Title: Biomarkers of Depression Symptoms Defined by Direct Intracranial Neurophysiology

Authors: Scangos KW[1,3]*, Khambhati AN[2,3], Daly PM[1,3], Owen LW[4], Manning JR[4], Ambrose JB[5], Austin E[5], Dawes HE[2,3], Krystal AD[1,3], Chang EF[2,3]

[1]Department of Psychiatry, University of California, San Francisco

[2]Department of Neurological Surgery, University of California, San Francisco

[3]Weill Institute for Neurosciences, University of California, San Francisco

[4]Department of Psychological and Brain Sciences, Dartmouth College

[5]Kaiser Redwood City, 100 Veterans Blvd, Redwood City, CA 94063

*** Corresponding Author:**

Katherine Scangos, MD, PhD

Assistant Professor of Psychiatry

University of California, San Francisco

katherine.scangos@ucsf.edu

(415) 476-7439

Abstract:

Quantitative biomarkers of depression are critical for development of rational therapeutics, but limitations of current low-resolution, indirect brain assays may impede their discovery. We applied graph theory and machine learning to a large unique dataset of intracranial electrophysiological recordings to generate a four-dimensional whole-brain model of neural activity. Using this model, we found patterns of network activity that correctly classified depression in over 80% of individuals. These complex patterns were especially evident in alpha and beta spectral power across frontal and occipital brain regions, respectively. Our findings reveal a widespread network of abnormal activity that may inform advanced personalized treatment.

Main:

Neurons work together in integrated brain networks to enable high-level functions such as human emotion and complex behavior[1,2]. Dysfunction in these networks can result in major depressive disorder, which is a leading cause of world-wide disability and lacks adequate treatment options (World Health Organization, 2012). A quantitative biomarker of depression that characterizes network dysfunction could enable early detection, objective disease classification, monitoring of disease progression, and an understanding of circuit physiology from which rational treatment design could evolve[3].

The majority of early studies seeking to characterize the biological substrate of depression have focused on specific brain regions (ex. subgenual anterior cingulate cortex[4-6]), distributed cognitive networks (ex. default mode network[2,7-10]), or univariate electrophysiological markers (ex. alpha asymmetry[11-16]). While these important studies set the framework for understanding the neural basis for depression, no consensus on a biomarker has been reached[10,17-21]. Recent computational advancements have introduced artificial intelligence in medical diagnostics with accuracies that rival physicians[22,23]. In psychiatry, studies have leveraged such data-driven approaches as well, applied to imaging modalities[24-32] or electrocorticography (EEG)[33,34]. However, the heterogeneity of depression and the inadequate resolution of current measurement tools may have impeded detection of reliable depression signals. Direct sampling from brain regions themselves through intracranial EEG (iEEG) may overcome such challenges and complement noninvasive techniques to provide insight into the architecture of abnormal depression circuits. These high-resolution recordings from cortical and subcortical brain regions can be performed for multi-day time periods in patients with epilepsy undergoing pre-surgical mapping. Such patients have high rates of co-morbid depression[35-40] that shares origin[41-45] and

treatment response[46] characteristics with primary depression. Findings in these patients therefore may inform understanding of depression circuits more generally.

Despite clear advantages of iEEG, population-level iEEG analyses of major depression have not been performed. One major barrier is the heterogeneous sampling of disease-relevant networks across subjects, owing to the clinical placement of electrodes for seizure monitoring. To overcome this challenge, we utilized graph theory and a novel mathematical approach that allowed us to create a model of multiregional iEEG activity for each individual based on information provided by the correlational structure of activity across our population (SuperEEG[47]). The resulting model provided a highly detailed representation of naturalistic brain activity across space and time, and its inherent organization into functional brain networks. By quantifying the multi-dimensional nature of corresponding brain dynamics, our study enabled us to better understand how patterns of rhythmic activity underlying functional connectivity differ in depressed and non-depressed brain networks[48].

We detected patterns of neural activity and connectivity that effectively distinguished depressed and non-depressed individuals. Yet the same network changes were not present in everyone with depression. Two depression-associated patterns emerged, and the presence of either one increases the likelihood of depression in any individual. One pattern was strongly characterized by excessive frontal cortical connectivity with decreased activity in the alpha spectral frequency band. The second was marked by decreased connectivity across the occipitotemporal region and dominant beta band activity.

Results:

Overall Approach

We took a two-step approach to quantify iEEG biomarkers of depression. With the first step ("Model building", **Fig. 1**), our overall goal was to identify large-scale functional networks across iEEG electrodes. To tackle inconsistent network sampling across individuals, we utilized a novel method called SuperEEG[47] that uses the correlational relationship of brain activity across electrodes to generate a whole-brain iEEG model for each subject. We next uncovered the inherent organization of the brain's correlational structure from the output of this model by examining network modularity, a graph theoretic concept that quantifies the organization of a system into constituent modules[49,50], and has been used to reveal system disruptions in disease states[51-57] including MDD[29]. This approach revealed 6 functional networks in our data, which we refer to as "modules."

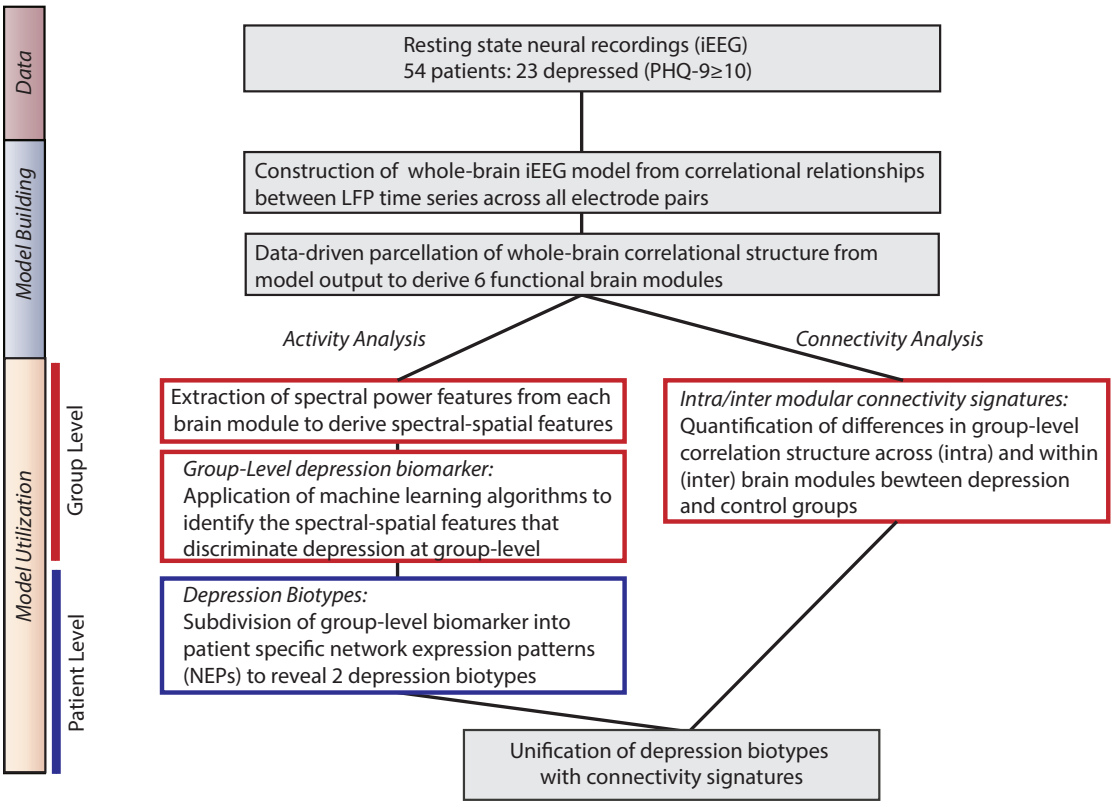

Fig. 1. Overall approach. *Model Building:* We utilized direct neural recordings from 54 patients to construct a whole-brain model of iEEG activity based on correlational relationships of neural LFP time series signals across all electrode pairs. We then parcellated this model into functional network modules using graph theory metrics. *Model Utilization:* We used the whole-brain iEEG model to study how brain activity and connectivity measures relate to depression status. We first defined spectral power features across network modules and applied supervised machine

learning to identify a group-level biomarker of depression (*Activity analysis*). In parallel, we identified alterations in functional network connectivity and organization between depressed and control groups (*Connectivity analysis*). Common features of the group-level biomarker expressed at the individual level clustered participants into neurophysiological biotypes with distinct patterns of altered activity and connectivity.

In the second step ("Model utilization", **Fig. 1**), we used the whole-brain model to relate the architecture and intrinsic neural activity of functional networks to depression status as defined by a validated self-report measure, the Patient Health Questionnaire (PHQ-9)[58-60]. Informed by previous work that identified spectral power features from quantitative EEG that were predictive of depression diagnosis and treatment response[11-16], we first performed an activity analysis, where we extracted spectral power features across each module, and then tested their ability to detect depressed from non-depressed patients at the group level using supervised machine learning. We next asked whether biomarkers defined at the group level account for inter-individual heterogeneity; a critical barrier to their clinical application. To address this question, we examined whether common features of the group-level biomarker clustered depressed patients into different neurophysiological biotypes. Finally, to assess the relationship between the brain's neural activity and its underlying functional connectivity, we performed a connectivity analysis in parallel to the activity analysis, where we quantified differences in the group-level correlation structure across depressed and non-depressed groups. By unifying results from both analyses, we linked distinct patterns of network disorganization with dysfunctional activity across depression subtypes.

Derivation of functional modules

Our total population consisted of 54 patients with intracranial electrodes placed for the purpose of pre-surgical mapping for treatment-resistant epilepsy. The number of electrodes per subject ranged from 33–201 (mean=120, SD=37). We utilized continuous iEEG recordings across this population to build a whole-brain, multi-subject model of iEEG recordings that served as the basis for studying system-level organization and signatures of depression (**Fig2a-c**, see *Methods*). In

brief, to generate this model, we first constructed subject-level full-brain correlational models. Interelectrode correlation matrices were constructed from activity where sensors were present and learned radial-basis function weighted averages were used to generate correlational information at locations where sensors were not present. The subject-level models were then averaged to generate a population-level model. We then used Gaussian process regression based on the population-level model and individual time series for each subject to reconstruct whole-brain local field potentials for each subject. The output of this SuperEEG[47] model is therefore an estimate of iEEG time series for a union set of electrodes across our total population (4,244 electrodes/per patient).

After construction of the model, we next utilized principles of graph theory to identify data-driven functional networks (modules) across it. Our rationale was that the inherent network structure present in the model could enhance discovery of depression circuitry over an atlas-based approach because atlases apply boundaries to brain regions based on structural rather than functional organization and thus, while they provide a useful comparison to a data-driven parcellation scheme, there is no reason to assume their boundaries will perfectly align. We applied a modularity optimization technique known as community detection which groups electrodes into non-overlapping modules by their correlational relationships[61,62] (**Fig. 2**d-e, and see *Methods*). Using this method, it is possible to study network organization at different levels of granularity. To objectively select the level of granularity, we compared the similarity of our data-driven partitions with those of a commonly employed Lausanne atlas, by calculating the adjusted Rand similarity index[63]. This method yielded parcellation into 6 stable modules (Jaccard index, $p<0.05$, permutation test, **Supplementary Fig. 2**). A graph of the network and subdivision into modules is shown in **Fig. 2e** where module membership is indicated by the color of nodes (iEEG electrodes) and edges (inter-electrode correlation from whole-brain model). To name each module, we pinpointed regions (hubs) with the most influential connectivity profile for each module

by calculating the participation coefficient metric[64-66] (see *Methods*) yielding the left dorsolateral prefrontal cortical (L-DLPFC), left occipitotemporal (L-OT), left orbitofrontal cortical (L-OFC), right frontotemporal (R-FT), right medial frontal (R-MF), and mid-hemispheric modules. **Fig. 2f** shows hub location by their mean Montreal Neurological Institute (MNI)[63] coordinates and associated Brodmann Areas.

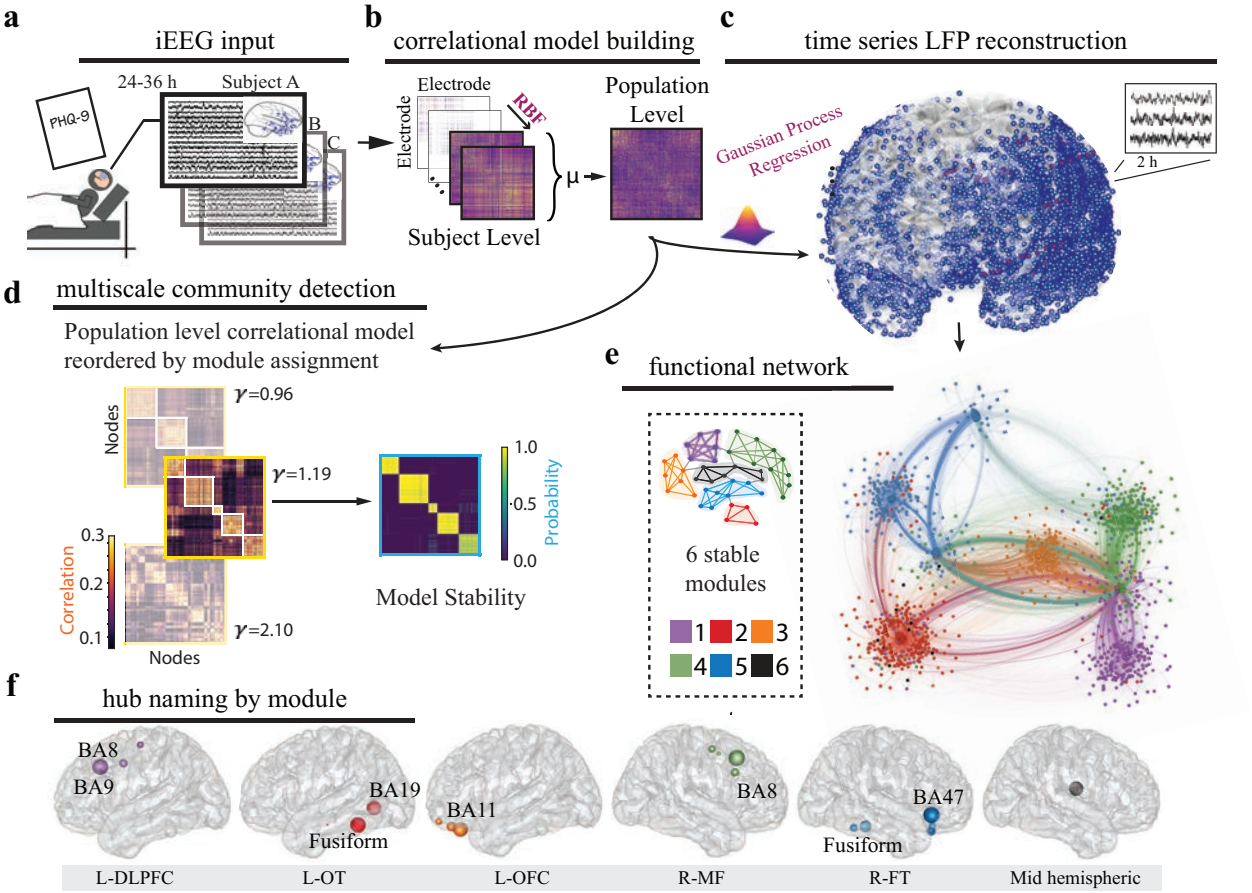

Fig. 2. **Identification of functional modules**. a. To generate a multi-subject whole-brain model of iEEG activity, patient's electrode locations across participants were first represented in a common space (Montreal Neurologic Institute (MNI) space). Electrode locations and sample recordings for a few example patients are shown. Activity was then randomly sampled in 1 min intervals across daytime hours to obtain a stable representation of resting-state brain activity across a 2-hour period. b. Individual inter-electrode correlation matrices were constructed for each participant at locations where electrodes were present. Subject-level full brain correlational models were then predicted using radial basis function (RBF)-weighted averages to estimate brain activity correlations at locations where sensors were not present (left). Subject-level correlational models were then averaged to generate a population level whole-brain correlational model (right). c. Local field potential activity for each of the 4,244 electrodes was then reconstructed using Gaussian process regression with the population-level model as a prior and

activity where electrodes were present as the marginal likelihood. d. Multiscale community detection was applied to the whole-brain model to group electrodes (nodes) into non-overlapping modules (communities) by their correlational relationships[61,62]. First, the population-level correlational model was reordered by the module assignment according to the modularity cost function. Network modules were identified at different levels of granularity by varying the tuning parameter, γ [29,49]. Increasing γ partitions the brain into increasing numbers of modules with a limit equal to the number of electrodes, as shown here for 3 values of γ (left). Next, the stability of this clustering at each value of γ was assessed by calculating module allegiance, which describes the probability that any two electrodes occupy the same module on repeated module detection[67] (right). A γ-value of 1.19 was selected by comparing the similarity of partitions generated by values of γ with those of a commonly used brain atlas[63], resulting in 6 modules. Of note, one of the modules is small and difficult to resolve in the figure. e. The graph of the large-scale network with module membership delineated by the color of the nodes and edges for selected γ of 1.19 is shown along with a schematic representation of the 6 modules. f. Hubs for each module were identified by selecting electrodes with the lowest 10% participation coefficients. Values were then averaged for each Lausanne brain region per module and weighted by the distribution of electrodes across Lausanne regions in all modules. Hub weight is indicated by the size of hub, and module assignment is indicated by hub color. Module 5 contained insufficient number of electrodes for hub identification (0.3% of total sample) and coefficients across all electrodes were utilized to name this module. L-DLPFC=left dorsolateral prefrontal cortex, L-OT=left occipitotemporal cortex, L-OFC=left orbitofrontal cortex, R-MF=right medial frontal cortex, R-FT=right frontotemporal cortex

Relationship of functional network identification to depression status

We next investigated the whole-brain iEEG model to study brain networks underlying depression. All patients in this study were undergoing surgical mapping with multi-channel iEEG for seizure localization as part of their standard medical care. Depression was measured coincident in time with iEEG recordings, using the validated PHQ-9 measure[58-60]. In accordance with literature-derived rates of depression in this population[35-40], 43% of our population had self-reported depression as defined by this measure (PHQ-9≥10, n=23), and 33% had mild or no symptoms of depression, which defined our control group (PHQ-9≤5, n=18). The two groups did not vary in age, sex, type of epilepsy, antidepressant usage, or anti-epileptic drug class (t-test, X^2, p>0.4, **Supplementary Table 1**).

Work over the past 50 years has highlighted the predictive potential of EEG to detect depression[68-71]. Disruption in frontal alpha power is one of the earliest and most-studied markers of depression[11-16], but its predictive potential remains unclear due to inconsistent findings across

studies[17-19]. We hypothesized that alpha power may be one part of a larger network story, and that by leveraging the high temporal resolution of iEEG, as well as the direct access to subcortical structures, we could perform a spectral analysis of brain activity across functional modules that would identify a network biomarker of depression. We calculated relative spectral power from the reconstructed time series in six frequency bands and averaged the resulting power values across time independently per module (see *Methods*). This process yielded 36 features per participant, where each feature contained information about a spectral power band across one functional module; we therefore collectively referred to these as "spectral-spatial features." We then utilized supervised machine learning to identify network activity that was predictive of depression. Our machine learning classification pipeline is shown in **Fig. 3a**. Neuronal activity is highly collinear, which suggests an underlying latent structure. We used principal component analysis to not only reveal these underlying hidden network variables but also to mitigate overfitting concerns by condensing the representation into fewer dimensions. We then utilized an ensemble of random forest and logistic regression machine learning classifiers with leave-one-out cross-validation to identify the spectral-spatial network features with the greatest discriminatory power. We found that four of the principal components had the strongest predictive ability to detect depressed from non-depressed subjects. Their component weights represent their contribution towards likelihood of depression. We show these weights in **Fig. 3b**. Utilizing the four most discriminative components alone, we achieved a mean classification accuracy of 82.7% on the training set and 82.4% on the test set (sensitivity=0.833, positive-predictive value = 0.87). In comparison, the same classification pipeline applied to a null model obtained from randomly permuting the target class labels led to a classification accuracy of 50.4%. The receiver operating curve (ROC) derived from cross-validation for the learned model in comparison to a null model is shown in **Fig. 3c** with an area under the curve (AUC) of 0.88. Together, these data suggest that a parsimonious model with four principal components, which capture major sources of variance in spectral-spatial

features, can detect subjects with depression from the control group significantly better than chance (see also **Supplementary Fig. 4**).

We next turned to an examination of the individual spectral-spatial features contained within the four components to identify the frequency bands and modules that comprise this putative biomarker. The component loadings on the individual spectral-spatial features within the four components in **Fig. 3b** represent the circuit activity that detects depression (for full description see **Supplementary Table 4**). To represent this biomarker spatially, we weighted the spectral distribution of the components with respect to their component loadings and component weights for each module (**Fig. 3d**). On visual inspection two gross patterns of spectral activity across the modules emerged. The first was high alpha power across the L-OT, R-FT, and mid-hemispheric modules (attention and default mode regions, modules 2,5, and 6 in **Fig. 3d**). The second was high delta and low alpha and theta power in the L-DLPFC and OFC modules (executive and limbic regions, modules 1 and 3 in **Fig. 3d**). These results suggest that low-to-mid- frequency activity across broad networks characterize depression at the group level. These observations motivated the subsequent statistical analysis to define the two patterns quantitatively.

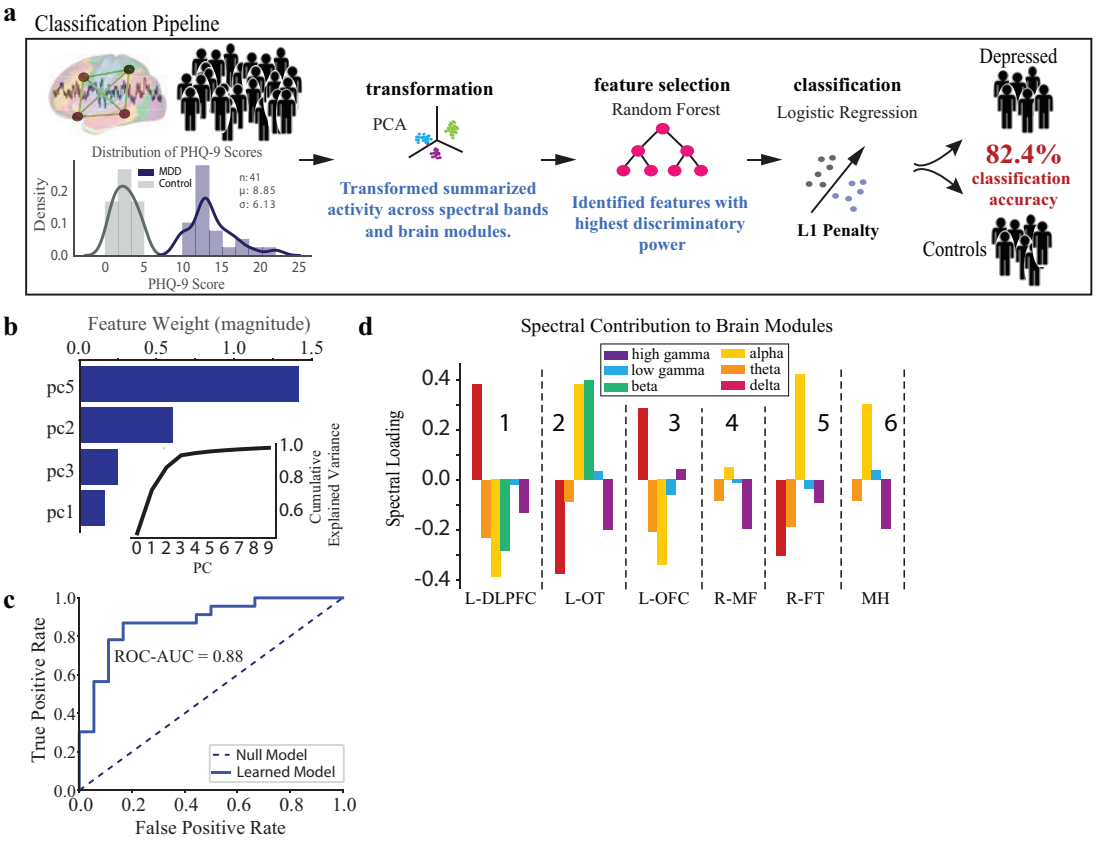

Fig. 3. Spectral-spatial features that discriminate depression at group level. a. Activity analysis pipeline showing steps including (i) power feature extraction, (ii) dimensionality reduction, (iii) transformation, (iv) feature selection, and (v) classification. The distribution of PHQ-9 scores across the depression (n=18, purple) and control groups (n=23, gray) is shown bottom left (mean PHQ-9 score 8.85, standard deviation 6.13). Power was extracted from the reconstructed time-series using the Morlet transformation in 30 s intervals across 6 frequency bands (delta = 1-4 Hz, theta = 5-8Hz, alpha = 9-12Hz, beta = 13-30Hz, low gamma (gammaL) = 31-70Hz, high gamma (gammaH) = 71-150Hz). This process yielded 25,464 spectral power features from our model (6 frequency bands x 4,244 electrodes x 2 hours). Z-scored relative power was calculated and averaged within each band across each of the 6 network modules. Power was then further averaged across time to yield 36 spectral-spatial features per participant. Principal component analysis was then used to transform the full spectral-spatial feature set into a lower dimensional space spanning multiple frequency bands. An iterative machine learning model that included random forest for feature selection and logistic regression for classification was performed with an exhaustive search of model parameters including component number. Feature selection yielded 7 lower dimensional components with the greatest predictive ability in discriminating depression, which was reduced to 4 features with the logistic L1 penalty. Utilizing these 4 components alone and leave-one-out cross validation yielded a classification accuracy of 82.7% on the training set and 82.4% on the test set (sensitivity=0.833, positive-predictive value = 0.87). b. The component weights of the four features with cumulative explained variance across the first 9 principal components shown in the inset. c. Receiver operating curve (ROC) derived from cross-validation for the learned model in comparison to a null model (AUC 0.88). d. Spectral distribution of the 4 components with respect to their component loadings and component weights (>0.2) for each of the six network modules. These spectral-spatial features represent the circuit activity that

detects depression. DLPFC = dorsolateral prefrontal cortex, OT = occipitotemporal, OFC = orbitofrontal cortex, MF = medialfrontal, FT = frontotemporal, MH = mid-hemispheric

Distinct Network Expression Patterns Define Depression Biotypes

We next asked how the biomarker that detected depression at the group level was expressed at the patient level. The rationale for this analysis is that depression is most likely to be a collection of several underlying disorders, which suggests there may be inter-individual heterogeneity in neural circuits involved. In moving toward a personalized approach to medicine, these individual biomarkers could identify and eventually track disease status in a continuous manner. To answer this question, we identified clusters of subjects who expressed different neurophysiological activity patterns. We first identified the log-odds impact of each original spectral-spatial feature from the group-level biomarker at the patient level. We then tested the distribution of feature impact on depression classification probability across participants by grouping covariates with similar feature impact scores using an agglomerative hierarchical clustering algorithm (see *Methods*)[72-75]. We found two distinct patterns of network activity (network expression patterns (NEPs)) that were distributed across two groups of patients (**Fig. 4a**). The first network (NEP1) was marked by increased beta power in the L-OT module, and increased alpha and decreased delta power over the L-OT and R-FT modules. The second network (NEP2) was marked by decreased theta in the L-DLPFC, L-OFC, and R-FT modules, and decreased alpha, beta power together with increased delta power within the L-DLPFC and L-OFC modules. These two NEPs were consistent with the core features of the group-level biomarker. However, their existence importantly demonstrated that different core features were relevant in different subjects.

We next quantified the impact of each NEP on each participant's probability of being classified as depressed using a sensitivity analysis. The difference in classification probability when one NEP was left out was attributed to the presence of that NEP, which quantified the relative importance

of that NEP toward the classification of depression. **Fig. 4b** shows the probability contribution of each NEP for each subject in the depressed group (top plot) and control group (bottom plot). While we anticipated that each individual would exhibit several NEPs with differing contributions to their depression classification, an alternate pattern emerged from the data. We found that increased activity in either NEP was correlated with depression, but that each patient exhibited activity in only one of the two NEPs. Thus, depressed participants fell into two groupings based on NEP activity which we refer to as biotype-1 (40% depressed subjects) and biotype-2 (35% depressed subjects). Biotype-1 was largely driven by NEP1 (n=8, mean probability contribution =0.44, SD=0.17) and biotype-2 was largely driven by NEP2 (n=7, mean probability contribution=0.40, SD=0.16, **Fig. 4c**). The remaining 25% of participants belonged to a group with mixed or no contribution from the two identified NEPs and may be evidence of additional biotypes that were not resolved in our dataset. Two distinct groups also emerged from the control participants. The first group (43% control patients) exhibited mixed or no contribution from the two depression NEPs, as we anticipated. In the second group, there was a strong decrease in one of the two NEPs that countered a smaller increase in the opposing NEP (57% of control group). This reciprocal activity was also observed in the depressed group, but in the opposing direction, with a strong increase in one NEP. We might speculate that relative NEP activity could represent either risky or conversely, protective activity profiles, and that NEP activity could be modulated in either direction to treat depression. The anatomical distribution of the two circuits underlying these depression biotypes is shown in **Fig. 4d**.

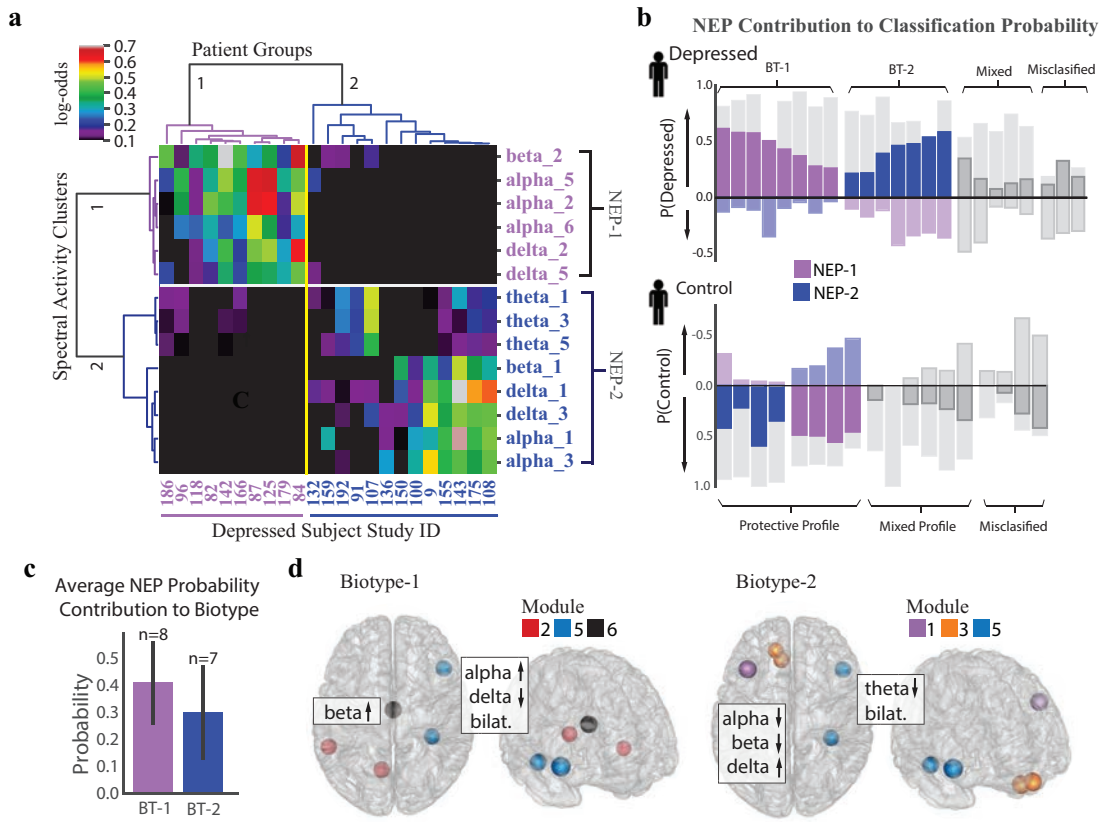

Fig. 4. Identification of patient-level depression biotypes. a. Hierarchical clustering on log-odds of spectral-spatial features at the individual patient level showing 2 patient groups (horizontal groupings) and 2 network expression patterns (NEPs) (vertical groupings). Columns represent individual patients with patient study number shown at bottom, and rows represent spectral power across one frequency band and module (ex. alpha_1=alpha power across module 1). Magnitude of log-odds represented by color of corresponding boxes (color-bar legend top right). Spectral-spatial features associated with NEP-1 represented in purple text and those associated with NEP-2 represented in blue text. b. NEP probability contribution for the depressed group (top plot) and control group (bottom plot) derived from a sensitivity analysis where the probability of depression for each individual was calculated in total and with a perturbation where each NEP was held out. The probability difference was attributed to the presence of the NEP. This probability contribution is represented by the colored bars overlaid over each patient's total probability of being depressed as derived from the machine learning classification model (gray bars, probability > 0.5 leads to classification of depression). The perturbations do not sum to produce the total classification probability; rather each quantifies the relative importance of that NEP toward depression. Bars in the positive direction indicates a positive contribution toward depression, and those in the negative direction indicate a protective contribution toward depression. Subjects where one NEP did not drive classification probability as evidenced by equal contribution of each NEP to classification in both the positive and negative direction shown with bars in dark gray (mixed profile). Subjects classified incorrectly shown on far right of each plot (misclassified). c. Mean probability contribution of each NEP to the associated biotype is shown. NEP-1 (purple bars) contributed most strongly to the probability of depression in biotype-1 (mean=44% probability contribution, SD=0.17) and NEP-2 (blue bars) contributed most strongly to biotype-2 (mean=40% probability contribution, SE=0.16). Number of participants who exhibit each NEP shown above

each bar. Error bar = standard deviation. d. Direction of activity and spatial distribution of activity changes within each biotype shown on glass-brain in several orientations. Hubs for each module within the two biotypes are designated by hub color.

Network Organization is Disrupted Across MDD Biotypes

In addition to alterations in the spectral content of network activity in depression, previous studies have observed distinct deficiencies in connectivity across depression networks[27,28,31-33]. A fundamental interest in neuroscience is the relationship between the brain's neural activity and its underlying functional and structural connectivity, which remains unknown. We expected that alterations in functional network topology would be present in our depressed population and that we could delineate new relationships between activity and functional connectivity with our high-resolution dataset to more fully characterize network markers of depression. We performed a connectivity analysis using correlation of local field potential activity across modules as an estimate of functional connectivity between electrodes. The graph of our whole-brain iEEG model (**Fig. 2e**) defines correlational relationships between electrodes across our total population. We thus examined these correlational relationships across control and depressed groups independently in order to measure the relative differences of functional network organization between the two groups. To do so, we examined the relative difference in connectivity strength across electrodes within and between functional modules. **Fig. 5a** shows the two-dimensional representation of the functional network structure for control (left) and depressed (right) groups. In comparison to the control group, we qualitatively observed an overall reduction in the segregation between modules in the depression network.

To quantify these differences, we first calculated the inter- and intra-modular connectivity strength by calculating the correlation between electrodes located in the same module (intra-modular) and across different modules (inter-modular). We next compared the distribution of correlation strengths between depressed and control groups for each possible module pair using a Cohen's

d effect size metric. We assessed significance by retaining the Cohen's d values that survived multiple comparisons testing with a null distribution of Cohen's d values generated by permuting nodal module assignments ($p<0.001$). **Fig. 5b** shows the heatmap of significant Cohen's *d* values, where a greater connectivity effect size for the depressed group is indicated in red, and lower connectivity effect size for the depressed group is indicated in blue. In the depressed group, there was overall greater frontal connectivity and weaker cross-hemispheric connectivity. Specifically, we observed greater intra-modular connectivity within L-DLPFC, L-OFC, and R-MFC modules, weaker intra-modular connectivity within L-OT and R-FT modules, and greater inter-modular connectivity between L-DLPFC, L-OFC and L-OT modules. Hubs in the insula, amygdala, temporal pole and fusiform gyrus drove the cross-module connectivity (top 10% participation coefficient, see *Methods*). We also observed a decrease in cross hemispheric connectivity in the depressed group compared to the control group (L-DLPFC/L-OFC to R-FT modules, and L-OT to R-FT/R-MFC modules), with hubs in the insula, temporal-parietal region and amygdala responsible for this decreased connectivity. The L-OFC module showed greater connectivity with the R-MFC module, and R-MFC module exhibited stronger connectivity with the R-FT module.

On the basis of the above analyses we were able to parse specific connectivity components that characterize the two depression biotypes (**Fig. 5c**), unifying both activity and connectivity analyses across cortical and deep structures in a way that has not previously been possible. In biotype-1 we observed increased beta power in the L-OT module, and right-left asymmetry in the alpha and delta bands over right frontal/L-OT modules with weaker intra- and inter-modular connectivity throughout. In biotype-2 we observed a hyperactive left frontal cortex that was more highly connected within itself but more weakly connected to R-FT module. Lower theta bilaterally was observed in this biotype.

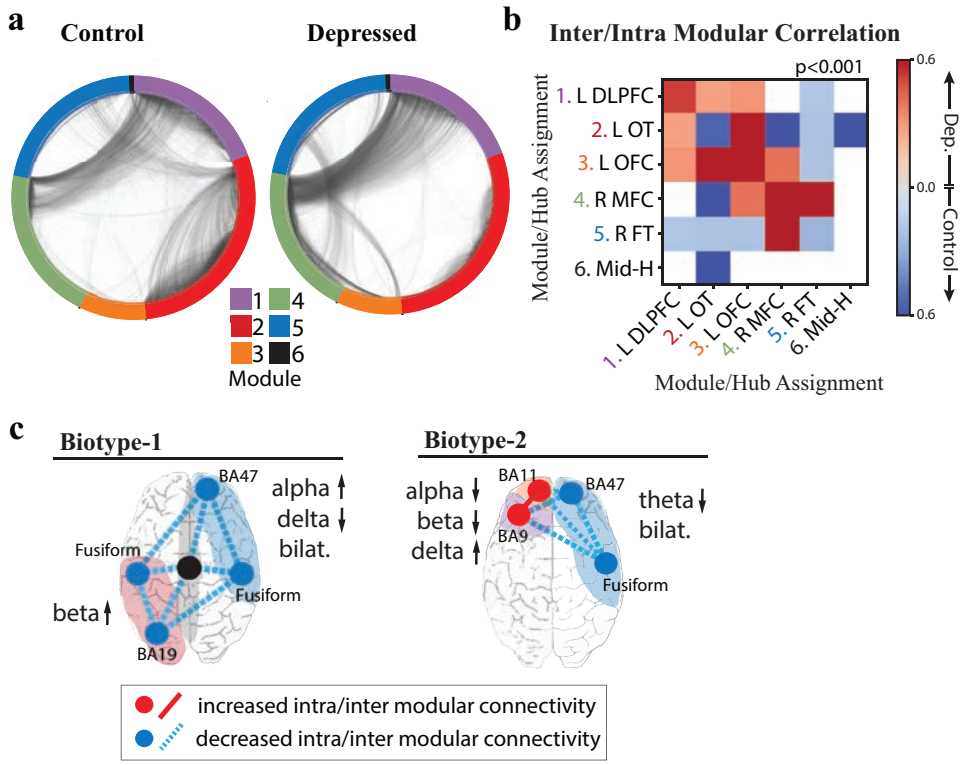

Fig. 5. Intra- and Inter-modular connectivity signatures of depression and control groups. a. Connectivity structure derived using whole-brain iEEG model recalculated for the control group (left) and depressed group (right) with module membership delineated by the color of the nodes (electrodes), and edges (connections between electrodes) delineated by the black interconnecting lines. b. Heatmap of significant Cohen's d values calculated from the distribution of correlation strengths between depressed and control groups for each possible module pair and compared to Cohen's d values for a null distribution derived from permuted nodal module assignment. Those that survived multiple comparison testing ($p<0.001$) were retained (red: increased connectivity for depressed group; blue: increased connectivity for control group; white: not significant). c. Schematic of biotype 1 (left) and biotype 2 (right) showing both connectivity and spectral power underlying each pattern. Increased connectivity strength shown in red, and decreased connectivity shown in blue (hub=intramodular, line=intermodular connectivity). Color of shaded area refers to module number as shown in color legend in panel a.

Discussion

The field of depression is limited by the absence of quantitative biomarkers, which are critical to objectively assess disease state and develop advanced treatments. In this report, we present the largest study of direct neural recordings aimed at identifying depression networks, made possible by iEEG recordings paired in time with a depression measure. The opportunity to directly record

semi-chronically from cortical and subcortical structures in this manner enabled us to build computational models of neural activity and incorporate both activity and connectivity analyses to resolve new biotypes of depression. We found a neurophysiological biomarker that was able to detect depression with 82% accuracy. We further found two distinct patterns of circuit dysfunction across this biomarker that were expressed in 84% of our depressed patients.

Overall, our findings suggest that two highly interconnected circuit activity patterns are involved in depression and that capturing this complexity is required for an objective diagnosis of depression. They unite previous fMRI and EEG studies that have individually found disruptions in some of the same measures, but have been inconsistent across studies, supporting the existence of several depression subtypes. For example, our results support disruptions in frontal theta, temporal beta[68], and alpha asymmetry[11,12,76-78], but in subsets of subjects and in combination with each other. Furthermore, they support studies that have shown decreased connectivity in the occipital, temporal, and right medial frontal regions in depression[24] and studies that suggest higher frontal connectivity[2,7,79-82]. The implications are that circuits could be amenable to interventions at many different potential control points along the network, and that interventions that change both connectivity and spectral power could be promising. For example, identifying an individual's prominent NEP could help practitioners chose between right and left DLPFC vs. OFC targets for repetitive transcranial magnetic stimulation[72,83]. We also found evidence of high activity in one network pattern, countered by an anti-weighting of the other pattern. This finding suggests the existence of protective or high-risk profiles and opens the possibility of developing preventative treatments. A library of new treatment targets and frequency-specific treatment parameters[84,85] points toward a future of increasingly personalized depression therapy.

Recently, Betzel and colleagues successfully applied a correlational network model similar to SuperEEG to multi-subject iEEG recordings, followed by community detection, and found network

organization to be representative of that obtained from DTI and fMRI[86]. We extended these findings, by applying the iEEG model to the study of disease status for the first time. Drysdale and colleagues used hierarchical clustering, similar to the clustering we utilized, on fMRI connectivity features and identified four depression biotypes [72]. One of these biotypes was dominated by globally decreased connectivity focused in the right OFC and left temporal-occipital region consistent with our biotype-1. Another exhibited increased frontal and decreased temporal connectivity consistent with our biotype-2. Of note, Drysdale and colleagues matched symptom profiles with fMRI data as a starting dimension for further exploration of depression biotypes. In contrast, in our method we sought to take those features that classified depression and determine if common activity patterns were observed independent of subject or physician input. Despite these differences, the similarity in biotypes – each developed in an empirical, data-driven approach from different modalities – is quite remarkable. The semi-chronic nature of our recordings enabled us to assess trait markers with neurophysiology alone and enhanced the connectivity analysis with an understanding of the neural activity that accompanied altered connectivity. Future analyses should evaluate how these putative biotypes map to symptom profiles of depression.

We acknowledge some weaknesses in the results presented. It remains unknown whether the depression networks we identified are related to the presence of epilepsy. Depression in epilepsy is thought to arise from similar origins to primary depression (stress[41], inflammation[42], medications [43,44], circuit dysfunction[45]), and is responsive to antidepressants[46] suggesting it can provide valuable insight into depression broadly. At minimum, the results presented here shed light on understanding depression in epilepsy patients. More importantly, the methods and findings provide a basis for applying similar approaches to broader depression populations and the potential for integrating such findings with non-invasive network features. While our whole-brain

iEEG model was extensive in coverage, we did not have electrodes placed in all brain regions, including some regions implicated in depression[87-91] and the density of electrode sampling varied across brain regions leading to uncertainty in the accuracy of estimation in sparsely sampled areas[47]. We dealt with this constraint by discounting the effect of each individual node degree before running community detection and comparing network measures to a null model that accounted for overall node density. At the time we carried out this analysis, SuperEEG was the only method developed to carry out a group-level analysis on this type of data. This approach bordered on computational infeasibility, even with the use of an 8-node high performance computing cluster. More methods are likely to emerge with advancements in data processing capabilities and accessibility. With such methods, we may be able to reduce assumptions and the estimation burden, extend coverage to more brain regions, and utilize larger samples.

In conclusion, using a valuable dataset of intracranial recordings, we identified novel biomarkers of depression. These results have important implications for disease subtyping, diagnosis, treatment planning, and monitoring of depression status. Quantitative markers of depression could enable a new wave of closed-loop interventional therapies that personalize treatment based on neurophysiological signals.

Methods:

Patient Characterization:

Participants included 54 adults (49% female) aged 20-67 who had medication-refractory epilepsy and were undergoing intracranial monitoring as part of their standard medical care. Neural data from these participants comprised our full dataset and was utilized to build the whole-brain iEEG model of LFP time-series. Participants were screened for depression following electrode implantation and concurrent with neural recordings using the Patient Health Questionnaire-9 (PHQ-9), a 9-item self-report instrument validated for depression screening [58-60]. A score ≥ 10 was

used to define the MDD group (moderate depression) and a score ≤ 5 defined the non-depressed control group generating a sample of 23 depressed subjects (56%) and 18 controls (44%). The remaining 13 patients were used in the first step of the study (model *building*) but not the second (model *utilization*). Data comprised a consecutive series of patients recruited from University of California, San Francisco and Kaiser Permanente, Redwood City, California over a 5-year period. This study was approved by the University of California, San Francisco Institutional Review Board with written informed consent provided by all subjects.

Electrode Implantation and Localization:

Subdural grid, strip, and depth electrodes (AdTech, Racine, WI, or Integra, Plainsboro, NJ) were implanted using standard neurosurgical techniques. Subjects underwent pre-operative 3 Tesla brain magnetic resonance imaging (MRI) and post-operative computed tomography (CT) scan to localize electrodes in patient centered coordinates using an open source python package for preprocessing imaging data for use in iEEG recordings[92]. The steps included warping brain reconstructions to a common Montreal Neurologic Institute (MNI) template and merging electrode locations across subjects. Surface warpings were then generated by projecting pial surfaces of the subject and template brains into a spherical coordinate space, and aligning the surfaces in that space. Depth warping was then performed using a combination of volumetric and surface warping[93]. For visualization, pre-operative T1-weighted MRI scans were pre-registered with the post-operative CT using Statistical Parametric Mapping software SPM12 and pial surface 3D reconstructions were generated using FreeSurfer[94].

Data Acquisition and Pre-Processing:

Data acquisition of iEEG recordings were acquired using the Natus EEG clinical recording system at a sampling rate of 1-2 kHz. Standard iEEG/ECoG pre-processing techniques were conducted in python including application of a 2-250Hz bandpass filter, notch filters at line noise frequency

and harmonics (60Hz, 120Hz, 180Hz, 240Hz), down sampling to 512Hz, and common average referencing to the mean of all channels.

Construction of Whole-Brain iEEG Model:

A functional connectivity imputation technique was utilized to estimate whole-brain iEEG activity for each subject (SuperEEG, developed by Jeremy Manning and Lucy Owens)[47]. This method involved four main steps outlined in **Supplementary Fig. 1**. First, pre-processed iEEG signals were chunked into 60s non-overlapping blocks and filtered for putative epileptiform activity or artifacts. We then randomly sampled the 60s intervals across daytime hours (8am-10pm) and concatenated them into 2 hours blocks, each representative of resting-state activity. To filter the activity of putative epileptiform activity or artifacts we identified signals that deviated from the norm using kurtosis, a measure of infrequent extreme peaked deviations[95]. Individual 60s blocks were retained if the number of artifactual signals (kurtosis>10) was less than 5% of the total sample. The same kurtosis threshold was then applied to each 2h block to further eliminate noisy electrodes. The kurtosis threshold of 10 was tested in Owen et al. on a similar intracranial dataset and verified by eye on a subset of samples by a provider trained in Clinical Neurophysiology. Second, the iEEG functional connectivity model was built by averaging patient level connectivity models. The patient level model leveraged known inter-electrode correlations to estimate temporal correlations at locations that were not present. This was accomplished by using the fact that the strength of relationships between sensors are proportional to neural proximity. As the distance between two sensors increases the correlation weakens. The rate of falloff was modeled using a gaussian radial basis function (RBF) at locations η of each known patient electrode to estimate the unknown correlations between x and η.

$$rbf(x|\eta,\lambda) = exp\left\{\frac{-\|x-\eta\|^2}{\lambda}\right\}$$

where λ was set to 20 based on prior work that sought to maximize both spatial resolution and generalizability to any brain location[64]. This relationship was utilized as the Weight matrix (W) and applied to every electrode set (i, j) for each patient as

$$W_s(i,j) \ = \ rbf(i|j,\lambda)$$

where the rbf function expands the single patient to include estimated temporal recording relationships from sensors in the full population. Third, the Weight matrix was then combined with the single patient correlation matrix ($C_s(i,j)$) to estimate each patient's full population anatomical connection structure using,

$$\widehat{N}_s(x,y) = \sum_{i=1}^{|R_s|} \sum_{j=1}^{i-1} W(x,i) \ \cdot W(y,j) \cdot z(C_s\,(i,j))$$

$$\widehat{D}_s(x,y) = \sum_{i=1}^{|R_s|} \sum_{j=1}^{i-1} W(x,i) \cdot W(y,j)$$

$$\hat{C}_s = r\left(\frac{\widehat{N}_s}{\widehat{D}_s}\right)$$

where r(·) represents the Fisher-z transformation. The patient specific estimates of full brain correlation matrix $\widehat{N}_s$ and the spatial proximity weighting matrix $\widehat{D}_s$ were combined in a weighted average to compute a single whole-brain expected correlation matrix $\widehat{K}$ as

$$\widehat{K} = r\left(\frac{\sum_{s=1}^{S} \widehat{N}_s}{\sum_{s=1}^{S} S\widehat{D}_s}\right)$$

Finally, estimated LFP activity for each electrode was reconstructed using Gaussian process regression according to

$$\hat{Y}_{s,\beta_s} = \left(\left(\hat{K}_{\beta_s,\alpha_s} \cdot \hat{K}^{-1}_{\alpha_s,\alpha_s}\right) \cdot \hat{Y}_{s,\alpha_s}\right)^T$$

where s represents the patient, α_s is the set of indices where we had observed recordings for patient s and β_s is the set of indices that were to be estimated. The submatrix $\hat{K}_{\beta_s,\alpha_s}$ represents the estimated correlations between locations of known activity and locations of reconstructed activity. The submatrix $\hat{K}_{\alpha_s,\alpha_s}$ represents correlations between known neural recordings.

Construction of the above model required extensive computational resources. Therefore, we sought to utilize the minimum required information required to obtain the majority of resting state information to enable computational feasibility. Using the 10h benchmark as the largest feasible model we could build, we compared 2, 4, 6 and 8 hour models to the 10 hour model and found that the difference in adding additional time beyond 2h was marginal and could be computed at a fraction of the computational cost. We therefore utilized the 2h model for further analysis (**Supplementary Fig. 2a**).

Signal Processing:

Standard signal processing techniques were applied to the time-series activity across all reconstructed electrodes. This included continuous wavelet transformation using the Morlet transform wavelet method (6-cycles)[96] performed in 30s intervals to obtain power spectra in 6 frequency bands (delta = 1-4Hz, theta = 5-8Hz, alpha = 9-12Hz, beta = 13-30Hz, low gamma (gammaL) = 31-70Hz, high gamma (gammaH) = 71-150Hz). Relative power was calculated by

dividing the power of each frequency band by the total power across the 6 frequency bands for each electrode. Signals were summarized by taking the mean power across time for each spectral band and were z-scored across patients.

Electrode Clustering into Functional Modules:

We used a well-validated technique, multiscale community detection, to identify distributed functional sub-networks (modules) across the model [61,62,97-100]. Individual functional connectivity models generated in the whole-brain iEEG reconstruction were used as a starting point in this analysis. Using the Louvain algorithm[101], we identified an optimal parcellation of electrodes into discrete functional modules by maximizing a modularity cost function defined by the following relationships,

$$P = \frac{2}{\|K\|_F} KJK^T$$

$$Q = \|(K - \gamma P) \circ G\|_F$$

where J is a ones matrix, $\circ$ is the Hadamard product and Gi,j is 0 if node pair (i, j) are assigned to different modules and 1 if the pair is assigned to the same module, Q is modularity, K is the connection weights (correlation) between node i and j, P is the Newman-Girvan null model[62] and γ is the weighting of that null model which is tuned to obtain network modules of different sizes. As the network fractures into many small modules for higher γ, the overall modularity Q decreases. We examined network modularity at values of γ between 0.5 and 2.1. We then assessed the stability of clustering at each value of γ by examining module allegiance[67]. It was calculated by repeating module detection 100 times and evaluating the probability that two electrodes occupied the same module.

To select an optimal γ, we compared partitions generated by each value of γ with partitions generated by a well-known validated atlas (the Lausanne atlas) by calculating the adjusted Rand score index [102], which is a measure of similarity between two clusterings (**Supplementary Fig. 2b-c**). Significance was assessed by a permutation test where the null model was generated by randomly assigning electrodes to each module and calculating the confidence interval of the similarity index generated from 1,000 random permutations and tested at significance level 0.05 for a 2-tailed test. Two similarity peaks were identified, with values of γ that generated 6 and 1,855 modules respectively. The peak with the highest modularity (lowest number of clusters) was selected for further analysis due to our goal of examining the brain at a low level of granularity. This selection enabled subsequent classification of activity across these clusters without overfitting our model.

Assigning Names to Modules:

We assigned a name to each module by examining the location of each module's most influential electrodes. We utilized the participation coefficient (PaC), which is a degree-based measure of network connectivity that describes a node's functional interaction within and across network modules[64-66]. This metric is typically utilized to identify influential hubs across a large-scale network. We utilized it in our study to identify the location of hubs that were most important for driving connectivity in each module identified through community detection. Groups of electrodes with low PaC values indicate hubs with high intramodular connectivity, also known as provincial hubs[103]. Similarly, connector hubs are those with high PaCs and drive intermodular connectivity. The PaC describes the weight of edges from node i to all other nodes in the same module relative to the weight of edges from that node to all nodes in the network according to

$$y_i = 1 - \sum_{c \in C} \left(\frac{k_i(c)}{k_i} \right)$$

where y_i is node i's participation coefficient, C is the set of all modules, $k_i(c)$ is the sum of all correlations between node i and other members of module C and k_i is the sum of all correlations between node i and members of all modules. We calculated the PaC for each electrode across our model, and then selected those with high and low participation values (top/bottom 10%). We then grouped these selected nodes by Lausanne atlas region, eliminating or combining a minority of regions due to having too few electrodes for analysis. We addressed the non-uniform distribution of electrodes across the model by then assigning each Lausanne region a score according to the following hub weight:

$$R_i = \frac{N_i M_j}{T_i}$$

where N_i is the number of selected electrodes (top/bottom 10%) in Lausanne region i, M_j is the number of selected electrodes in Lausanne region i of module j, and T_j is the number of total electrodes across modules in Lausanne region i. Hubs were those Lausanne regions with the highest hub weight. Hub location was identified by averaging the MNI coordinates of electrodes within each hub. The full list of Lausanne regions and hub weights is shown in **Supplementary Fig. 3 and Tables 2-3**. The purpose of the identified hubs in the present report was primarily descriptive and helped relate the computational model to known brain regions and structure; all subsequent analyses utilized the population set of electrodes across the full model.

Classification:

We utilized a machine learning algorithm to classify subjects with and without depression. Given our high dimensionality feature space generated by spectral decomposition across the whole-brain model (4,244 electrodes x 6 spectral bands), we recognized the need for dimensionality

reduction to learn a stable model. Our primary dimensionality reduction technique consisted of averaging spectral power across the electrodes within each module for each spectral frequency band. This yielded 36 features (6 frequency bands x 6 modules) that contained information both about the spectral power and the location of the activity. These features, referred to as spectral-spatial features, served as our starting feature space for entry into our classification pipeline. This data was then transformed using principal component analysis (PCA)[104]. PCA enabled us to reduce collinearity and to identify latent variables that described maximally distinct network-based features inclusive of multiple frequency bands and modules. Feature selection was performed by fitting a random forest classifier which evaluates covariates to identify the set of features with the highest discriminatory power[105]. A logistic regression classifier was then utilized to assess classification accuracy. Model parameterization was identified using leave-one-out cross validation[106]. A sparsity promoting regularization (L1) was picked as it selected fewer principal components while controlling for overfitting allowing better generalization. The mean accuracy, precision, recall and ROC-AUC was reported.

To further asses our model validity, we repeated our classification pipeline on a null model obtained from randomly permuting the target class labels. We also examined the stability of classification across models. **Supplementary Fig. 5a** shows the classification accuracies, precision and recall for the learned model in comparison to the null model and two nonlinear classifiers (random forest, support vector machine).

In order to control for possible differences in epileptiform activity residual to data-cleaning across the modules we calculated line-length, a commonly utilized measure for the detection of epileptiform activity[107] for each electrode and averaged across the modules. We used a logistic regression model to show that line-length across the six modules was not a significant predictor of depression status (R^2=0.15, p=0.13, logistic regression).

Hierarchical clustering to identify biotypes:

We reasoned that we could utilize the group-level biomarker to identify common features that defined depression at the individual level. To do so, we mapped the principal component values (feature loadings ≥0.2) back to the original space weighted by the logistic regression coefficients. This provided the log-odds impact of each original feature. We then tested the distribution of feature impact on depression classification probability across depressed participants by grouping similar log-odds impact covariates (thresholded at 0.15) utilizing an agglomerative hierarchical clustering algorithm[72-75]. A log-odds threshold of 0.15 was selected because it retained classification results for 98% of subjects while isolating the most contributory spectral-spatial features (see **Supplementary Fig. 5b** for non-thresholded model for comparison). The clustering yielded both patient and feature groupings that defined neurophysiological network expression patterns (NEPs) of depression. We quantified the impact of these NEPs on each participant's probability of being classified as depressed by performing a sensitivity analysis where we withheld each NEP and then attributed the probability decrement from the total classification probability to the withheld activity pattern. Biotypes were defined based on the relative contribution of each NEP to classification probability. Subjects in whom the contribution was not clear or was mixed across the NEPs were placed into a third mixed group.

Connectivity Analysis:

In order to understand the effect of depression on network topology we examined changes in connectivity across the control and depressed subjects. First, inter- and intramodular connectivity strengths were assessed by looking at the correlations between all electrodes within the same module (intramodular) and the correlations between electrodes across all pairs of modules (intermodular). Next, to assess the connections that significantly distinguished depressed subjects from controls, we used a Cohen's d effect size metric and compared the distribution of

correlation strengths across depressed and control groups for each possible module pair. To assess significance across these connections we generated a null distribution of Cohen's d values for each module pair, and retained the true Cohen's d values that survived multiple comparisons testing ($p<0.001$).

Data Availability

Raw data is provided in Supplementary Fig.5 for the hierarchical clustering and in Supplementary Tables 2 (participation coefficient) and 3 (principal component analysis). All datasets generated and analyzed in the current study are available from the corresponding authors upon request.

Code Availability

All custom code, statistical analysis and visualizations were performed in Python and are demonstrated in a set of Jupyter notebooks available for download. The following open-source code was used as described in the methods: SuperEEG (https://github.com/ContextLab/supereeg).

Acknowledgements

We thank the members of the Chang Lab for assistance. This work was supported by National Institutes of Health award K23NS110962. (K.W.S), NARSAD Young Investigator grant from the Brain and Behavioral Research Foundation (K.W.S), and a Ray and Dagmar Dolby Family Fund through the Department of Psychiatry at the University of California, San Francisco (K.W.S, A.D.K.) and a Brain Initiative grant (SUBNETS, E.F.C.). Dr Chang receives research support from National Institutes of Health, New York Stem Cell Foundation, the Howard Hughes Medical Institute, the McKnight Foundation, the Shurl and Kay Curci Foundation, and the William K. Bowes Foundation. Dr. Krystal receives support from National Institutes of Health, PCORI, Janssen, Jazz, Axsome, Reveal Biosensors.

Contributions

K.W.S., A.N.K., E.F.C., A.D.K. conceived the study. K.W.S., A.N.K., P.M.D. analyzed and interpreted the data. A.J.B and A.E. contributed to data collection. J.R.M and L.W.O contributed to data analysis methods. K.W.S wrote the manuscript with significant input from all authors. All authors reviewed and approved the manuscript.

Competing Interests statement

K.W.S., A.N.K., P.M.D., E.F.C., J.R.M and L.W.O report no competing interests. A.K. consults for Eisai, Evecxia, Ferring, Galderma, Harmony Biosciences, Idorsia, Jazz, Janssen, Merck, Neurocrine, Pernix, Sage, and Takeda.

References

1. Getting, P.A. Emerging principles governing the operation of neural networks. *Annu Rev Neurosci* **12**, 185-204 (1989).
2. Sheline, Y.I., Price, J.L., Yan, Z. & Mintun, M.A. Resting-state functional MRI in depression unmasks increased connectivity between networks via the dorsal nexus. *Proc Natl Acad Sci U S A* **107**, 11020-11025 (2010).
3. Insel, T.R. Next-generation treatments for mental disorders. *Sci Transl Med* **4**, 155ps119 (2012).
4. Botteron, K.N., Raichle, M.E., Drevets, W.C., Heath, A.C. & Todd, R.D. Volumetric reduction in left subgenual prefrontal cortex in early onset depression. *Biol Psychiatry* **51**, 342-344 (2002).
5. Kennedy, S.H.*, et al.* Changes in regional brain glucose metabolism measured with positron emission tomography after paroxetine treatment of major depression. *Am J Psychiatry* **158**, 899-905 (2001).
6. Yoshimura, S.*, et al.* Rostral anterior cingulate cortex activity mediates the relationship between the depressive symptoms and the medial prefrontal cortex activity. *J Affect Disord* **122**, 76-85 (2010).
7. Greicius, M.D.*, et al.* Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biol Psychiatry* **62**, 429-437 (2007).
8. Bluhm, R.*, et al.* Resting state default-mode network connectivity in early depression using a seed region-of-interest analysis: decreased connectivity with caudate nucleus. *Psychiatry Clin Neurosci* **63**, 754-761 (2009).
9. Grimm, S.*, et al.* Altered negative BOLD responses in the default-mode network during emotion processing in depressed subjects. *Neuropsychopharmacology* **34**, 932-943 (2009).
10. Zhu, X.*, et al.* Evidence of a dissociation pattern in resting-state default mode network connectivity in first-episode, treatment-naive major depression patients. *Biol Psychiatry* **71**, 611-617 (2012).
11. Gotlib, I.H., Ranganath, C. & Rosenfeld, J.P. Frontal EEG alpha asymmetry, depression, and cognitive functioning. *Cognition Emotion* **12**, 449-478 (1998).
12. Henriques, J.B. & Davidson, R.J. Left frontal hypoactivation in depression. *J Abnorm Psychol* **100**, 535-545 (1991).
13. Kemp, A.H.*, et al.* Disorder specificity despite comorbidity: resting EEG alpha asymmetry in major depressive disorder and post-traumatic stress disorder. *Biol Psychol* **85**, 350-354 (2010).
14. Diego, M.A., Field, T. & Hernandez-Reif, M. CES-D depression scores are correlated with frontal EEG alpha asymmetry. *Depress Anxiety* **13**, 32-37 (2001).
15. Kentgen, L.M.*, et al.* Electroencephalographic asymmetries in adolescents with major depression: influence of comorbidity with anxiety disorders. *J Abnorm Psychol* **109**, 797-802 (2000).

16. Jaworska, N., Blier, P., Fusee, W. & Knott, V. alpha Power, alpha asymmetry and anterior cingulate cortex activity in depressed males and females. *J Psychiatr Res* **46**, 1483-1491 (2012).
17. Gold, C., Fachner, J. & Erkkila, J. Validity and reliability of electroencephalographic frontal alpha asymmetry and frontal midline theta as biomarkers for depression. *Scand J Psychol* **54**, 118-126 (2013).
18. Reid, S.A., Duke, L.M. & Allen, J.J. Resting frontal electroencephalographic asymmetry in depression: inconsistencies suggest the need to identify mediating factors. *Psychophysiology* **35**, 389-404 (1998).
19. van der Vinne, N., Vollebregt, M.A., van Putten, M. & Arns, M. Frontal alpha asymmetry as a diagnostic marker in depression: Fact or fiction? A meta-analysis. *Neuroimage Clin* **16**, 79-87 (2017).
20. Wang, L., Hermens, D.F., Hickie, I.B. & Lagopoulos, J. A systematic review of resting-state functional-MRI studies in major depression. *J Affect Disord* **142**, 6-12 (2012).
21. Lui, S.*, et al.* Resting-state functional connectivity in treatment-resistant depression. *Am J Psychiatry* **168**, 642-648 (2011).
22. Esteva, A.*, et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115-118 (2017).
23. Gulshan, V.*, et al.* Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* **316**, 2402-2410 (2016).
24. Veer, I.M.*, et al.* Whole brain resting-state analysis reveals decreased functional connectivity in major depression. *Front Syst Neurosci* **4**(2010).
25. Zeng, L.L.*, et al.* Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis. *Brain* **135**, 1498-1507 (2012).
26. Liu, F.*, et al.* Abnormal amplitude low-frequency oscillations in medication-naive, first-episode patients with major depressive disorder: a resting-state fMRI study. *J Affect Disord* **146**, 401-406 (2013).
27. Lord, A., Horn, D., Breakspear, M. & Walter, M. Changes in community structure of resting state functional connectivity in unipolar depression. *PLoS One* **7**, e41282 (2012).
28. Zhang, J.*, et al.* Disrupted brain connectivity networks in drug-naive, first-episode major depressive disorder. *Biol Psychiatry* **70**, 334-342 (2011).
29. He, Y.*, et al.* Reconfiguration of Cortical Networks in MDD Uncovered by Multiscale Community Detection with fMRI. *Cereb Cortex* **28**, 1383-1395 (2018).
30. Sacchet, M.D., Prasad, G., Foland-Ross, L.C., Thompson, P.M. & Gotlib, I.H. Support vector machine classification of major depressive disorder using diffusion-weighted neuroimaging and graph theory. *Front Psychiatry* **6**, 21 (2015).
31. Chen, T.*, et al.* Anomalous single-subject based morphological cortical networks in drug-naive, first-episode major depressive disorder. *Hum Brain Mapp* **38**, 2482-2494 (2017).
32. Korgaonkar, M.S., Fornito, A., Williams, L.M. & Grieve, S.M. Abnormal structural networks characterize major depressive disorder: a connectome analysis. *Biol Psychiatry* **76**, 567-574 (2014).
33. Sun, Y., Hu, S., Chambers, J., Zhu, Y. & Tong, S. Graphic patterns of cortical functional connectivity of depressed patients on the basis of EEG measurements. *Conf Proc IEEE Eng Med Biol Soc* **2011**, 1419-1422 (2011).
34. Leuchter, A.F., Cook, I.A., Hunter, A.M., Cai, C. & Horvath, S. Resting-state quantitative electroencephalography reveals increased neurophysiologic connectivity in depression. *PLoS One* **7**, e32508 (2012).
35. Fuller-Thomson, E. & Brennenstuhl, S. The association between depression and epilepsy in a nationally representative sample. *Epilepsia* **50**, 1051-1058 (2009).

36. Gilliam, F., Hecimovic, H. & Sheline, Y. Psychiatric comorbidity, health, and function in epilepsy. *Epilepsy Behav* **4 Suppl 4**, S26-30 (2003).
37. Hermann, B.P. & Jones, J.E. Intractable epilepsy and patterns of psychiatric comorbidity. *Adv Neurol* **97**, 367-374 (2006).
38. Hermann, B.P., Seidenberg, M. & Bell, B. Psychiatric comorbidity in chronic epilepsy: identification, consequences, and treatment of major depression. *Epilepsia* **41 Suppl 2**, S31-41 (2000).
39. Rai, D.*, et al.* Epilepsy and psychiatric comorbidity: a nationally representative population-based study. *Epilepsia* **53**, 1095-1103 (2012).
40. Swinkels, W.A., Kuyk, J., van Dyck, R. & Spinhoven, P. Psychiatric comorbidity in epilepsy. *Epilepsy Behav* **7**, 37-50 (2005).
41. Wulsin, A.C., Solomon, M.B., Privitera, M.D., Danzer, S.C. & Herman, J.P. Hypothalamic-pituitary-adrenocortical axis dysfunction in epilepsy. *Physiol Behav* **166**, 22-31 (2016).
42. Vezzani, A., French, J., Bartfai, T. & Baram, T.Z. The role of inflammation in epilepsy. *Nat Rev Neurol* **7**, 31-40 (2011).
43. Mula, M. & Schmitz, B. Depression in epilepsy: mechanisms and therapeutic approach. *Ther Adv Neurol Disord* **2**, 337-344 (2009).
44. Schmitz, B. Effects of antiepileptic drugs on mood and behavior. *Epilepsia* **47 Suppl 2**, 28-33 (2006).
45. Gleichgerrcht, E., Kocher, M. & Bonilha, L. Connectomics and graph theory analyses: Novel insights into network abnormalities in epilepsy. *Epilepsia* **56**, 1660-1668 (2015).
46. Kanner, A.M. Depression in epilepsy: prevalence, clinical semiology, pathogenic mechanisms, and treatment. *Biol Psychiatry* **54**, 388-398 (2003).
47. Owen, L.W., Heusser A.C., Manning, J.R. . A Gaussian process model of human electrocorticographic data. *Preprint at https://www.biorxiv.org/content/10.1101/121020v2 (2018)*.
48. Gu, S.*, et al.* The Energy Landscape of Neurophysiological Activity Implicit in Brain Network Structure. *Sci Rep* **8**, 2507 (2018).
49. Garcia, J.O., Ashourvan, A., Muldoon, S.F., Vettel, J.M. & Bassett, D.S. Applications of community detection techniques to brain graphs: Algorithmic considerations and implications for neural function. *Proc IEEE Inst Electr Electron Eng* **106**, 846-867 (2018).
50. Bassett, D.S. & Sporns, O. Network neuroscience. *Nat Neurosci* **20**, 353-364 (2017).
51. Chen, Z.J., He, Y., Rosa-Neto, P., Gong, G. & Evans, A.C. Age-related alterations in the modular organization of structural cortical network by using cortical thickness from MRI. *Neuroimage* **56**, 235-245 (2011).
52. Bruno, J., Hosseini, S.M. & Kesler, S. Altered resting state functional brain network topology in chemotherapy-treated breast cancer survivors. *Neurobiol Dis* **48**, 329-338 (2012).
53. Cao, M., Shu, N., Cao, Q., Wang, Y. & He, Y. Imaging functional and structural brain connectomics in attention-deficit/hyperactivity disorder. *Mol Neurobiol* **50**, 1111-1123 (2014).
54. Sun, Y.*, et al.* Disrupted functional brain connectivity and its association to structural connectivity in amnestic mild cognitive impairment and Alzheimer's disease. *PLoS One* **9**, e96505 (2014).
55. Alexander-Bloch, A.F.*, et al.* Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Front Syst Neurosci* **4**, 147 (2010).
56. Yu, Q.*, et al.* Altered topological properties of functional network connectivity in schizophrenia during resting state: a small-world brain network study. *PLoS One* **6**, e25423 (2011).
57. Keown, C.L.*, et al.* Network organization is globally atypical in autism: A graph theory study of intrinsic functional connectivity. *Biol Psychiatry Cogn Neurosci Neuroimaging* **2**, 66-75 (2017).

58. Kroenke, K., Spitzer, R.L. & Williams, J.B. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* **16**, 606-613 (2001).
59. Spitzer, R.L., Kroenke, K. & Williams, J.B. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA* **282**, 1737-1744 (1999).
60. Spitzer, R.L., Williams, J.B., Kroenke, K., Hornyak, R. & McMurray, J. Validity and utility of the PRIME-MD patient health questionnaire in assessment of 3000 obstetric-gynecologic patients: the PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. *Am J Obstet Gynecol* **183**, 759-769 (2000).
61. Blondel, V.D., Guillaume, J.L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J Stat Mech-Theory E* (2008).
62. Newman, M.E. Finding community structure in networks using the eigenvectors of matrices. *Phys Rev E Stat Nonlin Soft Matter Phys* **74**, 036104 (2006).
63. Cammoun, L.*, et al.* Mapping the human connectome at multiple scales with diffusion spectrum MRI. *J Neurosci Meth* **203**, 386-397 (2012).
64. Bertolero, M.A., Yeo, B.T. & D'Esposito, M. The modular and integrative functional architecture of the human brain. *Proc Natl Acad Sci U S A* **112**, E6798-6807 (2015).
65. Guimera, R. & Amaral, L.A. Cartography of complex networks: modules and universal roles. *J Stat Mech* **2005**, nihpa35573 (2005).
66. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* **52**, 1059-1069 (2010).
67. Bassett, D.S., Yang, M., Wymbs, N.F. & Grafton, S.T. Learning-induced autonomy of sensorimotor systems. *Nat Neurosci* **18**, 744-751 (2015).
68. Newson, J.J. & Thiagarajan, T.C. EEG Frequency Bands in Psychiatric Disorders: A Review of Resting State Studies. *Front Hum Neurosci* **12**, 521 (2018).
69. Pollock, V.E. & Schneider, L.S. Quantitative, waking EEG research on depression. *Biol Psychiatry* **27**, 757-780 (1990).
70. Thibodeau, R., Jorgensen, R.S. & Kim, S. Depression, anxiety, and resting frontal EEG asymmetry: a meta-analytic review. *J Abnorm Psychol* **115**, 715-729 (2006).
71. Olbrich, S., van Dinteren, R. & Arns, M. Personalized Medicine: Review and Perspectives of Promising Baseline EEG Biomarkers in Major Depressive Disorder and Attention Deficit Hyperactivity Disorder. *Neuropsychobiology* **72**, 229-240 (2015).
72. Drysdale, A.T.*, et al.* Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med* **23**, 28-38 (2017).
73. Grisanzio, K.A.*, et al.* Transdiagnostic Symptom Clusters and Associations With Brain, Behavior, and Daily Function in Mood, Anxiety, and Trauma Disorders. *JAMA Psychiatry* **75**, 201-209 (2018).
74. Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N. & Barabasi, A.L. Hierarchical organization of modularity in metabolic networks. *Science* **297**, 1551-1555 (2002).
75. Rihel, J.*, et al.* Zebrafish behavioral profiling links drugs to biological targets and rest/wake regulation. *Science* **327**, 348-351 (2010).
76. Tomarken, A.J., Davidson, R.J., Wheeler, R.E. & Kinney, L. Psychometric properties of resting anterior EEG asymmetry: temporal stability and internal consistency. *Psychophysiology* **29**, 576-592 (1992).
77. Wheeler, R.E., Davidson, R.J. & Tomarken, A.J. Frontal brain asymmetry and emotional reactivity: a biological substrate of affective style. *Psychophysiology* **30**, 82-89 (1993).
78. Henriques, J.B. & Davidson, R.J. Regional brain electrical asymmetries discriminate between previously depressed and healthy control subjects. *J Abnorm Psychol* **99**, 22-31 (1990).

79. Alexopoulos, G.S.*, et al.* Functional connectivity in the cognitive control network and the default mode network in late-life depression. *J Affect Disord* **139**, 56-65 (2012).
80. Frodl, T.*, et al.* Functional connectivity bias of the orbitofrontal cortex in drug-free patients with major depression. *Biol Psychiatry* **67**, 161-167 (2010).
81. Nofzinger, E.A.*, et al.* Alterations in regional cerebral glucose metabolism across waking and non-rapid eye movement sleep in depression. *Arch Gen Psychiatry* **62**, 387-396 (2005).
82. Cheng, W.*, et al.* Medial reward and lateral non-reward orbitofrontal cortex circuits change in opposite directions in depression. *Brain* **139**, 3296-3309 (2016).
83. Feffer, K.*, et al.* 1Hz rTMS of the right orbitofrontal cortex for major depression: Safety, tolerability and clinical outcomes. *Eur Neuropsychopharmacol* **28**, 109-117 (2018).
84. Chanes, L., Quentin, R., Tallon-Baudry, C. & Valero-Cabre, A. Causal frequency-specific contributions of frontal spatiotemporal patterns induced by non-invasive neurostimulation to human visual performance. *J Neurosci* **33**, 5000-5005 (2013).
85. Cocchi, L. & Zalesky, A. Personalized Transcranial Magnetic Stimulation in Psychiatry. *Biol Psychiatry Cogn Neurosci Neuroimaging* **3**, 731-741 (2018).
86. Betzel, R.F.*, et al.* Structural, geometric and genetic factors predict interregional brain connectivity patterns probed by electrocorticography. *Nat Biomed Eng* (2019).
87. Mayberg, H.S.*, et al.* Cingulate function in depression: a potential predictor of treatment response. *Neuroreport* **8**, 1057-1061 (1997).
88. Riva-Posse, P.*, et al.* A connectomic approach for subcallosal cingulate deep brain stimulation surgery: prospective targeting in treatment-resistant depression. *Mol Psychiatry* **23**, 843-849 (2018).
89. Malone, D.A., Jr.*, et al.* Deep brain stimulation of the ventral capsule/ventral striatum for treatment-resistant depression. *Biol Psychiatry* **65**, 267-275 (2009).
90. Marchand, W.R.*, et al.* Aberrant functional connectivity of cortico-basal ganglia circuits in major depression. *Neurosci Lett* **514**, 86-90 (2012).
91. Hamani, C.*, et al.* The subcallosal cingulate gyrus in the context of major depression. *Biol Psychiatry* **69**, 301-308 (2011).
92. Hamilton, L.S., Chang, D.L., Lee, M.B. & Chang, E.F. Semi-automated Anatomical Labeling and Inter-subject Warping of High-Density Intracranial Recording Electrodes in Electrocorticography. *Front Neuroinform* **11**, 62 (2017).
93. Postelnicu, G., Zollei, L. & Fischl, B. Combined volumetric and surface registration. *IEEE Trans Med Imaging* **28**, 508-522 (2009).
94. Fischl, B. FreeSurfer. *Neuroimage* **62**, 774-781 (2012).
95. Akbarian, B. & Erfanian, A. Automatic Seizure Detection Based on Nonlinear Dynamical Analysis of EEG Signals and Mutual Information. *Basic Clin Neurosci* **9**, 227-240 (2018).
96. Schiff, S.J., Aldroubi, A., Unser, M. & Sato, S. Fast wavelet transformation of EEG. *Electroencephalogr Clin Neurophysiol* **91**, 442-455 (1994).
97. Antony, A.R.*, et al.* Simultaneous scalp EEG improves seizure lateralization during unilateral intracranial EEG evaluation in temporal lobe epilepsy. *Seizure* **64**, 8-15 (2019).
98. Mucha, P.J., Richardson, T., Macon, K., Porter, M.A. & Onnela, J.P. Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876-878 (2010).
99. Bassett, D.S.*, et al.* Robust detection of dynamic community structure in networks. *Chaos* **23**, 013142 (2013).
100. Sporns, O. & Betzel, R.F. Modular Brain Networks. *Annu Rev Psychol* **67**, 613-640 (2016).
101. Blondel, V.D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* **2008**, P10008 (2008).

102. Misic, B.*, et al.* Network-Level Structure-Function Relationships in Human Neocortex. *Cereb Cortex* **26**, 3285-3296 (2016).
103. van den Heuvel, M.P. & Sporns, O. Network hubs in the human brain. *Trends Cogn Sci* **17**, 683-696 (2013).
104. Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* **24**, 417-441, 498-520 (1933).
105. He, Z. & Yu, W. Stable feature selection for biomarker discovery. *Comput Biol Chem* **34**, 215-225 (2010).
106. Stone, M. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Socieity* **36**, 111-133 (1974).
107. Guo, L., Rivero, D., Dorado, J., Rabunal, J.R. & Pazos, A. Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. *J Neurosci Methods* **191**, 101-109 (2010).