone comprehensive scrutiny in scientific literature to assess brain function across a spectrum of processes, ranging from basic sensory to higher-level cognitive functions [9]. With over 150,000 peer-reviewed publications on ERPs and a robust scientific foundation dating back to 1934, there is substantial potential to integrate this ERP/EEG technology into clinical practice [9].

The extensive body of published scientific studies has not only paved the way for the translation of ERPs into a brain vital sign framework but has also laid the foundation for the innovative BrainView ERP platform. Meticulously designed by Medeia Inc., the BrainView platform aims to streamline the rapid recording and analysis of ERP responses using portable EEG devices, delivering automated, standardized, and clinically intuitive results. This pioneering platform has achieved FDA 510K clearance (K192753, K212684) and is fortified by international patents and trademarks. The development of BrainView commenced with the meticulous selection of three highly validated ERPs, each intricately linked to distinct cognitive processes: the N100 (associated with auditory sensation), P300 (related to basic attention), and N400 (linked to cognitive processing) [10,11]. These ERPs underwent rigorous validation across large cohorts of healthy individuals, with data on brain vital signs collected from thousands. The rapid testing process facilitated the swift establishment of normative ranges, ensuring reliability and validity, consistently aligning with existing research [11].

The N100, P300, and N400 ERPs are specifically triggered by standard video, auditory tone, and spoken word pair stimuli. These stimuli are designed to provoke sensory, attentional, and cognitive responses to unexpected events. The speed of these brain responses, measured in milliseconds, and their magnitude, measured in microvolts, are quantified and presented in a readily understandable standard report immediately following the scan. Importantly, these three ERP components can be elicited across different sensory modalities, throughout the lifespan, and repeatedly within individuals to monitor changes over time [10,12]. In the context of healthy aging and routine monitoring of cognitive function, these brain vital signs exhibit heightened sensitivity to subtle cognitive-process alterations that may escape detection through behavior-based tests [12].

The N100 manifests at around 100ms after the presentation of a tone, signifying the brain's acknowledgment that information has entered its auditory processing systems [13]. The P300, occurring at approximately 300ms after a tone, reflects an early stage of attentional processing—specifically, the discrimination between different events, such as discerning a deviant or unexpected sound or tone from a standard one [14]. The N400, peaking at approximately 400ms after the presentation of a word, comes into play when unexpected or incongruent word pairs are detected, serving as an index of one of the most advanced cognitive functions: language processing [15]. This cognitive response of the highest order, the N400, has been successfully validated through advanced neuroimaging techniques, involving comparisons with the underlying functional neuroanatomy [16]. The integration of these ERPs into BrainView provides a comprehensive and nuanced understanding of brain function, showcasing the platform's potential to contribute significantly to cognitive assessment and monitoring.

Medeia Inc., creator of BrainView, has been at the forefront of developing clinical ERP applications, emphasizing rapid and automated approaches for individual ERP recording. The aim of the study is to construct and validate a new candidate BrainView qEEG discriminant database, using patient data to create clinical profiles. Created based on the EEG norms, the BrainView

discriminant database incorporates a discriminant function for specific and definitive patient assessment and placement. The criteria checklist of endorsed EEG norms, providing practical guidance for understanding and evaluating qEEG normative database construction and discriminant databases for various brain disorders, will be discussed.

Materials and Methods

Discriminant databases serve the specific purpose of distinguishing between different groups or conditions and are often applied in clinical settings to differentiate between healthy and pathological states. The distinctions between normative and discriminant databases encompass their intended purpose, the subjects involved, and their utilization.

A normative database is designed to establish a baseline or 'normal' reference for a specific population, typically utilizing data from healthy individuals without known neurological or psychological disorders. Its purpose is to provide a standard for comparison and assessment of individual cases, determining how measurements deviate from the established norm.

In contrast, a discriminant or disease-specific database is focused on a particular disorder or condition. It involves collecting data from individuals diagnosed with a specific disorder, allowing for the identification of patterns or abnormalities associated with the targeted condition.

Despite these divergences in purpose, subjects, and use, both normative and discriminant databases follow a similar construction method, consisting of the same procedures. This method encompasses key procedures such as data collection and preprocessing, feature extraction and selection, as well as classification and validation.

In constructing and validating the new BrainView qEEG database, the following procedures were followed.

Subject and variable selection

Patient data acquisition occurred between 2018 and 2023 across multiple neurology offices. Resting EEG samples, obtained with eyes closed or open and free from artifacts, underwent analysis. Fast-Fourier Transformation (FFT) and direct Fourier Transform (Complex Demodulation) techniques were applied to extract at the spectral power resolution of 0.5 Hz the five primary frequency bands [Delta (0–4Hz), Theta (4–8Hz), Alpha (8–13Hz), Beta (low: 13-21; high: 21-30 Hz), and Gamma (30-45Hz)] and the standard ratios of frequency bands. Statistical analyses encompassed univariate, bivariate, and multivariate methods, presented in tables and topographical color maps for 19 monopolar and all 171 possible combinations of the 19 electrode bipolar derivations of the EEG.

Delta Waves: This type of brain wave has the highest amplitude and occurs at the slowest frequency. It is primarily observed during deep sleep.

Theta Waves: These brain waves are present when awake or in a light phase of sleep, such as when falling asleep. When occurring while awake, theta waves are associated with intense relaxation and are believed to play a crucial role in information processing and memory formation.

Alpha Waves: Produced when awake but in a very relaxed state, typically experienced when first waking up and not concentrating on anything specific.

Beta Waves: These brain waves are generated when the brain is fully awake, alert, and focused. They also occur during states of excitement or arousal.

Gamma Waves: Waves with the lowest amplitude but the fastest frequency among brain waves. They are generated when an individual is trying to solve a problem or intensely concentrating on a specific task, such as during learning.

Inclusion/exclusion criteria, demographics and gender

For subjects aged 4 to 18 years, parents completed a neurological history questionnaire for them, and psychometric evaluations were conducted. Adults (≥ 18 years) also completed a neurological questionnaire, and those deemed unhealthy were excluded based on questionnaire responses and/or physician comments. Physicians have access to the following questionnaires: GAD-7 (Anxiety Severity), DSM-5 Level 1 (Cross-Cutting Symptom Measures), PHQ-9 (Depression), PCL-C (PTSD Severity), GCS (Glasgow Coma Scale), and general neurological questionnaires. Inclusion required at least one questionnaire score below moderate and physician-verified health in that the patient was deemed healthy. Any patient

records or previously known medical records with questionnaire score of 'moderate' or 'severe' were excluded from the BrainView qEEG database, regardless of other information.

Demographic characteristics

It is crucial that the demographic mixture of males and females, various ethnic groups, and socioeconomic statuses be reasonably representative of the expected North American clientele. This diversity was derived from a large pool of subjects obtained from eight geographically dispersed sites, reflecting the North American demographics and addressing a wide range of ethnic and socioeconomic statuses found in the de-identified patient data before review.

Time of day and other miscellaneous factors

Due to numerous uncontrollable factors or confounders influencing the EEG frequency spectrum, statistical randomization was employed to address these variables. This approach acknowledges the impracticality of individually controlling each confounder, as doing so would be expensive, require a large sample size, and necessitate a precise match between the manner in which a patient's EEG was obtained and that in the database.

Client-based brainview qEEG database

Each client in the BrainView qEEG database completed a DSM-based questionnaire. Regression analysis was utilized to remove any psychopathology-related variance from the EEG data. This process ensures that the variance in the EEG of 'healthy' subjects, which is explained by the variance in the questionnaire, is removed to create a 'psychopathology-free' qEEG normative database or discriminant databases for various brain disorders.

Utilizing a client-based normative or discriminant database has its own set of advantages. Clients may harbor expectations distinct from those of 'healthy' subjects concerning EEG recordings. Given that it is common for clients to experience worry or stress during EEG sessions, research has demonstrated a significant correlation between anxiety levels and the power distribution of the frequency band spectrum [17,18]. In essence, profound differences may exist in the resting state EEG recordings of clients compared to 'healthy' subjects, differences unrelated to the psychological complaints of the clients. Therefore, comparing a client's EEG with a normative database comprising 'healthy' subjects without accounting for the aforementioned variations might lead to incorrect conclusions and render the treatment ineffective.

Discriminant databases for various brain disorders

Within discriminant databases, using discriminant functions with qEEG faces the challenge of determining the most appropriate functions tied to relevant features. This involves addressing three key issues [19]. First, there are numerous potential analytic approaches and data features, ranging from time-based measures to frequency-domain analytics and complex time-frequency hybrid features. Second, discrimination functions must be specific enough to accurately detect the targeted state, minimizing false identifications. For example, in mild traumatic brain injury (mTBI) detection, features and functions should distinguish not only mTBI from no trauma but also trauma from states like exhaustion or cognitive fatigue. Lastly, effective discrimination often requires ad hoc adjustments or comparisons against a known baseline due to individual and situational variability. Access to a 'ground truth' or 'baseline 0' state is crucial, though it may be challenging in real-world scenarios where patient access occurs post-trauma [19].

The term 'discriminant functions' encompasses a range of analysis methods where a specific model is employed to determine whether a set of data belongs to a particular group. These models use predefined rules to classify different classes based on qEEG variables or features. Developing machine learning models requires domain knowledge to decide which qEEG features are most relevant for detecting and classifying conditions [19]. Empowered by Medeia Inc.'s BrainView, machine learning, a subset of artificial intelligence, involves computer algorithms that leverage statistical methods and data to enhance performance automatically through experience.

Machine learning models or classifiers are trained on extracted features to identify patterns associated with distinct brain states or disorders. These algorithms learn to map input qEEG features to specific classes (healthy or diseased) based on training data. Rule-based machine learning involves identifying a set of rules representing the algorithm's learned knowledge, relying on the user's domain knowledge to determine essential input features for group discrimination. Feature selection is crucial not only for disease classification but also for eliminating outliers. In the upcoming discussion (see Discussion), exploration of discriminant

function-based features and classifiers tailored for the detection of various brain disorders will illuminate the process of constructing discriminant databases.

Digital electroencephalographic recording procedures

21-lead EEGs were recorded and digitized at 1000 Hz and 500 Hz using the International 10/20 system of electrode placement, with reference linked ear lobes and a single lead EEG configuration. This standardized system allows for consistent and comparable EEG recordings across different individuals and settings.

Impedance for each electrode was maintained at less than 5k ohms to 40k ohms for all subjects. Amplifiers were calibrated using sine wave calibration signals and standardized procedures. A permanent recording was made before and after each test session. The amplifier frequency response was approximately 3 db down at both 0.5 Hz and 40 Hz.

Artificial removal and quality control procedures

EEG recordings were screened for sharp waves, epileptogenic events, and artifacts (e.g., 'drowsiness']. EEG recording lengths varied from 300 seconds to 40 minutes. Artifact removal was as follows: 1 to 2 seconds of 'clean' or 'artifact-free' EEG recordings were selected as a template. This template was then used to compute matching amplitudes of EEG. The criteria were flexible, allowing for equal amplitudes or amplitudes that are 1.25 or 1.5 times larger. The final edited 'clean' or 'artifact-free' EEG recording varied in length from 120 seconds to 600 seconds.

Standardized de-artifacting procedures

Artifacts in EEG refer to unwanted signals or interference that are not directly related to neural activity. The artifacts found in resting-state EEG recordings include eye blinks, eye movements, movement of the head or body, line noise artifacts, and tonic or phasic muscle contractions. These artifacts can distort the EEG signal and make it challenging to interpret the underlying neural activity accurately. To minimize the impact of these artifacts on EEG recordings during data analysis, an automatic de-artifacting procedure was employed.

Manual de-artifaction is subjective, involving marking segments containing artifacts. The drawback is it can result in suboptimal inter- and intra-rater reliability. Automated de-artifacting methods can be either "semiautomatic" or "fully automatic," involving artifact "correction" or artifact "rejection" methods. Artifact rejection methods remove segments of EEG identified as being contaminated by artifacts, while artifact correction methods apply techniques that remove artifacts without removing the underlying EEG signal. One example of an artifact correction method is the use of "blind source separation (BSS)" that identifies different independent sources of variance in the EEG. The benefit of fully automatic de-artifacting methods is that they eliminate inter- and intra-rater variability and guarantee that each EEG will be de-artifacted using the exact same set of criteria.

Re-montage to the surface Laplacian and average reference

The "average reference method" involved summing the voltages across all 19 leads for each time point and dividing this value by the microvolt digital value from each lead at each time point. The reference-free surface "Laplacian or current source density (CSD)" was computed using the spherical harmonic Fourier expansion of the EEG scalp potentials to estimate the CSD directed at right angles to the surface of the scalp in the vicinity of each scalp location [20]. The Laplacian is reference-free in that it is only dependent upon the electrical potential gradients surrounding each electrode. Both methods used a digital EEG time series that was then submitted to the same age groupings, power spectral analysis, and Gaussian evaluations as the Linked Ears method.

FFT linked ears, average reference and Laplacian

The sampling rate was 500 samples per second; the EEG recordings were high-pass filtered at 40 Hz, and the FFT Power Spectral Density was computed as follows. A Hanning window was used for each four-second epoch, resulting in a 0.5 Hz resolution. The 75% sliding window method of Kaiser and Sterman was used to compute the FFT database for linked ears, average reference, and Laplacian estimator of CSD in both the eyes closed and eyes open conditions [21]. Successive four-second epochs were advanced by 500-millisecond steps to minimize the effects of the FFT windowing procedure. The FFT Power Spectral Density, with 512 points and 2.5-second epochs, thus produced a total of six different 80 frequency values in μV2/Hz from 0 to 40 Hz in 0.5 Hz increments. These values were then used to compute means and standard deviations for different age groups, as described.

Amplifier and digital matching

As the frequency characteristics of all amplifiers differ (<3 Hz and >20 Hz), and there are no universal standards that all EEG amplifier manufacturers must abide by, their filter and gain characteristics must be equilibrated to match those of the normative EEG amplifiers and those of various brain disorders' EEG. To achieve this, we injected microvolt sine waves from 0 to 40 Hz in 1 Hz steps into each amplifier system. The ratio of the frequency response characteristics between the normative EEG amplifiers and the amplifier characteristics used for EEG recording in a patient were then used as equilibration factors. A note of caution: It may not be possible to equilibrate some frequencies that are severely attenuated by the amplifier filters. For example, ratios greater than 5.0 will significantly amplify the noise of the amplifiers where little or no EEG signal is present, rendering the Z-scores invalid.

Step-by-step data processing

Below is a step-by-step data processing procedure for validating a normative EEG database or discriminant databases for various brain disorders and calculating sensitivity (Figure 1) [4].

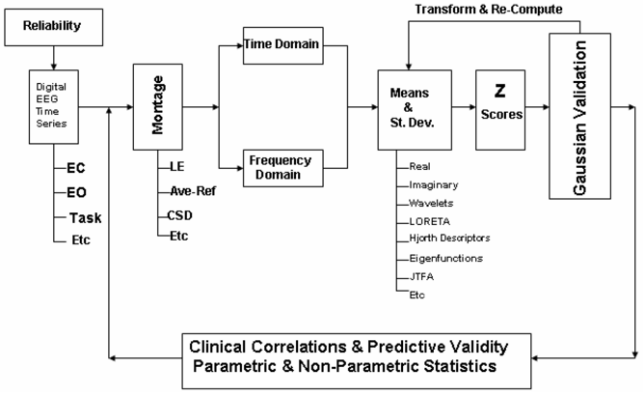

Figure 1: An illustration of a step-by-step procedure by which any normative EEG database can be validated, and sensitivities calculated. The left side of figure is the edited and artifact clean and reliable digital EEG time series which may be re-referenced or re-montaged, which is then analyzed in either the time domain or the frequency domain (4).

- Patient Inclusion and Exclusion
- Include only artifact free EEG data recordings
- EEG data filtering
  - Power noise filtering: 60Hz or 50Hz (4th order)
  - LPF: 70Hz (4th order)
  - HPF: 0.5Hz (4th order)
- Amplifier and Digital Matching
- EEG data Re-Montage: Surface Laplacian, Average Reference and Linked ears
- FFT Spectral analysis on 0.5 to 40 Hz
- Data Grouping
- Age group
- EEG Channel
- Montage Type
- Patient States: Eye Open (EO) or Eye Closed (EC)
- Frequency
- Gaussian Distribution Analysis
- Mean, SD
- Z-Scoring Data Samples
- Gaussian Validation
- Exclude extremes (from step 10) and Re-Computing (go to step 6)

Statistical foundations and performance validation

- **Validation by Clinical Correlations:** Validity concerns the relationship between what is being measured and the nature and use to which the measurement is being applied. Hypothesis formation and testing, as emphasized in Daubert, are important aspects of determining the validity of a scientific measure [3,4].
- **Predictive Validity of Normative/Discriminant Databases:** Nunally (1978) defined predictive validity as follows: 'When the purpose is to use an instrument to estimate some important form of behavior that is external to the measuring instrument itself, the latter is referred to as criterion [predictive] validity' [22]. For example, science 'validates' the clinical usefulness of a measure by assessing its false positive and false negative rates, as well as by examining statistically significant correlations with other clinical measures and, especially, with clinical outcomes.

To prepare the patient for a brainview assessment

To perform a reliable BrainView assessment, it is essential to observe the following patient preparations: patients should abstain from consuming caffeine at least 2 hours before the assessment, avoid taking any new medications or supplements unless directed by a healthcare provider, and refrain from using alcohol, marijuana, or other recreational drugs at least 6 hours prior to the assessment. Patients with pacemakers should not undergo testing during the visit and are required to complete a brief neuropsychological questionnaire about their symptoms before testing.

During the testing, ensure the patient is comfortably seated in a chair while brain behavioral measurements and activity recordings are conducted with both eyes open and eyes closed. Electrogel will be applied to establish the necessary contact between the scalp and the EEG cap electrodes, facilitating the recording of brain waves (Figure 2). The test duration is approximately 25 to 35 minutes.

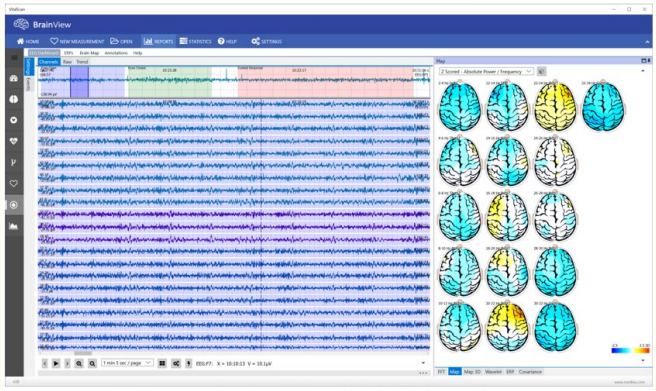

Figure 2: An illustration of an EEG recording of brain waves and qEEG topographical maps obtained using the BrainView Neural Scan System containing the Brainview normative database. Figure 2 depicts a common modern qEEG analysis where EEG traces are displayed on the left and quantitative analysis showing the brain maps on the right.

Brief guide to operate the brainview neural scan system for patient assessment

The BrainView System comprises a workstation with a 21-channel EEG amplifier, assessment and treatment technologies, cloud-based reports, two reusable EEG caps, an SpO2 finger sensor, a headset with three-lead ECG sensors, and a response button for ERP auditory, visual, and motor tests (Figure 3). The BrainView technologies include EEG, ECG, ERP (visual, auditory, no-go), frequency-based analysis of EEG data, qEEG, functional EEG, brain 3-D mapping (eLORETA Source Analysis), behavioral metrics, subjective neuropsychological surveys, and heart rate variability analysis (HRV).

Figure 3: An image of the BrainView Neural Scan System developed by Medeia Inc. The BrainView system is portable, easy-to-use, and non-invasive. The BrainView system is a 21-channel EEG/ERP amplifier with a dedicated laptop and testing supplies. The system utilizes high-quality circuit boards and components to allow for high-quality brain measurements, as well as essential heart data (HRV).

To operate the BrainView Neural Scan System, follow these steps: Turn on your laptop, open the BrainView software, and ensure that the EEG amplifier device's USB is properly connected. To confirm the connection, click on the settings button and press "Check Device Connection." Position the patient comfortably in a chair facing the laptop screen at eye level. Choose the appropriate EEG cap size based on the patient's head measurements, ensuring the cap is washed and fully dry before use.

For patient preparation, apply the three-lead ECG sensors-place the red lead under the right clavicle, the black lead under the

left clavicle, and the yellow lead below the last left rib. Attach the pulse Ox SpO2 finger sensor to the non-dominant hand and place the response button in the dominant hand. In the software, select "New Measurement" and then "Neurofunctional Response Test." Choose options for eyes open, eyes closed, evoked response mixed, and evoked response auditory tests. Avoid selecting shorter testing times; the test should last at least 20 minutes for accurate cognitive assessment.

Proceed to the patient information section, select "New" or "Existing Patient," and enter the patient's details, including name, date of birth, gender, weight, height, medications, symptoms, or previous diagnoses. Progress to the patient questionnaire, guiding the patient through detailed answers—an essential step.

In the pre-test screen, check signal quality. Place the EEG cap on the patient's head, ensuring it fits snugly but not too tight. Use a gel-blunt needle to inject electrogel on every EEG electrode, ensuring it pops up approximately with 1 milliliter. Wiggle in a circular motion to move hair out of the way. Begin with the green ground and reference electrodes, as they are crucial. Check the connection quality in the BrainView software, aiming for all indicators to be green or blue. Wait for any orange dots to change before proceeding.

Once all signals are good, start the test, checking waveforms to ensure all 21 channels are functioning correctly. Place headphones over the patient's cap and ears, connecting the other end to the device amplifier's audio port. All test instructions will be played over the laptop speakers, with ERP special noises played to the patient through the headphones.

Ensure a quiet testing environment. Instruct the patient to avoid talking, muscle movement, or eye blinking. Use the response button to start the test, which includes a baseline reading with eyes open and closed, as well as auditory and visual exercises. Midway, if data quality is insufficient, the software will prompt a recheck.

The eyes-closed test follows, and the final stage involves ERP visual and auditory stimuli. The patient responds to specific instructions, allowing assessment of evoked responses and working memory.

After a successful test, view the results on the overview page, disconnect the patient, and wash and dry the cap before testing the next patient. The software results, starting with the neurofunctional test option, provide a general summary with scales ranging from red (abnormal) to green (healthy), helping diagnose and assess the patient's cognitive health. Light green is borderline, while yellow and orange indicate areas of concern.

Results and Discussion

The BrainView database includes data from 28,283 subjects (14,165 males, 14,118 females) for eyes-open resting EEGs and eyes-closed condition, as depicted in figure 4 and detailed in table 1. The database comprises 50.1% males and 49.9% females, spanning an age range from 4 to 85 years.

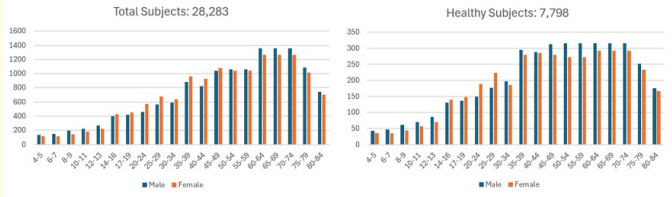

Figure 4: The BrainView database was represented by a similar number of female and male subjects ranging in age from 4 to 85 years.

	Total Subjects			Healthy Subjects		
Age (years)	**Total**	**Male**	**Female**	**Total**	**Male**	**Female**
4-5	254	139	115	78	43	35
6-7	269	151	118	83	47	36
8-9	338	193	145	106	61	45
10-11	404	223	181	128	71	57
12-13	491	268	223	157	86	71
14-16	829	399	430	271	131	140
17-19	872	418	454	284	136	148
20-24	1034	460	574	338	150	188
25-29	1242	565	677	401	177	224
30-34	1230	590	640	383	197	186
35-39	1845	885	960	575	295	280
40-44	1755	825	930	574	288	286
45-49	2122	1042	1080	593	313	280
50-54	2097	1058	1039	587	315	272
55-59	2097	1058	1039	587	315	272
60-64	2620	1355	1265	608	315	293

65-69	2620	1355	1265	608	315	293
70-74	2620	1355	1265	608	315	293
75-79	2096	1084	1012	486	252	234
80-84	1448	742	706	343	176	167
Total	28,283	14,165	14,118	7,798	3,998	3,800

Table 1: The Brain View database comprises subjects, ranging from four years of age as the youngest to 85 years as the oldest, in both eyes open and eyes closed resting EEGs. The distribution is tabulated as 48.5% males and 51.5% females.

When examining the sensitivities of the BrainView qEEG database, table 2 illustrates that the age-dependent data are modeled under a normal/Gaussian distribution, with over 95% of healthy subjects falling within +/-2 standard deviations, and more than 97% falling within +/-3 standard deviations. Hence, BrainView can aid in identifying and distinguishing abnormal EEG values in clinical patients, and can transform EEG data into various Z-scores, encompassing absolute power, relative power, power ratio, asymmetry, coherence, and phase.

AGE	±2SD	>=2SD	<=-2SD	±3SD	>=3SD	<=-3SD
04-10	0.95441	0.9772	0.97723	0.997442	0.99872	0.998723
10-15	0.95439	0.97734	0.97714	0.997439	0.998733	0.998714
15-20	0.95442	0.97733	0.97712	0.997448	0.998732	0.998712
20-30	0.95434	0.97723	0.97722	0.997439	0.998725	0.998728
30-40	0.95435	0.97734	0.97731	0.997432	0.998732	0.998735
40-50	0.95443	0.97725	0.97723	0.997445	0.998729	0.998723
50-85	0.95445	0.97723	0.97716	0.997443	0.998729	0.998712
ALL	0.95448	0.9772	0.97727	0.997444	0.998726	0.998725

Table 2: Brain View qEEG database sensitivities are modeled under a normal/Gaussian distribution, with over 95% of healthy subjects falling within +/-2 standard deviations, and greater than 97% falling within +/-3 standard deviations in an age-dependent analysis.

Normative Database Sensitivities

FP = TP/(TP + FP) or FN = TP/(TP + FN)

Furthermore, the sensitivity and specificity of the database for various clinical conditions. To assess the database's sensitivity and specificity to a particular clinical condition, a total of 20,485 patients were examined. Each control patient in the study suffers from a brain disorder, with Anxiety being the most prevalent condition and Addiction being the least prevalent. The database identified Alzheimer's Disease with a sensitivity of ~86%, and a specificity of ~83% among the 1,098 control patients. This was closely followed by PTSD, which showed ~89% sensitivity but only 79% specificity. The database demonstrated the lowest sensitivity (~60%) and specificity (~59%) in identifying patients with Addiction (n = 820). Although the number of control patients with Anxiety was the highest (n = 3,906), the resulting levels of sensitivity (~65%) and specificity (~64%) were not the highest. This suggests that sample size does not appear to significantly impact the database's sensitivity and specificity for a particular clinical condition. Overall, these findings suggest a strong ability of the database to detect various brain disorders.

The Z-score plays a critical role in a database's identification of various biomarkers linked to specific clinical conditions (Table 3). By examining sensitivity and specificity under different parameter types, such as Z-scores, across various testing conditions, one can assess, validate, and subsequently employ the association of certain parameters or biomarkers with a particular brain disorder in clinical settings, thereby enhancing diagnosis. Apart from Z-scores, both relative and absolute powers are significant and distinguishing features utilized in qEEG databases for discerning various clinical brain conditions. Additionally, frequency band imbalances linked with diverse clinical conditions may serve as

biomarkers for disease identification. Refer to table 4 for a summary of delta/theta, alpha, and beta imbalances associated with various clinical conditions, thus potentially acting as additional biomarkers for individual clinical conditions.

Clinical Condition	Indicator Code	Test Stage	Indicator Frequency	Type	Nb of Patients	Sensitivity	Specificity
ADD	OC-Ref-L6-Alpha-PO-Asymmetry	EyeOpen vs EyeClosed	Alpha	Relative Power Ratio	240	95.75	59.77
ADD	EC-Avg-L4-Beta3a-ZO	EyeClosed	Beta3a	Relative Z-Score	130	84.37	70.97
ADD	ERP-WhiteNoise-tP2	ERP	White Noise - P2 latency	Time	210	82.25	61.34
ADD	EC-Avg-L4-Beta-ZO	EyeClosed	Beta	Relative Z-Score	130	74.64	68.38
ADD	EC-Avg-L1-Beta3a-ZO	EyeClosed	Beta3a	Relative Z-Score	170	74.05	65.88
ADD	EC-Ref-L1-Alpha-PO	EyeClosed	Alpha	Relative Power	120	73.96	65.49
ADHD	EC-Ref-T3-Beta-Peak-PR	EyeClosed	Beta - Peak	Relative Power	120	75.42	70.87
ADHD	ERP-WhiteNoise-tP3b	ERP	White Noise - P3b latency	Time	470	73.22	74.28
ADHD	EC-Avg-L6-Alpha-PA	EyeClosed	Alpha	Absolute Power	810	72.29	74.50
ADHD	EC-Avg-L6-F4T20-PA	EyeClosed	Total Power	Absolute Power	810	71.92	73.93
ADHD	EO-Avg-L1-ThetaBeta3-PR-Ratio	EyeOpen	Theta - Beta3	Relative Power Ratio	910	71.46	69.35
Alzheimer	EC-Ref-L4-Alpha-PO	EyeClosed	Alpha	Relative Power	210	96.36	61.56
Alzheimer	EO-Avg-L6-Beta2a-PR	EyeOpen	Beta2a	Relative Power	240	93.79	75.93
Alzheimer	EO-Avg-L6-Theta-PO	EyeOpen	Theta	Relative Power	240	93.47	78.31
Alzheimer	EO-Avg-L1-Delta1-PO	EyeOpen	Delta1	Relative Power	210	89.53	68.71

Anxiety	EC-Ref-L4-Beta2b-ZA	EyeClosed	Beta2b	Absolute Z-Score	610	80.10	78.83
Anxiety	EC-Ref-L4-Beta3a-ZA	EyeClosed	Beta3a	Absolute Z-Score	610	79.83	78.67
Anxiety	EO-Ref-L1-Beta3a-ZA	EyeOpen	Beta3a	Absolute Z-Score	640	79.32	78.20
Anxiety	EC-Avg-L6-Alpha2-PA	EyeClosed	Alpha2	Absolute Power	500	77.78	84.88
Autism	EO-Avg-L6-Alpha2-PA	EyeOpen	Alpha2	Absolute Power	340	72.86	77.48
Autism	EO-Avg-L4-Delta1-PA	EyeOpen	Delta1	Absolute Power	300	72.59	74.71
Autism	EO-Avg-L6-F4T20-PA	EyeOpen	Total Power	Absolute Power	340	71.72	73.57
Autism	EO-Avg-L4-Theta-PA	EyeOpen	Theta	Absolute Power	300	71.37	70.93
Depression	EO-Avg-L4-F3T8F12T26-PA	EyeOpen	Total Power	Absolute Power	160	89.77	59.76
Depression	EO-Ref-L4-Theta2-PA	EyeOpen	Theta2	Absolute Power	170	83.86	56.70
Depression	ERP-WhiteNoise-vN4	ERP	White Noise - N4	Amplitude	100	80.66	66.00
Depression	EO-Avg-L4-Delta1-ZA	EyeOpen	Delta1	Absolute Z-Score	160	70.12	62.57
Depression	OC-Avg-L4-Gamma1-ZR-Asymmetry	EyeOpen vs EyeClosed	Gamma1	Relative Z-Score Ratio	140	70.01	66.71
Depression	EO-Avg-L4-ThetaBeta3b-PR-Ratio	EyeOpen	Theta - Beta3b	Relative Power Ratio	160	66.29	58.34
Memory Disorder	EC-Avg-L1-Theta-PA	EyeClosed	Theta	Absolute Power	630	73.29	75.47
PTSD	EC-Avg-L6-Beta3a-ZR	EyeClosed	Beta3a	Relative Z-Score	220	88.59	82.65
PTSD	EO-Ref-L1-Delta-ZR	EyeOpen	Delta	Relative Z-Score	230	83.98	78.22
PTSD	EC-Avg-L6-Alpha1-ZR	EyeClosed	Alpha1	Relative Z-Score	220	81.01	76.08
PTSD	EO-Ref-L1-F3T8F12T26-ZR	EyeOpen	Total Power	Relative Z-Score	230	80.86	80.68

PTSD	EC-Avg-L4-Theta-PR	EyeClosed	Theta	Relative Power	250	80.55	77.56
Schizophrenia	OC-Avg-L6-Alpha1-PO-Asymmetry	EyeOpen vs EyeClosed	Alpha1	Relative Power Ratio	210	98.76	75.44
Schizophrenia	EC-Avg-L6-Theta1-ZO	EyeClosed	Theta1	Relative Z-Score	240	92.21	64.46
Schizophrenia	EC-Avg-L6-Delta2-ZR	EyeClosed	Delta2	Relative Z-Score	240	91.95	66.90
Schizophrenia	EC-Avg-L6-Beta2a-PO	EyeClosed	Beta2a	Relative Power	240	90.92	60.36
mTBI	EO-Avg-L4-Theta1-PO	EyeOpen	Theta1	Relative Power	200	77.62	75.85
mTBI	EO-Avg-L1-Beta3b-PR	EyeOpen	Beta3b	Relative Power	250	72.99	70.46
mTBI	EO-Avg-L1-Gamma1-ZR	EyeOpen	Gamma1	Relative Z-Score	250	72.91	71.70
mTBI	ERP-ReactionVariance	ERP	GoNoGo - Reaction Variance	Time	280	71.00	75.71
mTBI	ERP-GoNoGo-tP3	ERP	GoNoGo - P3 latency	Time	210	69.63	67.95

Table 3: The parameters and conditions employed in examining various clinical brain disorders is displayed. The accuracy of identifying individual clinical conditions are depicted by the levels of sensitivity and specificity resulting from the testing conditions. For instance, specific parameter codes, frequencies, and parameter types were utilized to test, identify, and correlate these factors with particular clinical conditions, yielding a wide range of sensitivities and specificities.

Delta/Theta Imbalance	Alpha Imbalance	Beta Imbalance
Cognitive Impairment	Depression	Anxiety
Impulsivity	Victim Mentality	Obsessive Compulsive Disorder (OCD)
Hyperactivity	Excessive Self-Concern	Migraine
Focus and Attention Issues	Passive Aggressive	Tension Headaches
ADHD	Irritability	Insomnia
Socially Inappropriate	Avoidance Behavior	Obsessive Thinking
Easily Distracted	Rumination	Excessive Rationalization
Excessive Speech	Anger	Poor Emotional Self-Awareness
Disorganized	Self-Depreciation	Panic Attacks
Hyper-Emotional	Agitation	Worry
Traumatic Brain Injury (TBI)	Fibromyalgia	Chronic Pain
Dementia	Withdrawal Behavior	Hyper-Vigilant
Learning Disorders		Dislike Change
Autisim/Asperger's		Restless

Table 4: The association of various clinical conditions with specific frequency band imbalances, including delta/theta, alpha, and beta imbalances, is illustrated. In addition to other complementary assessments, the correlation between frequency band imbalances and clinical conditions may serve as an additional biomarker for identifying the condition.

Discussion

In constructing a new BrainView qEEG discriminant database, the study follows the criteria checklist of endorsed EEG norms. These EEG norms provide practical guidance on database construction, highlighting a common set of scientific, methodological, and statistical standards adhered to by both normative and discriminant databases. Furthermore, the discussion includes ERP biomarkers, qEEG features, and machine learning classifiers to further elucidate the process involved in discriminant database construction for various brain disorders.

EEG criteria checklist to guide normative/discriminant database construction

Depending on the disorder or purpose (diagnosis or prognosis), the selection of a sample for database creation involves careful consideration of factors such as a well-defined and disclosed set of inclusion/exclusion criteria. The determination of the sample size takes into account variables like age, gender, socioeconomic status, and geographical distribution where relevant [23]. Matousek and Petersen measured qEEG in 401 subjects (218 females), ranging in age from 2 months to 22 years, residing in Stockholm, Sweden, all without negative clinical histories and performing at grade level [2]. The sample size varied from 18 to 49 per one-year age grouping. Similar inclusion/exclusion criteria were later used in constructing the NYU normative database, the University of Maryland (UM) database, and Brain Resource International Database [2,4,24]. Rigorous screening of subjects in a representative normative 'healthy' database is crucial to include samples of healthy individuals, including normally functioning individuals, and exclude those with a history of neurological or psychiatric problems, school failure, and other deviant behaviors. "Representative sampling" entails a demographically balanced sample concerning gender, ethnic background, socioeconomic status, and age to minimize/prevent sampling bias.

Study subjects participate in active task tests involving recording EEG, EPs, and ERPs while a subject performs a perceptual or cognitive task. Such studies report reproducible task-dependent changes in brain dynamics, essential for understanding normal and pathological brain processes governing perceptual and cognitive function. In contrast, an eyes-closed or eyes-open EEG state, commonly used in developing reference normative EEG databases, involves an alert subject sitting quietly without movement. The eyes-closed and eyes-open conditions are preferred for their simplicity and the relative uniformity of EEG recording conditions, enabling reliable cross-laboratory and cross-population comparisons.

Following data acquisition, artifact cleaning, and reliable digital EEG data conversion to time series, which may be re-referenced or re-montaged, the data are then analyzed in either the time domain or the frequency domain. The selected normal subjects are grouped by age, with a sufficiently large sample size. The means and standard deviations of the EEG time series and/or frequency domain analyses are computed for each age group. Transforms are applied to approximate a Gaussian distribution of the EEG measures that comprise the means. Once approximation to Gaussian is completed, Z-scores are computed for each subject in the database, and leave-one-out Gaussian cross-validation is computed to arrive at optimum Gaussian cross-validation sensitivity. Finally, the Gaussian validated norms are subjected to content and predictive validation procedures, such as correlation with neuropsychological test scores and intelligence, discriminant analyses, neural networks, and outcome statistics. Content validation is carried out with respect to clinical measures, such as intelligence, neuropsychological test scores, school achievement, and so forth. Predictive validation is carried out with respect to discriminative, statistical, or neural network clinical classification accuracy. Both parametric and non-parametric statistics are used to determine the content and predictive validity of a normative EEG database.

There is no absolute sample size considered optimal for a qEEG database, as statistically, sample size is related to "effect size" and "power" [2,23]. The smaller the effect size, the larger the necessary sample size to detect that effect. Another consideration related to sample size is the degree to which a sample approximates a Gaussian distribution. Increased sample size is often necessary to achieve a Gaussian distribution and cross-validation accuracy. An "adequate" sample size is one that enables a Gaussian distribution and cross-validation accuracy, considering the varying human development and maturation at different ages [2,4].

Not only does an adequate sample size matter, but the quality of the sample is also crucial. Sample adequacy in a qEEG normative database necessitates the strict removal of artifacts and measures to ensure high test-retest reliability. Historically, multiple trained individuals visually examined EEG samples from each subject in the

database. Manual artifact removal is necessary despite any digital signal processing methods used. Measures of split-half reliability and test-retest reliability (> 0.9) are crucial to demonstrate the internal consistency and reliability of the normative database [2,4].

The clinical sensitivity and specificity of qEEG are directly related to the stability and reliability upon repeat testing. QEEG has proven to be highly reliable and reproducible [2]. The inherent stability and reliability of qEEG can be demonstrated even when sampling/acquisition time frames are small, with 82% reliability following a 20-second EEG data acquisition, 90% reliability at a 40-second acquisition time frame, and 92% reliability at 60 seconds [2]. Hamilton-Bruce*., et al*. found EEG recordings to be highly reliable even when the same EEG was independently analyzed by three different individuals [25]. Recommendations suggest at least 60 seconds, and preferably 2 to 5 minutes, of artifact-free EEG recordings for clinical evaluation [2]. Predictive accuracy and error rates of any EEG-based prediction or analyses depend on the data constituting the EEG database and the statistical methods used.

To assess the robustness of a database, various statistical tests are employed, including those found in peer-reviewed publications and tests for statistical validity, reliability, and cross-validation. Adherence to scientific standards in EEG machines and recordings is crucial, encompassing aspects such as amplifier matching (critical for normative databases but relatively less critical in standard "control group" studies), meticulous calibration, artifact elimination, and compliance with standards during acquisition, analysis, and approximation of EEG data to a Gaussian/normal distribution. E. Roy John and colleagues (1982 to 1988) formed a consortium of universities to address the 'standardization' need [2]. In the 1980s, matching different EEG systems was technically challenging due to primitive analytic software. To overcome this, qEEG users relied on relative power, as absolute power was not comparable between different EEG machines. The absence of frequency response standardization between different EEG machines meant no cross-platform standardization of qEEG. It wasn't until the mid-1990s that computer speed and software development made amplifier matching and normative database amplifier equilibration possible.

The first statistics evaluating replication and independent cross-validation of normative qEEG databases were applied by E. Roy John and collaborators from 1974 to 1977 [2]. They compared EEG from a sample of New York Harlem black children with the Matousek and Petersen normative database (correlation > 0.8). Emphasis on the approximation to a Gaussian distribution was underscored by both Dr. E. Roy John and Dr. Frank Duffy in the 1970s and 1980s [2,26,27]. In 1994, the American EEG Association produced a position paper reiterating the statistical standards of replication, cross-validation, reliability, and Gaussian approximation as acceptable basic standards for any normative qEEG database [2]. The American EEG Society adopted the same standards. Dr. John and colleagues from the 1980s to the 1990s continued to evaluate and analyze the statistical properties of normative qEEG databases, including EEG samples obtained from different laboratories worldwide [2,4,28].

Comparative analysis using Z-scores was first applied by Matousek and Petersen [2]. They computed means and standard deviations in one-year age-groups and were the first to use t-tests and Z-scores to compare an individual to the normative database means and standard deviations. John*., et al*. expanded on the use of the Z-score for clinical evaluation, including multivariate measures such as the Mahalanobis distance metric [2,4]. Direct normalization of the Gaussian distribution using Z-scores is useful for comparing individuals to a qEEG normative database [4]. The standard-score equation, where the mean is 0 and standard deviation is 1, is employed to cross-validate a normative database, emphasizing the importance of approximation to a Gaussian/normal normative qEEG database. Deviations to the right of the mean are positive, and those to the left are negative. Different values of Z allow for the calculation of different values of Y. For assessing deviation from normal, the values of Z above and below the mean, encompassing 95% of the area of the Gaussian, are often used as a level of confidence necessary to minimize Type-I and Type-II errors [2].

Cross-validation is critical in determining the sensitivity, false positives, and false negatives of a normative database. Due to the expense of acquiring independent data, most cross-validations are computed using a leave-one-out cross-validation procedure [2,4]. Briefly, the procedure involves injecting microvolt calibration sine waves into the input of the amplifiers of different EEG machines and then injecting the same microvolt signals into the normative database amplifiers, obtaining two frequency response curves [4]. Equilibration of a normative qEEG database to a different EEG machine is the ratio of the frequency response curves of the two

amplifiers, used as amplitude scaling coefficients in the power spectral analysis. This step was crucial, enabling absolute power Z-scores and normative database comparisons. Relative power is only used when there is no equilibration of absolute amplitude, as relative power distorts the spectrum and depends on absolute power for its interpretation.

Predictive accuracy and error rates depend on the data that make up a given EEG database and the statistics pertinent to the database. The Supreme Court addressed the statistical foundations of the scientific method in Daubert, 1993, regarding the admissibility of scientific evidence [3]. The Four Daubert Factors for scientific standards of admissibility in Federal Courts include i) hypothesis testing, ii) error estimates of reliability and validity, iii) peer-reviewed publications, and iv) general acceptance [2,3]. Other factors pertinent to the admissibility in Federal Courts of scientific evidence obtained using EEG normative databases include a) inclusion/exclusion criteria, b) methods to remove artifacts and adequate sample sizes per age groups, c) demographic representativeness (e.g., balanced gender, ethnicity, socioeconomic status, etc.), d) means and standard deviations as being normally distributed or "Gaussian", including Gaussian Cross-Validation, and e) content validity by correlations with neuropsychological test scores and IQ achievement scores, as validation.

The criteria checklist of EEG norms guiding normative database construction were meticulously adhered to in this study to develop the BrainView qEEG discriminant database, which follows a Gaussian distribution with its sample size. Developed using patient data, BrainView's database discriminant function determines the likelihood that a patient belongs to a specific clinical group, helping to narrow the assessment to a specific clinical category. Moreover, BrainView surpasses the 'normality' databases in size, incorporating a larger sample size with a broader representation of both genders and similarly covering the 'lifespan' of the individuals. The larger sample size provides more data points to robustly support qEEG analysis. Only with a substantial number of subjects can one confidently depend on the accuracy of parameters such as predicted mean and predicted standard deviation. Besides applying spectral analysis to the EEG data to obtain quantitative metrics associated with behavorial-cognitive brain functions, another body of metrics is derived by using LORETA (Low-Resolution Electromagnetic Tomography), a source localization technique [29]. As a result, qEEG metrics are represented as two- or three-dimensional brain maps for expert interpretation. QEEG, with its advanced digital analysis, facilitates detailed, user-independent assessments of functional abnormalities in the brain.

ERP biomarker applications in discriminant database construction

While ERPs have been a known concept for many, their potential and cost-effective advantages were traditionally limited to specialized laboratory settings [7]. Recent engineering advancements, however, have brought portable EEG systems into the mainstream, making them more widely available and recognized. Simultaneously, historical challenges related to the complexity and variability of EEG data have been effectively addressed through advancements in signal processing and classification [9,30]. Consequently, over the past decade, extensive research and development efforts have culminated in the creation of a vital sign framework for managing brain health. This marks a significant shift, with ERPs transitioning from a specialized laboratory tool to a potentially transformative clinical asset. The potential role of ERP biomarkers in various brain disorders is discussed below.

ERP biomarkers in TBI/concussion

The primary objective of incorporating ERPs into clinical practice is to achieve a quantifiable measure of cognitive function. ERPs as neurophysiological metrics exhibit remarkable sensitivity in detecting alterations in cognitive processing across a diverse array of neurological, developmental, and mental health conditions. Notably, the N100 (auditory sensation), P300 (basic attention), and N400 (cognitive processing) are among the most extensively studied ERPs, covering a broad spectrum of information processing stages. ERPs are highly responsive tools with established reliability for assessing cognitive dysfunction, especially in concussion-related impairments. Research in youth contact sports, like football and hockey, has recorded ERPs at key stages: baseline, injury, return-to-play, and season conclusion [11,31]. Notably, post-concussion changes were detectable in N100, P300, and N400 ERPs, emphasizing their sensitivity. P300 amplitudes remained elevated post-return-to-play, indicating it as a sensitive biomarker to lingering impairment, applicable beyond diagnosed concussions to detect sub-concussive changes [31].

Imaging methods like CT and MRI are effective in the initial stages of head injuries but often yield normal results later. ERPs, particularly the extensively researched P300, prove valuable in detecting subtle alterations in information processing due to diffuse axonal injury, especially in intensive care units [32,33]. Numerous studies consistently show reduced N100 and P300 amplitudes, along with delayed P300 latency, in traumatic brain injuries (TBI) [32,33]. Additionally, delayed N100 latency and decreased auditory N100 amplitude are observed [32]. P300 latency is delayed in TBI-injured individuals, and correlates significantly with standard clinical measures in disorders of consciousness [34]. Reduced N400 amplitudes are noted days after injury [34]. Adults with childhood brain injuries also exhibit reduced N400 amplitude [35]. Rehabilitation, including intensive speech therapy, shows the re-emergence of components like N400. In conclusion, combining ERPs with clinical assessments provides valuable insights into neuropsychological mechanisms post-traumatic brain injuries.

ERP biomarkers in stroke

For stroke, the P300 effectively indicates cognitive recovery post-stroke. Within four weeks, P300 latency is notably delayed, and amplitude reduced [36]. Average P300 latency improves within 12 months, but amplitude does not progress [36]. Studies show a significant reduction in P300 latency 24 months post-stroke [37]. In addition, N400 peak latency correlates with post-stroke aphasia progression [38]. Research has shown a strong link between the ERPs and standardized neuropsychological tests in both healthy individuals and those with neurological acquired brain injuries [7].

ERP biomarkers in mild cognitive impairment (MCI)/ Alzheimer's disease (AD) transition

Multiple clinical studies have investigated the P300 as an indicator of cognitive functions, especially attention and working memory, with relevance to conditions like MCI, dementia, and AD [39]. Literature indicates that delays in P300 latency correlate with impaired memory processes [39,40]. P300 latency increases in dementia, correlating with disease severity [39]. A critical review by Horvath*., et al*. suggests that abnormalities in P300 latency and amplitude (delay and reduction) serve as sensitive tools for detecting cognitive decline in AD [41]. P300 alterations may distinguish MCI from healthy controls and AD patients and aid in detecting the MCI-to-AD transition [41]. Similarly, the N400, another ERP component, is applied in MCI and dementia contexts. The N400, valuable in MCI and dementia, indicates reduced or abnormal responses, predicting MCI progression to dementia and serving as AD biomarkers [41]. Patients with dementia or MCI may exhibit reduced amplitude or absent N400s, useful in predicting MCI conversion to dementia, making the N400 a valuable biomarker for early AD detection and staging [41].

ERP biomarkers in multiple sclerosis (MS)

Cognitive dysfunction affects 30-70% of people with MS [42]. Research on the P300 reveals delayed latency and reduced amplitude in MS-related cognitive impairment [43]. Recent findings suggest P300 latency as a sensitive prognostic indicator for disability progression over 15 years [44]. Combining P300 studies with advanced neuroimaging, such as MRI, uncovered a positive correlation between lesion volume in the frontal horn and brain stem with P300 latency, emphasizing significant links between neurophysiological measures and brain anatomy alterations [45]. This integrated approach enhances our understanding of cognitive dysfunction in MS by examining P300 latency alongside structural changes in the brain.

ERP biomarkers in alcoholism

The literature consistently emphasizes a decrease in N100 amplitude caused by alcohol consumption [46]. The P300 response holds promise as an endophenotypic marker for alcohol dependence. Individuals with alcoholism often display diminished P300 amplitudes, and this occurrence is not solely attributable to alcohol's adverse effects on the brain [46,47]. Significantly, numerous clinical and electrographic characteristics linked to alcohol dependence may return to normal levels following a period of abstinence. However, it is noteworthy that the reduction in P300 amplitude persists, even after prolonged abstinence [33].

ERP biomarkers in attention deficit/hyperactivity disorder (ADHD)

For developmental conditions such as Autism Spectrum Disorder and ADHD, the N100 and P300 are also crucial biomarkers. A thorough 10-year review of ERP research on ADHD uncovered significant correlations between various ERPs, including N100 and P300, and the disorder [48]. Notably, distinctions from healthy individuals were observed in early orienting, inhibitory control, and error-processing components. In a meta-analysis focused on P300 characteristics in adults with ADHD, a consistent pattern

emerged, revealing significantly reduced P300 amplitudes in individuals with ADHD compared to controls [49]. This trend was reliably replicated across multiple studies included in the review. This finding aligns with existing research involving children with ADHD, and intriguingly, the amplitude reduction becomes more pronounced as individuals with ADHD age [49].

ERP biomarkers in autism spectrum disorder (ASD)

Children diagnosed with ASD display distinctive ERP patterns characterized by reduced allocation and engagement of attention resources during the processing of visual stimuli [50]. ERP outcomes have consistently unveiled notable differences in N100 and P300 latency and amplitude when compared to their neurotypical counterparts. Specifically, studies involving adolescents and young adults with ASD have highlighted N100 differences, offering additional insights into potential sensory gating deficits within the ASD population [51]. In-depth analytical reviews focused on ASD and ERPs have thoroughly examined both visual and auditory-elicited ERPs, shedding light on impairments in lower and higher-level visual and auditory functioning within the ASD population [52].

ERP biomarkers in depression

The challenge of achieving objective psychiatric diagnoses has led to a focus on identifying biomarkers. ERPs, explored for their clinical utility, show potential biomarker roles in various psychiatric or mental health conditions. Depression is a prevalent and debilitating global illness, impacting various aspects of life [53]. Despite the effectiveness of antidepressants and cognitive-behavioral therapies, a substantial percentage of patients do not respond adequately [54,55]. The complexity of this disorder, characterized by diverse clinical presentations and a lack of informative biomarkers, poses challenges for accurate diagnosis and successful treatment. The auditory P300 response has garnered attention as a potential biomarker for depression-related neural alterations [56]. Diminished P300 amplitude is observed in individuals with depression, especially those with suicidal ideation, psychotic features, or severe depression [57]. Findings by Key*., et al*. suggest accelerated cognitive aging in major depressive patients, highlighting the frontal P300 latency as a potential biomarker [58].

ERP Biomarkers in Schizophrenia

The latency of P300 is delayed in individuals with schizophrenia, and there is a notable sensitivity of P300 amplitudes to fluctuations in symptom severity [34,35]. Particularly, the auditory P300 amplitude has been identified as a specific trait marker for schizophrenia [59]. This consistent observation of reduced P300 amplitude in individuals with schizophrenia, compared to healthy controls, corresponds with the frequently observed fronto-temporal atrophy in those with compromised attentional processing [60]. In addition to P300 abnormalities, other research has documented a decrease in N100 amplitude and an increase in N400 latency within the schizophrenia patient population [60,61]. These findings collectively contribute to a comprehensive understanding of the neurophysiological alterations associated with schizophrenia, shedding light on the intricate interplay between cognitive processing and structural changes in brain anatomy.

ERP biomarkers in post-traumatic stress disorder (PTSD)

In PTSD, a comprehensive review revealed reduced P300 amplitude in response to stimuli, with context-dependent information processing dissociation [62]. The review also highlighted inconsistent findings regarding changes in N100 response amplitude and latency. Additional research suggests that P300 results indicate a context-dependent information processing dissociation in PTSD, leading to reduced processing of neutral stimuli but enhanced processing of trauma-related stimuli or neutral stimuli in the context of trauma-related patterns [63]. PTSD is a persistent condition affecting an individual's overall well-being [62,63]. Healthcare professionals on the COVID-19 frontline face high rates of stress, anxiety, depression, and sleep problems, emphasizing the urgent need for effective strategies [64]. In PTSD research, the extensively studied P300 frequently exhibits abnormalities correlating with illness severity [62]. Differences in ERP features, especially improvements in basic attention and cognitive processing, are observed. Quantifying treatment impact is crucial for evidence-based interventions in individuals with PTSD.

Feature selection and classification for discriminant database construction

As previously mentioned, discriminant function-based features and classifiers play a crucial role in the diagnosis of various brain disorders. Examining how these tools are tailored for specific conditions sheds light on the intricacies of constructing discriminant databases for enhanced disease detection, for AD,

MCI, MS, ASD, ADHD, alcoholism, depression, PTSD, schizophrenia, stroke, TBI/concussion.

Features and classifiers used in TBI/concussion discriminant databases

The Brainscope model incorporates a discriminant function specifically designed for studying TBI [65]. Another notable qEEG discriminant function, developed by Thatcher*., et al*. involved 20 and 16 qEEG features, respectively [66,67]. The original function consisted of 20 measures, including coherence, phase, amplitude asymmetry, and relative power at various frequency bands [66]. It demonstrated accuracy in discriminating mTBI (aka. concussion) patients from healthy controls. In a subsequent study in 2001, another discriminant function, utilizing 16 measures of EEG coherence, phase, and amplitude asymmetry, was used in classifying mild from severe TBI [67].

QEEG-based discriminant functions, incorporating multivariate features, show promise for TBI detection, enhancing accuracy, specificity, and sensitivity. However, caution is warranted, as some studies indicate a return of prominent qEEG features to normal within a few days after a concussion [68, 69]. To improve the discriminative power of classification algorithms, combined analyses, more selective features, or larger sample sizes may be considered [65,67]. Careful consideration is necessary when dealing with a high computational cost, especially with numerous features, where a reduced number of features can expedite computation and reduce noise [70,71].

Discriminant functions within the TBI domain can be employed to classify patients based on severity or injury presence, utilizing various metrics (e.g., spectral or functional connectivity features) within a single model. However, relying on a single type of analysis may lack sufficient discriminating power, especially considering the altered alpha, delta, beta, and theta power reported in mTBI [65,66]. Improved accuracy is observed when a combination of qEEG features is utilized. For example, using spectral analysis alone may be insufficient, but when combined with coherence, it becomes a viable tool for mTBI detection, showing adequate discriminative power.

Prichep*., et al*. presented data from a large mTBI group of 633 patients, utilizing a significantly more complex discriminant function with variables such as spectral analysis measures, information theory-based measures, scale-free brain activity measures, fractal dimension measures, functional connectivity measures, and various multivariate measures [65]. The improved discriminant function could differentiate between mTBI and moderate/severe TBI patients, indicating its ability to detect mTBI and distinguish it from more severe forms. Overall, these studies suggest that a well-designed discriminant function can serve as a practical tool for mTBI detection.

Features and classifiers used in stroke discriminant databases

Several studies have assessed strokes using various EEG features, including band power changes, brain symmetry index, and spatiotemporal measures, yielding diverse outcomes [72,73]. Caiola*., et al*. developed feature-based models employing statistical, spatiotemporal, and connectivity EEG measures to classify normal, TBI, and stroke patients, generating 1406 features for each 3-minute EEG segment [73]. Feature selection through Linear Discriminant Analysis (LDA) and ReliefF methods identified 192 and 100 most crucial features, respectively. Machine learning models trained with these feature sets, including Decision Trees, Support Vector Machines (SVMs), and k-Nearest Neighbors (KNNs), demonstrated optimal performance with the medium gaussian SVM model for LDA features and the cubic SVM model for ReliefF features, outperforming the full feature set.

Utilizing EEG data-driven machine learning, Vivaldi*., et al*. also supported TBI and stroke classification [74]. Their analysis revealed distinctive EEG patterns in TBI and stroke patients compared to normal subjects, showcasing changes in coherence and relative Power Spectral Density (PSD), particularly in fronto-temporal and parietal regions. LDA feature selection and SVM models consistently performed well across both classifications and validation methods. Compared to normal controls, both TBI and stroke patients exhibited an overall reduction in coherence and relative PSD in delta frequency, with stroke patients displaying more significant changes and a global decrease in theta power. The study suggests EEG-based machine learning models as promising tools for TBI and stroke detection and classification.

Features and classifiers used in schizophrenia discriminant databases

Schizophrenia, a severe psychiatric disorder affecting approximately 1% of the global population, is characterized

by persistent debilitation [75]. Commonly utilized features for classification involve amplitude and latency components such as N100, P300, P50, and N100, with various classifiers tested [76]. Santos Febles*., et al*. explored the effectiveness of Multiple Kernel Learning (MKL) for classifying schizophrenia based on ERP measures extracted from auditory and visual P300 and mismatch negativity (MMN) [76]. To manage the extensive feature set, the Boruta method, a Random Forest (RF)-based feature selection algorithm, was applied, categorizing features into peak-related, peak-to-peak related, and signal-related features, resulting in 726 features for classification (282 for auditory P300, 282 for visual P300, and 162 for MMN]. The classification accuracy reached 83% with the entire dataset and increased to 86% after applying Boruta feature selection, emphasizing the auditory P300 paradigm's significant contribution. Incorporating MKL and Boruta into the analysis of these neurophysiological biomarkers can enhance the diagnosis of schizophrenia.

Features and classifiers used in PTSD discriminant databases

Recent studies have explored the use of microstate characteristics and functional connectivity in machine learning models to predict PTSD [77,78]. Non-spectral features such as reduced frontal to posterior right hemispheric alpha Granger causality (GC) and reduced theta orthogonalized power envelope correlations (PECs) have been investigated by Clancy*., et al*. and Toll*., et al*. respectively [77,78]. Most EEG studies on PTSD traditionally focus on individual feature types. Kim*., et al*. stands out as one of the few studies exploring combinations of different resting-state EEG features, including spectral power, spatial covariance, and network metrics [79]. Using a Riemannian geometry-based classifier, the Fisher geodesic minimum distance to the mean (FgMDM), Kim*., et al*. compared it with conventional classifiers such as LDA, SVM, and RF [79]. The FgMDM classifier demonstrated an average classification accuracy of 75.24%, outperforming LDA, SVM, and RF classifiers with maximum accuracies of 66.54%, 61.11%, and 60.99%, respectively. This study emphasizes the potential of the FgMDM framework in significantly improving the diagnostic accuracy of PTSD when utilizing resting-state EEG data.

Li*., et al*. presented a study that computed multiple EEG features commonly used in EEG research [80]. These features fall into three categories: spectral features (power, asymmetry, frontal theta/beta ratio, peak alpha frequency, and 1/f exponent), functional connectivity features (Imcoh, wPLI, PEC, and GC), and features capturing the temporal dynamics of EEG (microstates and DFA exponents). The investigation focused on distinguishing veterans with probable PTSD from combat-exposed controls using feature selection and machine learning classification. The best-performing classifier, an SVM using all features, achieved a balanced test accuracy of 62.9%. Functional connectivity features were identified as the most crucial for classifications, with SVM using all features showing the highest accuracy. The selected features were predominantly connectivity features, particularly Imcoh and GC, along with wPLI. The only consistently selected non-connectivity feature was the 1/f exponent. Notably, classifiers using specific features like 1/f exponents, GC, Imcoh, and wPLI had slightly lower but still respectable balanced test accuracies compared to using all features. Subtyping within PTSD revealed distinct patterns, leading to an improved classifier accuracy of 79.4% for a subtype characterized by hyperconnectivity in parietal, temporal, visual areas, and the posterior cingulate cortex. The study's novel framework combining subtyping and machine learning offers valuable insights into potential quantifiable biomarkers for PTSD subtypes.

Features and classifiers used in depression discriminant databases

In Cai*., et al*. discriminant EEG analysis for depression detection, linear and nonlinear features were utilized, categorized into Time [e.g., peak, variance, skewness, kurtosis, Hjorth parameters (activity, mobility, complexity)] and Frequency (e.g., relative and absolute centroid frequency, relative power, and absolute power) domains [81]. Nonlinear features included C0-complexity, Kolmogorov Entropy, Shannon Entropy, Correlation Dimension, and Power-Spectral Entropy, extracted from various EEG waves and electrodes. A total of 270 features were obtained for feature selection using the MRMR technique. Four classification algorithms [KNN, SVM, Classification Trees (CT), and Artificial Neural Networks (ANN)] were compared, with KNN achieving the best performance at 79.27% accuracy. "Absolute power of theta wave" consistently stood out as a strong performer, suggesting a robust link between theta wave power and depression—a crucial characteristic for detection.

Higuchi's Fractal Dimension (HFD) and Sample Entropy (SampEn) were also found effective in detecting depressive disorders [82]. These measures, extracted from EEG signals, were employed with various machine learning algorithms, achieving an average accuracy ranging from 90.24% to 97.56%. SampEn demonstrated better performance among the two measures. Using HFD and SampEn alongside machine learning techniques allows for accurate discrimination between patients diagnosed with depression and controls, serving as a highly sensitive and clinically relevant marker for depressive disorder diagnosis.

Deslandes*., et al*. further aimed to distinguish depression from dementia using five qEEG variables for discriminant analysis [83]. These variables included normed monopolar relative power theta for Cz, normed monopolar relative power alpha for P3, normed bipolar relative power theta for the head, normed bipolar relative power total for T3-F7, and normed bipolar coherence delta for fronto-temporal. The discriminant analysis effectively distinguished between dementia and depression, showing a high level of agreement (91.2%) with clinical diagnoses (DSM-IV]. The qEEG variables also demonstrated a high level of concordance (90.4%) with clinical diagnoses, highlighting the accuracy of this method in distinguishing between Primary Degenerative Dementia and Major Depressive Disorder. Deviations in qEEG variables associated with slow rhythms and alpha rhythm further supported the discriminant accuracy of the method.

Features and classifiers used in alcoholism discriminant databases

Alcohol Use Disorder (AUD) poses a global social and health challenge, complicating screening due to the subjectivity of self-reports [84]. Mumtaz*., et al*. focused on developing a machine learning method to classify alcohol abusers from healthy controls and distinguish among healthy controls, alcohol abusers, and alcoholics [85]. QEEG features like absolute power (AP) and relative power (RP) were extracted and selected using methods like t-test and principal component analysis (PCA). LDA, SVM, Multilayer Back-Propagation Network (MLP), and Logistic Model Trees (LMT) were employed for classification, with LMT achieving the best performance with 96% accuracy, 97% sensitivity, and 93% specificity. Subgroup classification for AUD patients also yielded accuracy exceeding 90%. Results highlight significant neurophysiological differences among alcohol abusers, alcoholics, and controls, emphasizing decreased theta in AUD patients compared to healthy controls.

Spectral power analysis, particularly focusing on higher theta power, has been a popular EEG method to discriminate between alcoholics and control groups [86]. Machine learning techniques have shown promise in clinical applications for screening alcoholic subjects from healthy controls, offering potential solutions to these challenges [85,86].

Features and classifiers used in ADHD discriminant databases

Researchers have extensively explored the potential of EEG measures in diagnosing ADHD. Nonlinear features have been extracted from EEG signals for ADHD detection using classifiers like SVM, multilayer perceptron, and KNN [87]. Higuchi and Katz fractional dimension-based feature extraction methods are commonly employed due to the complex and nonlinear nature of EEG signals [87]. Studies, such as Joy*., et al*. have used these methods, with 112 features, to discriminate between ADHD and normal subjects, achieving a maximum classification accuracy of 100% with an ANN classifier [87].

ERP features have also been explored for ADHD diagnosis, with Merzagora*., et al*. finding that non-linear classifiers outperformed linear ones, achieving an accuracy of over 90% [88]. Mueller*., et al*. accurately classified ADHD patients and controls using independent ERP components, achieving a 92% classification accuracy with a non-linear SVM classifier [89]. Challenges persist in identifying the optimal feature extraction technique and applying the most effective classifier algorithm for achieving maximum classification accuracy in ADHD diagnostic methods.

Features and classifiers used in ASD discriminant databases

The conventional EEG measures utilized in ASD discriminant function analysis since 1986 include PSD and coherence [90]. Instead, Ahmadlou*., et al*. used complexity and chaos theory to unveil a nonlinear feature space for investigating EEG signals in children with ASD [91]. Fractal Dimension was proposed to explore the complexity and dynamical changes in the ASD brain, with Higuchi's Fractal Dimension (HFD) and Katz's Fractal Dimension (KFD) investigated as computation methods. The study presented a wavelet-chaos-neural network methodology for an automated EEG-based diagnosis of ASD, tested on a dataset from two groups: nine

children with ASD (aged 6 to 13) and eight non-ASD children (aged 7 to 13). Using a radial basis function classifier, the model achieved an impressive accuracy of 90%, based on the most significant features identified through analysis of variation statistical tests. Specifically, three KFDs in delta (loci Fp2 and C3) and gamma (locus T6) EEG sub-bands demonstrated high significance ($P < 0.001$). Significant differences between ASD and control groups were observed, particularly in gamma, beta, and alpha bands for HFD, and in gamma, beta, and delta bands for KFD [90]. The study highlighted the effectiveness of KFD as a discriminating tool between ASD and control groups. Overall, this research emphasizes how fractal dimension, by providing additional information about EEG signals, can serve as an important instrument for identifying brain abnormalities in ASD.

Features and classifiers used in MS discriminant databases

The McDonald criteria, involving clinical features, cerebrospinal fluid analyses, imaging techniques, and blood tests, is widely used for diagnosing MS due to the absence of distinct markers [92]. Nonlinear EEG analysis in MS is a relatively novel area, with fractal dimension, recurrence quantification analysis, mutual information, and coherence being commonly used for dynamics analysis [93].

Mohseni*., et al*. proposed a hybrid approach to MS diagnosis, aiming to reduce classification error rates [94]. The method involves analyzing EEG descriptors in both time and frequency domains. A modified ant colony optimization method (m-ACO) is used for feature selection, followed by a SVM classifier to determine the presence of the disease. A metaheuristic algorithm adjusts SVM parameters to counter overfitting. The study achieves significant classification accuracy, exceeding 98.5%, particularly in alpha, beta, and gamma bands. Features extracted include statistical measures (integral of the primary signal, absolute mean value, root mean squares, waveform length, zero-crossing, etc.) in the time domain and fractal dimension-based features (Katz dimension, Higuchi dimension, Petrosian dimension, correlation dimension, self-similar fractal). A total of 31 features are generated, and the final selected vector exhibits strong classification accuracy, considering only 25 to 40% of the total features. The study also highlights the importance of considering different EEG signal characteristics, such as linear frequency and temporal features, for a comprehensive analysis. The proposed methodology showcases the effectiveness of combining linear and nonlinear signal descriptors, feature selection through m-ACO, and an optimized SVM classification algorithm. Thus, the hybrid approach of integrating both linear and nonlinear features has achieved superior accuracy levels and illustrated its potential as a valuable tool for improving MS diagnostic outcomes.

Features and classifiers used in MCI/AD discriminant databases

While CSF and neuroimaging markers are gold standards for *in vivo* AD assessment, their invasiveness and cost limit their frontline screening utility [95]. EEG has been extensively studied as a non-invasive alternative for AD analysis [96]. Linear and nonlinear features have been utilized to diagnose AD in recent years. For AD detection from EEG findings, researchers have recommended diverse features representing EEG complexity, synchrony, and regularity. During the feature extraction phase, discriminant features are extracted from EEG signals. Feature selection or reduction methods can be applied to decrease the number of features, making them independent and reducing computational complexity.

Ge., *et al*. aimed to develop a robust discriminant system based on time-frequency features of qEEG integrated with machine learning techniques [97]. Four wavelet features and Permutation Entropy were extracted for classification using eight supervised learning classifiers (LDA, Logistic Regression (Logreg), KNN, SVM, RF, Naïve Bayes (Nbayes), Ensemble Methods (Adaboost), NN). The proposed routine achieved an average accuracy of 93.18% for differential diagnosis of AD patients and normal controls. The study revealed that combinations of parametric and nonparametric features provided high accuracy in discriminating between AD patients and normal controls, with the best accuracy of 93.18% achieved using all five features (Variance, Pearson Correlation Coefficient, Interquartile Range, Hoeffding's D Measure, Permutation Entropy]. The core features indicating AD included decreased alpha power frequency and a general increase in delta and theta rhythms [98]. Huang*., et al*. proposed that combined alpha and theta global field power were the best discriminating variables between AD patients and controls (84% accuracy) and AD and MCI subjects (78% accuracy) [98].

In another study, a RF model effectively predicted the conversion of Early Mild Cognitive Impairment (EMCI) patients to

AD with 93.6% accuracy [99]. Removing certain features improved accuracy, highlighting the importance of feature selection. The RF consistently outperformed SVM implementations, emphasizing its effectiveness for individualized MCI to AD conversion prediction. The analysis with Logreg found that individual versions of 6 (age, race, FAQ, ADAS13, ADAS11, MMSE), 9 (age, race, APOE4, hippocampal and ventricular volume, ADAS13, ADAS11, FAQ, MMSE), and 13 (age, race, APOE4, hippocampal and ventricular volume, ADAS13, ADAS11, FAQ, MMSE, #words memorized, learned, forgotten, %words forgotten) features exhibited lower accuracy than the RF model. Although Logreg had a lower area under the curve (AUC) than RF and XGBoost, it outperformed the best SVM model. Nonetheless, the best model, a 9-feature RF implementation, achieved an accuracy of 93.6% for predicting conversion from EMCI to AD. This model using EMCI patients can predict conversion 5-7 years prior to AD onset.

Feature selection and classifier choice depend on the data nature, discrimination task complexity, and result interpretability. It's essential to tailor these elements to the specific requirements of the discriminant database construction and analysis. In both cases, validation methods such as cross-validation and external validation are crucial to database performance assessment.

Future studies on brainview discriminant databases for various brain disorders

Normality data for all three ERP (N100, P300, N400) responses in healthy individuals across the lifespan are available in literature sources and the BrainView reference database. The next objective is to establish discriminant databases for various clinical conditions, including neurological (AD, MCI, MS), developmental (ADHD, ASD), and mental health (PTSD, schizophrenia, depression), as well as alcoholism and acquired brain injury. Medeia Inc. aims to make BrainView qEEG discriminant databases the gold standard for diagnosing and predicting brain disorders, employing discriminant analyses to match patient qEEG profiles with specific clinical profiles. The creation of the normative database marked a significant milestone, inspiring the development of databases to enhance patient assessment across a range of neurological disorders. Comprehensive validation is essential to ensure the reliability and applicability of BrainView in diverse clinical contexts.

Conclusion

The application of qEEG and ERP in clinical practice holds significant promise as a tool offering insights into the neurophysiological aspects of psychological disorders. QEEG has the potential to blend a high level of standardization with a personalized medicine approach to mental health care. However, the effectiveness of qEEG in clinical settings relies on the advancement of automated and standardized processing methodologies. Past research utilized resting-state qEEG for biomarker discovery, facing challenges due to non-standardized databases and analytical complexities. Standardization is crucial for qEEG to establish itself among established biomarker development methods, particularly given its wide variability across individuals. Substantial technical and statistical enhancements in the field since the inception of qEEG have considerably contributed to its clinical viability.

To unlock its full potential in clinical practice, qEEG necessitates the integration of standardized de-artifacting techniques, qEEG databases, and interpretation methods. The cross-validated and reliable BrainView qEEG database emerges as a promising tool to unlock and utilize biomarkers for various brain disorders, aiming to enhance the quality of life for many. Future studies will include the development and validation of BrainView discriminant databases for various brain disorders, paving the way for the adoption of BrainView as the gold standard for neurophysiological and neuropsychological assessments in clinical settings.

Acknowledgements

None.

Conflict of Interest

None.

Bibliography

1. Berger H. "Uber das Electrenkephalogramm des Menschen". *Archiv Fur Psychiatrie und Neverkrankheiten* 87 (1929): 527-570.

2. Miranda P., *et al*. "Technical and statistical milestones and standards for construction, validation and/or comparison of Quantitative Electroencephalogram (QEEG) normative databases". *Journal of Systems and Integrative Neuroscience* 7 (2021) : 1-13.

3. Daubert v Merrell Dow Pharmaceuticals Inc., 509 U.S. 579 (1993).

4. Thatcher RW., *et al*. "Quantitative EEG normative databases: Validation and clinical correlation". *Journal of Neurotherapy* 7.3-4 (2003): 87-121.

5. Yin-Chen W and I-Mei L. "The Resting State of Taiwan EEG Normative Database: Z-Scores of Patients with Major Depressive Disorder as the Cross-Validation". *Brain Science* 13 (2023): 351.

6. Perrault A. "One in three Canadians will be affected by a brain disease, disorder, or injury: NeuroScience Canada calls for increased investment in neuroscience research". *Neuroscience Canada* (2006).

7. Gawryluk JR., *et al*. "Improving the clinical assessment of consciousness with advances in electrophysiological and neuroimaging techniques". *BMC Neurology* 10.11 (2010).

8. Luck SJ. "An Introduction to the Event-Related Potential Technique (2nd edition), MIT Press". (2014).

9. Adrian ED and Matthews BHC. "The interpretation of potential waves in the cortex". *Journal of Physiology* 81.4 (1934): 440-471.

10. Pawlowski GM., *et al*. "Brain Vital Signs: Expanding From the Auditory to Visual Modality". *Frontiers in Neuroscience* 12 (2019): 968.

11. Carrick FR., *et al*. "Brain Vital Signs in Elite Ice Hockey: Towards Characterizing Objective and Specific Neurophysiological Reference Values for Concussion Management". *Frontiers in Neuroscience* 15 (2021): 952.

12. Ghosh HS., *et al*. "Developing Brain Vital Signs: Initial Framework for Monitoring Brain Function Changes Over Time". *Frontiers in Neuroscience* 10 (2016): 211.

13. Davis PA. "Effects of Acoustic Stimuli on the Waking Human Brain". *Journal of Neurophysiology* 2.6 (1939): 494-499.

14. Sutton S., *et al*. "Information delivery and the sensory evoked potential". *Science* 155.3768 (1967): 1436-1439.

15. D'Arcy RCN., *et al*. "Separating phonological and semantic processing in auditory sentence processing : A high-resolution event-related brain potential study". *Humam Brain Mapping* 22.1 (2004): 40-51.

16. Ghosh HS., *et al*. "Multimodal characterization of the semantic N400 response within a rapid evaluation brain vital sign framework". *Journal of Translational Medicine* 16.1 (2018): 151.

17. Knyazev GG. "Cross-frequency coupling of brain oscillations: An impact of state anxiety". *International Journal of Psychophysiology* 80.3 (2011): 236-245.

18. Andersen SB., *et al*. "Electrophysiological correlates of anxious rumination". *International Journal of Psychophysiology* 71.2 (2009): 156-169.

19. Rapp PE., *et al*. "Traumatic brain injury detection using electrophysiological methods". *Frontiers in Human Neuroscience* 9 (2015): 11.

20. Pascual-Marqui., *et al*. "Current source density estimation and interpolation based on the spherical harmonic Fourier expansion". *International Journal of Neuroscience*19.3-4 (1988): 237-249.

21. Kaiser DA and Sterman MB. "Automatic artifact detection, overlapping windows, and state transitions". *Journal of Neurotherapy* 4.3 (2001): 85-92.

22. Nunnally JC. "Psychometric theory". New York: McGraw-Hill (1978).

23. Cohen J. "Statistical power analysis for the behavioral sciences". New York: Academic Press (1977).

24. Gordon E., *et al*. "Integrative neuroscience: The role of a standardized database". *Clinical EEG and Neuroscience* 36.2 (2005): 64-75.

25. Hamilton-Bruce MA., *et al*. "Interoperator variability in quantitative electroencephalography". *Clinical and Experimental Neurology* 28 (1991): 219-224.

26. John ER., *et al*. "Developmental equations for the electroencephalogram". *Science* 210 (1980): 1255-1258.

27. Duffy FH., *et al*. "Brain electrical activity mapping (BEAM): A method for extending the clinical utility of EEG and evoked potential data". *Annals of Neurology* 5.4 (1979): 309-321.

28. Prichep LS. "Use of normative databases and statistical methods in demonstrating clinical utility of QEEG: Importance and cautions". *Clinical EEG and Neuroscience* 36.2 (2005): 82-87.

29. Pascual-Marqui RD. "Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details". *Methods and Findings in Experimental and Clinical Pharmacology* 24.D (2002): 5-12.

30. Martinek R., *et al*. "Advanced Bioelectrical Signal Processing Methods: Past, Present and Future Approach - Part II: Brain Signals". *Sensors* 21.19 (2021): 6343.

31. Fickling SD., *et al*. "Subconcussive changes in youth football players: objective evidence using brain vital signs and instrumented accelerometers". *Brain Communication* 4.2 (2022).

32. Duncan CC., *et al*. "Closed head injury-related information processing deficits: An event-related potential analysis". *International Journal of Psychophysiology* 58.2-3 (2005): 133-157.

33. Duncan CC., *et al*. "Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400". *Clinical Neurophysiology* 120.11 (2009): 1883-1908.

34. Lew HL., *et al*. "Temporal stability of auditory event-related potentials in healthy individuals and patients with traumatic brain injury". *Journal of Clinical Neurophysiology* 24.5 (2007): 392-397.

35. Knuepffer C., *et al*. "Reduced N400 semantic priming effects in adult survivors of paediatric and adolescent traumatic brain injury". *Brain Language* 123.1 (2012): 52-63.

36. Dejanović M., *et al*. "The role of P300 event-related potentials in the cognitive recovery after the stroke". *Acta Neurologica Belgica* 115.4 (2015): 589-595.

37. Stahlhut L., *et al*. "The impact of stroke on cognitive processing - A prospective event-related potential study". *Journal of the Neurological Sciences* 339.1-2 (2014): 157-163.

38. Marchand Y., *et al*. "Linking neurophysiological and neuropsychological measures for aphasia assessment". *Clinical Neurophysiology* 113.11 (2002): 1715-1722.

39. Polich J. "Clinical application of the P300 event-related brain potential". *Physical Medicine and Rehabilitation Clinics of North America* 15.1 (2004): 133-161.

40. Braverman ER and Blum K. "P300 (Latency) Event-Related Potential: An Accurate Predictor of Memory Impairment". *Clinical Electroencephalography* 34.3 (2003): 124-139.

41. Horvath A., *et al*. "EEG and ERP biomarkers of Alzheimer's disease: A critical review". *Frontiers in Bioscience-Landmark* 23.2 (2018): 183-220.

42. Rao SM., *et al*. "Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction". *Neurology* 41.5 (1991): 685-691.

43. Gil R., *et al*. "Event-related auditory evoked potentials and multiple sclerosis". *Electroencephalography and Clinical Neurophysiology* 88.3 (1993): 182-187.

44. Szilasiová J., *et al*. "Cognitive Event-Related Potentials-The P300 Wave Is a Prognostic Factor of Long-Term Disability Progression in Patients with Multiple Sclerosis". *Journal of Clinical Neurophysiology* 39.5 (2020): 390-396.

45. Piras MR., *et al*. "Longitudinal study of cognitive dysfunction in multiple sclerosis: Neuropsychological, neuroradiological, and neurophysiological findings". *Journal of Neurology, Neurosurgery and Psychiatry* 74.7 (2003): 878-885.

46. Miyazato Y and Ogura C. "Abnormalities in event-related potentials: N100, N200 and P300 topography in alcoholics". *Journal of Psychiatry and Neuroscience* 47 (1993): 853-862.

47. Patrick CJ., *et al*. "P300 amplitude as an indicator of externalizing in adolescent males". *Psychophysiology* 43.1 (2006): 84-92.

48. Johnstone SJ., *et al*. "Ten years on: A follow-up review of ERP research in attention-deficit/hyperactivity disorder". *Clinical Neurophysiology* 124.4 (2013): 644-657.

49. Szuromi B., *et al*. "P300 deficits in adults with attention deficit hyperactivity disorder: A meta-analysis". *Psychological Medicine* 41.7 (2011): 1529-1538.

50. Wang S., *et al*. "Early and late stage processing abnormalities in autism spectrum disorders: An ERP study". *PLoS One* 12.5 (2017): e0178542.

51. Chien YL., *et al*. "P50-N100-P200 sensory gating deficits in adolescents and young adults with autism spectrum disorders". *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 95 (2019).

52. Jeste SS and Nelson CA. "Event Related Potentials in the Understanding of Autism Spectrum Disorders: An Analytical Review". *Journal of Autism and Developmental Disorders* 39 (2008): 495.

53. Proudman D., *et al*. "The growing burden of major depressive disorders (MDD): for researchers and policy makers". *Pharmacoeconomics* 39 (2021): 619-625.

54. Joffe R., *et al*. "Antidepressant treatment of depression: a meta-analysis". *The Canadian Journal of Psychiatry* 41 (1996): 613-616.

55. Kessler RC., *et al*. "The epidemiology of major depressive disorder: results from the national comorbidity survey replication (NCS-R)". *JAMA* 289.23 (2003): 3095-3105.

56. Mumtaz W., *et al*. "Review on EEG and ERP predictive biomarkers for major depressive disorder". *Biomed Signal Process Control* 22 (2015): 85-98.

57. Hansenne M., *et al*. "Suicidal behavior in depressive disorder: An event-related potential study". *Biological Psychiatry* 40.2 (1996): 116-122.

58. Key AP., *et al*. "Electrophysical biomarkers and age characterize phenotypic heterogeneity among individuals with major depressive disorder". *Frontiers in Human Neuroscience* 16 (2022).

59. Mathalon DH., *et al*. "Trait and state aspects of P300 amplitude reduction in schizophrenia: A retrospective longitudinal study". *Biological Psychiatry* 47.5 (2000): 434-449.

60. O'Donnell BF., *et al*. "Auditory event-related potential abnormalities in bipolar disorder and schizophrenia". *International Journal of Psychophysiology* 53.1 (2004): 45-55.

61. Sur S and Sinha VK. "Event-related potential: An overview". *Industrial Psychiatry Journal* 181 (2009): 70-73.

62. Javanbakht A., *et al*. "Event-related potential studies of posttraumatic stress disorder: A critical review and synthesis". *Biology of Mood and Anxiety Disorders* 1 (2011): 5.

63. Karl A., *et al*. "Meta-analytic review of event-related potential studies in post-traumatic stress disorder". *Biological Psychology* 71.2 (2006): 123-147.

64. Marvaldi M., *et al*. "Anxiety, depression, trauma-related, and sleep disorders among healthcare workers during the COVID-19 pandemic: A systematic review and meta-analysis". *Neuroscience and Biobehavioral Reviews* 126 (2021): 252-264.

65. Prichep LS., *et al*. "Classification of traumatic brain injury severity using informed data reduction in a series of binary classifier algorithms". *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 20.6 (2012): 806-822.

66. Thatcher RW., *et al*. "EEG discriminant analyses of mild head trauma". *Electroencephalography and Clinical Neurophysiology* 73.2 (1989): 94-106.

67. Thatcher R., *et al*. "Estimation of the EEG power spectrum using MRI T2 relaxation time in traumatic brain injury". *Clinical Neurophysiology* 112.9 (2001): 1729-1745.

68. Jacquin A., *et al*. "A multimodal biomarker for concussion identification, prognosis and management". *Computers in Biology and Medicine* 102 (2018): 95-103.

69. Dow RS., *et al*. "Electroencephalographic studies immediately following head injury". *American Journal of Psychiatry* 101.2 (1944): 174-183.

70. Dosenbach NUF., *et al*. "Prediction of individual brain maturity using fMRI". *Science* 329 (2010): 1358-1361.

71. De Carvalho JB., *et al*. "A combined Fourier analysis and support vector machine for EEG classification". *Chilean Journal of Statistics* 10.2 (2019): 3.

72. Doerrfuss JI., *et al*. "Quantitative and qualitative EEG as a prediction tool for outcome and complications in acute stroke patients". *Clinical EEG and Neuroscience* 51.2 (2020): 121-129. Top of Form

73. Caiola M., *et al*. "EEG classification of traumatic brain injury and stroke from a nonspecific population using neural networks". *PLOS Digital Health* 2.7 (2023): e0000282.

74. Vivaldi N., *et al*. "Evaluating performance of EEG data-driven machine learning for traumatic brain injury classification". *IEEE Transactions on Biomedical Engineering* 68.11 (2021): 3205-3216.

75. McGrath J., *et al*. "A systematic review of the incidence of schizophrenia: The distribution of rates and the influence of sex, urbanicity, migrant status and methodology". *BMC Medicine* 2 (2004): 1-22.

76. Santos Febles E., *et al*. "Machine learning techniques for the diagnosis of schizophrenia based on event-related potentials". *Frontiers in Neuroinformatics* 16 (2022).

77. Clancy K., *et al*. "Restless 'rest': intrinsic sensory hyperactivity and disinhibition in post-traumatic stress disorder". *Brain* 140 (2017): 2041-2050.

78. Toll RT., *et al*. "An electroencephalography connectomic profile of posttraumatic stress disorder". *American Journal of Psychiatry* 177 (2020): 233-243.

79. Kim Y-W., *et al*. "Riemannian classifier enhances the accuracy of machine-learning-based diagnosis of PTSD using resting EEG". *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 102 (2020): 109960.

80. Li Q., *et al*. "Resting-state EEG functional connectivity predicts post-traumatic stress disorder subtypes in veterans". *Journal of Neural Engineering* 19.6 (2022).

81. Cai H., *et al*. "A pervasive approach to EEG-based depression detection". *Complexity* (2018): 13.

82. Čukić M., *et al*. "The successful discrimination of depression from EEG could be attributed to proper feature extraction and not to a particular classification method". *Cognitive Neurodynamics* 14.4 (2020): 443-455.

83. Deslandes A., *et al*. "Quantitative electroencephalography (qEEG) to discriminate primary degenerative dementia from major depressive disorder (depression)". *Arquivos de Neuro-Psiquiatria* 62.1 (2004).

84. Popham RE and Schmidt W. "Words and deeds: the validity of self-report data on alcohol consumption". *Journal of Studies on Alcohol* 82.3 (1981): 355-358.

85. Mumtaz W., *et al*. "Automatic diagnosis of alcohol use disorder using EEG features". *Knowledge-Based Systems* 105 (2016): 48-59.

86. Rangaswamy M., *et al*. "Theta power in the EEG of alcoholics". *Alcohol, Clinical and Experimental Research* 27.4 (2003): 607-615.

87. Catherine Joy R., *et al*. "Detection and classification of ADHD from EEG signals using tunable Q-factor wavelet transform". *Journal of Sensors* (2022): 17.

88. Merzagora AC., *et al*. "Model comparison for automatic characterization and classification of average ERPs using visual oddball paradigm". *Clinical Neurophysiology* 120 (2009): 264-274.

89. Mueller A., *et al*. "Classification of ADHD patients on the basis of independent ERP components using a machine learning system". *Nonlinear Biomedical Physics* 4.1 (2010): S1.

90. Billeci L., *et al*. "On the Application of Quantitative EEG for Characterizing Autistic Brain: A Systematic Review". *Frontiers in Human Neuroscience* 7 (2013): 442.

91. Ahmadlou M., *et al*. "Fractality and a wavelet-chaos-neural network methodology for EEG-based diagnosis of autistic spectrum disorder". *Journal of Clinical Neurophysiology* 27 (2010): 328-333.

92. Ömerhoca S., *et al*. "Multiple sclerosis: Diagnosis and differential diagnosis". *Archives of Neuropsychiatry* 55.S1 (2018).

93. Hernandez CI., *et al*. "Examining electroencephalogram signatures of people with multiple sclerosis using a nonlinear dynamics approach: A systematic review and bibliographic analysis". *Frontiers in Computational Neuroscience* 17 (2023).

94. Mohseni E and Moghaddasi SM. "A hybrid approach for MS diagnosis through nonlinear EEG descriptors and metaheuristic optimized classification learning". *Computational Intelligence and Neuroscience* (2022): 5430528.

95. Laske C., *et al*. "Innovative diagnostic tools for early detection of Alzheimer's disease". *Alzheimer's Dement* 11 (2015): 561-578.

96. Cassani R., *et al*. "Systematic review on resting-state EEG for Alzheimer's disease diagnosis and progression assessment". *Disease Markers*2018 (2018): 241-250.

97. Ge Q., *et al*. "A Robust Discriminant Framework Based on Functional Biomarkers of EEG and Its Potential for Diagnosis of Alzheimer's Disease". *Healthcare (Basel)* 8.4 (2020): 476.

98. Huang C., *et al*. "Discrimination of Alzheimer's disease and mild cognitive impairment by equivalent EEG sources: a cross-sectional and longitudinal study". *Clinical Neurophysiology* 111.11 (2000): 1961-1967.

99. Velazquez M and Lee Y. "Random Forest model for feature-based Alzheimer's disease conversion prediction from early mild cognitive impairment subjects". *PloS One* 16.4 (2021): e0244773.