

NIH Public Access

Author Manuscript

Neuroimage. Author manuscript; available in PMC 2009 October 1.

Published in final edited form as:

Neuroimage. 2008 October 1; 42(4): 1560-1568. doi:10.1016/j.neuroimage.2008.05.065.

Hybrid ICA-Bayesian Network approach reveals distinct effective connectivity differences in schizophrenia

D. Kim¹, J. Burge^{4,5}, T. Lane⁴, G. D Pearlson^{3,6}, K. A Kiehl⁷, and V. D. Calhoun^{1,2,3}

1 The Mind Research Network, Albuquerque, NM 87131

2 Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM 87131

3 Dept. of Psychiatry, Yale University, New Haven, CT 06520

4 Department of Computer Science, University of New Mexico, Albuquerque, NM 87131

5 Google Inc, Mountain View, CA 94043

6 Olin Neuropsychiatry Research Center, Hartford, CT 06106

7 Department of Psychiatry, University New Mexico, Albuquerque, NM 87131

Abstract

We utilized a discrete dynamic Bayesian network (dDBN) approach (Burge et al., 2007) to determine differences in brain regions between patients with schizophrenia and healthy controls on a measure of effective connectivity, termed the approximate conditional likelihood score (ACL) (Burge and Lane, 2005). The ACL score represents a class-discriminative measure of effective connectivity by measuring the relative likelihood of the correlation between brain regions in one group versus another. The algorithm is capable of finding non-linear relationships between brain regions because it uses discrete rather than continuous values and attempts to model temporal relationships with a first-order Markov and stationary assumption constraint (Papoulis, 1991). Since Bayesian networks are overly sensitive to noisy data, we introduced an independent component analysis (ICA) filtering approach that attempted to reduce the noise found in fMRI data by unmixing the raw datasets into a set of independent spatial component maps. Components that represented noise were removed and the remaining components reconstructed into the dimensions of the original fMRI datasets. We applied the dDBN algorithm to a group of 35 patients with schizophrenia and 35 matched healthy controls using an ICA filtered and unfiltered approach. We determined that filtering the data significantly improved the magnitude of the ACL score. Patients showed the greatest ACL scores in several regions, most markedly the cerebellar vermis and hemispheres. Our findings suggest that schizophrenia patients exhibit weaker connectivity than healthy controls in multiple regions, including bilateral temporal and frontal cortices, plus cerebellum during an auditory paradigm.

Correspondence: Dae II Kim, The Mind Research Network, 1101 Yale Boulevard NE, Albuquerque, NM 87131, PH: 646-675-9193, Email: dkim@mrn.org, Vince D. Calhoun, The Mind Research Network, 1101 Yale Boulevard NE, Albuquerque, NM 87131, PH: 860-545-7768, E-mail: vcalhoun@unm.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

INTRODUCTION

fMRI data analysis often focuses on the study of *functional specialization*, which is the inquiry of particular brain regions and their localized functions (Cohen and Tong, 2001). Currently, the most popular method of analyzing and understanding fMRI data is the general linear model (GLM) (Friston, 1994). In the context of fMRI analyses, the GLM represents an application of studies geared towards functional specialization. It effectively finds where task-related regions of the brain might be, but does not provide answers to the inter-relationships between these regions or in other words, their *effective connectivity* (Petersson et al., 1999a, b). The study of schizophrenia can benefit from this approach. Schizophrenia is a disorder with an unknown etiology and pathophysiology that is further complicated by its symptomatic heterogeneity (Pearlson, 2000). The substantial base of neuroimaging literature that deals with this disorder suggests that its associated cognitive deficits are not localized to a single brain region, but represent a more diffuse cognitive dysfunction (Ross et al., 2006). Thus, locating specific brain regions that behave abnormally in schizophrenia might be insufficient to understand the underlying basis.

To determine effective connectivity within fMRI images, tools such as structural equation modeling and dynamic causal modeling (Friston et al., 2003; McIntosh and Gonzalez-Lima, 1994) can be used to identify interactions between different brain regions. However, most of these implementations require the selection of a limited number of brain regions for analysis and are constrained by fundamental assumptions regarding the form of correlations between these regions of interest (Burge et al., 2007). In order to overcome these obstacles, a recently developed approach was created using discrete dynamic Bayesian networks (dDBN) applied to fMRI data (Burge and Lane, 2005; Burge et al., 2007). Bayesian approaches have been used with fMRI data to answer questions related to causality, connectivity, and optimal model selection (Friston, 2005; Friston et al., 2002). For our purposes, the dDBN approach was utilized to determine a locally optimal structure that discriminated patients with schizophrenia from healthy controls based on a measure of effective connectivity termed the approximate conditional likelihood score (ACL) (Burge and Lane, 2005).

Although the resulting structure from the dDBN analysis should accurately represent classdiscriminative structures, this process is often complicated by physiological/artifactual noise in fMRI data. The advantages of using ICA as a preprocessing step have been shown in previous fMRI studies to be a beneficial tool in reducing noise-related artifacts. These types of noise can be found using independent component analysis (ICA) (Calhoun et al., 2001b). ICA is a statistical and computational data-driven technique that attempts to discover hidden factors underlying sets of random variables, measurements or signals. It assumes that the fMRI data are linear mixtures of independent source signals and attempts to extract maximally independent signals and their mixing coefficients. The driving principle behind ICA is that these independent source signals represent coherent groupings of fMRI activations, often referred to as component maps, that are considered to be functionally relevant (Calhoun et al., 2001c). Another advantage of ICA is the ability to reconstruct the fMRI data from a subset of these component maps. Thus, by implementing ICA on our fMRI dataset, we were able to locate components that represented noise/artifacts, remove them from our analysis and reconstruct an fMRI dataset. This new dataset was analyzed using our dDBN algorithm and its results were compared with our original analysis that did not include the ICA filtering approach.

To determine effective connectivity differences between patients and controls, we analyzed a dataset consisting of 35 subjects with paranoid schizophrenia and 35 matched healthy controls, performing an auditory oddball task. Our hypothesis was that the ACL scores for healthy controls would be higher than that of patients, based on previous studies that have shown

aberrant connectivity and synchrony in schizophrenia patients (Friston and Frith, 1995; Garrity et al., 2007; Kim et al., 2007; Schlosser et al., 2003). We further hypothesized that the classdiscriminative families found from both the original analyses and the ICA filtered approach would represent significant ROIs that characterize schizophrenia during this auditory paradigm and that these regions would be consistent with previous studies that have implicated areas such as the bilateral temporal lobes, bilateral parietal lobes, thalamus, and cerebellum (Kiehl et al., 2005; Kim et al., 2007; Laurens et al., 2005).

METHODS

Participants

Thirty-five outpatients with schizophrenia (30 males) and thirty-five matched healthy controls (30 males) provided written informed consent and volunteered for the study at Hartford Hospital. Healthy controls were free from any Axis I disorder, as assessed with the SCID (Structured Clinical Interview for DSM-IV-TR screening device (First et al., 1995). Patients met criteria for paranoid schizophrenia (subtype 295.30) in the DSM-IV based on a structured clinical interview and review of the case file. All participants were right handed and there were no significant group differences in age (patients, 38 ± 11 years, range 18–59; controls, 37 ± 12 years, range 18-55). Intelligence estimates were determined from NART (National Adult Reading Test (Nelson, 1982)) scores; healthy controls ($n=17, 22 \pm 7$ points) scored higher than patients (n=26, 35 ± 15 points; t(41)=3.1323, p < .005). To determine the presence/absence of psychotic symptoms, the mean PANSS (Positive and Negative Syndrome Scale (Kay et al., (1987)) scores for patients were determined (n=28, 66 ± 19.6). Medication information was available for 24 patients; 13 were taking atypical antipsychotic medications, 4 were on typical antipsychotic medications, 2 were taking both atypical and typical medications, and three were on no medications at all. Four participants from the patient group were omitted from analysis, as they demonstrated extremely poor performance on the auditory oddball task (more than ten total incorrect responses in either targets or novels for both sessions). Two additional participants were omitted for excessive head motion (greater than one and a half voxel-length (6mm) in translation or (4 degrees) in rotation). All participants had normal hearing (assessed by self-report) and were able to carry out both tasks successfully during practice, and during the scanning session.

Task: Auditory Oddball

Two runs of auditory stimuli were presented to each participant by a computer stimulus presentation system (VAPP: http://nilab.psychiatry.ubc.ca/vapp/) via insert earphones embedded within 30 dB sound attenuating MR compatible headphones. The standard stimulus was a 500 Hz tone, the target stimulus a 1000 Hz tone, and the novel stimuli consisted of non-repeating random digital noises (e.g., tone sweeps, whistles). The target and novel stimuli each occurred with a probability of 10%. The non-target stimuli occurred with a probability of 80%. The stimulus duration was 200 ms with a random 1000, 1500, or 2000 ms inter-stimulus interval. All stimuli were presented at approximately 80 decibels and all participants reported that they could hear the stimuli and discriminate them from the background scanner noise. The headphones were designed to work with the head restraint system in order to minimize head movement.

An MRI compatible fiber-optic response device (Lightwave Medical, Vancouver, B.C.) was used to acquire behavioral responses. Prior to entry into the scanning room, each participant performed a practice block of 10 trials to ensure understanding of the instructions. The participants were instructed to respond as quickly and accurately as possible with their right index finger when they heard the target stimulus and not to respond to the non-target stimuli or the novel stimuli.

Imaging Parameters

Imaging was implemented on a 3T Siemens Allegra MR system. Conventional spin-echo T1 weighted sagittal localizers were acquired to view the positioning of the participant's head in the scanner and to graphically prescribe the functional image volumes. Functional image volumes were collected with a gradient-echo sequence (TR=1500 ms, TE=27 ms, FA=60°, FOV= 22×22 cm, 64×64 matrix, 4 kHz bandwidth, 3.44 by 3.44 mm in plane resolution, 4 mm slice thickness, 1 mm gap, 29 slices acquired axially) effectively covering the entire brain (145 mm) in an ascending manner. There were two runs of 255 time points each, prefaced by a 9 second rest block allowing T1 effects to stabilize.

Data Analysis: Pre-processing

Data was preprocessed using SPM2 (http://www.fil.ion.ucl.ac.uk/spm/). Images were realigned using INRIalign – a motion correction algorithm unbiased by local signal changes (Freire and Mangin, 2001; Freire et al., 2002). Data was spatially normalized (Ashburner and Friston, 1999) into the standard Montreal Neurological Institute space and spatially smoothed with a $10 \times 10 \times 10$ mm³ full width at half-maximum Gaussian kernel. The data (originally collected at $3.44 \times 3.44 \times 5$ mm³) was slightly sub-sampled to 3 mm³, resulting in $53 \times 63 \times 46$ voxels and a fifth-order infinite impulse response Butterworth low-pass filter of 0.25 Hz was applied to remove high-frequency noise.

Discrete Dynamic Bayesian Structure Search

Here we provide a general overview of the dDBN approach to fMRI data. See (Burge and Lane, 2005; Burge et al., 2007) for more details. BNs are graphical models that concisely represent joint probability dependencies among a set of variables (RVs): X_1, X_2, X_n . These variables are depicted as nodes within the Bayesian graphical model and links between these nodes represent dependencies between these variables. In our analyses, the variables that pertain to our nodes are all the neuroanatomical regions of interests (ROIs) determined by the automated anatomical labeling (AAL) atlas, a total of 116 distinct brain regions that include right/left associations (Tzourio-Mazoyer et al., 2002). Relationships between nodes are between *parent* and *child* nodes, where a single child node and a set of parents nodes $Pa(X_i)$, constitute a *family*. Families can be uniquely identified by their child node, so that the number of families resulting from the analyses is equal to the number of ROIs specified by the atlas. Since our analysis consisted of 116 ROIs, this resulted in 116 families where ROIs could appear as parent nodes more than once between families, but were uniquely constrained by the child nodes within our atlas. The maximum number of parent nodes is set *a priori* to *n* and for our analysis we chose n = 3.

In a discrete Bayesian approach (dBN), the values that RVs can represent are limited to discrete values. To perform this discretization, we averaged the voxels composing with each ROI together to represent the ROI for that particular time point. These values were quantized to 2-bit values, corresponding to one of four possible states (Very low, Low, High, and Very high) (Burge et al., 2007). Thus, a subject's dataset would be represented as a discretized set of 116 ROI time courses. The advantages of a discrete approach allows for a conditional probability table to capture all the possible relationships found between the parent and child nodes. The disadvantage is seen in the loss of precision due to the discretization of our RVs. Thus a sacrifice in specificity is made to allow for the expression of complex non-linear RV dynamics.

A dDBN is a specialization of a dBN that models temporal processes and explicitly models time points in an fMRI dataset. Although it would be ideal to model every time point in this dataset, the task quickly becomes computationally infeasible. We make the Markov order 1 and assume stationarity to deal with this intractability (Papoulis, 1991). The resulting graphical topology of a dDBN then is a topology that consists of two columns of ROIs where each column

represents a time frame (t and t+1). The time frames are averaged across the entire time series and the columns become a succinct representation of temporally consecutive RV dynamics. Parent nodes occupy time column t while child nodes occupy time t + 1 with the constraint that links point only forward in time. In Figure 1, we show a simplified hypothetical representation of several 1-bit discrete ROIs, their graphical topology, and an example of their associated conditional probability tables.

When the structure of a dDBN (i.e., the set of links in a dDBN) is unknown, it must be searched for. Determining an optimal dynamic Bayesian structure is an NP-hard (Chickering, 1994) problem and a locally optimal structure is often determined instead. The initial structure is frequently determined by utilizing a greedy search algorithm that iteratively pairs a single parent with a child node. When the maximum number of parents have been added (i.e., for our analysis this was set to 3), the final structure is determined. To score these structures, we utilize the ACL score developed by Burge et. al. The ACL score represents a quantitative measure of the class-discriminative qualities found between two classes of data (e.g., patients and controls). For example, a strong dependency found between two variables within the patient group would have a high ACL score only if those same variables had a weak dependency in the healthy control group and vice versa. For further details, a summary of the dDBN algorithm is listed in Table 1. Formulas related to this particular scoring metric as well as determining the optimal graphical topologies are listed in Table 2.

Filtering with Independent Component Analysis

The ICA was performed using the GIFT toolbox, version 1.3c (http://icatb.sourceforge.net). For a full discussion of the methods implemented in ICA we refer the reader to these articles (Calhoun and Adali, 2006; Calhoun et al., 2003; Calhoun et al., 2001a; Calhoun et al., 2001b). The goal of ICA is to split the fMRI datasets into a number of independent components that represent a linear combination of significant signal sources. The optimal number that ICA used to split the fMRI datasets into a final set of spatially independent components was first determined using a modified minimum description length (MDL) algorithm (Li et al., 2007), which was found to be 29. Since group ICA requires that all subjects are analyzed at once, a method for data compression using principal component analysis (PCA) allowed the datasets to be loaded into memory at one time. This allowed us to concatenate the subjects into a large matrix where a group spatial ICA is then performed using the infomax algorithm (Bell and Sejnowski, 1995).

The key point here for our analysis is that ICA is capable of finding noise and movement related signal sources which are known to show characteristics of independence (McKeown et al., 1998). To determine which components were related to noise we correlated each ICA component spatial map with prior probabilistic maps of grey matter, white matter, and cerebral spinal fluid (CSF) within a standardized brain space provided by MNI templates in SPM5. Those that had high correlation values (R²) with these maps were considered noise-related components and removed from the analysis in the same process used in a previous study (Stevens et al., 2006). The ability of ICA to reconstruct these datasets allowed the creation of a new fMRI dataset that contained none of the noise components that we removed. These datasets were then analyzed again using the dDBN algorithm and its results were compared with our first analysis with the unmodified datasets.

Score Validation and Robustness

Fourier Bootstrapping Score Validation—It is important to consider that the correlations between our random variables (i.e., parent to child) might be due to chance. To remedy this, we utilize a Fourier bootstrapping method (Prichard and Theiler, 1994) to create a surrogate ROI time series that contains higher order correlations found in the original fMRI data (first

and second order moments-nonlinear correlations are not preserved). This validation method was applied as a measure of confidence in the previous paper by Burge et. al. The surrogate time series represents a non-parametric null hypothesis against which to test the likelihood that a given ACL score occurred by chance. In other words, if the ACL score is significantly diminished over multiple datasets, then the score on the actual data is most likely not due to chance. The parent and child nodes of interest from the surrogate datasets are the high scoring ACL families that were found from the original datasets. Each dDBN family's score (non-filtered and ICA filtered) on the surrogate data is inferred not by individual parent-child relationships, but the entire family as a whole. Their representative z-score is the number of standard deviations of their ACL score on the true data versus the surrogate datasets. For our method, 35 surrogate datasets were generated by our Fourier bootstrapping procedure to calculate a set of z-scores for each family and repeated over 50 iterations. A family was considered to be significantly relevant to our analyses if its z-score was greater than 2. Families that did not meet this threshold were removed and not considered for further discussion.

Structural Robustness—Generally, a method can be considered robust if by applying a similar dataset, the results are also similar. In our approach, a structure can be considered robust if after removing a number of datasets and running another structure search, the same parent-child relationships continue to appear. To determine robustness, we chose a sample size of 5 subject's datasets to remove from the original 35 datasets for both groups and repeated this process six more times, removing an entirely different subset each time. This process is similar to k-fold cross validations that utilize a leave-one-out method of removing a dataset to testing a given hypothesis. Here, the robustness value was calculated based off the parent-child links found from the original results. If that same parent-child relationship was seen throughout all seven reductions, the link was then considered 100% robust. If that same link was seen only in three of the seven reductions, the link would then be considered 43% robust.

RESULTS

High scoring families that did not pass a confidence measure of z = 2 (i.e., p < .05) from the Fourier bootstrapping procedure were considered insignificant and removed for further discussion. The top five families with the highest ACL scores that passed our confidence measure were considered to be representative of significant differences between patients and healthy controls. An example of the two highest scoring families for both groups are shown in Figure 2, where t represents the parent nodes and t+1 refer specifically to child nodes. The parent that is listed first among the family represents the ROI that contributed most followed by the second and third respectively (see Table 3). Also of interest was the number of times any one specific ROI appeared as a parent among the families that were found from each analysis. For example, if a particular parent ROI appeared multiple times among the top five highest scoring families, we believed that ROI played a substantial role in the characterization of differences between parent ROIs and child ROIs were found to be stronger going forward in time, which suggests causal relationships, but to determine explicit causality would require further validation measures that we did not perform.

The ICA filtering approach produced a total of 29 components where 14 out of those 29 components were considered noise/artifacts and removed before back-reconstruction of the fMRI data. Five of the spatial components maps that were removed are shown in Figure 3, to show an example of the various patterns of noise activity found using ICA.

Patients – ACL families and parents

The unfiltered approach found the children of the five highest scoring families to be regions associated with the cerebellar vermis and cerebellar hemispheres. The top scoring family was composed entirely of regions within the vermis while the second family consisted mostly of regions within the cerebellum. The ROI that appeared most as a parent for this analysis was the vermis_10, the anterior region of the vermis followed by the left cerebellum_10, located within the anterior end of the cerebellum. The top ten ROIs that appeared most often were mostly regions found within the cerebellum and more specifically areas such as the vermis.

The ICA filtered approach found more variation in the children of the top scoring families than the non-filtered approach. The left superior temporal gyrus was the child for the highest scoring family followed by the left cerebellum_9 gyrus. The ROI that appeared most often for this analysis was the vermis_7, in the mid-region of the vermis followed by the right cerebellum_10, similar to the second-place ROI found in the unfiltered approach, but on the opposite side of the brain. Bilateral amygdala was identified among this group of highly occurring parents, which was not the case using the unfiltered approach.

Controls – ACL families and parents

The unfiltered approach found the children of the top two highest scoring families to be associated with bilateral regions within the rectus while the other three high scoring families were also associated with various frontal and temporal regions. The ROI that appeared most consistently as a parent was left Heschl's gyrus, followed by the left putamen.

The ICA filtered approach found the children of the top five families to be similar to the unfiltered approach with the left medial frontal orbital region to be the only differing child. The parents of the common families between the two approaches were also identical, except for the family that consisted of the right cerebellum_crus1 which replaced the left superior frontal gyrus with the left supramarginal gyrus. The ROI that appeared most often as a parent was the left pre-central gyrus, followed by the right pallidum. The highest appearing parent ROIs for this analysis differed significantly from the unfiltered approach as well as their respective number of appearances, which was significantly lower.

DISCUSSION

We hypothesized that the ACL score would be significantly lower for schizophrenia patients than controls based on previous research showing aberrant connectivity and synchrony in this population. A high ACL score implies that the relationship found in one class of data (patients) is due to the weakness of that same relationship in another class of data (controls) and vice versa. Our results are consistent with our hypothesis, where we found large differences in the ACL scores for all between-group comparisons of their families. Healthy control ACL scores exceeded those of patients by a factor of four. The large score difference between the two groups indicates that healthy controls exhibit this particular measure of effective connectivity more consistently than do schizophrenia patients.

The specificity of this aberrant connectivity in schizophrenia might be elucidated further by looking at the high-scoring families and parent appearances for healthy controls. Many of the high-scoring families consisted of child ROIs with connections to bilateral frontal and temporal lobes, which have been shown using different connectivity measures to be abnormal in schizophrenia patients (Calhoun et al., 2004; Garrity et al., 2007; Winterer et al., 2003). After ICA filtering, the second highest ACL score was found between the cerebellum and dorsal lateral prefrontal cortex, which suggests that this connection might be impaired in patients with schizophrenia. An interesting finding that emerged from both approaches (unfiltered and ICA

filtered), was that the top three families were almost identical and consisted of regions mainly in the gyrus rectus and cerebellum. The gyrus rectus is located in the medial orbital gyrus and is considered to be an extension of the anterior cingulate onto the frontal cortex (Morecraft et al., 1992). To our knowledge, few neuroimaging studies have associated schizophrenia with this region with the bulk of studies focusing on volumetric differences rather than functional ones (Nakamura et al., 2008; Spalletta et al., 2003). A study involving elderly depressed patients found significant volumetric differences in this region, possibly suggesting that it might be associated with the regulation of emotional affect. This might then be possibly related to the negative symptoms of schizophrenia, though further studies will be needed to ascertain this relationship. The cerebellum is well known from structural and functional studies to be affected in schizophrenia patients. As well, it plays a major role in the cortico-cerebellarthalamus-cortical circuit (CCTCC), hypothesized (Andreasen et al., 1998; Martin and Albers, 1995) to play an important role in schizophrenia. The consistency of these three families across both approaches implies that their relationship is not indicative of noise/artifact signals and could represent a significant between group difference.

The class-discriminative families found in patients with schizophrenia show a large number of regions within the cerebellum. Both approaches (filtered ICA and unfiltered), found that the parent ROIs which appeared the most times were regions within the vermis and cerebellum, suggesting that they play a disproportionate role in their relationships to other regions. The unfiltered approach found almost all of its high scoring families within regions of the cerebellum including the cerebellar vermis, while the ICA filtered approach resulted in the appearance of a number of different regions such as the left superior temporal lobe, left precentral gyrus, and the right superior occipital lobe in its resulting families. This is in marked contrast to the high-scoring families identified in the control group, where the filtering approach produced families that were almost identical to the non-filtered approach. This difference might be due to the possibility that healthy controls are less sensitive to an ICA filtering procedure due to the strength and consistency of their cognitive networks. However, in the context of parent ROIs that appear most often, the cerebellum and vermis regions represent the top two parent ROIs for both methods. Previous studies have shown the vermis to be implicated in schizophrenia from structural and functional MRI analyses, where it was found to be significantly reduced in patients (Okugawa et al., 2003). Functional studies have also implicated the cerebellum as a dysfunctional area in schizophrenia (Andreasen et al., 1998; Ende et al., 2005; Kim et al., 2007), though the vermis is not often discussed due to the frequent use by researchers of the Talairach atlas, which does not distinguish the vermis as a distinct brain region (Talairach and Tournoux, 1988).

Noise in fMRI datasets are considered to be an unfortunate, but consistent marker that can confound the results of any analysis. The sensitivity of a dDBN analysis motivated us to use ICA as a method to remove noise-related artifacts. Our results suggest that the ICA filtering approach seems effective in dealing with patient populations such as schizophrenia, where variability within BOLD activity tends to be greater than healthy controls (Kim et al., 2007). The class-discriminative measures for patients with schizophrenia showed a significantly lower score than controls and filtering seemed to have improved these scores significantly, but not for controls. This discrepancy might be due to the possibility that noise in populations that are highly variable (i.e., schizophrenia) have a more profound effect on the underlying BOLD activation of interest if these underlying correlations are relatively weak. Thus, this might account for the increased variability of top-scoring child ROIs found after ICA filtering for patients with schizophrenia.

Though the dDBN algorithm is capable of determining non-linear relationships between these ROIs, it does so at the cost of precision due to the quantization of the fMRI datasets. The sheer complexity of fMRI data does not lend itself well to simple models and thus this tradeoff of

precision for tractability presents itself as a possible confound in our analysis. Also, the temporal relationships that the dDBN algorithm attempts to model is limited to two consecutive time points and can miss fundamental relationships that might exist in larger time windows, even though these consecutive relationships are averaged across all time points. Again, this limitation is due to the computational infeasibility of modeling a probabilistic framework that can account for every time point within the fMRI dataset and a compromise needs to be made between precision and tractability. Another final limitation that needs to be mentioned is the accuracy of ICA in filtering out noise/artifactual signals. The removal of Heschl's gyrus and putamen as the two highest appearing parents after ICA filtering suggests that the exclusion of certain components believed to be related to noise might concomitantly diminish important correlations that were part of that component. It is difficult to assess fully the accuracy of this filtering approach, but its use here was mainly a vehicle to determine which ROIs were consistently high-scoring structures and their particular effects on measures of effective connectivity.

To further elucidate effective connectivity relationships in schizophrenia, future studies might focus on implementing the dDBN algorithms on larger datasets or across multiple sites. A dDBN analysis of subjects performing multiple tasks might also determine common families that exist regardless of the experimental task at hand. Ultimately, the dDBN method is only one of many Bayesian approaches to fMRI data. Algorithms that can better account for the limitations mentioned above are needed in order to determine a more accurate characterization of effective connectivity in patients with schizophrenia.

CONCLUSION

The strength of the analysis methods used here lies in the class-discriminative nature of our nonlinear effective-connectivity measure, the ACL score, and our ability to determine these relationships for all ROIs within the AAL atlas. We introduced an ICA filtering approach to remove independent noise signals that were assumed to be irrelevant to our datasets and analyzed the resulting filtered datasets to determine the consistency of our results. We found that healthy controls have generally stronger measures of effective connectivity, as assessed using the ACL score, compared to patients with paranoid schizophrenia, especially in bilateral temporal and frontal lobes and that filtering using ICA improved the ACL score for patients. We recommend the use of ICA in filtering noise components from fMRI datasets and suggest that the selection process should be conservative so that meaningful activation is not lost. Furthermore, we believe that ICA filtering plays a more significant role when trying to determine higher-level correlations (i.e, effective connectivity) which are much more sensitive to noisy datasets. The dDBN algorithm in conjunction with ICA filtering is a powerful tool to determine non-linear relationships between brain regions. In our study, we attempted to dilineate meaningful structures that represent a class-discriminative measure of effective connectivity in an attempt to better characterize the nature of schizophrenia.

Acknowledgements

This work was supported by NIH grant 1R01EB #000840. John Burge and Terran Lane's contribution to this work was supported by NIMH grant number 1R01MH076282-01 as part of the NSF/NIH Collaborative Research in Computational Neuroscience Program.

References

Andreasen NC, Paradiso S, O'Leary DS. "Cognitive dysmetria" as an integrative theory of schizophrenia: a dysfunction in cortical-subcortical-cerebellar circuitry? Schizophr Bull 1998;24:203–218. [PubMed: 9613621]

Kim et al.

- Ashburner J, Friston KJ. Nonlinear spatial normalization using basis functions. Hum Brain Mapp 1999;7:254–266. [PubMed: 10408769]
- Bell AJ, Sejnowski TJ. An information maximisation approach to blind separation and blind deconvolution. Neural Comput 1995;7:1129–1159. [PubMed: 7584893]
- Burge, J.; Lane, T. Learning Class-Discriminative Dynamic Bayesian Networks. Proceedings of the International Conference on Machine Learning; Bonn, Germany. 2005. p. 97-104.
- Burge J, Lane T, Link H, Qiu S, Clark VP. Discrete dynamic Bayesian network analysis of fMRI data. Hum Brain Mapp. 2007
- Calhoun VD, Adali T. 'Unmixing' Functional Magnetic Resonance Imaging with Independent Component Analysis. IEEE Eng in Medicine and Biology 2006;25:79–90.
- Calhoun VD, Adali T, Hansen JC, Larsen J, Pekar JJ. ICA of fMRI: An Overview. Proc Int Conf on ICA and BSS. 2003
- Calhoun VD, Adali T, Pearlson GD. Independent Components Analysis Applied To fMRI Data: A Natural Model And Order Selection. Proc NSIP. 2001a
- Calhoun VD, Adali T, Pearlson GD, Pekar JJ. Group ICA of Functional MRI Data: Separability, Stationarity, and Inference. Proc Int Conf on ICA and BSS 2001b:155–160.
- Calhoun VD, Adali T, Pearlson GD, Pekar JJ. Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. Hum Brain Map 2001c;13:43–53.
- Calhoun VD, Kiehl KA, Liddle PF, Pearlson GD. Aberrant localization of synchronous hemodynamic activity in auditory cortex reliably characterizes schizophrenia. Biol Psychiatry 2004;55:842–849. [PubMed: 15050866]
- Chickering, G.; Heckerman. Learning Bayesian Networks is NP-Hard. Technical Report MSR-TR-94– 17; Microsoft. 1994.
- Cohen JD, Tong F. Neuroscience. The face of controversy. Science 2001;293:2405–2407. [PubMed: 11577224]
- Ende G, Hubrich P, Walter S, Weber-Fahr W, Kammerer N, Braus DF, Henn FA. Further evidence for altered cerebellar neuronal integrity in schizophrenia. Am J Psychiatry 2005;162:790–792. [PubMed: 15800155]
- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. Structured Clinical interview for DSM-IV axis I disorders-patient edition (SCID-I/P, Version 2.0). Biometrics Research Department; New York State Psychiatric Institute New York: 1995.
- Freire L, Mangin JF. Motion correction algorithms may create spurious brain activations in the absence of subject motion. Neuroimage 2001;14:709–722. [PubMed: 11506543]
- Freire L, Roche A, Mangin JF. What is the best similarity measure for motion correction in fMRI time series? IEEE Trans Med Imaging 2002;21:470–484. [PubMed: 12071618]
- Friston K. Statistical Parametric Mapping. 1994
- Friston KJ. Models of brain function in neuroimaging. Annu Rev Psychol 2005;56:57–87. [PubMed: 15709929]
- Friston KJ, Frith CD. Schizophrenia: a disconnection syndrome? Clin Neurosci 1995;3:89–97. [PubMed: 7583624]
- Friston KJ, Glaser DE, Henson RN, Kiebel S, Phillips C, Ashburner J. Classical and Bayesian inference in neuroimaging: applications. Neuroimage 2002;16:484–512. [PubMed: 12030833]
- Friston KJ, Harrison L, Penny W. Dynamic causal modelling. Neuroimage 2003;19:1273–1302. [PubMed: 12948688]
- Garrity AG, Pearlson GD, McKiernan K, Lloyd D, Kiehl KA, Calhoun VD. Aberrant "default mode" functional connectivity in schizophrenia. Am J Psychiatry 2007;164:450–457. [PubMed: 17329470]
- Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 1987;13:261–276. [PubMed: 3616518]
- Kiehl KA, Stevens MC, Celone K, Kurtz M, Krystal JH. Abnormal hemodynamics in schizophrenia during an auditory oddball task. Biological Psychiatry 2005;57:1029–1040. [PubMed: 15860344]
- Kim D, Pearlson GD, Kiehl KA, Bedrick E, Demirci O, Calhoun VD. A method for multi-group interparticipant correlation: Abnormal synchrony in patients with schizophrenia during auditory target detection. Neuroimage. 2007

Kim et al.

- Laurens KR, Kiehl KA, Ngan ET, Liddle PF. Attention orienting dysfunction during salient novel stimulus processing in schizophrenia. Schizophr Res 2005;75:159–171. [PubMed: 15885507]
- Li YO, Adali T, Calhoun VD. Estimating the number of independent components for functional magnetic resonance imaging data. Hum Brain Mapp 2007;28:1251–1266. [PubMed: 17274023]
- Martin P, Albers M. Cerebellum and schizophrenia: a selective review. Schizophr Bull 1995;21:241–250. [PubMed: 7631171]
- McIntosh AR, Gonzalez-Lima F. Structural Equation Modeling and Its Application to Network Analysis in Functional Brain Imaging. Hum Brain Map 1994;2:2–22.
- McKeown MJ, Jung TP, Makeig S, Brown G, Kindermann SS, Lee TW, Sejnowski TJ. Spatially independent activity patterns in functional MRI data during the stroop color-naming task. Proc Natl Acad Sci 1998;95:803–810. [PubMed: 9448244]
- Morecraft RJ, Geula C, Mesulam MM. Cytoarchitecture and neural afferents of orbitofrontal cortex in the brain of the monkey. J Comp Neurol 1992;323:341–358. [PubMed: 1460107]
- Nakamura M, Nestor PG, Levitt JJ, Cohen AS, Kawashima T, Shenton ME, McCarley RW. Orbitofrontal volume deficit in schizophrenia and thought disorder. Brain 2008;131:180–195. [PubMed: 18056163]
- Nelson, HE. National Adult Reading Test (NART). NFER-Nelson; Windsor: 1982.
- Okugawa G, Sedvall GC, Agartz I. Smaller cerebellar vermis but not hemisphere volumes in patients with chronic schizophrenia. Am J Psychiatry 2003;160:1614–1617. [PubMed: 12944335]
- Papoulis, A. Probability, random variables, and stochastic processes. McGraw-Hill; New York: 1991.
- Pearlson GD. Neurobiology of schizophrenia. Ann Neurol 2000;48:556–566. [PubMed: 11026439]
- Petersson KM, Nichols TE, Poline JB, Holmes AP. Statistical limitations in functional neuroimaging. I. Non-inferential methods and statistical models. Philos Trans R Soc Lond B Biol Sci 1999a;354:1239– 1260. [PubMed: 10466149]
- Petersson KM, Nichols TE, Poline JB, Holmes AP. Statistical limitations in functional neuroimaging. II. Signal detection and statistical inference. Philos Trans R Soc Lond B Biol Sci 1999b;354:1261–1281. [PubMed: 10466150]
- Prichard D, Theiler J. Generating surrogate data for time series with several simultaneously measured variables. Phys Rev Lett 1994;73:951–954. [PubMed: 10057582]
- Ross CA, Margolis RL, Reading SA, Pletnikov M, Coyle JT. Neurobiology of schizophrenia. Neuron 2006;52:139–153. [PubMed: 17015232]
- Schlosser R, Gesierich T, Kaufmann B, Vucurevic G, Hunsche S, Gawehn J, Stoeter P. Altered effective connectivity during working memory performance in schizophrenia: a study with fMRI and structural equation modeling. Neuroimage 2003;19:751–763. [PubMed: 12880804]
- Spalletta G, Tomaiuolo F, Marino V, Bonaviri G, Trequattrini A, Caltagirone C. Chronic schizophrenia as a brain misconnection syndrome: a white matter voxel-based morphometry study. Schizophr Res 2003;64:15–23. [PubMed: 14511797]
- Stevens M, Kiehl KA, Pearlson GD, Calhoun VD. Functional neural circuits for mental timekeeping. Hum Brain Map. 2006
- Talairach, J.; Tournoux, P. A co-planar sterotaxic atlas of a human brain. Thieme; Stuttgart: 1988.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 2002;15:273–289. [PubMed: 11771995]
- Winterer G, Coppola R, Egan MF, Goldberg TE, Weinberger DR. Functional and effective frontotemporal connectivity and genetic risk for schizophrenia. Biol Psychiatry 2003;54:1181–1192. [PubMed: 14643085]



Figure 1.

An example representation of a single dDBN family. For simplicity, the ROIs are considered to be discretized into binary values of high and low. The second column shows an example of what a graphical topology for a single family could look like where three parent nodes (left thalamus, right amygdala, left insula) have the strongest correlation with the unique child node (left rectus). The third column represents the conditional probability tables for these three parent-child links.



Figure 2.

The figure above represents the two top scoring families for patients and healthy controls with or without ICA for a total of eight families. ROIs within the same color represent nodes within the same family. The ICA suffix within the legend refers to an ICA filtered approach and the number following it refers to whether it was the highest or second highest scoring family. It is important to note that the top two families for healthy controls without filtering and the top family for healthy controls with filtering had the same identical parents. (Pat = Patients, Ctrl = Healthy Controls, #1 = Highest scoring, #2 = Second highest scoring, ICA = ICA filtered approach)



Figure 3.

An overlay of 5 out of 14 components that were selected to represent the various types of noise found within fMRI data using ICA. For presentation purposes, only five components were used due to overlapping spatial patterns often found throughout these noise components. Each color refers to a specific component that was removed for ICA filtering.

Summary of the dDBN algorithm using an ACL scoring metric.

1.	Initialize structure B_s with no links connecting nodes in column t to nodes in column $t+1$
2.	Repeat <i>numParents</i> times
3.	Repeat for every node, X_i in column $t+1$
4.	Repeat for every node, X_k , in column t
5.	Add X_k as a parent of X_i , i.e., insert X_k into $Pa(X_i)$
6.	Calculate the ACL score for X_i
7.	Remove X_k from $Pa(X_i)$
8.	End 4.
9.	Permanently add the parent X_k to node X_n with highest ACL score
10.	End 3.
11.	End 2.
12.	Empty the parent set of all but <i>numBestToKeep</i> of the highest scoring <i>families</i>
13.	Return high scoring and high confidence structures

numParents refer to the chosen number of maximum parent nodes any particular child node can have. For our analysis this was set at 3. *numBestToKeep* is the maximum number of families that are allowed to keep their particular structure. For our analysis, this was set to allow for all 116 ROIs to have their own families.

=

Table 2Equations for the dDBN algorithms.

$$P(D \square B) = \prod_{m=1}^{\square D \square} P(X_1 = Val_D(X_1, m), \dots, X_N = Val_D(X_N, m))$$

$$\prod_{m=1}^{\square D \square} \prod_{i=1}^{N} P(X_i = Val_D(X_i, m) \square Pa_{i,1} = Val_D(Pa_{i,1}, m), \dots, Pa_{i,p} = Val_D(Pa_{i,p}, m))$$

$$P(X_i = Val_D(X_i, m) \square Pa_{i,1} = Val_D(Pa_{i,1}, m), \dots, Pa_{i,p} = Val_D(Pa_{i,p}, m))$$

Formula 2.1 is the data likelihood equation, the probability that a specific Bayesian network *B* generated an observed data point. Formula 2.1 requires that the parameters for the dDBN have already been computed with maximum likelihood estimates. $P(X_i)$ represents the probability that a particular child ROI (X_i) takes on the mth data-point (where a datapoint represents an entire time-series) within dataset *D* (the total number of datapoints) given its parent ROI, $Pa_{i,p}$ where *p* refers to the maximum number of parents. In other words, the data likelihood can be considered a multiplicative series of probabilities that represent the relationship between a parent and a child for a given Bayesian network (BN).

$$ACL(M \Box D) = \frac{\prod_{d_1} \Box_{D_1} P(X_{d_1} \Box B_1)}{\prod_{d_2} \Box_{D_2} P(X_{d_2} \Box B_1)} \cdot \frac{\prod_{d_2} \Box_{D_2} P(X_{d_2} \Box B_2)}{\prod_{d_1} \Box_{D_1} P(X_{d_1} \Box B_2)}$$

The ACL score is a metric that represents a measure of difference between two Bayesian networks (BN) B_1 and B_2 , which in our case refers to the BNs for healthy controls and patients with schizophrenia, respectively. *M* refers to the model that contains both BNs. *X* refers to the entire set of random variables for subject *d* within the larger set of datasets D_1 and D_2 that refer to healthy controls and patients with schizophrenia respectively. It is important to note that $P(Xd_1|B_1)$ for example is equivalent to P(D|B) from formula 2.1 above.

<u> </u>
σ
+
~
\leq
=
-
÷
U.
$\mathbf{\Sigma}$
2
E
5
0
-
~
\leq
B
5
Ę
10
S.
9
0
_

NIH-PA Author Manusci

NIH-PA Author Manuscript

Kim	et	al.

Child/ACL Score	Patients - NO ICA FILTER Parents	Robust	z/p-values	Child/ACL Score	Patients - ICA FILTER Parents	Robust	z/p-values
Vermis_9 22.103	Vermis_3 Vermis_8 Vermis_0	71% 57% 71%	3.21/.0007	Temporal_Sup_L 32.867	Cerebellum_10_L Cerebellum_3_L Himoreannue_1	71% 86% 86%	3.87/10 ⁻⁵
Cerebellum_4_5_L 19.606	Cerebellum Cerebellum10_L	43% 100%	3.26/.0006	Cerebellum_9_L 31.121	Cerebellum_10_L Vermis_8	71%	4.43/10 ⁻⁶
Cerebellum_3_L 18.749	Cerebellum_Crus2_R	100% 43% 23%	2.44/.0073	Precentral_L 30.182	Cerebellum_4.5_R Rolandic_Oper_R Vi.	100% 29% 57%	2.60/.0047
Vermis_8 18.648	Cerebellum_9_R V	43% 43%	3.47/.0003	Vermis_10 28.099	Supramarginal_L Vermis_4_5	51% 57%	3.08/.0010
Vermis_1_2 17.916	Temporal_Pole_Sup_R Cerebellum_9_L Frontal_Sup_Oth_L	0% 14% 14%	$5.18/10^{-7}$	Occipital_Sup_R 26.107	Cerebellum_10_L Vermis_10 Vadate_R	80% 14% 86% 43%	5.86/10 ⁻⁹
Child/ACL Score Rectus_L	Parents Parents Rectus_R	Robust 100%	z/p-values 7.33/10 ⁻¹³	Child/ACL Score Rectus_L	Parents Parents Rectus R	Robust 100%	z/p-values 6.71/10 ⁻¹²
152.534 Rectus_R 115.469	Kectus_L Rectus_L Rectus_L	100% 100% 100%	6.49/10 ⁻¹¹	Cerebellum_Crus1_R 99.511	Kectus_L Supramarginal_L Frontal_Mid_R	100% 29% 71%	3.55/.0002
Cerebellum_Crus1_R 114.821	Frontal_Sup_L Temporal_Sup_L	86% 100%	2.06/.0197	Rectus_R 115.469	1emporal_bup_L Rectus_R Rectus_L	100% 100% 100%	4.21/10 ⁻⁵
Frontal_Sup_Orb_L 111.264	Frontal_Inf_Orb_L Rectus_L	100% 71% 100%	2.868/.0021	Frontal_Mid_Orb2_L 86.422	Rectus_L Frontal_Mid_Orb2_L	0% 100%	3.30/.0005
Cerebellum_Crus2_R 94.722	rronat_sup_Urb_L Frontal_Mid_R Temporal_Sup_L Frontal_Sup_L	100% 57% 100% 100%	3.69/.0001	Cerebellum_Crus2_R 85.999	Frontal_Mid_R Temporal_Sup_L Frontal_Sup_L	100% 29% 43%	4.42/10 ⁻⁶
		:					

The top five highest scoring families are listed here for patients and healthy controls with and without ICA filtering. Child nodes are represented within the first column followed by their associated parent nodes. Robustness measures indicate the consistency in which these parent nodes appeared across seven iterations and the z/p-values refer to the Fourier bootstrapping results.

Table 4

Top ten parent ROIs by appearances

Page	18
------	----

Patients - No ICA Filter	Appearances	Patients - ICA Filter	Appearances
cerebellar vermis_10	33	vermis_7	18
cerebellum_10_1	32	cerebelum_10_r	16
rectal gyrus_r	32	cerebelum_3_r	14
cerebellar vermis_1_2	19	cerebellar vermis_10	13
cerebellar vermis_8	16	cerebellum_10_1	11
cerebellum_9_r	12	amygdala_1	10
cerebellum_crus2_1	12	caudate_r	10
rectal gyrus_1	11	amygdala_r	9
cerebellum_10_r	10	cerebellar vermis_6	9
cerebellum_8_1	9	angular gyrus_1	8
Controls - No ICA Filter	Appearances	Controls - ICA Filter	Appearances
Heschl_l	47	precentral_1	15
putamen_1	21	pallidum_r	12
precentral_1	19	occipital_sup_1	11
cingulum_mid_r	18	temporal_sup_1	9
temporal_sup_1	16	temporal_mid_1	8
occipital_mid_1	11	lingual_1	8
cingulum_post_1	10	insula_1	8
pallidum_r	10	temporal_pole_sup_r	7
insula_1	8	frontal_inf_orb_1	7
putamen_r	8	supp_motor_area_1	7

The top ten most frequently appearing parent nodes are listed here for both patients and healthy controls using a filtered ICA and unfiltered approach. Substantial reductions of ROIs between the filtered and unfiltered approach suggests the a significant sensitivity for noisy data using a dDBN approach.