(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2018 (27.12.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/235076 A1

(51) International Patent Classification:
G16H 50/20 (2018.01)

(21) International Application Number:
PCT/IL2018/050675

(22) International Filing Date:
18 June 2018 (18.06.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/522,780 21 June 2017 (21.06.2017) US

(71) Applicant: **HADASIT MEDICAL RESEARCH SERVICES AND DEVELOPMENT LTD.** [IL/IL]; Jerusalem BioPark, Hadassah Ein Kerem Medical Center, P.O. Box 12000, 9112001 Jerusalem (IL).

(72) Inventor: **ARZY, Shahar**; 29A Neve Shaanan Street, Apt. 10, 9370821 Jerusalem (IL).

(74) Agent: **EHRLICH, Gal** et al.; G. E. EHRLICH (1995) LTD., 11 Menachem Begin Road, 5268104 Ramat Gan (IL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *with international search report (Art. 21(3))*

(54) Title: METHOD AND SYSTEM FOR PREDICTING RESPONSE TO PHARMACOLOGICAL TREATMENT FROM EEG

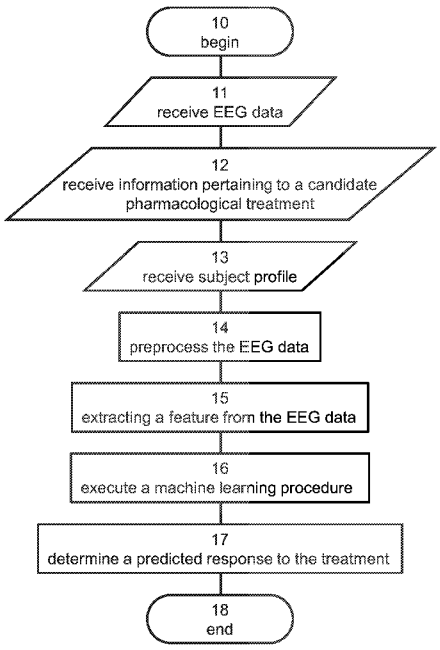


FIG. 1

(57) Abstract: A method of analyzing EEG data collected from a brain of a subject diagnosed with a neuropsychiatric disorder, comprises extracting a feature from the EEG data and receiving information pertaining to a candidate pharmacological treatment for the neuropsychiatric disorder. The method also comprises executing a machine learning procedure, wherein the extracted feature and the received information are inputs of the machine learning procedure, and determining a predicted response of the subject to the pharmacological treatment based on an output of the machine learning procedure.

WO 2018/235076 A1

METHOD AND SYSTEM FOR PREDICTING RESPONSE TO PHARMACOLOGICAL TREATMENT FROM EEG

This application claims the benefit of priority of U.S. Provisional Patent Application No. 62/522,780 filed 21 June 2017, the contents of which are incorporated herein by reference in their entirety.

FIELD AND BACKGROUND OF THE INVENTION

The present invention, in some embodiments thereof, relates to neurophysiology and, more particularly, but not exclusively, to method and system for predicting a response to pharmacological treatment from EEG.

The brain is a complex structure of nerve cells that produce signals called excitatory post synaptic potentials (EPSP). These potentials summate in the cortex and extend through the coverings of the brain to the scalp, where they can be measured using appropriate electrodes. Rhythmical measured activity represents postsynaptic cortical neuronal potentials which are synchronized by the complex interaction of large populations of cortical cells. Electroencephalography (EEG), a noninvasive recording technique, is one of the commonly used systems for monitoring brain activity. EEG data is simultaneously collected from a multitude of channels at a high temporal resolution, yielding high dimensional data matrices for the representation of brain activity. In addition to its unsurpassed temporal resolution, EEG is non-invasive, wearable, and more affordable than other neuroimaging techniques.

Several techniques have been proposed for utilizing EEG for determining reaction to drugs. For example, U.S. Patent No. 7,725,174 discloses a method of indicating a subject's reaction to different agents administered to induce anesthesia using two sets of EEG signals are received from two sets of electrodes located on two regions of a head of a subject. U.S. Patent No. 8,632,750 discloses a method for recommending therapy for a patient having a behaviorally diagnosed psychiatric condition, using patient univariate measures extracted from recorded EEG.

Additional background art includes Khodayari-Rostamabad *et al.* (2010) Clinical Neurophysiology, Volume 121(12), 1998-2006, Khodayari-Rostamabad et al. "A machine learning approach using EEG data to predict response to SSRI treatment for major depressive disorder", Clinical Neurophysiology 124: 1975-1985, 2013, and Arzy et al (2010), European Journal of Neurology, 17(10): 1308-1312.

SUMMARY OF THE INVENTION

According to an aspect of some embodiments of the present invention there is provided a method of analyzing EEG data collected from a brain of a subject diagnosed with a neuropsychiatric disorder. The method comprises: extracting a feature from the EEG data; receiving information pertaining to a candidate pharmacological treatment for the neuropsychiatric disorder; executing a machine learning procedure wherein the extracted feature and the received information are inputs of the machine learning procedure; and determining a predicted response of the subject to the pharmacological treatment based on an output of the machine learning procedure.

According to some embodiments of the invention the execution of the machine learning and the determination of the predicted response is by a server system and the method comprises receiving the EEG data and the information by the server system over a communication network.

According to some embodiments of the invention the method comprises transmitting the predicted response to a client computer.

According to some embodiments of the invention the method comprises collecting EEG signals from the brain of the subject, and digitizing the EEG signals to provide the EEG data.

According to some embodiments of the invention the feature comprises covariance between data channels.

According to some embodiments of the invention the method comprises separating the data into an absolute component and a phase component prior to the feature extraction, wherein the covariance is calculated separately for at least one of the components.

According to some embodiments of the invention the method comprises decomposing the data into frequency bands, wherein the covariance is calculated separately for at least one of the frequency bands.

According to some embodiments of the invention the further comprises decomposing the data into frequency bands, wherein the extracting the feature comprises calculating covariance between data channels separately for at least two of the frequency bands, and wherein the feature comprises a mean of the covariance over the at least two of the frequency bands.

According to some embodiments of the invention the feature comprises a coherence level of the data.

According to some embodiments of the invention the machine learning procedure comprises at least one procedure selected from the group consisting of k-nearest neighbors analysis, support vector machine, clustering, linear modeling, decision tree learning, ensemble

learning procedure, neural networks, probabilistic model, graphical model, Bayesian network, and association rule learning.

According to some embodiments of the invention the neuropsychiatric disorder is a disorder affecting the central nervous system.

According to some embodiments of the invention the neuropsychiatric disorder comprises at least one disorder selected from the group consisting of dementia, movement disorder, depression, anxiety disorder, bipolar disorder, schizophrenia, obsessive-compulsive disorder, Parkinson's disease, and chronic pain.

According to some embodiments of the invention the neuropsychiatric disorder comprises epilepsy.

According to some embodiments of the invention the candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of an antiepileptic drug.

According to some embodiments of the invention the candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of a combination of at least two antiepileptic drugs.

According to some embodiments of the invention the candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of a drug selected from the group consisting of valproic acid, levetiracetam, lamotrigine, ezogabine, pregabalin, topiramate, rufinamide, gabapentin, clonazepam, oxcarbazepine, phenobarbital, phenytoin, sultiam and felbamate, or salts or prodrugs or derivatives of any purity or sustained release formulation thereof.

According to some embodiments of the invention the candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of carbamazepine or salts or prodrugs or derivatives of any purity or sustained release formulation thereof.

According to some embodiments of the invention the neuropsychiatric disorder is epilepsy, wherein the candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of valproic acid, wherein the method comprises decomposing the data into frequency bands, wherein the extracting the feature comprises calculating covariance between data channels separately for at least two of the frequency bands, wherein the feature comprises a mean of the covariance over the at least two of the frequency bands, and wherein the machine learning procedure comprises k-nearest neighbors analysis.

According to some embodiments of the invention the neuropsychiatric disorder is epilepsy, wherein the candidate pharmacological treatment comprises administration of a

predetermined therapeutic amount of levetiracetam, wherein the method comprises, prior to the feature extraction, decomposing the data into frequency bands, and separating the data into an absolute component and a phase component, wherein the feature comprises covariance between data channels calculated separately for each component and each frequency band, and wherein the machine learning procedure comprises k-nearest neighbors analysis.

According to some embodiments of the invention the neuropsychiatric disorder is epilepsy, wherein the candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of levetiracetam, wherein the method comprises, prior to the feature extraction, decomposing the data into frequency bands, and extracting a phase component of the data for each frequency band, wherein the feature comprises covariance between data channels calculated separately for at least two of the frequency bands, and wherein the machine learning procedure comprises k-nearest neighbors analysis.

According to some embodiments of the invention the neuropsychiatric disorder is epilepsy, wherein the candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of lamotrigine, wherein the method comprises, prior to the feature extraction, decomposing the data into frequency bands, and extracting an absolute component of the data for each frequency band, wherein the extracting the feature comprises calculating covariance between data channels for the absolute component separately for at least two of the frequency bands, wherein the feature comprises a mean of the covariance over the at least two of the frequency bands, and wherein the machine learning procedure comprises k-nearest neighbors analysis.

According to some embodiments of the invention the EEG data are collected after at least one session of the candidate pharmacological treatment, and the method comprises determining a response of the subject to the at least one session of the pharmacological treatment and generating output including a recommendation for future sessions of the pharmacological treatment based on an output of the machine learning procedure.

According to an aspect of some embodiments of the present invention there is provided a computer software product, comprises a computer-readable medium in which program instructions are stored, which instructions, when read by a data processor, cause the data processor to receive EEG data collected from a brain of a subject diagnosed with a neuropsychiatric disorder, and information pertaining to a candidate pharmacological treatment for the neuropsychiatric disorder, and to execute the method as delineated above and optionally as further detailed below.

According to an aspect of some embodiments of the present invention there is provided a system for analyzing EEG data, the system comprises a processor arranged to perform code instructions, comprises: code instructions for receiving EEG data collected from a brain of a subject diagnosed with a neuropsychiatric disorder; code instructions for extracting a feature from the EEG data; code instructions for receiving information pertaining to a candidate pharmacological treatment for the neuropsychiatric disorder; code instructions for executing a machine learning procedure wherein the extracted feature and the received information are inputs of the machine learning procedure; and code instructions for determining a predicted response of the subject to the pharmacological treatment based on an output of the machine learning procedure.

According to some embodiments of the invention the system comprises a transceiver arranged to receive and transmit information on a communication network; wherein the processor is arranged to communicate with the transceiver over the communication network; and wherein the code instructions comprise code instructions for transmitting the predicted response to a client computer.

According to some embodiments of the invention the determining the predicted response comprises classifying whether the treatment is effective, partially effective or not effective.

According to some embodiments of the invention the code instructions comprise code instructions for generating an output indicating a recommended change in the treatment.

According to some embodiments of the invention the code instructions comprise code instructions for generating an output pertaining to potential side effects cause by the treatment.

According to some embodiments of the invention the code instructions comprise code instructions for receiving blood test data of the subject, following the treatment, analyzing the blood test data to determine a response of the subject to the treatment, and comparing the response to the predicted response.

According to some embodiments of the invention the system wherein the analysis comprises determining vitamin deficiency.

Unless otherwise defined, all technical and/or scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of embodiments of the invention, exemplary methods and/or materials are described below. In case of conflict, the patent specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and are not intended to be necessarily limiting.

Implementation of the method and/or system of embodiments of the invention can involve performing or completing selected tasks manually, automatically, or a combination thereof. Moreover, according to actual instrumentation and equipment of embodiments of the method and/or system of the invention, several selected tasks could be implemented by hardware, by software or by firmware or by a combination thereof using an operating system.

For example, hardware for performing selected tasks according to embodiments of the invention could be implemented as a chip or a circuit. As software, selected tasks according to embodiments of the invention could be implemented as a plurality of software instructions being executed by a computer using any suitable operating system. In an exemplary embodiment of the invention, one or more tasks according to exemplary embodiments of method and/or system as described herein are performed by a data processor, such as a computing platform for executing a plurality of instructions. Optionally, the data processor includes a volatile memory for storing instructions and/or data and/or a non-volatile storage, for example, a magnetic hard-disk and/or removable media, for storing instructions and/or data. Optionally, a network connection is provided as well. A display and/or a user input device such as a keyboard or mouse are optionally provided as well.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

Some embodiments of the invention are herein described, by way of example only, with reference to the accompanying drawings. With specific reference now to the drawings in detail, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of embodiments of the invention. In this regard, the description taken with the drawings makes apparent to those skilled in the art how embodiments of the invention may be practiced.

In the drawings:

FIG. 1 is a flowchart diagram of a method of analyzing EEG data collected from a brain of a subject diagnosed with a neuropsychiatric disorder, according to various exemplary embodiments of the present invention;

FIG. 2 is a schematic illustration of a machine learning system suitable for use according to some embodiments of the present invention; and

FIG. 3 is a schematic illustration of a computing platform that can be used according to some embodiments of the present invention.

DESCRIPTION OF SPECIFIC EMBODIMENTS OF THE INVENTION

The present invention, in some embodiments thereof, relates to neurophysiology and, more particularly, but not exclusively, to method and system for predicting a response to pharmacological treatment from EEG.

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not necessarily limited in its application to the details of construction and the arrangement of the components and/or methods set forth in the following description and/or illustrated in the drawings and/or the Examples. The invention is capable of other embodiments or of being practiced or carried out in various ways.

The present inventor devised a technique suitable for determining a predicted response of a subject diagnosed with a neuropsychiatric disorder to a pharmacological treatment.

The term "neuropsychiatric disorder" refers to any disorder affecting the central nervous system (CNS), and may include, but is not limited to, one or more of epilepsy, movement disorder, depression, dementia, anxiety disorder, bipolar disorder, schizophrenia, obsessive-compulsive disorder, Parkinson's disease, and chronic pain.

While the embodiments below are described with a particular emphasis to epilepsy, it is to be understood that the technique of the present embodiments is useful for any of the above neuropsychiatric disorders or any other disease treated with psychotropic or CNS related drug.

Epilepsy is a chronic neurological condition that is characterized by recurrent seizures. The primary indication of epilepsy is synchronized electrical activity between large numbers of brain neurons. During a seizure, a group of neurons in the brain demonstrate a highly synchronized firing pattern. Epileptic seizures are characterized by symptoms ranging from a major motor convulsion to a brief period of lack of awareness. Epileptic seizures are typically classified as being either generalized and or partial. A partial seizure commences from a single location within one of the hemispheres and may result from a brain lesion. Generalized seizures involve the whole cortex simultaneously and are assumed to have a genetic basis.

The majority of epileptic seizures are controlled pharmaceutically by antiepileptic drugs (AEDs), which reduce neuronal burst firing and synchronization in localized neuronal ensembles and inhibit the spread of abnormal firing to distant brain sites via three main molecular mechanisms: (i) modulation of voltage-gated ion channels (*e.g.*, sodium, GABA, and calcium); (ii) inhibition of synaptic excitation; and (iii) enhancement of synaptic inhibition.

It was realized by the present inventor, that while the condition of the vast majority of epilepsy patients is successfully controlled by AEDs, the process of determining an adequate treatment, including determining an appropriate dosage (also referred to as "drug titration") for a

specific epileptic subject can be difficult. Typically, treatment is prescribed using trial and error and is supplemented by a clinical evaluation that includes reports provided by the subject and/or physician, and neurological and blood tests. It was additionally realized by the present inventor that such a procedure is very slow (oftentimes prolongs over several months), and is also inaccurate and biased. The present inventor realized that one of the reasons for the drawbacks of a conventional process of determining an adequate treatment for a specific subject is that the conventional process does not measure the effect of the drug on the brain. The technique of the present embodiments therefore improves efficacy and reduces side-effects of the treatment.

The technique of the present embodiments utilizes EEG data for predicting a response of a subject to pharmacological treatment and can therefore provide healthcare providers, *e.g.*, physicians and patients, as well as to healthcare reimbursement providers, e.g., providers of prescription drug insurance, with insightful information useful for clinical decision making. In some embodiments of the present invention the method utilizes EEG data for determining a response of the subject to previous pharmacological treatments and so as to allow correcting a dosage and/or drug regimen for future treatments. The technique of the present embodiments can optionally and preferably classify whether the treatment is effective, partially effective or not effective. Based on this classification, an output indicating a recommended change in the treatment can be generated. Also contemplated are embodiments in which an output pertaining to potential side effects cause by the treatment is generated. In some embodiments of the present invention, a subject is treated by the treatment, and blood test data of the subject is obtained after the treatment. The blood test data can optionally and preferably be analyzed (*e.g.*, to identify vitamin deficiency) so as to determine a response of the subject to the treatment. This response can be compared to the predicted response.

The Examples section that follows describes experiments which successfully classified EEG recordings of patients from diverse age groups according to their response to AED.

FIG. 1 is a flowchart diagram of a method of analyzing EEG data collected from a brain of a subject diagnosed with a neuropsychiatric disorder, according to various exemplary embodiments of the present invention. It is to be understood that, unless otherwise defined, the operations described hereinbelow can be executed either contemporaneously or sequentially in many combinations or orders of execution. Specifically, the ordering of the flowchart diagrams is not to be considered as limiting. For example, two or more operations, appearing in the following description or in the flowchart diagrams in a particular order, can be executed in a different order (*e.g.*, a reverse order) or substantially contemporaneously. Additionally, several operations described below are optional and may not be executed.

At least part of the operations can be can be implemented by a data processing system, *e.g.*, a dedicated circuitry or a general purpose computer, configured for receiving data and executing the operations described below. At least part of the operations can be implemented by a cloud-computing facility at a remote location. The data processing system or cloud-computing facility can serve, at least for part of the operations as an image processing system, wherein the data received by the data processing system or cloud-computing facility include image data.

Computer programs implementing the method of the present embodiments can commonly be distributed to users on a distribution medium such as, but not limited to, a floppy disk, a CD-ROM, a flash memory device and a portable hard drive. From the distribution medium, the computer programs can be copied to a hard disk or a similar intermediate storage medium. The computer programs can be run by loading the computer instructions either from their distribution medium or their intermediate storage medium into the execution memory of the computer, configuring the computer to act in accordance with the method of this invention. All these operations are well-known to those skilled in the art of computer systems.

The method of the present embodiments can be embodied in many forms. For example, it can be embodied in on a tangible medium such as a computer for performing the method operations. It can be embodied on a computer readable medium, comprising computer readable instructions for carrying out the method operations. In can also be embodied in electronic device having digital computer capabilities arranged to run the computer program on the tangible medium or execute the instruction on a computer readable medium.

Referring now to FIG. 1, the method begins at **10** optionally and preferably continues to **11** at which EEG data are received. The data can be received by recording EEG signals directly from the subject by placing a set of EEG electrodes on the scalp of the subject, and digitizing the EEG signals to provide the data. The EEG data can be collected from the brain of the subject before or after the subject is treated by a pharmacological treatment of interest.

As used herein, “EEG signals” refers to recordings of cerebral electrical activity, or other types of brain electrically activity, recorded from any location on the cranium. The portion of the data that corresponds to signals acquired by each EEG electrode is referred to as a "data channel."

As used herein, “EEG electrode” refers to any electrode placed on a person's head and capable of detecting brain electrical activity. EEG electrodes can be placed according to known positioning systems, such as but not limited to, the expanded international 10/20 placement system. In addition, as used throughout, EEG can include cerebral electrical activity or other types of brain electrical activity, and it is to be understood that the technique of the present disclosure can be applied to any type of brain electrical activity signal.

The EEG data can alternatively be received from an external source, such as a computer readable memory medium on which the EEG data are stored, in which case the EEG signals are recorded and digitized by a separate system.

The method can proceed to **12** at which information pertaining to a candidate pharmacological treatment for the neuropsychiatric disorder is received. The information typically includes a candidate drug or a combination of two or more drugs to be administered to the subject, and optionally and preferably also full or partial information regarding a candidate dosage regimen associated with the candidate drug or drugs (*e.g.*, formulation, route of administration, unit dose, frequency of treatment, duration of treatment, *etc.*).

The candidate drug is optionally and preferably, but not necessarily, an antiepileptic drug or a combination of two or more antiepileptic drugs. These embodiments are particularly useful when the neuropsychiatric disorder includes epilepsy, but are also useful for treating other neuropsychiatric disorders. Representative examples of drugs that are contemplated according to some embodiments of the present invention include, without limitation, valproic acid, levetiracetam, lamotrigine, carbamazepine, ezogabine, pregabalin, topiramate, rufinamide, gabapentin, clonazepam, oxcarbazepine, phenobarbital, phenytoin, sultiam and felbamate, or salts or prodrugs or derivatives of any purity or sustained release formulation thereof. Other drugs, not necessarily drugs considered antiepileptic, are also contemplated in some embodiments of the present invention.

In some embodiments of the present invention the method proceeds to **13** at which subject profile is received. The subject profile typically includes one or more of weight, body mass index, gender, age, ethnicity, race, allergies, clinical history, family history. In some embodiments of the present invention subject profile also includes a genetic profile, which can encompass the genes in an entire genome of the subject, or it can encompass a specific subset of genes. The genetic profile may include genomic profile, a proteomic profile, an epigenomic profile and/or a transcriptomic profile.

Any of operations **11**-**13** can be executed digitally by transmitting digital data to a data processor that executes the method, or by instructing data processor to access a computer readable medium and retrieve the data digitally from the medium.

In some embodiments of the present invention the method proceeds to **14** at which the EEG data are preprocessed. The preprocessing operation or set of operations are preferably applied digitally after the EEG signals are digitized.

One type of preprocessing operation that can be employed according to preferred embodiments of the invention is frequency decomposition. In these embodiments, the data are

decomposed into two or more frequency bands. Frequency bands suitable for the present embodiments including, without limitation, the EEG δ band, which is typically less than 4 Hz, but not less than 1 Hz, the EEG θ band, which is typically from about 4 Hz to about 7 Hz, the EEG α band, which is typically from about 8 Hz to about 15 Hz, the EEG β band, which is typically from about 16 Hz to about 31 Hz, and the EEG γ band, which is typically from about 32 Hz to about 80 Hz. Other frequency bands, which may or may not overlap with these bands (*e.g.*, the EEG μ range, typically from about 8 Hz to about 12 Hz, and/or high and very high frequencies, for example, up to 1000 Hz or more) are also contemplated. Each of the above bands can be subdivided into two or more sub-bands. Frequency bands that do not necessarily match the aforementioned EEG bands are also contemplated.

As used herein the term "about" refers to ± 10 %.

Another type of preprocessing operation that can be employed according to preferred embodiments of the invention includes removal of artifacts from the EEG data.

As used herein, "artifacts" refers to any electrical potential recorded while obtaining an EEG or other brain electrical activity signal that is not of cerebral origin or is the result of abnormal brain activity. EEG artifacts are oftentimes also referred to as "noise."

Representative examples of types of artifacts that can optionally and preferably be removed at **14** include, without limitation, non-physiological artifacts, such as, but not limited to, cable or electrode movement causing over-range artifacts, and impulse artifacts (for example due to due to an abrupt change in the impedance of an electrode); non-brain-generated physiological artifacts, such as, but not limited to, horizontal/lateral eye movements, vertical eye movements (*e.g.*, blinks), and electromyographic activity; and brain-generated artifacts, such as, but not limited to, significantly low amplitude signal, and atypical electrical activity pattern (for example due to paroxysmal brain activity).

Lateral eye movements can be identified, as waveforms of 1 Hz or less that have opposite polarity at EEG channels F7 and F8, and can be removed using a FIR filter with passband of about 0.5-3 Hz. Vertical eye movement can be identified by locating large excursions on the Fp1 and Fp2 leads, and can be removed a low-pass filter in the range of about 0.5-5 Hz. Cable or electrode movement artifacts can be identified when the signal amplitude is greater than a threshold which can be, for example, between about 50 μV and about 250 μV. Impulse artifacts can be identified as follows. A frontal EEG channel is first high-pass filtered with cutoff frequency to remove an alpha band from the signal in that channel. In some embodiments, the cut-off frequency is 15 Hz. Next, high-frequency variations of signal amplitude in successive segments with overlap are examined. A segment is identified as containing an impulse artifact

when the difference between the maximal and minimal levels value exceeds a threshold. In certain embodiments, the threshold can be between 25 μV and 250 μV. Electromyographic activity can be identified when a relative energy between two sub-bands of an EEG β band exceeds a threshold. Data segments containing significantly low amplitude signal can be identified by searching for signal epochs with mean-square energy below a threshold, *e.g.*, between 1 μV^2 and 25 μV^2. Atypical electrical activity pattern can be identified using a combination of wavelet analysis and fractal dimension computation.

Another type of preprocessing operation that can be employed according to preferred embodiments of the invention includes dimensionality reduction. This can be achieved using any dimensionality reduction technique known in the art, including, without limitation, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Independent Component Analysis (ICA), kernel PCA, and feature selection.

At **15** the method optionally and preferably extracts one or more features from the EEG data.

As used herein "feature of EEG" refers to any quantitative measure that can be extracted from EEG data. Representative examples include, without limitation, source localization, cross-correlation, covariance, coherence, coupling, cross-covariance, cross-power, symmetry, correlation, autoregressive coefficients, entropy, spectral entropy, energy, energy derivative, filtered amplitude squared, fractal dimension, mean frequency, nonlinear decorrelation lag, nonlinear energy operator, number of zero crossings, Pisarenko harmonic decomposition, power distribution in frequency bands, principal components, principal Lyapunov exponent, cepstrum, spike (occurrence, amplitude, curvature), wavelet subband energy, wavelet compression coefficients, epileptiform discharge complexity (a measure of number of peaks, amplitude, frequency content and morphology of spike waveforms), amount of background disruption (amount of deviation from baseline time and frequency characteristics of electrical signals), regional coherence (coherence of activity in a focal brain region compared to that of other regions in the brain) and zero crossings of energy derivative.

Source localization is a feature that identifies the sources in the brain for each activity as measured by the EEG electrodes. Source localization conceptually starts with a forward problem statement of the form $y = Ax$, wherein an independent source potential variable x produces an effect measured as the dependent measured potential variable y in accord with a transformation operator A. The source potential variable x is an ordered set corresponding to a number of discrete sources of electromagnetic radiation, and likewise the measure potential variable y is an ordered set corresponding to a number of EEG electrodes placed at discrete locations. The

transformation operator A is a matrix that specifies, generally, the physics of propagating electromagnetic radiation from points inside the body to points outside the body. The operator A can be represented as a matrix and can be determined by modeling or measurement.

In source localization the inverse of the forward problem is solved ("inverse problem"). The solution to the inverse problem works backward from the sensed potentials to infer the set of source potentials that produced them. The inverse problem is of the form $A^{-1}y = x$, which is to induce or infer the independent variable x from knowledge of the dependent variable y. There are a number of possible sets of circumstances x that are consistent with the observed results y, and identifying the correct set of circumstances can be done using additional criteria. Such additional criteria are typically provided in the form of one or more assumptions regarding the expected behavior of the source variable x. A typical assumption is that the source variable x can exhibit spatial smoothness. For example, in what is known as the "minimum norm" approach, the source variable is constrained to minimum variance. Alternatively, Low Resolution Electromagnetic Tomographic Analysis (LORETA) minimizes the second derivative of the source variable in three dimensions, which smooths the potential in three-space.

Utilizing additional criteria allows selection of a solution to the inverse problem in which the criteria are best met, and rejection of alternative solutions in which the criteria are less well met, or are not met.

When source localization is employed, the results of the source localization are supplemented to, or replace the EEG data. For example, the identified sources of brain activity can be used as channels of the EEG data.

In embodiments in which the data are decomposed into frequency bands or sub-bands, one or more features of EEG can be extracted for one or more particular frequency bands and/or a particular frequency sub-bands, and more preferably separately for each frequency band and sub-band.

In some embodiments of the present invention, covariance between data channels is used as a feature of the EEG data. The covariance is optionally and preferably calculated with respect to one or more components of the data. For example, in some embodiments, the data of each channel are separated into an absolute component and a phase component, and the covariance is calculated separately for one of these components or both these components. When the data is decomposed into frequency bands and/or sub-bands, the covariance of one or both of the absolute and phase components is optionally and preferably calculated separately for at least one of the frequency bands and sub-bands. Use of covariance between data channels calculated separately for each component and each frequency band is particularly useful when the neuropsychiatric

disorder is epilepsy, and the pharmacological drug comprises levetiracetam. Use of covariance between channels calculated for the absolute component of the data for each frequency band, is particularly useful when the neuropsychiatric disorder is epilepsy, and the pharmacological drug comprises lamotrigine.

In some embodiments, a covariance quantity, or a covariance quantity of one or both of the absolute and phase components, is/are calculated separately for more than one frequency band and/or sub-band, and the mean of these calculated covariance quantities over these frequency bands and/or sub-bands. The mean of the calculated covariance quantities is referred to a mean covariance. In some embodiments, the mean covariance is used as a feature of the EEG data.

Use of the mean of the calculated covariance quantities is useful, for example, when the neuropsychiatric disorder is epilepsy, and the candidate drug comprises valproic acid.

Another feature that can be used according to some embodiments of the present invention is a coherence level of the data. This feature can be calculated separately for one or more of the frequency bands and sub-bands, or for the full range of frequencies. The coherence level typically characterizes the level by which the same frequency bands or sub-band of two data channels preserve their relative phase.

In some embodiments of the present invention the method proceeds to **16** at which a machine learning procedure is executed. As used herein the term "machine learning" refers to a procedure embodied as a computer program configured to induce patterns, regularities, or rules from previously collected data to develop an appropriate response to future data, or describe the data in some meaningful way.

A machine learning system **20** suitable for use according to some embodiments of the present invention is schematically illustrated in FIG. 2. Machine learning system **20** comprises a processing system **22**, which can be a part of a central processing unit (CPU) circuit, an input **24** and an output **26**. Input **24** can be fed by more than one source. For example, input **24** can receive the extracted EEG features, the information that pertains to the candidate pharmacological treatment, and optionally also the subject profile. Machine learning system **20** is preferably pre-trained in a manner that the received inputs are processes by processing system **22** to provide, at **28**, output that is indicative of the predicted response of the subject to the pharmacological treatment.

In some embodiments of the invention the machine learning procedure comprises, or is, a supervised learning procedure. In supervised learning, global or local goal functions are used to

optimize the structure of the learning system. In other words, in supervised learning there is a desired response, which is used by the system to guide the learning.

In some embodiments of the invention the machine learning procedure comprises, or is, an unsupervised learning procedure. In unsupervised learning there are typically no goal functions. In particular, the learning system is not provided with a set of rules. One form of unsupervised learning according to some embodiments of the present invention is unsupervised clustering in which the data objects are not class labeled, *a priori*.

In some embodiments of the invention the machine learning procedure comprises, or is, a semi-supervised learning procedure. In a semi-supervised learning the learning system is first provided with set of an annotated data on which a desired response is expected. This set is referred to as supervised data. Then the learning system uses the set of rules extracted from the supervised data to learn another set of data to which set of rules is not provided.

The output provided by machine learning system **20** can be a binary output that confirms or disconfirms a predicted existence of a response. For example, output can include one of two values, *e.g.*, a first output value (*e.g.*, "1") when machine learning system **20** predicts that the subject will be responsive to the candidate pharmacological treatment, and second output value (*e.g.*, "0") when machine learning system **20** predicts that the subject will be non-responsive to the pharmacological treatment. Alternatively, and more preferably, the output can be a non-binary output. In this case, it can include one of a plurality of discreet values or one of a range of values indicative of the predicted level or intensity by which the subject will be responsive. Still alternatively, the output can include a set of values, each pertaining to different property of the predicted response, such as, but not limited to, level or intensity of the response, polarization of the response, speed of the response, *etc*. When a therapeutic dose of the pharmacological treatment can inhibit, slow, or reverse the progression of more than one condition, the set of output values can include one or more values for each of the conditions.

Representative examples of machine learning procedures suitable for the present embodiments, include, without limitation, clustering, association rule algorithms, feature evaluation algorithms, subset selection algorithms, support vector machines, classification rules, cost-sensitive classifiers, vote algorithms, stacking algorithms, Bayesian networks, decision trees, neural networks, instance-based algorithms, linear modeling algorithms, k-nearest neighbors (KNN) analysis, ensemble learning algorithms, probabilistic models, graphical models, logistic regression methods (including multinomial logistic regression methods), gradient ascent methods, singular value decomposition methods and principle component analysis. Among neural network models, the self-organizing map and adaptive resonance theory are commonly used unsupervised

learning algorithms. The adaptive resonance theory model allows the number of clusters to vary with problem size and lets the user control the degree of similarity between members of the same clusters by means of a user-defined constant called the vigilance parameter.

Following is an overview of some machine learning procedures suitable for the present embodiments.

Support vector machines are algorithms that are based on statistical learning theory. A support vector machine (SVM) according to some embodiments of the present invention can be used for classification purposes and/or for numeric prediction. A support vector machine for classification is referred to herein as "support vector classifier," support vector machine for numeric prediction is referred to herein as "support vector regression".

An SVM is typically characterized by a kernel function, the selection of which determines whether the resulting SVM provides classification, regression or other functions. Through application of the kernel function, the SVM maps input vectors into high dimensional feature space, in which a decision hyper-surface (also known as a separator) can be constructed to provide classification, regression or other decision functions. In the simplest case, the surface is a hyper-plane (also known as linear separator), but more complex separators are also contemplated and can be applied using kernel functions. The data points that define the hyper-surface are referred to as support vectors.

The support vector classifier selects a separator where the distance of the separator from the closest data points is as large as possible, thereby separating feature vector points associated with objects in a given class from feature vector points associated with objects outside the class. For support vector regression, a high-dimensional tube with a radius of acceptable error is constructed which minimizes the error of the data set while also maximizing the flatness of the associated curve or function. In other words, the tube is an envelope around the fit curve, defined by a collection of data points nearest the curve or surface.

An advantage of a support vector machine is that once the support vectors have been identified, the remaining observations can be removed from the calculations, thus greatly reducing the computational complexity of the problem. An SVM typically operates in two phases: a training phase and a testing phase. During the training phase, a set of support vectors is generated for use in executing the decision rule. During the testing phase, decisions are made using the decision rule. A support vector algorithm is a method for training an SVM. By execution of the algorithm, a training set of parameters is generated, including the support vectors that characterize the SVM. A representative example of a support vector algorithm suitable for the present embodiments includes, without limitation, sequential minimal optimization.

In KNN analysis, the affinity or closeness of objects is determined. The affinity is also known as distance in a feature space between objects. Based on the determined distances, the objects are clustered and an outlier is detected. Thus, the KNN analysis is a technique to find distance-based outliers based on the distance of an object from its kth-nearest neighbors in the feature space. Specifically, each object is ranked on the basis of its distance to its kth-nearest neighbors. The farthest away object is declared the outlier. In some cases the farthest objects are declared outliers. That is, an object is an outlier with respect to parameters, such as, a k number of neighbors and a specified distance, if no more than k objects are at the specified distance or less from the object. The KNN analysis is a classification technique that uses supervised learning. An item is presented and compared to a training set with two or more classes. The item is assigned to the class that is most common amongst its k-nearest neighbors. That is, compute the distance to all the items in the training set to find the k nearest, and extract the majority class from the k and assign to item.

KNN is useful, for example, when the neuropsychiatric disorder is epilepsy, the candidate drug comprises valproic acid, and the extracted feature is a mean of covariance values over two or more frequency bands or sub-bands. KNN is also useful when the neuropsychiatric disorder is epilepsy, the candidate drug comprises levetiracetam, and the EEG feature comprises covariance between channels calculated separately for each component and each frequency band. KNN is also useful when the neuropsychiatric disorder is epilepsy, the candidate drug comprises lamotrigine and the EEG feature comprises covariance between channels calculated for the absolute component of the data for each frequency band.

Association rule algorithm is a technique for extracting meaningful association patterns among features.

The term "association", in the context of machine learning, refers to any interrelation among features, not just ones that predict a particular class or numeric value. Association includes, but it is not limited to, finding association rules, finding patterns, performing feature evaluation, performing feature subset selection, developing predictive models, and understanding interactions between features.

The term "association rules" refers to elements that co-occur frequently within the datasets. It includes, but is not limited to association patterns, discriminative patterns, frequent patterns, closed patterns, and colossal patterns.

A usual primary step of association rule algorithm is to find a set of items or features that are most frequent among all the observations. Once the list is obtained, rules can be extracted from them.

The aforementioned self-organizing map is an unsupervised learning technique often used for visualization and analysis of high-dimensional data. Typical applications are focused on the visualization of the central dependencies within the data on the map. The map generated by the algorithm can be used to speed up the identification of association rules by other algorithms. The algorithm typically includes a grid of processing units, referred to as "neurons". Each neuron is associated with a feature vector referred to as observation. The map attempts to represent all the available observations with optimal accuracy using a restricted set of models. At the same time the models become ordered on the grid so that similar models are close to each other and dissimilar models far from each other. This procedure enables the identification as well as the visualization of dependencies or associations between the features in the data.

Feature evaluation algorithms are directed to the ranking of features or to the ranking followed by the selection of features based on their impact.

Information gain is one of the machine learning methods suitable for feature evaluation. The definition of information gain requires the definition of entropy, which is a measure of impurity in a collection of training instances. The reduction in entropy of the target feature that occurs by knowing the values of a certain feature is called information gain. Information gain may be used as a parameter to determine the effectiveness of a feature in explaining the response to the treatment. Symmetrical uncertainty is an algorithm that can be used by a feature selection algorithm, according to some embodiments of the present invention. Symmetrical uncertainty compensates for information gain's bias towards features with more values by normalizing features to a [0,1] range.

Subset selection algorithms rely on a combination of an evaluation algorithm and a search algorithm. Similarly to feature evaluation algorithms, subset selection algorithms rank subsets of features. Unlike feature evaluation algorithms, however, a subset selection algorithm suitable for the present embodiments aims at selecting the subset of features with the highest impact on the subject's response to the treatment, while accounting for the degree of redundancy between the features included in the subset. The benefits from feature subset selection include facilitating data visualization and understanding, reducing measurement and storage requirements, reducing training and utilization times, and eliminating distracting features to improve classification.

Two basic approaches to subset selection algorithms are the process of adding features to a working subset (forward selection) and deleting from the current subset of features (backward elimination). In machine learning, forward selection is done differently than the statistical procedure with the same name. The feature to be added to the current subset in machine learning is found by evaluating the performance of the current subset augmented by one new feature using

cross-validation. In forward selection, subsets are built up by adding each remaining feature in turn to the current subset while evaluating the expected performance of each new subset using cross-validation. The feature that leads to the best performance when added to the current subset is retained and the process continues. The search ends when none of the remaining available features improves the predictive ability of the current subset. This process finds a local optimum set of features.

Backward elimination is implemented in a similar fashion. With backward elimination, the search ends when further reduction in the feature set does not improve the predictive ability of the subset. The present embodiments contemplate search algorithms that search forward, backward or in both directions. Representative examples of search algorithms suitable for the present embodiments include, without limitation, exhaustive search, greedy hill-climbing, random perturbations of subsets, wrapper algorithms, probabilistic race search, schemata search, rank race search, and Bayesian classifier.

A decision tree is a decision support algorithm that forms a logical pathway of steps involved in considering the input to make a decision.

The term "decision tree" refers to any type of tree-based learning algorithms, including, but not limited to, model trees, classification trees, and regression trees.

A decision tree can be used to classify the datasets or their relation hierarchically. The decision tree has tree structure that includes branch nodes and leaf nodes. Each branch node specifies an attribute (splitting attribute) and a test (splitting test) to be carried out on the value of the splitting attribute, and branches out to other nodes for all possible outcomes of the splitting test. The branch node that is the root of the decision tree is called the root node. Each leaf node can represent a classification (e.g., whether a particular portion of the group dataset matches a particular portion of the subject-specific dataset) or a value. The leaf nodes can also contain additional information about the represented classification such as a confidence score that measures a confidence in the represented classification (*i.e.*, the likelihood of the classification being accurate). For example, the confidence score can be a continuous value ranging from 0 to 1, which a score of 0 indicating a very low confidence (e.g., the indication value of the represented classification is very low) and a score of 1 indicating a very high confidence (e.g., the represented classification is almost certainly accurate). Decision tree can be applied to the extracted EEG feature(s) and/or the possible clinical decisions that may be taken by the clinician. This allows the technique of the present embodiments to aid the clinician in determining the desired AED and/or its dosage.

Regression techniques which may be used in accordance with the present invention include, but are not limited to linear Regression, Multiple Regression, logistic regression, probit regression, ordinal logistic regression ordinal Probit-Regression, Poisson Regression, negative binomial Regression, multinomial logistic Regression (MLR) and truncated regression.

A logistic regression or logit regression is a type of regression analysis used for predicting the outcome of a categorical dependent variable (a dependent variable that can take on a limited number of values, whose magnitudes are not meaningful but whose ordering of magnitudes may or may not be meaningful) based on one or more predictor variables. Logistic regression may also predict the probability of occurrence for each data point. Logistic regressions also include a multinomial variant. The multinomial logistic regression model is a regression model which generalizes logistic regression by allowing more than two discrete outcomes. That is, it is a model that is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, given a set of independent variables (which may be real-valued, binary-valued, categorical-valued, etc.). For binary-valued variables, a cutoff between the 0 and 1 associations is typically determined using the Yuden Index.

A Bayesian network is a model that represents variables and conditional interdependencies between variables. In a Bayesian network variables are represented as nodes, and nodes may be connected to one another by one or more links. A link indicates a relationship between two nodes. Nodes typically have corresponding conditional probability tables that are used to determine the probability of a state of a node given the state of other nodes to which the node is connected. In some embodiments, a Bayes optimal classifier algorithm is employed to apply the maximum a posteriori hypothesis to a new record in order to predict the probability of its classification, as well as to calculate the probabilities from each of the other hypotheses obtained from a training set and to use these probabilities as weighting factors for future predictions of the subject's response to treatment. An algorithm suitable for a search for the best Bayesian network, includes, without limitation, global score metric-based algorithm. In an alternative approach to building the network, Markov blanket can be employed. The Markov blanket isolates a node from being affected by any node outside its boundary, which is composed of the node's parents, its children, and the parents of its children.

Instance-based techniques generate a new model for each instance, instead of basing predictions on trees or networks generated (once) from a training set.

The term "instance", in the context of machine learning, refers to an example from a dataset.

Instance-based techniques typically store the entire dataset in memory and build a model from a set of records similar to those being tested. This similarity can be evaluated, for example, through nearest-neighbor or locally weighted methods, *e.g.*, using Euclidian distances. Once a set of records is selected, the final model may be built using several different techniques, such as the naive Bayes.

Neural networks are a class of algorithms based on a concept of inter-connected "neurons." In a typical neural network, neurons contain data values, each of which affects the value of a connected neuron according to connections with pre-defined strengths, and whether the sum of connections to each particular neuron meets a pre-defined threshold. By determining proper connection strengths and threshold values (a process also referred to as training), a neural network can achieve efficient recognition of images and characters. Oftentimes, these neurons are grouped into layers in order to make connections between groups more obvious and to each computation of values. Each layer of the network may have differing numbers of neurons, and these may or may not be related to particular qualities of the input data.

In one implementation, called a fully-connected neural network, each of the neurons in a particular layer is connected to and provides input value to those in the next layer. These input values are then summed and this sum compared to a bias, or threshold. If the value exceeds the threshold for a particular neuron, that neuron then holds a positive value which can be used as input to neurons in the next layer of neurons. This computation continues through the various layers of the neural network, until it reaches a final layer. At this point, the output of the neural network routine can be read from the values in the final layer. Unlike fully-connected neural networks, convolutional neural networks operate by associating an array of values with each neuron, rather than a single value. The transformation of a neuron value for the subsequent layer is generalized from multiplication to convolution.

Referring now again to FIG. 1, the method proceeds to **17** at which the predicted response of the subject to the pharmacological treatment is determined based on the output of the machine learning procedure. This can be done by associating the output value or values with a human readable descriptor or descriptors. The method can display, print or store in a computer readable medium a report regarding the predicted response.

When the EEG data are collected after one or more previous sessions of the pharmacological treatment, the method optionally and preferably determines a response of the subject to these previous sessions based on the output of the machine learning procedure. In these embodiments the method can generate an output including a recommendation to correcting future sessions of the treatment, based on the output. For example, when the subject is non-

responsive, to a particular dosage regimen of a particular drug, the output can include a recommendation to increase the dosage of the drug or to change a dosage ratio among two or more drugs, or the like.

The method ends at **18**.

FIG. 3 illustrates a computing platform that can be used according to some embodiments of the present invention for executing method **10**. Shown is a client computer **30** having a hardware processor **32**, which typically comprises an input/output (I/O) circuit **34**, a hardware central processing unit (CPU) **36** (*e.g.*, a hardware microprocessor), and a hardware memory **38** which typically includes both volatile memory and non-volatile memory. CPU **36** is in communication with I/O circuit **34** and memory **38**. Client computer **30** preferably comprises a graphical user interface (GUI) **42** in communication with processor **32**. I/O circuit **34** preferably communicates information in appropriately structured form to and from GUI **42**. An EEG system **66**, configured for collecting EEG signals using an arrangement of EEG electrodes (not shown), optionally and preferably communicates with client computer **30**, particularly with hardware processor.

Also shown is a server computer **50** which can similarly include a hardware processor **52**, an I/O circuit **54**, a hardware CPU **56**, a hardware memory **58**. I/O circuits **34** and **54** of client **30** and server **50** computers preferable operate as transceivers that communicate information with each other via a wired or wireless communication. For example, client **30** and server **50** computers can communicate via a network **40**, such as a local area network (LAN), a wide area network (WAN) or the Internet. Server computer **50** can be in some embodiments be a part of a cloud computing resource of a cloud computing facility in communication with client computer **30** over the network **40**.

GUI **42** and processor **32** can be integrated together within the same housing or they can be separate units communicating with each other. GUI **42** can optionally and preferably be part of a system including a dedicated CPU and I/O circuits (not shown) to allow GUI **42** to communicate with processor **32**. Processor **32** issues to GUI **42** graphical and textual output generated by CPU **36**. Processor **32** also receives from GUI **42** signals pertaining to control commands generated by GUI **42** in response to user input. GUI **42** can be of any type known in the art, such as, but not limited to, a keyboard and a display, a touch screen, and the like. In preferred embodiments, GUI **42** is a GUI of a mobile device such as a smartphone, a tablet, a smartwatch and the like. When GUI **42** is a GUI of a mobile device, processor **32**, the CPU circuit of the mobile device can serve as processor **32** and can execute the code instructions described herein.

Client **30** and server **50** computers can further comprise one or more computer-readable storage media **44**, **64**, respectively. Media **44** and **64** are preferably non-transitory storage media storing computer code instructions as further detailed herein, and processors **32** and **52** execute these code instructions. The code instructions can be run by loading the respective code instructions into the respective execution memories **38** and **58** of the respective processors **32** and **52**. Storage media **64** preferably also store a library of reference data as further detailed hereinabove.

In operation, EEG system **66** optionally and preferably collects EEG signals from the brain of the subject, and transmits the signals to processor **32**. Processor **32** can digitize the EEG signals to provide the EEG data as further detailed hereinabove. Alternatively, the EEG data can be stored on storage **44**, in which case processor **32** reads the EEG data from storage **44** instead of receiving it from EEG system **66**. Processor **32** of client computer **30** receives, via GUI **42**, information pertaining to a candidate pharmacological treatment for the neuropsychiatric disorder, and optionally also a subject profile, as further detailed hereinabove. Processor **32** can transmit the EEG data and the input from GUI **42** over the network **40** to server computer **50**. Computer **50** can extract a feature from the EEG data, and executes a machine learning procedure wherein the extracted feature, the received information and optionally the subject profile are inputs of the machine learning procedure, as further detailed hereinabove. Computer **50** can also determining the predicted response of the subject to the pharmacological treatment based on the output of the machine learning procedure, and transmit the predicted response to client computer **32**. Alternatively, one or more of the operations, or all the operations, can be executed by computer **32**.

The word "exemplary" is used herein to mean "serving as an example, instance or illustration." Any embodiment described as "exemplary" is not necessarily to be construed as preferred or advantageous over other embodiments and/or to exclude the incorporation of features from other embodiments.

The word "optionally" is used herein to mean "is provided in some embodiments and not provided in other embodiments." Any particular embodiment of the invention may include a plurality of "optional" features unless such features conflict.

The terms "comprises", "comprising", "includes", "including", “having” and their conjugates mean "including but not limited to".

The term “consisting of” means “including and limited to”.

The term "consisting essentially of" means that the composition, method or structure may include additional ingredients, steps and/or parts, but only if the additional ingredients, steps

and/or parts do not materially alter the basic and novel characteristics of the claimed composition, method or structure.

As used herein, the singular form "a", "an" and "the" include plural references unless the context clearly dictates otherwise. For example, the term "a compound" or "at least one compound" may include a plurality of compounds, including mixtures thereof.

Throughout this application, various embodiments of this invention may be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 3, 4, 5, and 6. This applies regardless of the breadth of the range.

Whenever a numerical range is indicated herein, it is meant to include any cited numeral (fractional or integral) within the indicated range. The phrases “ranging/ranges between” a first indicate number and a second indicate number and “ranging/ranges from” a first indicate number “to” a second indicate number are used herein interchangeably and are meant to include the first and second indicated numbers and all the fractional and integral numerals therebetween.

It is appreciated that certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable subcombination or as suitable in any other described embodiment of the invention. Certain features described in the context of various embodiments are not to be considered essential features of those embodiments, unless the embodiment is inoperative without those elements.

Various embodiments and aspects of the present invention as delineated hereinabove and as claimed in the claims section below find experimental support in the following examples.

EXAMPLES

Reference is now made to the following examples, which together with the above descriptions illustrate some embodiments of the invention in a non limiting fashion.

Example 1

Predicting Efficiency of Anti-Epileptic Drugs using Machine Learning and EEG

Epilepsy is a chronic disease, affecting approximately 1% of the population. Most patients start suffering from the disease in childhood and adolescence or when they are over 65. The majority of epileptic seizures are controlled pharmaceutically by antiepileptic drugs (AEDs), which reduce neuronal burst firing and synchronization in localized neuronal ensembles and inhibit the spread of abnormal firing to distant brain sites via three main molecular mechanisms: (i) modulation of voltage-gated ion channels (mostly calcium and sodium); (ii) inhibition of synaptic excitation; and (iii) enhancement of synaptic inhibition. There are several different types of AEDs, four of which are considered primary, as they cover together roughly 80% of the cases.

In this Example a tool for measuring the effects of certain AEDs on brain activity is demonstrated.

EEG recordings of patients from diverse age groups have been collected, and different machine learning procedures were employed to classify the recordings into two groups: recordings of patients successfully medicated by one AED, and recordings of patients that were not medicated at the time of the EEG recording.

In this Example, the task of predicting the effectiveness of an AED is formulated into a supervised binary classification problem. A machine learner was given a training set with example from two classes, along with labels that indicates from which class each example was sampled. The task of the learner was to find a prediction rule that given an unlabeled example predicts its label. The performance of the learner was measured on a new unseen data set, referred to as the test set. In the present Example, the two classes were epileptic patients with no medicine, referred to herein as class D0, and epileptic patients that are stable with a treatment of a single AED, where "stable" was defined as either having no seizure over the past year, referred to herein as subclass D4, or stable clinically with no symptoms in the EEG but for less than a year, referred to herein as subclass D3.

The experiment was performed for three types of AED: Depalept™, Keppra™ and Lamictal™. For each AED a different learner was trained and its performance was evaluated separately from the other two. Because of the difficulty in collecting a large data set of patients, a leave-one-out method was employed where at each iteration a learner is trained on the entire dataset minus one patient that might have one or more records, and then tested on that patient. The average accuracy of the learner on the patients was calculated to provide an estimation of the learner's performance.

A preprocessing was applied the EEG data. The preprocessing consisted of high pass filter, which passed only frequencies higher than 0.1 Hz; a low pass filter which passed only frequencies lower than 70 Hz; a notch filter at 50 Hz; and an electrode interpolation using spherical spline interpolating method. Good segments, at different lengths, were selected from the data by a human expert.

Two different features of the EEG data were tested: covariance and coherence. For the covariance feature, the covariance of the phase and the power in different frequency bands, in each pair of electrodes. A Fourier Transform of the signal was computed at each electrode by FFT, and the data were separated into the following frequency bands 1-4 Hz, 4-8 Hz, 8-13 Hz, 13-30 Hz and 30-50 Hz, referred to in this example as delta, theta, alpha, beta and gamma, respectively. The power covariance of the FFT result (the covariance applied to the absolute component) and the phase covariance (the covariance applied to the angle component) were computed at each band. Covariance of absolute component and covariance of the phase component were tested separately as well as in combination.

For the coherence feature, the following expression was calculated:

Matlab® FFT algorithm was applied to each 2-second epoch of data yielding Fourier coefficients with frequency resolution $\Delta f = 0.5$ Hz at each of the $n = 1, 2,..., N$ channels for epochs $k = 1, 2,..., K$, with $K = 120$ epochs in each brain state. The cross spectrum $G_{uv}(f_n)$ estimate based on the complex valued coefficients $F_{uk}(f_n)$, $F_{uk}(f_n)$ for channel pair (u,v) is:

$$G_{uv}(f_n) = \frac{2}{K}\sum_{k=1}^{K} F_{uk}(f_n)\, F_{vk}^{*}(f_n) \qquad n = 1, N$$

For single channels ($u = v$) the cross-spectrum reduces to the usual power spectrum $P_u(f_n) = G_{uu}(f_n)$. The coherence γ_{uv} between channels u and v at each frequency f is the normalized cross spectrum, given by:

$$\gamma_{uv}^{2}(f) = \frac{|G_{uv}(f)|^2}{G_{uu}(f)\, G_{vv}(f)}$$

The coherence was calculated between all pairs of data channels.

Two different learners were tested. KNN and SVM.

The KNN learner stored, at the training stage, all the training data into the memory. When a new example was presented, the learner found the K closest examples from the training set according to a similarity measure, and predicts the most common label among those examples. Both L1 and L2 norms were used for distance measure, with similar results.

The SVM learner found a maximal margin hyperplane that divided the training data into two classes, where the margin was defines as the minimal distance of an example from the hyperplane.

Each patient in the data set had one or more records, which might have different labels. From each record, a number of clean segments were extracted by the human expert. The learner was trained on all of the segments from the current training set. When a new record was presented for classification, the learner predicted a label for each of the record's segments. In order to predict the label of the record, voting over the labels of the different segments was employed. Several voting methods were tested, all providing similar results. In this example, a simple majority vote is presented.

Weighted accuracy was used as performance measure. The weighted accuracy was defined as (TP/No. of Positive + TN/No. of Negative)/2, where TP and TN are true positive and true negative, respectively.

The data set used in the analysis is summarized in Table 1.

Table 1

	Patients	Records	Segments
class D0	38	48	836
Depalept	24	37	518
Keppra	10	20	363
Lamictal	38	53	1523

In one analysis, phase and power frequency covariance were used and the performances of SVM and KNN were compared on the different data set. The K parameter for KNN was chosen empirically. The obtained predicted accuracy values are summarized in Table 2.

Table 2

		phase cov	power cov	power + phase cov	coherence
Depalept	SVM	0.58	0.66	0.61	0.61
	KNN(K=7)	0.73	0.67	0.69	0.71
Keppra	SVM	0.65	0.69	0.64	0.6
	KNN(K=20)	0.78	0.7	0.76	0.68
Lamictal	SVM	0.55	0.5	0.58	0.52
	KNN(K=20)	0.57	0.69	0.56	0.54

As a baseline, a simple feature which calculates the phase and power covariance for all of the frequencies was used, as summarized in Table 3.

Table 3

		phase cov	power cov	power + phase cov
Depalept	SVM	0.68	0.65	0.66
	KNN(K=7)	0.72	0.63	0.73
Keppra	SVM	0.64	0.67	0.57
	KNN(K=20)	0.76	0.68	0.72
Lamictal	SVM	0.54	0.59	0.59
	KNN(K=20)	0.58	0.53	0.67

The mean covariance provided better results for both Lamictal™ and Depalept™. A possible explanation for this is that one of the frequency bands is more dominant and relevant for the predictions. In order to test this hypothesis, a KNN learner was trained using the power and phase covariance at one frequency band only. The obtained predicted accuracy values are summarized in Table 4.

Table 4

	delta	theta	alpha	beta	gamma
Depalept(K=7)	0.49	0.63	0.58	0.68	0.71
Keppra(K=20)	0.76	0.7	0.71	0.74	0.76
Lamictal(K=20)	0.51	0.5	0.52	0.48	0.56

Table 4 demonstrates one or two frequency bands with a better prediction accuracy than the others for Depalept™ and Lamictal™, but a small difference between the prediction accuracy for each frequency band for Keppra™.

PCA was employed to reduce the data dimension of the frequency representation (both power and phase) from 3610 to 100. The obtained predicted accuracy values for band frequency covariance are summarized in Table 5.

Table 5

	Depalept	Keppra	Lamictal
KNN with PCA	0.67	0.75	0.59
KNN without PCA	0.63	0.76	0.55

Table 5 demonstrates that PCA improves the prediction accuracy of the frequency band covariance for Depalept™ and Lamictal™, which is in line with the results presented in Table 4, above.

PCA was also applied for the mean of covariance values as calculated per frequency band. In this case, the dimension was reduced from 722 to 100. The predicted accuracy values for the mean of covariance values are summarized in Table 6.

Table 6

	Depalept	Keppra	Lamictal
KNN with PCA	0.73	0.74	0.66
KNN without PCA	0.73	0.76	0.66

Table 6 demonstrates that the prediction accuracy of the mean of covariance values is not affected by the dimension of the data.

Dimensionality was also reduced by forward greedy feature selection. For each feature f_i, the conditioned probability density function (PDF) of the feature was estimated. Two probability density functions were defined, referred to herein as Q^+ and Q^-, where Q^+ denotes the conditional probability density that f_i describes a data sample belonging to the class of stable patients, and Q^- denotes the conditional probability density that f_i describes a data sample belonging to the class of patients that where no administered with AED. Formally:

$$Q^+ = P(f_i \mid \text{stable}), \text{ and}$$

$$Q^- = P(f_i \mid \text{no medicine}).$$

The PDFs were estimated using kernel density estimation which allows estimating the PDFs without making assumptions about the distribution of the data. The following two Kullback Leibler divergences were calculated for Q^+ and Q^-, $D_KL(Q^- \parallel Q^+)$ and $D_KL(Q^+ \parallel Q^-)$. The average of these divergences was computed and used as a score for selecting the features. This allowed selecting the feature or features that best separated between the two classes, since the Kullback Leibler divergence between the conditional probability densities can serve as a measure of separateness. The obtained predicted accuracy values of the forward greedy feature selection for the band frequency covariance are summarized in Table 7. In Table 7, S denotes the number of features (with highest score) that were selected.

Table 7

S	Depalept	Keppra	Lamictal
100	0.66	0.73	0.59
50	0.65	0.78	0.65
20	0.63	0.78	0.69

In two additional tests, the KNN learner was trained on the D0, and on the D4 subclass. KNN and mean frequency covariance were used, with the same parameters as above. The obtained predicted accuracy values are summarized in Table 8.

Table 8

	Depalept	Keppra	Lamictal
D0 vs D3+D4	0.73	0.72	0.67
D0 vs D4	0.69	0.88	0.6

The present Example demonstrates the ability of the technique of the present embodiments to classify EEG data into two classes: stable patients medicated by a single medicine, and unstable patients, with no medicine. KNN analysis was found to be a better learner for this task than SVM. The best prediction accuracy for Keppra™ was 78 % achieved using KNN with K=20, band frequency covariance (power and phase) and phase covariance (see Table 2), and was also confirmed in the feature selection test (see Table 7). The best prediction accuracy for Depalept™ was 73 %, achieved using KNN with K=20, and mean frequency covariance. The best prediction accuracy for Lamictal™ was 69 %, achieved using KNN with K=20 and mean power covariance.

The features which provided better results were frequency covariance in the higher frequency bands.

Some other future direction might include features with larger interpretability – for instance, it might be useful to look at the signal at the brain sources rather than the electrodes. This can be done by finding the sources of the activity recorded in the EEG using inverse solution methods.

Although the invention has been described in conjunction with specific embodiments thereof, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims.

All publications, patents and patent applications mentioned in this specification are herein incorporated in their entirety by reference into the specification, to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated herein by reference. In addition, citation or identification of any reference in this application shall not be construed as an admission that such reference is available as prior art to the present invention. To the extent that section headings are used, they should not be construed as necessarily limiting.

WHAT IS CLAIMED IS:

1. A method of analyzing EEG data collected from a brain of a subject diagnosed with a neuropsychiatric disorder, the method comprising:

extracting a feature from the EEG data;

receiving information pertaining to a candidate pharmacological treatment for said neuropsychiatric disorder;

executing a machine learning procedure wherein said extracted feature and said received information are inputs of said machine learning procedure; and

determining a predicted response of the subject to said pharmacological treatment based on an output of said machine learning procedure.

2. The method according to claim 1, wherein said execution of said machine learning and said determination of said predicted response is by a server system and the method comprises receiving the EEG data and said information by said server system over a communication network,.

3. The method according to claim 2, further comprising transmitting said predicted response to a client computer.

4. The method according to claim 1, further comprising collecting EEG signals from the brain of the subject, and digitizing said EEG signals to provide said EEG data.

5. The method according to any of claims 1-4, wherein said feature comprises covariance between data channels.

6. The method of claim 5, further comprising separating the data into an absolute component and a phase component prior to said feature extraction, wherein said covariance is calculated separately for at least one of said components.

7. The method according to any of claims 5 and 6, further comprising decomposing the data into frequency bands, wherein said covariance is calculated separately for at least one of said frequency bands.

8. The method of claim 1, further comprising decomposing the data into frequency bands, wherein said extracting said feature comprises calculating covariance between data channels separately for at least two of said frequency bands, and wherein said feature comprises a mean of said covariance over said at least two of said frequency bands.

9. The method according to any of claims 5-8, wherein said feature comprises a coherence level of the data.

10. The method according to any of claims 1-9, wherein said machine learning procedure comprises at least one procedure selected from the group consisting of k-nearest neighbors analysis, support vector machine, clustering, linear modeling, decision tree learning, ensemble learning procedure, neural networks, probabilistic model, graphical model, Bayesian network, and association rule learning.

11. The method according to any of claims 1-10, wherein the neuropsychiatric disorder is a disorder affecting the central nervous system.

12. The method according to any of claims 1-11, wherein the neuropsychiatric disorder comprises at least one disorder selected from the group consisting of dementia, movement disorder, depression, anxiety disorder, bipolar disorder, schizophrenia, obsessive-compulsive disorder, Parkinson's disease, and chronic pain.

13. The method according to any of claims 1-12, wherein the neuropsychiatric disorder comprises epilepsy.

14. The method according to any of claims 1-13, wherein said candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of an antiepileptic drug.

15. The method according to any of claims 1-13, wherein said candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of a combination of at least two antiepileptic drugs.

16. The method according to any of claims 1-13, wherein said candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of a drug selected from the group consisting of valproic acid, levetiracetam, lamotrigine, ezogabine, pregabalin, topiramate, rufinamide, gabapentin, clonazepam, oxcarbazepine, phenobarbital, phenytoin, sultiam and felbamate, or salts or prodrugs or derivatives of any purity or sustained release formulation thereof.

17. The method according to any of claims 1-16, wherein said candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of carbamazepine or salts or prodrugs or derivatives of any purity or sustained release formulation thereof.

18. The method according to claim 1, wherein the neuropsychiatric disorder is epilepsy, wherein said candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of valproic acid, wherein the method comprises decomposing the data into frequency bands, wherein said extracting said feature comprises calculating covariance between data channels separately for at least two of said frequency bands, wherein said feature comprises a mean of said covariance over said at least two of said frequency bands, and wherein said machine learning procedure comprises k-nearest neighbors analysis.

19. The method according to claim 1, wherein the neuropsychiatric disorder is epilepsy, wherein said candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of levetiracetam, wherein the method comprises, prior to said feature extraction, decomposing the data into frequency bands, and separating the data into an absolute component and a phase component, wherein said feature comprises covariance between data channels calculated separately for each component and each frequency band, and wherein said machine learning procedure comprises k-nearest neighbors analysis.

20. The method according to claim 1, wherein the neuropsychiatric disorder is epilepsy, wherein said candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of levetiracetam, wherein the method comprises, prior to said feature extraction, decomposing the data into frequency bands, and extracting a phase component of the data for each frequency band, wherein said feature comprises covariance between data

channels calculated separately for at least two of said frequency bands, and wherein said machine learning procedure comprises k-nearest neighbors analysis.

21. The method according to claim 1, wherein the neuropsychiatric disorder is epilepsy, wherein said candidate pharmacological treatment comprises administration of a predetermined therapeutic amount of lamotrigine, wherein the method comprises, prior to said feature extraction, decomposing the data into frequency bands, and extracting an absolute component of the data for each frequency band, wherein said extracting said feature comprises calculating covariance between data channels for said absolute component separately for at least two of said frequency bands, wherein said feature comprises a mean of said covariance over said at least two of said frequency bands, and wherein said machine learning procedure comprises k-nearest neighbors analysis.

22. The method according to any of claims 1-21, wherein said EEG data are collected after at least one session of said candidate pharmacological treatment, and the method comprises determining a response of the subject to said at least one session of said pharmacological treatment and generating output including a recommendation for future sessions of said pharmacological treatment based on an output of said machine learning procedure.

23. A computer software product, comprising a computer-readable medium in which program instructions are stored, which instructions, when read by a data processor, cause the data processor to receive EEG data collected from a brain of a subject diagnosed with a neuropsychiatric disorder, and information pertaining to a candidate pharmacological treatment for said neuropsychiatric disorder, and to execute the method according to any of claims 1-21.

24. A system for analyzing EEG data, the system comprising a processor arranged to perform code instructions, comprising:

code instructions for receiving EEG data collected from a brain of a subject diagnosed with a neuropsychiatric disorder;

code instructions for extracting a feature from the EEG data;

code instructions for receiving information pertaining to a candidate pharmacological treatment for said neuropsychiatric disorder;

code instructions for executing a machine learning procedure wherein said extracted feature and said received information are inputs of said machine learning procedure; and

code instructions for determining a predicted response of the subject to said pharmacological treatment based on an output of said machine learning procedure.

25. The system according to claim 24, further comprising a transceiver arranged to receive and transmit information on a communication network;

wherein said processor is arranged to communicate with said transceiver over said communication network; and

wherein said code instructions comprise code instructions for transmitting said predicted response to a client computer.

26. The system according to any of claims 24 and 25, wherein said determining said predicted response comprises classifying whether said treatment is effective, partially effective or not effective.

27. The system according to any of claims 24-26, wherein said code instructions comprise code instructions for generating an output indicating a recommended change in said treatment.

28. The system according to any of claims 24-27, wherein said code instructions comprise code instructions for generating an output pertaining to potential side effects cause by said treatment.

29. The system according to any of claims 24-28, wherein said code instructions comprise code instructions for receiving blood test data of said subject, following said treatment, analyzing said blood test data to determine a response of said subject to said treatment, and comparing said response to said predicted response.

30. The system of claim 29, wherein said analysis comprises determining vitamin deficiency.

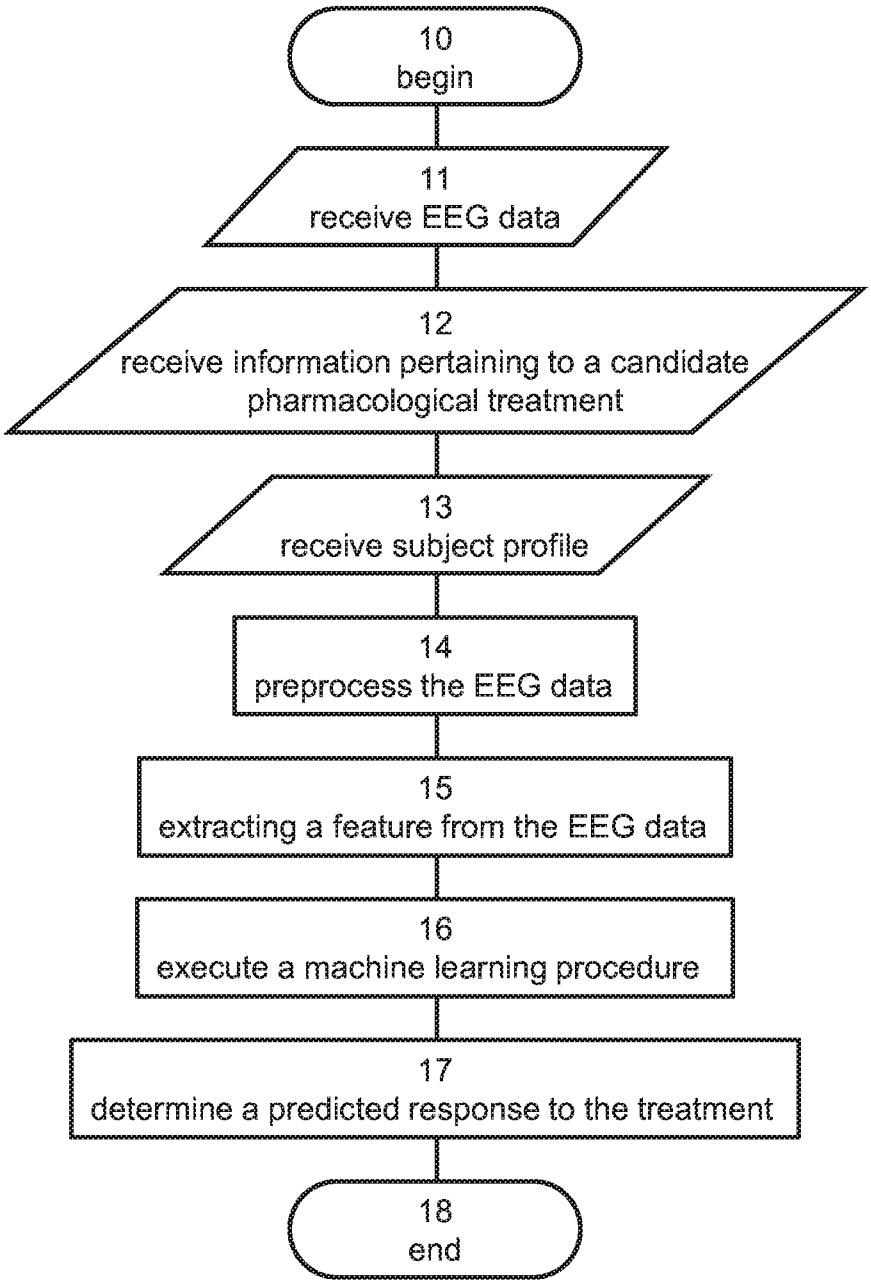

10
begin
11
receive EEG data
12
receive information pertaining to a candidate pharmacological treatment
13
receive subject profile
14
preprocess the EEG data
15
extracting a feature from the EEG data
16
execute a machine learning procedure
17
determine a predicted response to the treatment
18
end

FIG. 1

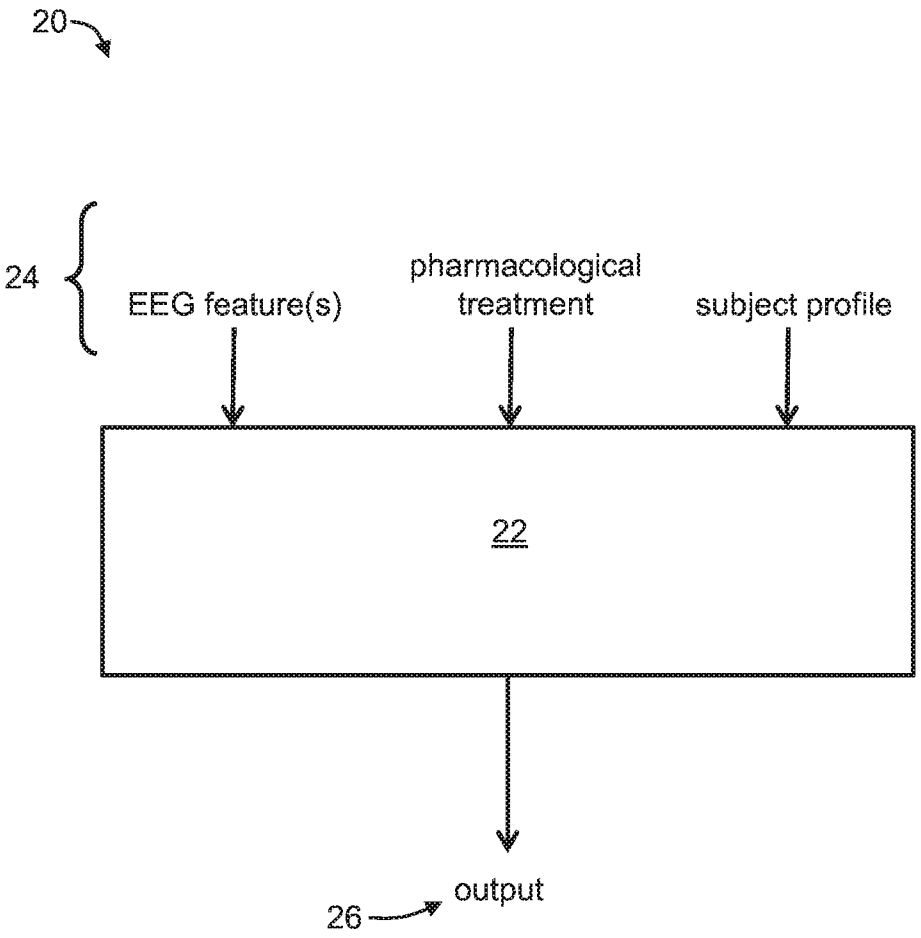

20
24
EEG feature(s)
pharmacological treatment
subject profile
22
26
output

FIG. 2

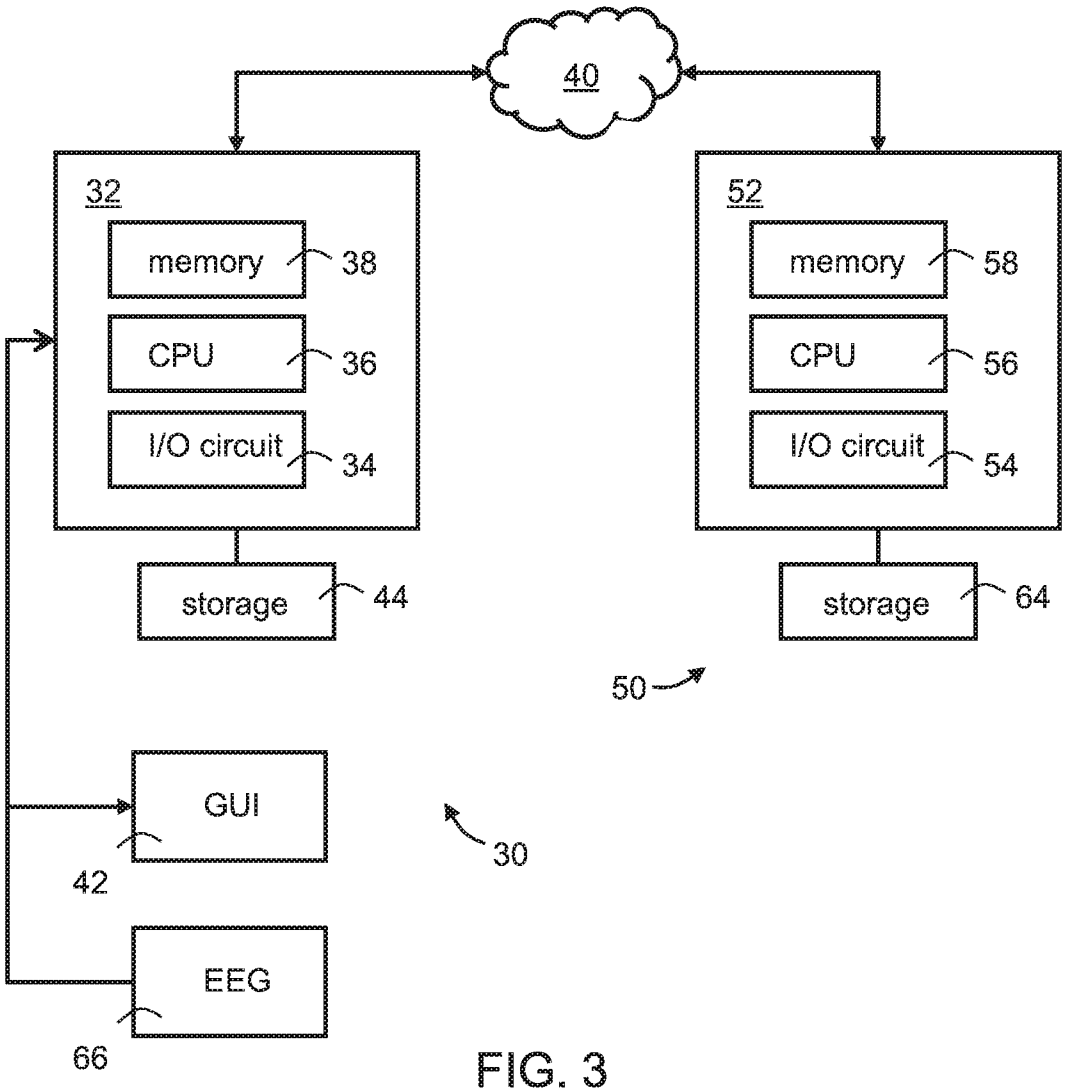

40
32
memory
38
CPU
36
I/O circuit
34
storage
44
52
memory
58
CPU
56
I/O circuit
54
storage
64
50
GUI
42
30
EEG
66

FIG. 3

INTERNATIONAL SEARCH REPORT

International application No
PCT/IL2018/050675

A. CLASSIFICATION OF SUBJECT MATTER
INV. G16H50/20
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G16H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SONJA SIMPRAGA: "EEG biomarker integration for better decision making in clinical trials", INTERNATIONAL PHARMACO-EEG SOCIETY, 26 October 2016 (2016-10-26), pages 41-42, XP55503693, http://www.ipeg-society.org/userfiles/files/IPEG%202016-.pdf	1,12,23, 24
Y	the whole document	1-30
	----- -/--	

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 September 2018	18/09/2018

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Vogt, Titus

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	SONJA SIMPRAGA ET AL: "EEG machine learning for accurate detection of cholinergic intervention and Alzheimer's disease", SCIENTIFIC REPORTS, vol. 7, no. 1, 18 July 2017 (2017-07-18), page 5775, XP55503684, DOI: 10.1038/s41598-017-06165-4	1,12,23, 24
Y,P	abstract	1-30

X	KHODAYARI-ROSTAMABAD A ET AL: "Using pre-treatment EEG data to predict response to SSRI treatment for MDD", 2010 ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY : (EMBC 2010) ; BUENOS AIRES, ARGENTINA, 31 AUGUST - 4 SEPTEMBER 2010, IEEE, PISCATAWAY, NJ, USA, 31 August 2010 (2010-08-31), pages 6103-6106, XP032109523, DOI: 10.1109/IEMBS.2010.5627823 ISBN: 978-1-4244-4123-5	1,4,5, 7-12,23, 24,26
Y	Abstract, Introduction, Merthods, Results; table II	1-30

X	KHODAYARI-ROSTAMABAD A ET AL: "A pilot study to determine whether machine learning methodologies using pre-treatment electroencephalography can predict the symptomatic response to clozapine therapy", CLINICAL NEUROPHYSIOLOGY, ELSEVIER SCIENCE, IE, vol. 121, no. 12, 1 December 2010 (2010-12-01), pages 1998-2006, XP027448967, ISSN: 1388-2457, DOI: 10.1016/J.CLINPH.2010.05.009 [retrieved on 2010-06-17]	1,4,5, 7-12,23, 24,26
Y	Abstract; Methods;; figure 1; table 5	1-30

X	KHODAYARI-ROSTAMABAD AHMAD ET AL: "Latent Variable Dimensionality Reduction Using a Kullback-Leibler Criterion and Its Application to Predict Antidepressant Treatment Response", 2013 INTERNATIONAL WORKSHOP ON PATTERN RECOGNITION IN NEUROIMAGING, IEEE, 22 June 2013 (2013-06-22), pages 148-151, XP032486013, DOI: 10.1109/PRNI.2013.46 [retrieved on 2013-09-18]	1,4,5, 7-12, 22-24,26
Y	Abstract, Methods, Results, Conclusions; table I	1-30

	-/--	

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RAVAN MARYAM ET AL: "A machine learning approach using auditory odd-ball responses to investigate the effect of Clozapine therapy", CLINICAL NEUROPHYSIOLOGY, ELSEVIER SCIENCE, IE, vol. 126, no. 4, 27 August 2014 (2014-08-27), pages 721-730, XP029203467, ISSN: 1388-2457, DOI: 10.1016/J.CLINPH.2014.07.017	1,4,5, 7-12, 22-24,26
Y	tables 1-9	1-30

X	SINISA COLIC ET AL: "Prediction of antiepileptic drug treatment outcomes using machine learning", JOURNAL OF NEURAL ENGINEERING, INSTITUTE OF PHYSICS PUBLISHING, BRISTOL, GB, vol. 14, no. 1, 30 November 2016 (2016-11-30), page 16002, XP020312943, ISSN: 1741-2552, DOI: 10.1088/1741-2560/14/1/016002 [retrieved on 2016-11-30]	1,4, 10-14, 16, 22-24,26
Y	Abstract, Methods, Results, Conclusions.; figures 3,6,7	1-30

X	WAJID MUMTAZ ET AL: "A wavelet-based technique to predict treatment outcome for Major Depressive Disorder", PLOS ONE, vol. 12, no. 2, 2 February 2017 (2017-02-02), page e0171409, XP55503720, DOI: 10.1371/journal.pone.0171409	1,4,5, 7-12,23, 24
Y	Abstract, Materials and Methods.; figure 3; tables 11,13,14	1-30

1