Title

Big Data Science: Applying Unsupervised and Supervised Machine Learning Algorithms to Predict and Differentiate Between Vulvodynia and Healthy Controls Using High Dimensional Neuroimaging Data

Author

Gordon, David

Publication Date

2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Big Data Science: Applying Unsupervised and Supervised Machine Learning Algorithms to Predict and Differentiate Between Vulvodynia and Healthy Controls Using High Dimensional Neuroimaging Data

A thesis submitted in partial satisfaction of the

requirements for the degree Master of Science

in Clinical Research

by

David Gordon

2018

ABSTRACT OF THE THESIS

Big Data Science: Applying Unsupervised and Supervised Machine Learning Algorithms to Predict and Differentiate Between Vulvodynia and Healthy Controls Using High Dimensional Neuroimaging Data

by

David Gordon

Master of Science in Clinical Research

University of California, Los Angeles, 2018

Professor Janet Sinsheimer, Chair

Purpose: Due to the high-dimensionality and multicollinearity of brain morphometric and functional network features, as well as the small sample size in our study, we utilized a sparse partial least squares discriminatory analysis (sPLSDA) algorithm to deal with these challenges and select a subset of the original features to explore the underlying mechanisms of vulvodynia that differentiate affected individuals from healthy controls. To the best of our knowledge, this is the first study to perform unsupervised and supervised machine learning on neuroimaging data among individuals diagnosed with vulvodynia. **Methods:** We used a holdout procedure and performed a random 70/30 split for both case and healthy control data. This resulted in a training set N=86 ($N_{controls}$=26, N_{cases}=60) and a test set N=37 ($N_{controls}$=11, N_{cases}=26). We computed principal component analysis (PCA), partial least squares discriminatory analysis (PLSDA), and sPLSDA, to extract and select features from the original set of features that differentiate patients with vulvodynia from healthy controls. Furthermore, we applied a 10-fold cross validation

approach to split the observations into 10 sets and repeatedly train the model on 9 sets and evaluate its performance on the 10^{th} set. Class prediction was determined using the Mahalanobis distance metric, which utilizes a majority vote algorithm. **Results:** The sPLSDA algorithm selected 30 features from the 2768 original features to differentiate vulvodynia from healthy controls. The specificity, sensitivity, and predictive accuracy for the sPLSDA algorithm was found to be 89%, 73%, and 86%, respectively. The most influential selected features that differentiate patients with vulvodynia from healthy controls were functional network features, specifically of the within-module degree z score and participation related coefficient metrics. **Discussion:** By visualizing the sPLSDA, PLSDA, and PCA algorithms, we were able to examine how each algorithm performed on the discrimination, which in turns reveals potential insight into underlying mechanisms of vulvodynia, such as the important selected features. The predictive accuracy of the sPLSDA in our study was comparable with the predictive accuracy in previous neuroimaging studies utilizing sPLSDA and support vector machines in conditions often comorbid with vulvodynia.

The thesis of David Gordon is approved.

Jennifer Labus

Victor Chaban

David Elashoff

Janet Sinsheimer, Committee Chair

University of California, Los Angeles

2018

Table of Contents

List of Figures and Tables

Introduction

The Pain and Interoception Imaging Network (PAIN) is a multimodal, multisite, brain-imaging repository for chronic somatic and visceral pain disorders and aims to combine medical imaging with advanced physiological, genetics, and omics measures (Labus et al., 2016), (Clemens et al., 2014), (Landis et al., 2014). This archive provides a unique opportunity for the application of Big Data Science to further understanding of the central mechanism in chronic pain. According to Van Horn and Toga, Big Data Science refers to multifactorial and broad ranging data that is accumulating exponentially and increasing the demand for data sharing, multisite studies, multi-datatype archiving, and exploring and mining the data (Van Horn & Toga, 2014). Moreover, Big Data Science requires novel thinking about how data is generated, managed, integrated, and leveraged to advance scientific understanding and healthcare (Bui, Van Horn, & Consortium, 2017). The conversion of medical images into mineable high-dimensional data is known as medical imaging computing and is supported by the concept that medical images contain information regarding underlying pathophysiology and that these relationships can be demonstrated through quantitative image analysis (Gillies, Kinahan, & Hricak, 2016). Furthermore, medical imaging machine learning algorithms have been integrated with clinical variables to predict new cases and healthy controls (Labus et al., 2015), (Gupta, Mayer, et al., 2015). The use of structural, functional, and anatomical neuroimaging markers could support new gene discoveries and a better understanding of underlying mechanisms of disease as evident in current studies integrating neuroimaging endophenotypes and genetic data to lead to new treatments (Glahn, Thompson, & Blangero, 2007), (Thompson et al., 2014), (Blokland et al., 2017). Furthermore, it has been suggested that medical imaging data can be combined with transcriptomic data to bring further insight into molecular intricacies of painful diseases (Katrib,

Hsu, Bui, & Xing, 2016). The integration of such complex and heterogeneous Big Data offers unmatched opportunities to study chronic pain conditions (CPCs) (Labus et al., 2016), (Clemens et al., 2014), (Landis et al., 2014), (Van Horn & Toga, 2014) (K. A. Le Cao, Rossouw, Robert-Granie, & Besse, 2008), (K. A. Le Cao, Boitard, & Besse, 2011).

Patterns of brain connectivity can be measured using neuroimaging and can be indexed quantitatively and visualized using bioinformatics tools (Irimia, Goh, Torgerson, Vespa, & Van Horn, 2014). Previous studies have shown that there are two main domains of brain networks; functional and structural networks, that reflect different underlying neurobiological mechanisms and are important for identifying and interpreting functional and structural quantitative measures (van den Heuvel & Sporns, 2013). More specifically, functional networks are derived from statistical descriptions of time series data, which in resting-state functional MRI (rs-fMRI) studies are expressed as Pearson correlations. Whereas, structural networks demonstrate anatomical connectivity, which in diffusion tensor imaging (DTI) are expressed as probabilistic and deterministic tractography (van den Heuvel & Sporns, 2013). According to Sporns et al., the building blocks for organizing brain networks can be thought of as modules (Sporns & Betzel, 2016). Methodological approaches for detecting functional and structural modules, are of particular interest, as they reveal subnetworks that are particularly densely connected and correspond to specialized functional components (Sporns & Betzel, 2016).

Functional magnetic resonance imaging (fMRI) is one of the most widely accepted methods using blood oxygenation level dependent (BOLD) signal contrast, because of its non-invasive and high spatial resolution (Matthews & Jezzard, 2004). fMRI is widely used to detect brain

regions that change their level of activation in response to tasks, but can also measure correlations in spontaneous fluctuations of signals between different brain regions, and reveal functionally connected circuits. More recently, the functional activity and connectivity of the brain during the resting state have attracted increasing attention (Tunbridge, Farrell, Harrison, & Mackay, 2013). Resting-state functional MRI studies have identified associated fluctuations in brain regions involved in motor, auditory, visual, and language function (Craig, 2013).

Vulvodynia is a chronic vulvar pain condition that is often under treated and under diagnosed and lasts at least three months, in which women experience spontaneous unprovoked pain in the vulvar vestibule, pain provoked by mechanical stimulation including sexual intercourse, tampon insertion, or both (Henzell, Berzins, & Langford, 2017; Wesselmann, Bonham, & Foster, 2014). The type of pain reported includes neuropathic descriptors (such as hot-burning, itching, tingling or "pins and needles", and light touch) as well as continuous descriptors (such as throbbing and tender) and intermittent descriptors (such as sharp and stabbing) (Dargie, Gilron, & Pukall, 2017), (Reed, Harlow, Plegue, & Sen, 2016). Previous studies have shown an increased familial risk among patients treated for vulvodynia (Morgan et al., 2016).

A recent vulvodynia study using neuroimaging data found evidence that regions comprising the sensorimotor, salience, and default mode network showed substantial alterations in comparison to healthy controls (Gupta, Rapkin, et al., 2015). Moreover, Schweinhardt et al. found that patients with vulvodynia had increased gray matter in pain modulatory and stress-related areas (Schweinhardt, Kuchinad, Pukall, & Bushnell, 2008). In a study applying mild to moderate pressure to the vulvar vestibule, it was found that patients with vulvodynia had significantly

higher activation levels in the insular and frontal cortical regions than did healthy controls (Pukall et al., 2005). Furthermore, in patients with vuldoynia undergoing evoked pain at the thumb during fMRI, it was found that there was augmented brain activation, which suggests central neural pathology in vuvlodynia (Hampson et al., 2013).

The aim of this study is to apply big data methodologies to differentiate patients with vulvodynia from healthy controls using functional network measures and gray matter morphometry (volume, cortical thickness, surface area, and mean curvature). Although we may perform unsupervised machine learning, the extracted features may not clearly differentiate vulvodynia from healthy controls and supervised machine learning may need to be utilized to select the important features that differentiate vulvodynia from healthy controls. To perform the computational analysis, we will use the programming language R and the R package mixOmics version 6.2.0 (Rohart, Gautier, Singh, & Cao, 2017). This study was approved by the UCLA Institutional Review Board.

Methods

Participants

This case-control study sample consisted of a total of 123 female subjects (37 healthy controls and 86 patients with vulvodynia) enrolled in studies at the Oppenheimer Center for Neurobiology of Stress and Resilience at UCLA (Labus et al., 2016). The diagnosis of vulvodynia was made during a clinical examination by an OB/GYN. The data used in this study was obtained from the PAIN repository and CNS repository (Labus et al., 2016), (Clemens et al., 2014). To test the generalizability of the predictive model, we used a holdout procedure through random 70/30 split for both case and healthy control data. This resulted in a training set N=86 ($N_{controls}$=26, N_{cases}=60) and a test set N=37 ($N_{controls}$=11, N_{cases}=26) (Table 1).

Structural and Functional Imaging Acquisitions

Whole brain functional resonance imaging (fMRI) data was acquired using a 3.0T MRI scanner (Siemens Trio; Siemens, Erlangen, Germany) (Gupta, Rapkin, et al., 2015). A high resolution structural image was acquired from each subject for registration purposes with a magnetization-prepared rapid acquisition gradient-echo sequence, repetition time = 2200ms, echo time = 3.26ms, structural acquisition time =5m 12s, slice thickness = 1mm, 176 slices, 256*256 voxel matrix, 1mm voxel size. Resting state scans were acquired using the following parameters: 40-slice whole brain volumes, slice thickness = 4mm, repetition time = 2000ms, echo time= 28ms, resting acquisition time = 10ms, flip angle = 77°, field of view = 220, 2×2×2 mm voxel size. Noise reducing headphones were used. Subjects rested with eyes closed while functional blood oxygen-level dependent images were acquired.

T-1 Image Segmentation

Image segmentation and regional parcellation of gray matter images was performed using the FreeSurfer software (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999; Fischl, Sereno, Tootell, & Dale, 1999) and workflow pipelines from the UCLA Laboratory of Neuroimaging pipeline using the Destrieux atlas and the Harvard-Oxford atlas (Destrieux, Fischl, Dale, & Halgren, 2010). This parcellation yielded 74 cortical structures, 7 subcortical structures, the cerebellum, and the brainstem, for a complete set of 165 parcellations for the entire brain. Gray matter volume was computed for all 165 regions. In addition, for cortical regions , cortical thickness, surface area, and mean curvature were computed.

Functional Brain Network Construction

The parcellation and the functional connectivity results were combined to produce a 165x165 weighted, undirected connectivity matrix. Resting-state image preprocessing was performed using the SPM8 software (Friston). Images were transformed from Digital Imaging and Communications in Medicine (DICOM) into Neuroimaging Informatics Technology Initiative (NIfTI), slice-time corrected, co-registered with the high-resolution structural images, spatially normalized to the Montreal Neurological Institute (MNI) space, and resampled to a voxel size of 2 X 2 X 2 mm. Normalized functional images were further preprocessed and analyzed using the SPM-based CONN toolbox version 13 (Whitfield-Gabrieli & Nieto-Castanon, 2012). The resting-state images were filtered using a band-pass filter ($0.008/s<f<0.08/s$) to reduce the low- and high-frequency noises. A component based noise-correction method (Whitfield-Gabrieli & Nieto-Castanon, 2012), was applied to remove nuisances for better sensitivity and specificity of the analysis. Six motion realignment parameters and confounds for white matter and CSF were

removed using regression. ROI-to-ROI functional connectivity, cross-correlations between the blood oxygenated level dependent time series, computed between all the brain regions in CONN toolbox. The connectivity correlation coefficients were then used to construct the final functional network setting negative values to zero. The magnitude of the correlation represents the weights in the functional network.

Computing Functional Network Metrics

The Graph Theoretic General Linear Model tool and in-house matlab scripts were used to calculate and analyze the brain network properties and organization from the subject-specific functional brain networks (Spielberg, 2015). A brain network is defined by a collection of nodes (brain regions) and links (functional connections) between pairs of nodes (brain regions) (Rubinov & Sporns, 2017). Several local weighted network metrics indexing centrality, segregation and integration were computed. Regions with high centrality are highly influential, communicate with many other regions, facilitate functional integration, and play a key role in network resilience. Three general measures of connectivity were assessed: Centrality, Segregation, and Integration; which are explained in detail below as well outlined in Table 2.

A. Centrality: 1) Degree Strength reflects the number of other regions a region interacts with functionally, 2) Betweenness centrality, reflecting the ability of a region to control information flow between two other regions, 3) Eigenvector Centrality, where higher values region is connected to other highly connected regions reflective of the global (vs. local) prominence of a region, 4) Pagerank Centrality, which is a variant of eigenvector centrality, 5) Within-Module Degree Z Score, is a within-module version of degree centrality used to determine nodes with a

large number of nodes, 6) Participation Coefficient, which is a measure of diversity of intermodular connections of individual nodes, 7) Diversity Coefficient, which is a related measure to the participation coefficient based on Shannon entropy.

B. Segregation: Segregation reflects the presence of densely connected brain regions forming an infrastructure having the ability to perform specialized processes and indicating functional segregation. Here we compute the clustering coefficient, where a higher value indicates the presence of cluster connectivity around an individual brain region.

C. Integration: Integration quantifies the brain's ability to expeditiously combine specialized information from distributed brain regions. This is characterized by the ease by which brain regions can communicate with every other brain region in the whole brain network. A primary measure of integration is a region's average characteristic path length or average minimal travel distance between regions, shorter characteristic path length indicates greater ability to integrate information.

Modules and Important Related Metrics

According to Sporns et al., modules can be thought of as building blocks in the organization of brain networks as well as correspond to clusters of nodes that are connected (Sporns & Betzel, 2016). Functional connectivity metrics of particular importance to modules are the within-module degree z score and participation coefficient, which are mathematically defined below (Sporns & Betzel, 2016).

Within-module degree z score is defined as:

$$z_i = \frac{k_i - \bar{k}_{s_i}}{\sigma_{k_{s_i}}}$$

Where ki is the number of links of node i to other nodes in its module si, k̄si is the average of k over all the nodes in si, and σksi is the standard deviation of k in si, then zi is the within-module degree z-score and measures how well-connected node i is to other nodes in the module (Guimerà & Nunes Amaral, 2005).

Participation coefficient is defined as:

$$y_i = 1 - \sum_{m \in M} \left(\frac{k_i(m)}{k_i} \right)^2$$

Where M is the set of modules and ki(m) is the number of links between i and all nodes in module m.

Overview of Analysis

We will apply sPLSDA, PLSDA, and PCA to perform dimension reduction and uncorrelate the independent variables to determine the optimal model for differentiating patients with vulvodynia from healthy controls. Moreover, we will select the model that best differentiates the study populations to predict new cases and healthy controls utilizing functional network measures and gray matter morphometry (volume, cortical thickness, surface area, and mean curvature). To perform the computational analysis, we will use the programming language R and the R package mixOmics version 6.2.0 (Rohart et al., 2017).

Sparse Partial Least Squares Discriminatory Analysis

One of the statistical methodologies we will utilize to explore the relationship between the high dimensionality of neuroimaging data is supervised machine learning. sPLSDA is a variant of partial least squares that is a feature selection technique and has been used to analyze brain networks in IBS and other conditions (Labus et al., 2015), (Labus et al., 2009), (Labus et al., 2008). We chose sPLSDA based on its ability to deal with a large number of predictors, multicollinearity with predictors, small sample size, and ability to select important features (Rohart et al., 2017). The sPLSDA has shown good classification performance in comparison with other methods and can be used to input data and predict new cases (Rohart et al., 2017), (K. A. Le Cao et al., 2011), (K. A. Le Cao et al., 2008).

sPLSDA utilizes shrinkage methodologies to help yield a less complex model by shrinking the regression coefficients (Clemmensen, Hastie, Witten, & Ersboll, 2012). This process is known as penalization or regularization. The least absolute shrinkage and selection operator (lasso), L1 penalty, has the effect of forcing some of the coefficient estimates to be exactly equal to zero when the tuning parameter λ is sufficiently large. Hence, much like best subset selection, the lasso performs variable selection. Furthermore, Rohart et al. suggests incorporating soft-thresholding to the L1 penalty for signal denoising, which produces smoother results and selects the number of features desired on each dimension (Rohart et al., 2017). The R Package mixOmics, replaces the λ parameter with a function called keepX for soft-thresholding (Rohart et al., 2017). Thus, sPLSDA has an L1 penalization on the loading vector a_h to shrink some coefficients to zero. Thus, for each dimension h = 1,.., H, sPLSDA solves:

$$\max_{(a_h, b_h)} \operatorname{cov}(X_h a_h, Y_h b_h), \qquad \text{s.t.} \quad ||a_h||_2 = ||b_h||_2 = 1 \text{ and } ||a_h||_1 \leq \lambda_h$$

where λ_h is a non negative parameter that controls the amount of shrinkage in a_h.

Ultimately, the algorithm uses a weighted loadings filter method to selects only a few of the most important features that contributes to the final model. Each feature selected for a component has an associated loading, which is a measure of the importance of that feature in the component for the discrimination into the two groups (Mehmood, Liland, Snipen, & Saebo, 2012) (Gidskehaug, Anderssen, Flatberg, & Alsberg, 2007). To avoid potentially overfitting the model, we did not utilize the mixOmics optimal constraint function (Rohart et al., 2017).

Partial Least Squares Discriminatory Analysis

The PLSDA algorithm applies a dimension reduction technique to obtain a set of principle values (Mizuta, 2012). PLSDA identifies a new set of features in a supervised way by making use of the outcome variable in order to identify new features that not only approximate old features well, but also that are related to the outcome variable, which is a limitation of other dimension reduction techniques such as PCA.

Principal Component Analysis

We will explore the performance of an unsupervised learning method, PCA, which is a feature extraction technique and discovers the signal in features without the outcome variable by using an orthogonal transformation to convert a set of potentially correlated variables into a set of linearly uncorrelated variables called principal components. Principal components are a function of all the features.

Predicting New Case and Control by Distance Metric

In this study, the Mahalanobis distance was used to classify selected features as being of the healthy control or vulvodynia class. The Mahalanobis distance is a commonly used metric in machine learning to measure distance between observations, as it controls for correlation between the different features, which is a limitation of the standard Euclidean distance (Shen, Kim, & Wang, 2010). According to Rosenbaum and Rubin, using the Mahalanobis distance may slightly improve the matches and reduce bias (Rosenbaum & Rubin, 1983), (Rubin, 1980). Moreover, the Mahalanobis distance is recommended for use among study populations that may have an imbalance between study groups (Rohart et al., 2017). The R package mixOmics version 6.2.0 uses the function called majority vote, based on the Mahalanobis distance, to choose the class, defined as the centroid distance, that is most common among the selected observations (K.-A. Le Cao, Rohart, Gonzalez, & Dejean, 2017). The Mahalanobis distance formula is shown below, where Gk is the centroid:

$$\left(\text{dist}(x, G_k) = \sqrt{(x - G_k)^\top S^{-1}(x - G_k)}, \text{ where } S \text{ is the variance-covariance matrix of } x - G_k\right).$$

K-Fold Cross-Validation

K-fold cross validation is a model validation method, which is used to show how well the results from a training set generalize to an independent test set. For this study, we will set k=10 and repeat 10 times, so it resamples 100 times. When k = 10, it splits the observations into 10 sets. This will repeatedly train the model on 9 sets and evaluate its performance on the 10^{th} set.

Performance Evaluations

We will compute the specificity, sensitivity, precision, recall, observed accuracy, expected accuracy, kappa statistics, f-measure, and balanced error rate (BER) to evaluate the performance of the binary classifier. Kappa statistic is a good measure to inspect classifications that may be due to chance. As kappa statistic value calculated for classifiers closer to 1, the performance of classifier is assumed to be more realistic rather than being by chance. Thus, kappa statistic value is a recommended metric to consider for evaluation in the performance analysis of classifiers. Observed accuracy and expected accuracy are commonly used metrics to determine the performance/ discriminative ability of binary classifiers. Furthermore, the f-measure is a measure of a test's accuracy, which is usually used as performance evaluation metric to assess the performance of binary classifier, based on the harmonic mean for the classifier's precision and recall.

Results

We first performed a preliminary PCA analysis on the brain morphometric and functional network measures data (Figure 1), which allows us to explore the sources of variation in the data in an unsupervised analysis. In order to understand the amount of variation explained, we set number of components to a rather large number, 10. In Figure 1, the PCA numerical output shows that about 26% of the total variance is explained with 3 principal components. Figure 1 shows the variance explained per component.

In Figure 2, the samples are represented on the first two principal components and colored according to the condition group, vulvodynia is red and healthy control is blue. We observe that group types may not explain the major source of variation. Note that since PCA is unsupervised, we only take into account the sample type information after the PCA, for visualization purposes. Confidence ellipses for each group are plotted to highlight the strength of the discrimination (confidence level set to 95%).

In Figure 3, we see the supervised analysis helped refine the clusters of samples in a supervised fashion. For discriminant analysis, we set up the Y as a factor indicating the group membership of each condition. We fit a PLS-DA model with two components; the samples are then projected in the subspace spanned by these two components. We observe a fairly clear separation of the two groups compared to an unsupervised PCA sample plot; however, this model is built on all 2768 features, which makes it difficult to examine potential underlying mechanism of vulvodynia. Confidence ellipses for each group are plotted to highlight the strength of the discrimination (confidence level set to 95%).

In Figure 4, the balanced error rate (12%) is shown for the PLSDA model. The PLSDA model is built on all 2768 features, many of which may be uninformative to characterize the different groups. Using the tune.splsda function, we set ncomp=3, as we saw from the PCA and BER that after the first two components, the error rate does not significantly improve. We selected 10-fold CV repeated 10 times, which resamples 100 times, and specified a Mahalanobis distance to predict the group membership across all CV runs. BER is appropriate for an unbalanced number of samples per group as it calculates the average proportion of wrongly classified samples in each group, weighted by the number of samples in each group. The BER indicates a small gain in adding more than 10 selected variables for component 1 and 20 selected variables for component 2. Furthermore, the classification performance displayed suggests that ncomp=2 is sufficient to achieve good performance.

Figure 5 shows the final sPLSDA model, which is built on 30 selected features across two component; 10 selected features in component 1 and 20 selected features in component 2. The sample plots on the first two components shows that vulvodynia is fairly well separated on the first component, while the second component discriminated vulvodynia slightly better. Also, there seems to be clusters within the vulvodynia group that may be indicative of subtypes for vulvodynia. Confidence ellipses for each group are plotted to highlight the strength of the discrimination (confidence level set to 95%).

In Figure 6, we computed a confusion matrix for the test set. The positive predictive value/specificity was found to be 89%. Furthermore, the negative predictive value/sensitivity was found to be 73%. The precision was found to be 89%. The recall was found to be 92%.

The f-measure was found to be 91%. Moreover, the observed accuracy of the binary classifier was found to be 86%; whereas, the expected accuracy (due to chance) was found to be 57%. Therefore the kappa statistic can be calculated using the observed accuracy (86.4%) and the expected accuracy (57.4%) = 0.864 - 0.574 / 1 - 0.574 = 0.29 / 0.426 = 0.68.

In Figure's 7 and 8, we see the feature visualization for loading weights for component's 1 and 2, respectively. The plotLoading function was used, where color indicates the group for which the selected variable has a maximal mean value. The most important variables are ordered from bottom to top.

In Table 3, the features compromising each brain signature are displayed. The two most important features on the first brain signature and higher values in the vulvodynia group compared to the healthy control group are the 1) Subparietal sulcus within-module degree z-score and 2) Posterior-ventral part of the cingulate gyrus within-module degree z-score. The two most important features on the first brain signature and lower values in the vulvodynia group compared to the healthy control group are the 1) Inferior part of the precentral sulcus diversity coefficient and 2) Inferior frontal sulcus diversity coefficient. Similarly, the three most important features on the second brain signature and lower values in the vulvodynia group compared to the healthy control group are the 1) Superior frontal gyrus participation coefficient, 2) Superior temporal sulcus participation coefficient, and 3) Transverse frontopolar gyri and sulci participation coefficient.

Discussion

The aim of this study was to employ unsupervised and supervised machine learning algorithms to differentiate patients with vulvodynia from healthy controls using high-dimensional neuroimaging data. We were able to demonstrate the benefits of sPLSDA over PLSDA and PCA to differentiate vulvodynia from healthy controls to include being able to select 30 important features from the 2768 original features and predict new cases and controls.

The sensitivity (73%), specificity (89%), and predictive accuracy (86%) achieved with the sPLSDA in our study was comparable to the sensitivity (68%), specificity (71%), and predictive accuracy (70%), obtained using sPLSDA on gray matter morphometry data for discriminating healthy controls from subjects with a condition often comorbid with vulvodynia, irritable bowel syndrome (IBS) (Labus et al., 2015) as well as comparable to the predictive accuracy obtained using support vector machines and whole-brain voxel-based morphometry data for discriminating healthy controls from subjects with interstitial cystitis/bladder pain syndrome, which is also a condition comorbid with vulvodynia and IBS (Berman et al., 2002).

Previous studies have proposed that the functional role of brain regions could be determined, largely, by its within-module degree z score (which expresses the number of connections a brain region makes to other brain regions in the module) and its participation coefficient (which expresses the degree to which a brain region's connections are distributed across modules) that define how the brain region is positioned in its own module and with respect to other modules (Sporns & Betzel, 2016), (Guimerà & Nunes Amaral, 2005), (Rives & Galitski, 2003), (Han et al., 2004). Moreover, modular structures have been reported in biochemical networks and stated

to be crucial to functionality (Hartwell, Hopfield, Leibler, & Murray, 1999), (Ravasz, Somera, Mongru, Oltvai, & Barabási, 2002), (Ravasz, 2009), (Holme, Huss, & Jeong, 2003), (Papin, Reed, & Palsson, 2004). Furthermore, previous studies have reported the need to identify modules accurately (Girvan & Newman, 2002), (Newman & Girvan, 2004), (Eriksen, Simonsen, Maslov, & Sneppen, 2003), (Bui-Xuan & Jones, 2014), (Newman, 2004), (Radicchi, Castellano, Cecconi, Loreto, & Parisi, 2004), (Radicchi, 2018).

In our study, functional features dominated morphometric features and the most influential functional features on brain signature 1 were of the within-module degree z score (which were higher in the vulvodynia group compared to the healthy control group) and diversity coefficient metrics (which is a participation coefficient related measure and was lower in the vulvodynia group compared to the healthy control group). Moreover, on brain signature 2, the most influential functional features were of the participation coefficient metrics (which were lower in the vulvodynia group compared to the healthy control group). Therefore, the most influential selected features in our study were a combination of the within-module degree z score and participation related coefficient metrics. The most influential within-module degree z score metrics in brain signature 1 are located in the cingulate and parietal regions, which are regions known to be pain-related networks (Walitt, Ceko, Gracely, & Gracely, 2016). Furthermore, the cingulate region has been associated with the default mode network (DMN) and it is hoped that measurements of the constitution of the DMN and how easily it deactivates during tasks in the different central sensitivity syndromes will provide more insight into these disorders (Walitt et al., 2016).

While PCA is able to reduce the dimensionality of high dimensional data and uncorrelate independent variables, it is not a feature selection technique rather it is a feature extraction technique. Therefore, performing PCA gives us principal components that are a function of all the features and did not reveal a clear pattern based on group type, which was used for visualization purposes only after PCA was performed. Therefore, we explored the use of feature selection techniques that were able to reduce the dimensionality of data, uncorrelate the independent variables, and select important features. First, we utilized PLSDA, which selected important features that differentiate the patient populations; however, it included all of the 2768 original features, which makes it difficult to examine the underlying mechanisms of vulvodynia. Therefore, we introduced a sparse penalty function to PLSDA. Using sPLSDA, we were able to build a prediction algorithm on 30 features from the 2768 original features.

As mentioned, the sPLSDA method utilizes penalization to shrink regression coefficients toward zero and some exactly to zero in order to identify the most key features (Clemmensen et al., 2012). Selected key features reveal what the algorithm learned to extract as meaningful information and visualizing these key features helps represent that and perhaps lays a foundation for future work (Ho, Speier, El-Saden, & Arnold, 2017). We considered utilizing ridge regression; however ridge includes all predictors in the final model and we wanted a model that selects a subset of the old features. As a result, models generated from the lasso are generally much easier to interpret than those produced by ridge regression. Similarly, Lange et al. suggests the use of penalization methods for high dimensional data to help with identifying key features (Lange, Papp, Sinsheimer, & Sobel, 2014). Furthermore, Chi et al. demonstrated the use of penalization methods for selecting key features among integrated imaging and genetics data (Chi

et al., 2013). To assess the model performance we examined the test error rate, particularly the BER. For a good model performance, we would expect a low BER, which was achieved. To examine the BER, we employed the k-fold cross validation approach, where k=10. We double-checked the BER accuracy by also performing a permutation test, where we resampled the training set 5 times. The BER was found to be higher for the permutation test (30%) than the k-fold cross validation (12%) approach; therefore, we suspect the model is a good fit.

Limitations

Criteria from a Receiver Operating Characteristic (ROC) curve or Area Under the Curve (AUC) may not be particularly insightful since the prediction threshold in the mixOmics package is based on a specified distance; in the case of our study the Mahalanobis distance. (Rohart et al., 2017). The findings from this study are based on data collected from one site; therefore, these findings may not generalize to other sites. Furthermore, data on ethnicity/race was not available. Moreover, ROI data was based on only one atlas. Also, this study only used two types of data. Additionally, the study sample size was small and may not generalize.

Future Directions

Medical imaging machine learning algorithms that differentiate irritable bowel syndrome have been applied to clinical variables, such as bowel symptoms, to examine its potential use in drug discovery and clinical decision support (Labus et al., 2015). Similarly, the algorithm in our study, that utilizes functional network measures and gray matter morphometry, may provide

insight when applied to clinical variables, such as pain duration and level of pain intensity. Furthermore, recent studies are attempting to integrate anatomical and functional neuroimaging endophenotypes with SNP biomarkers and genetic variants to explore underlying mechanisms of disease (Thompson et al., 2014), (Blokland et al., 2017), (Davies et al., 2015), (Adams et al., 2016), (Hibar et al., 2015), (Trampush et al., 2017). A recent study reported that chronic pain conditions are moderately heritable (McIntosh et al., 2016). The sPLSDA method is robust and could be used to integrate neuroimaging data with genetic data (Rohart et al., 2017).

Table 1. Study Population

Training Set (N=86)	Test Set (N=37)	Total (N=123)
Vulvodynia **N=60** **Healthy Controls** **N=26**	**Vulvodynia** **N=26** **Healthy Controls** **N=11**	**Vulvodynia** **N=86** **Healthy Controls** **N=37**
Mean Age: 28 **SD: 6.5**	**Mean Age: 28** **SD: 7.4**	**Mean Age: 28** **SD: 6.8**

Table 2. Functional Network Measures

Functional (RS-fMRI) Network Measures for 165 Brain Regions
Centrality 1.Degree strength 2.Betweenness centrality 3.Eigenvector centrality 4.Pagerank centrality 5.Closeness centrality 6.Within-module degree z-score 7.Participation coefficient 8.Diversity coefficient
Integration 1. Average characteristic path length
Segregation 1.Clustering coefficient 2.Local efficiency 3.Global efficiency
Assortativity 1. Local assortativity

Figure 1. Principal Component Analysis

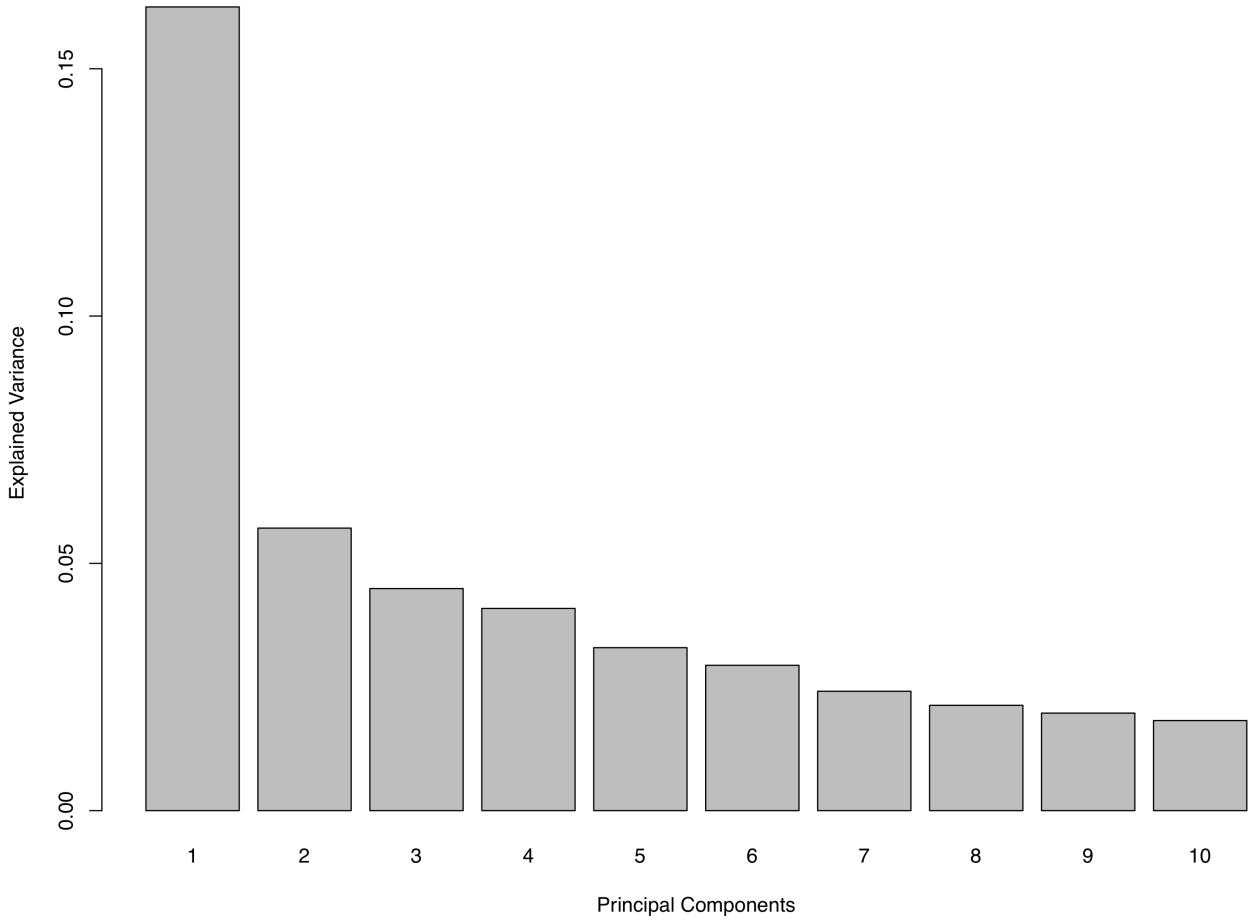

Figure 2. PCA on Brain Morphometric Measurements and Functional Network Measures

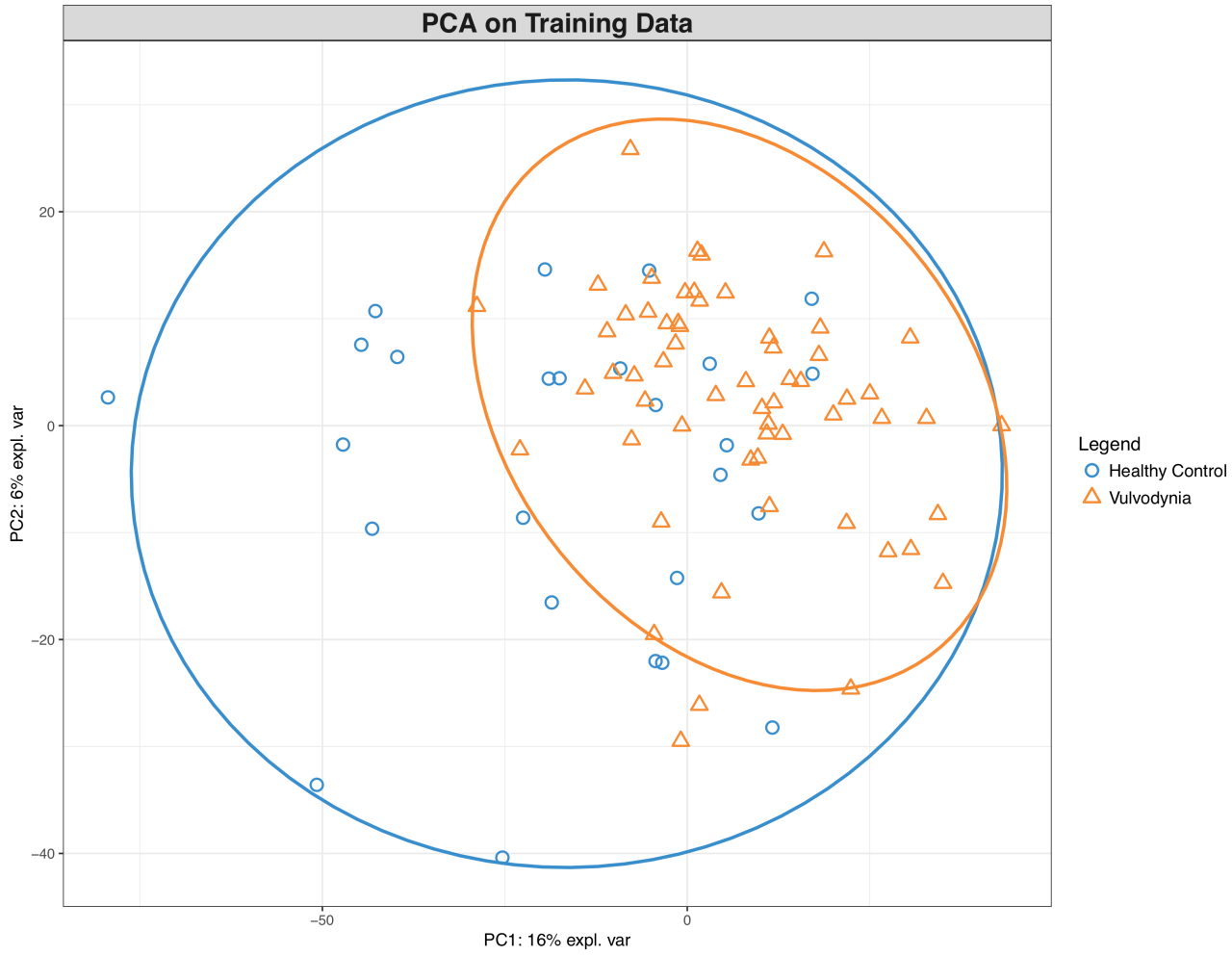

Figure 3. Partial Least Squares Discriminatory Analysis on Brain Morphometric Measurements and Functional Network Measures

Figure 4. Balanced Error Rate for Tuning sPLSDA Model

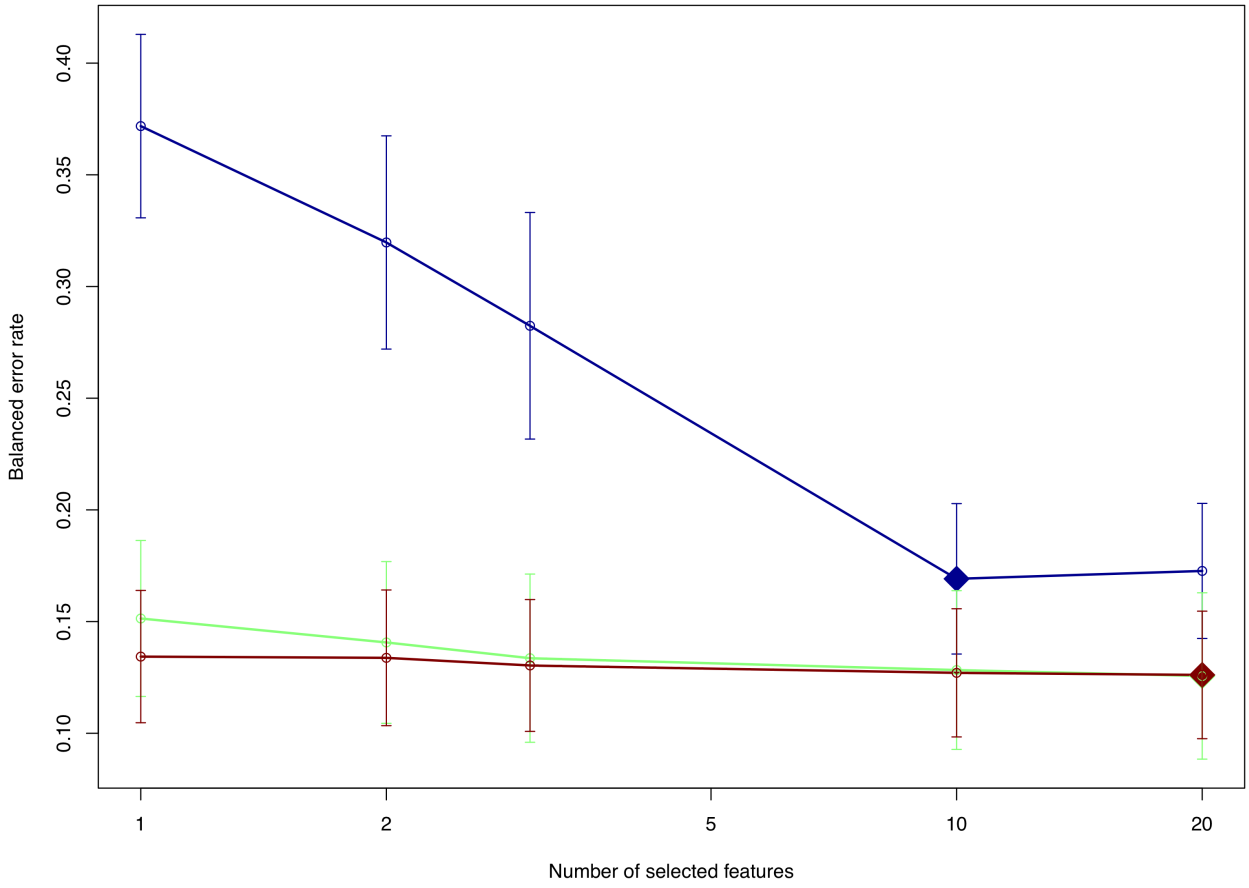

Figure 5. Final Model - Sparse Partial Least Squares Discriminatory Analysis on Brain Morphometric Measurements and Functional Network Measures

Figure 6. Confusion Matrix for Test Set

	True Condition		
Predicted Condition		Vulvodynia	Healthy Control
	Vulvodynia	24 (True Positive)	3 (False Positive)
	Healthy Control	2 (False Negative)	8 (True Negative)

Figure 7. Feature Visualization for Loadings on Component 1

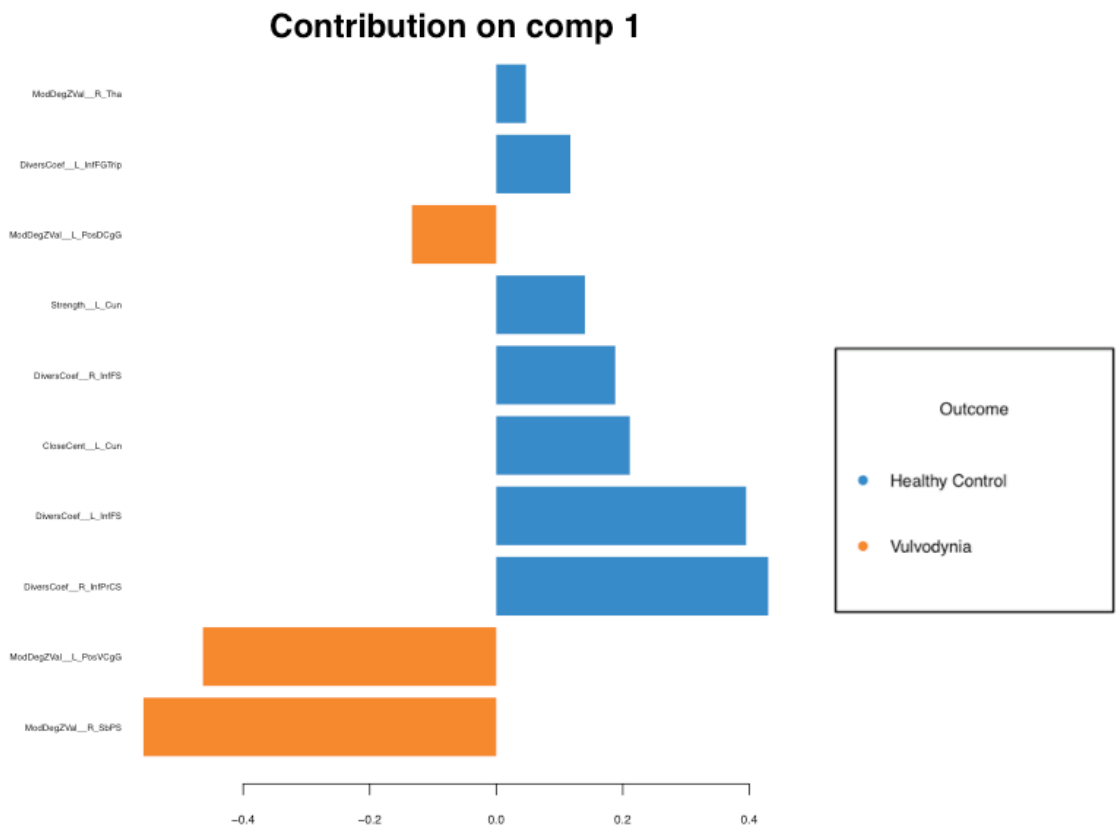

Figure 8. Feature Visualization for Loadings on Component 2

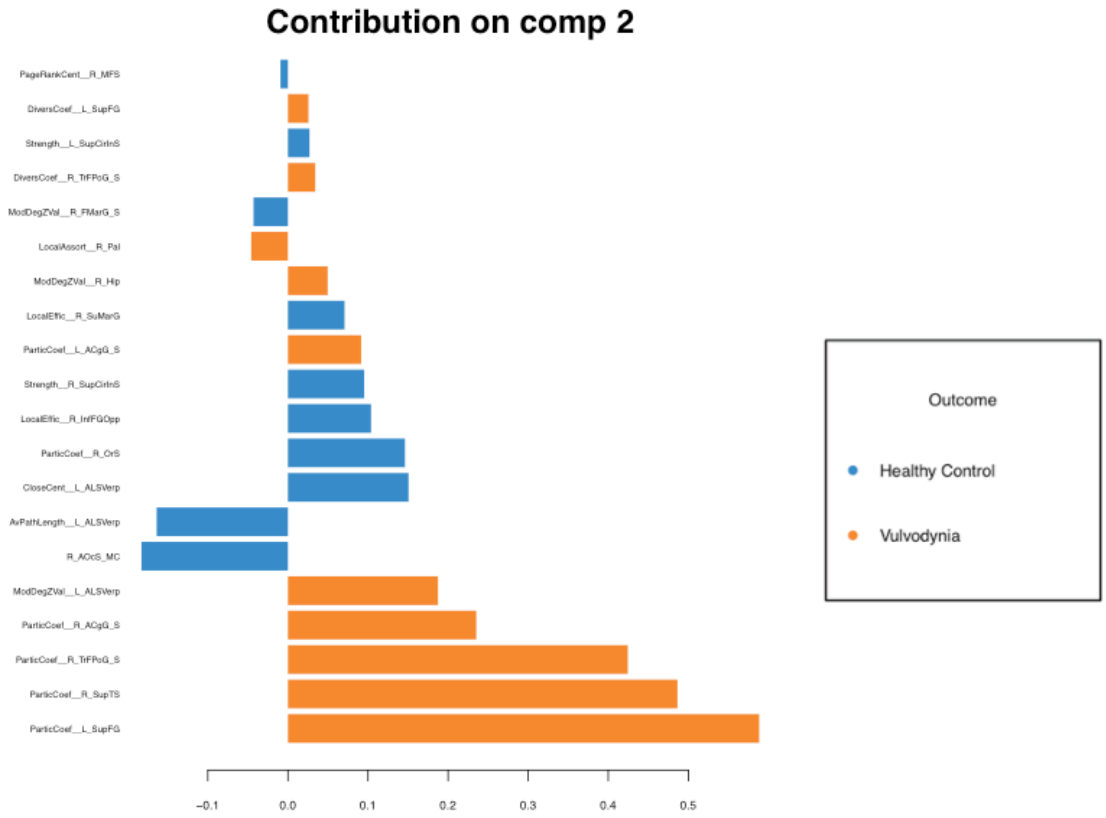

Table 3. Features Comprising Each Brain Signature

Brain Region	Description	Hemisphere	Metric	Loading Comp 1	Loading Comp 2	Network
Brain Signature 1						
Parietal	Subparietal sulcus	R	Within-module degree z-score	-0.56		Executive control
Cingulate	Posterior-ventral part of the cingulate gyrus (vPCC,isthmus of the cingulate gyrus)	L	Within-module degree z-score	-0.46		DMN
Parietal	Inferior part of the precentral sulcus	R	Diversity Coefficient	0.43		Somatosensory
Frontal	Inferior frontal sulcus	L	Diversity Coefficient	0.40		Executive control
Occipital	Cuneus(O6)	L	Closeness Centrality	0.21		Other
Frontal	Inferior frontal sulcus	R	Diversity Coefficient	0.19		Executive control
Occipital	Cuneus(O6)	L	Strength	0.14		Other
Cingulate	Posterior-dorsal part of the cingulate gyrus(dPCC)	L	Within-module degree z-score	-0.13		DMN
Frontal	Triangular part of the inferior frontal gyrus	L	Diversity Coefficient	0.12		Emotional Arousal
Subcortical	Thalamus	R	Within-module degree z-score	0.05		Sensorimotor (reward)
Brain Signature 2						
Frontal	Superior frontal gyrus(F1)	L	Participation Coefficient		0.59	Somatosensory
Temporal	Superior temporal sulcus(parallel sulcus)	R	Participation Coefficient		0.49	Other
Frontal	Transverse frontopolar gyri and sulci	R	Participation Coefficient		0.42	Emotional Arousal/Central Autonomic Network/Salience
Cingulate	Anterior part of thecingulate gyrus and sulcus(ACC)	R	Participation Coefficient		0.24	Emotional Arousal
Insular	Vertical ramus of theanterior segment of the lateral sulcus(orfissure)	L	Within-module degree z-score		0.19	Salience/Central Automatic Network
Occipital	Anterior occipital sulcus andpreoccipital notch(temporo-occipital incisure)	R	Mean Curvature		-0.18	Other
Insular	Vertical ramus of theanterior segment of the lateral sulcus(orfissure)	L	Average Path Length		-0.16	Salience/Central Automatic Network

Insular	Vertical ramus of theanterior segment of the lateral sulcus(orfissure)	L	Closeness Centrality		0.15	Salience/Central Automatic Network
Frontal	Orbital sulci(H-shaped sulci)	R	Participation Coefficient		0.15	Salience/Central Automatic Network/Reward (sensorimotor)
Frontal	Opercular part of the inferior frontal gyrus	R	Local Efficacy		0.10	Emotional Arousal
Insular	Superior segment of thecircular sulcus of the insula	R	Strength		0.10	Salience/Central Automatic Network
Cingulate	Anterior part of thecingulate gyrus and sulcus(ACC)	L	Participation Coefficient		0.09	Emotional Arousal
Parietal	Supramarginal gyrus	R	Local Efficacy		0.07	DMN
Subcortical	Hippocampus	R	Within-module degree z-score		0.05	Emotional Arousal/Reward (sensorimotor)
Subcortical	Pallidum	R	Local Assortativity		-0.05	Sensorimotor (reward)_
Frontal	Fronto-marginal gyrus (of Wernicke) and sulcus	R	Within-module degree z-score		-0.04	Executive control
Frontal	Transverse frontopolar gyri and sulci	R	Diversity Coefficient		0.03	Salience/Emotiona l Arousal/Central Automatic Network
Insular	Superior segment of thecircular sulcus of the insula	L	Strength		0.03	Salience/Central Automatic Network
Frontal	Superior frontal gyrus(F1)	L	Diversity Coefficient		0.03	Somatosensory
Frontal	Middle frontal sulcus	R	Page Rank Centrality		-0.01	Executive control

[1]Negative loadings for each component are associated with higher values and a group contribution/importance in the vulvodynia group compared to the healthy control group.
[2]Positive loadings for each component are associated with higher values and a group contribution/importance in the healthy control group compared to the vulvodynia group.

References

1. Adams, H. H., Hibar, D. P., Chouraki, V., Stein, J. L., Nyquist, P. A., Rentería, M. E., . . . Thompson, P. M. (2016). Novel genetic loci underlying human intracranial volume identified through genome-wide association. *Nat Neurosci, 19*(12), 1569-1582. doi:10.1038/nn.4398
2. Berman, S. M., Naliboff, B. D., Chang, L., Fitzgerald, L., Antolin, T., Camplone, A., & Mayer, E. A. (2002). Enhanced preattentive central nervous system reactivity in irritable bowel syndrome. *Am J Gastroenterol, 97*(11), 2791-2797. doi:10.1111/j.1572-0241.2002.07024.x
3. Blokland, G. A. M., Del Re, E. C., Mesholam-Gately, R. I., Jovicich, J., Trampush, J. W., Keshavan, M. S., . . . Petryshen, T. L. (2017). The Genetics of Endophenotypes of Neurofunction to Understand Schizophrenia (GENUS) consortium: A collaborative cognitive and neuroimaging genetics project. *Schizophr Res*. doi:10.1016/j.schres.2017.09.024
4. Bui, A. A. T., Van Horn, J. D., & Consortium, N. B. K. C. (2017). Envisioning the future of 'big data' biomedicine. *J Biomed Inform, 69*, 115-117. doi:10.1016/j.jbi.2017.03.017
5. Bui-Xuan, B. M., & Jones, N. S. (2014). How modular structure can simplify tasks on networks: parameterizing graph optimization by fast local community detection. *Proc Math Phys Eng Sci, 470*(2170), 20140224. doi:10.1098/rspa.2014.0224

6. Chi, E. C., Allen, G. I., Zhou, H., Kohannim, O., Lange, K., & Thompson, P. M. (2013). IMAGING GENETICS VIA SPARSE CANONICAL CORRELATION ANALYSIS. *Proc IEEE Int Symp Biomed Imaging, 2013*, 740-743. doi:10.1109/ISBI.2013.6556581
7. Clemens, J. Q., Mullins, C., Kusek, J. W., Kirkali, Z., Mayer, E. A., Rodriguez, L. V., . . . Group, M. R. N. S. (2014). The MAPP research network: a novel study of urologic chronic pelvic pain syndromes. *BMC Urol, 14*, 57. doi:10.1186/1471-2490-14-57
8. Clemmensen, L., Hastie, T., Witten, D., & Ersboll, B. (2012). Sparse Discriminant Analysis. In: Technometrics.
9. Craig, A. D. (2013). An interoceptive neuroanatomical perspective on feelings, energy, and effort. *Behav Brain Sci, 36*(6), 685-686; discussion 707-626. doi:10.1017/S0140525X13001489
10. Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage, 9*(2), 179-194. doi:10.1006/nimg.1998.0395
11. Dargie, E., Gilron, I., & Pukall, C. F. (2017). Self-Reported Neuropathic Pain Characteristics of Women With Provoked Vulvar Pain: A Preliminary Investigation. *J Sex Med, 14*(4), 577-591. doi:10.1016/j.jsxm.2017.02.008
12. Davies, G., Armstrong, N., Bis, J. C., Bressler, J., Chouraki, V., Giddaluru, S., . . . Scotland, G. (2015). Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). *Mol Psychiatry, 20*(2), 183-192. doi:10.1038/mp.2014.188

13. Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage, 53*(1), 1-15. doi:10.1016/j.neuroimage.2010.06.010
14. Eriksen, K. A., Simonsen, I., Maslov, S., & Sneppen, K. (2003). Modularity and extreme edges of the internet. *Phys Rev Lett, 90*(14), 148701. doi:10.1103/PhysRevLett.90.148701
15. Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage, 9*(2), 195-207. doi:10.1006/nimg.1998.0396
16. Fischl, B., Sereno, M. I., Tootell, R. B., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp, 8*(4), 272-284.
17. Friston, K. Statistical Parametric Mapping Software. In.
18. Gidskehaug, L., Anderssen, E., Flatberg, A., & Alsberg, B. K. (2007). A framework for significance analysis of gene expression data using dimension reduction methods. *BMC Bioinformatics, 8*, 346. doi:10.1186/1471-2105-8-346
19. Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images Are More than Pictures, They Are Data. *Radiology, 278*(2), 563-577. doi:10.1148/radiol.2015151169
20. Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proc Natl Acad Sci U S A, 99*(12), 7821-7826. doi:10.1073/pnas.122653799

21. Glahn, D. C., Thompson, P. M., & Blangero, J. (2007). Neuroimaging endophenotypes: strategies for finding genes influencing brain structure and function. *Hum Brain Mapp, 28*(6), 488-501. doi:10.1002/hbm.20401

22. Guimerà, R., & Nunes Amaral, L. A. (2005). Functional cartography of complex metabolic networks. *Nature, 433*(7028), 895-900. doi:10.1038/nature03288

23. Gupta, A., Mayer, E. A., Sanmiguel, C. P., Van Horn, J. D., Woodworth, D., Ellingson, B. M., . . . Labus, J. S. (2015). Patterns of brain structural connectivity differentiate normal weight from overweight subjects. *Neuroimage Clin, 7*, 506-517. doi:10.1016/j.nicl.2015.01.005

24. Gupta, A., Rapkin, A. J., Gill, Z., Kilpatrick, L., Fling, C., Stains, J., . . . Labus, J. S. (2015). Disease-related differences in resting-state networks: a comparison between localized provoked vulvodynia, irritable bowel syndrome, and healthy control subjects. *Pain, 156*(5), 809-819. doi:10.1097/01.j.pain.0000461289.65571.54

25. Hampson, J. P., Reed, B. D., Clauw, D. J., Bhavsar, R., Gracely, R. H., Haefner, H. K., & Harris, R. E. (2013). Augmented central pain processing in vulvodynia. *J Pain, 14*(6), 579-589. doi:10.1016/j.jpain.2013.01.767

26. Han, J. D., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., . . . Vidal, M. (2004). Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature, 430*(6995), 88-93. doi:10.1038/nature02555

27. Hartwell, L. H., Hopfield, J. J., Leibler, S., & Murray, A. W. (1999). From molecular to modular cell biology. *Nature, 402*(6761 Suppl), C47-52. doi:10.1038/35011540

28. Henzell, H., Berzins, K., & Langford, J. P. (2017). Provoked vestibulodynia: current perspectives. *Int J Womens Health, 9*, 631-642. doi:10.2147/IJWH.S113416

29. Hibar, D. P., Stein, J. L., Renteria, M. E., Arias-Vasquez, A., Desrivières, S., Jahanshad, N., . . . SYS. (2015). Common genetic variants influence human subcortical brain structures. *Nature, 520*(7546), 224-229. doi:10.1038/nature14101

30. Ho, K., Speier, W., El-Saden, S., & Arnold, C. (2017). Classifying Acute Ischemic Stroke Onset Time using Deep Imaging Features. In: AMIA Annu Symp Proc.

31. Holme, P., Huss, M., & Jeong, H. (2003). Subnetwork hierarchies of biochemical pathways. *Bioinformatics, 19*(4), 532-538.

32. Irimia, A., Goh, S. Y., Torgerson, C. M., Vespa, P., & Van Horn, J. D. (2014). Structural and connectomic neuroimaging for the personalized study of longitudinal alterations in cortical shape, thickness and connectivity after traumatic brain injury. *J Neurosurg Sci, 58*(3), 129-144.

33. Katrib, A., Hsu, W., Bui, A., & Xing, Y. (2016). "RADIOTRANSCRIPTOMICS": A synergy of imaging and transcriptomics in clinical assessment. *Quant Biol, 4*(1), 1-12. doi:10.1007/s40484-016-0061-6

34. Labus, J. S., Naliboff, B., Kilpatrick, L., Liu, C., Ashe-McNalley, C., dos Santos, I. R., . . . Mayer, E. A. (2016). Pain and Interoception Imaging Network (PAIN): A multimodal, multisite, brain-imaging repository for chronic somatic and visceral pain disorders. *Neuroimage, 124*(Pt B), 1232-1237. doi:10.1016/j.neuroimage.2015.04.018

35. Labus, J. S., Naliboff, B. D., Berman, S. M., Suyenobu, B., Vianna, E. P., Tillisch, K., & Mayer, E. A. (2009). Brain networks underlying perceptual habituation to repeated aversive visceral stimuli in patients with irritable bowel syndrome. *Neuroimage, 47*(3), 952-960. doi:10.1016/j.neuroimage.2009.05.078

36. Labus, J. S., Naliboff, B. N., Fallon, J., Berman, S. M., Suyenobu, B., Bueller, J. A., . . . Mayer, E. A. (2008). Sex differences in brain activity during aversive visceral stimulation and its expectation in patients with chronic abdominal pain: a network analysis. *Neuroimage, 41*(3), 1032-1043. doi:10.1016/j.neuroimage.2008.03.009
37. Labus, J. S., Van Horn, J. D., Gupta, A., Alaverdyan, M., Torgerson, C., Ashe-McNalley, C., . . . Mayer, E. A. (2015). Multivariate morphological brain signatures predict patients with chronic abdominal pain from healthy control subjects. *Pain, 156*(8), 1545-1554. doi:10.1097/j.pain.0000000000000196
38. Landis, J. R., Williams, D. A., Lucia, M. S., Clauw, D. J., Naliboff, B. D., Robinson, N. A., . . . Group, M. R. N. S. (2014). The MAPP research network: design, patient characterization and operations. *BMC Urol, 14*, 58. doi:10.1186/1471-2490-14-58
39. Lange, K., Papp, J. C., Sinsheimer, J. S., & Sobel, E. M. (2014). Next Generation Statistical Genetics: Modeling, Penalization, and Optimization in High-Dimensional Data. *Annu Rev Stat Appl, 1*(1), 279-300. doi:10.1146/annurev-statistics-022513-115638
40. Le Cao, K.-A., Rohart, F., Gonzalez, I., & Dejean, S. (2017). Omics Data Integration Project: Predict Method for (mint).(block).(s)pls(da) methods
41. p.p1 {margin: 0.0px 0.0px 0.0px 0.0px; font: 24.3px Helvetica; -webkit-text-stroke: #000000}
42. span.s1 {font-kerning: none}. In.
43. Le Cao, K. A., Boitard, S., & Besse, P. (2011). Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics, 12*, 253. doi:10.1186/1471-2105-12-253

44. Le Cao, K. A., Rossouw, D., Robert-Granie, C., & Besse, P. (2008). A sparse PLS for variable selection when integrating omics data. *Stat Appl Genet Mol Biol, 7*(1), Article 35. doi:10.2202/1544-6115.1390

45. Matthews, P. M., & Jezzard, P. (2004). Functional magnetic resonance imaging. *J Neurol Neurosurg Psychiatry, 75*(1), 6-12.

46. McIntosh, A. M., Hall, L. S., Zeng, Y., Adams, M. J., Gibson, J., Wigmore, E., . . . Hocking, L. J. (2016). Genetic and Environmental Risk for Chronic Pain and the Contribution of Risk Variants for Major Depressive Disorder: A Family-Based Mixed-Model Analysis. *PLoS Med, 13*(8), e1002090. doi:10.1371/journal.pmed.1002090

47. Mehmood, T., Liland, K., Snipen, L., & Saebo, S. (2012). A review of variable selection methods in Partial Least Squares Regression. In: Chemometrics and Intelligent Laboratory Systems.

48. Mizuta, M. (2012). Dimension Reduction Methods. In: Handbook of Computational Statistics.

49. Morgan, T. K., Allen-Brady, K. L., Monson, M. A., Leclair, C. M., Sharp, H. T., & Cannon-Albright, L. A. (2016). Familiality analysis of provoked vestibulodynia treated by vestibulectomy supports genetic predisposition. *Am J Obstet Gynecol, 214*(5), 609.e601-607. doi:10.1016/j.ajog.2015.11.019

50. Newman, M. E. (2004). Fast algorithm for detecting community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys, 69*(6 Pt 2), 066133. doi:10.1103/PhysRevE.69.066133

51. Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys, 69*(2 Pt 2), 026113. doi:10.1103/PhysRevE.69.026113

52. Papin, J. A., Reed, J. L., & Palsson, B. O. (2004). Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. *Trends Biochem Sci, 29*(12), 641-647. doi:10.1016/j.tibs.2004.10.001

53. Pukall, C. F., Strigo, I. A., Binik, Y. M., Amsel, R., Khalifé, S., & Bushnell, M. C. (2005). Neural correlates of painful genital touch in women with vulvar vestibulitis syndrome. *Pain, 115*(1-2), 118-127. doi:10.1016/j.pain.2005.02.020

54. Radicchi, F. (2018). Decoding communities in networks. *Phys Rev E, 97*(2-1), 022316. doi:10.1103/PhysRevE.97.022316

55. Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., & Parisi, D. (2004). Defining and identifying communities in networks. *Proc Natl Acad Sci U S A, 101*(9), 2658-2663. doi:10.1073/pnas.0400054101

56. Ravasz, E. (2009). Detecting hierarchical modularity in biological networks. *Methods Mol Biol, 541*, 145-160. doi:10.1007/978-1-59745-243-4_7

57. Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., & Barabási, A. L. (2002). Hierarchical organization of modularity in metabolic networks. *Science, 297*(5586), 1551-1555. doi:10.1126/science.1073374

58. Reed, B. D., Harlow, S. D., Plegue, M. A., & Sen, A. (2016). Remission, Relapse, and Persistence of Vulvodynia: A Longitudinal Population-Based Study. *J Womens Health (Larchmt), 25*(3), 276-283. doi:10.1089/jwh.2015.5397

59. Rives, A. W., & Galitski, T. (2003). Modular organization of cellular networks. *Proc Natl Acad Sci U S A, 100*(3), 1128-1133. doi:10.1073/pnas.0237338100
60. Rohart, F., Gautier, B., Singh, A., & Cao, K.-A. L. (2017). mixOmics: an R package for 'omics feature selection and
61. multiple data integration. In: bioRxiv.
62. Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika, 70*(1), 41-55. doi:10.2307/2335942
63. Rubin, D. B. (1980). Bias Reduction Using Mahalanobis-Metric Matching. *Biometrics, 36*(2), 293-298. doi:10.2307/2529981
64. Rubinov, M., & Sporns, O. (2017). Brain Connectivity Toolbox. In.
65. Schweinhardt, P., Kuchinad, A., Pukall, C. F., & Bushnell, M. C. (2008). Increased gray matter density in young women with chronic vulvar pain. *Pain, 140*(3), 411-419. doi:10.1016/j.pain.2008.09.014
66. Shen, C., Kim, J., & Wang, L. (2010). Scalable large-margin Mahalanobis distance metric learning. *IEEE Trans Neural Netw, 21*(9), 1524-1530. doi:10.1109/TNN.2010.2052630
67. Spielberg, J. (2015). Graph Theoretic General Linear Model. In.
68. Sporns, O., & Betzel, R. F. (2016). Modular Brain Networks. *Annu Rev Psychol, 67*, 613-640. doi:10.1146/annurev-psych-122414-033634
69. Thompson, P. M., Stein, J. L., Medland, S. E., Hibar, D. P., Vasquez, A. A., Renteria, M. E., . . . Alzheimer's Disease Neuroimaging Initiative, E. P. I. C., I. M. A.GEN Consortium, S.guenay Youth Study (SYS) Group. (2014). The ENIGMA Consortium: large-scale

collaborative analyses of neuroimaging and genetic data. *Brain Imaging Behav, 8*(2), 153-182. doi:10.1007/s11682-013-9269-5

70. Trampush, J. W., Yang, M. L., Yu, J., Knowles, E., Davies, G., Liewald, D. C., . . . Lencz, T. (2017). GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium. *Mol Psychiatry, 22*(3), 336-345. doi:10.1038/mp.2016.244
71. Tunbridge, E. M., Farrell, S. M., Harrison, P. J., & Mackay, C. E. (2013). Catechol-O-methyltransferase (COMT) influences the connectivity of the prefrontal cortex at rest. *Neuroimage, 68*, 49-54. doi:10.1016/j.neuroimage.2012.11.059
72. van den Heuvel, M. P., & Sporns, O. (2013). Network hubs in the human brain. *Trends Cogn Sci, 17*(12), 683-696. doi:10.1016/j.tics.2013.09.012
73. Van Horn, J. D., & Toga, A. W. (2014). Human neuroimaging as a "Big Data" science. *Brain Imaging Behav, 8*(2), 323-331. doi:10.1007/s11682-013-9255-y
74. Walitt, B., Ceko, M., Gracely, J. L., & Gracely, R. H. (2016). Neuroimaging of Central Sensitivity Syndromes: Key Insights from the Scientific Literature. *Curr Rheumatol Rev, 12*(1), 55-87.
75. Wesselmann, U., Bonham, A., & Foster, D. (2014). Vulvodynia: Current state of the biological science. *Pain, 155*(9), 1696-1701. doi:10.1016/j.pain.2014.05.010
76. Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect, 2*(3), 125-141. doi:10.1089/brain.2012.0073