Functional connectivity decreases in autism in emotion, self, and face circuits identified by Knowledge-based Enrichment Analysis

Wei Cheng, Edmund T. Rolls, Jie Zhang, Wenbo Sheng, Liang Ma, Lin Wan, Qiang Luo, Jianfeng Feng

PII: S1053-8119(16)30796-0
DOI: http://dx.doi.org/10.1016/j.neuroimage.2016.12.068
Reference: YNIMG13689

To appear in: NeuroImage

Received date: 9 July 2016
Revised date: 9 December 2016
Accepted date: 22 December 2016

Cite this article as: Wei Cheng, Edmund T. Rolls, Jie Zhang, Wenbo Sheng, Liang Ma, Lin Wan, Qiang Luo and Jianfeng Feng, Functional connectivity decreases in autism in emotion, self, and face circuits identified by Knowledge based Enrichment Analysis, NeuroImage
http://dx.doi.org/10.1016/j.neuroimage.2016.12.068

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 galley 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.
Functional connectivity decreases in autism in emotion, self, and face circuits identified by Knowledge-based Enrichment Analysis

Wei Cheng\textsuperscript{a, 1}, Edmund T. Rolls\textsuperscript{b,c, 1}; Jie Zhang\textsuperscript{a, 1}; Wenbo Sheng\textsuperscript{a}; Liang Ma\textsuperscript{f}; Lin Wan\textsuperscript{f}; Qiang Luo\textsuperscript{a,e}; Jianfeng Feng\textsuperscript{a, b, d, *}

\textsuperscript{a}School of Mathematical Sciences and Centre for Computational Systems Biology, Fudan University, Shanghai, 200433, PR China

\textsuperscript{b}Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK

\textsuperscript{c}Oxford Centre for Computational Neuroscience, Oxford, UK

\textsuperscript{d}School of Mathematical Sciences, School of Life Science and the Collaborative Innovation Center for Brain Science, Fudan University, Shanghai, 200433, PR China

\textsuperscript{e}School of Life Sciences, Fudan University, Shanghai, 200433, PR China

\textsuperscript{f}Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China.

\textsuperscript{g}National Center of Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.

*Corresponding author. Professor Jianfeng Feng, Centre for Computational Systems Biology, School of Mathematical Sciences, Fudan University, Shanghai 200433, China, and Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK. jianfeng.feng@warwick.ac.uk

ABSTRACT

A powerful new method is described called Knowledge based functional connectivity Enrichment Analysis (KEA) for interpreting resting state functional connectivity, using circuits that are functionally identified using search terms with the Neurosynth database. The method derives its power by focusing on neural circuits, sets of brain regions that share a common biological function, instead of trying to interpret single functional connectivity links. This provides a novel way of investigating how task- or function-related related networks have resting state functional connectivity differences in different psychiatric states, provides a new way to bridge the gap between task and resting-state functional networks, and potentially

\textsuperscript{1} These authors contributed equally to this work.
helps to identify brain networks that might be treated. The method was applied to interpreting functional connectivity differences in autism. Functional connectivity decreases at the network circuit level in 394 patients with autism compared with 473 controls were found in networks involving the orbitofrontal cortex, anterior cingulate cortex, middle temporal gyrus cortex, and the precuneus, in networks that are implicated in the sense of self, face processing, and theory of mind. The decreases were correlated with symptom severity.

**Key Words:** autism spectrum disorder; Neurosynth database; orbitofrontal cortex; precuneus; resting state fMRI

**INTRODUCTION**

Analysis of functional connectivity with resting state fMRI (functional magnetic resonance imaging) data has considerably advanced in the last decade and the advances have led to many investigations of altered functional connectivity between brain areas in mental disorders. The approach enables examination of how functional dysconnectivity relates to neurodegenerative and psychiatric brain diseases (Menon, 2011; Whitfield-Gabrieli and Ford, 2012), including depression (Lui et al., 2011; Tao et al., 2013), Alzheimer's disease (Sheline and Raichle, 2013), schizophrenia (Fitzsimmons et al., 2013), autism (Müller et al., 2011) and attention-deficit/hyperactivity disorder (ADHD) (Posner et al., 2014; Tomasi and Volkow, 2012). An advantage of resting state fMRI is that there is no task being performed, so possible differences in task performance between the groups are not a limitation. Of course, if differences are found in resting state functional connectivity, it is of great interest to investigate in which tasks these differences are most evident.

With increasing amounts of fMRI data, however, the challenge lies not only in identifying significantly altered functional connectivities, but also in interpreting the results to gain insights into the underlying neurobiological systems and their contributions to the disorders. (The functional connectivity (FC) between two brain areas is defined as the (Pearson) correlation between the BOLD signal measured across the time series. The
functional connectivity is also referred to as a link, where link is just the strength of the functional connectivity between a pair of brain regions.) One common approach is to order the links according to their differential strength between the patients and controls, and then to focus on a handful of links at the top of the list (Tao et al., 2013). The method has some limitations. First, the large number of functional connectivity links identified may not have an underlying neurobiological theme, and the interpretation of the functions of these links and how these may be related to the disorder can be difficult. Second, the large number of functional connectivities (FC) produces a great burden on correction for multiple comparisons, especially in the case of whole brain analysis and small, for example voxel-level, regions. Third, the approach considers each link as independent and ignores the relationship between two links, which can result in missing important effects evident at the brain circuit level. A similar issue applies to genetic studies, where it is a challenging task to interpret the results of some significant single nucleotide polymorphisms (SNPs) or genes. To examine the association between identified biological categories or pathways and outcomes, a number of enrichment tools have been developed, one of which is Gene Set Enrichment Analysis (GSEA) which evaluates microarray data at the level of gene sets and has proved to be a seminal approach (Subramanian et al., 2005).

An important and natural way to interpret the biological meaning of functional connectivity is finding the function of the regions involved using task-based fMRI. There is evidence that the major functional networks found with task-based fMRI correspond to the networks identified by correlated spontaneous fluctuations observed with resting-state fMRI (Buckner et al., 2013; Cole et al., 2014; Fair et al., 2007; Hermundstad et al., 2013; Smith et al., 2009). This consistency of brain networks revealed by task-based and resting-state fMRI leads us to a novel approach in this paper in which brain networks related to different tasks and functions are identified by combining data from many separate studies using the Neurosynth approach which uses search terms to identify regions of interest (Yarkoni et al., 2011), and then the identification of which of the networks involved is different between healthy controls and a patient group using resting-state fMRI. Many studies have analyzed the
functional connectivities between ROIs that are defined by the activation profiles of one or two specific tasks, such as reading competence (Koyama et al., 2011) and self-evaluation (Burrows et al., 2016). However, focusing on specific tasks in specific studies may limit what can be discovered. The method described here on the other hand provides a more holistic perspective, by considering activations found in a wide range of relevant tasks and conditions, and by allowing the data to be drawn from thousands of neuroimaging investigations. The approach described here also has the advantage that it analyzes whole networks or circuits, and not single connectivity links between a pair of brain regions.

In this paper, we develop an approach inspired by GSEA for the analysis of functional connectivity in the brain and its relationship to mental disorders. An important aim is to test whether functionally identified networks have altered connectivity in patients vs controls using resting state neuroimaging data measured when no task is being performed. Further, the aim is to analyze the resting state functional connectivity at the network circuit level rather than at the level of individually analyzed functional connectivity links between areas. Key steps in the analysis are that task-related Regions of Interest (ROIs) are identified in a large dataset, then functional connectivity between those ROIs is calculated from resting state fMRI in patient and control groups, and then a Kolmogorov-Smirnov-like statistic, the enrichment score ($ES$), is used to identify which functional connectivity links are different between the patient (or more generally experimental) and control groups.

To implement this novel approach described in this paper, we developed a method called Knowledge based functional connectivity Enrichment Analysis (KEA) to explore how resting state functional connectivity at the brain circuit level in patient groups vs controls can be related to the functions of the circuits identified based on their activation in a given task. The approach is modeled along the lines of Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005), but with modifications to improve the statistical power. We also illustrate the utility of KEA by applying it to autism with a large sample size of resting state fMRI data (a total of more than 900 individuals, but we also validate our results by using dataset with small sample size, Table S8). A flow chart of KEA is shown in Fig. 1 and an outline of the
implementation follows. The Neurosynth database approach (Yarkoni et al., 2011) was used to identify the coordinates of activation foci in task- or function-related fMRI of the human brain using 219 search terms that we defined to search through more than 11,000 fMRI investigations. Those coordinates were used to define regions of interest (about 25 voxels) for each task search term. These regions of interest define circuits. Differences in the functional connectivities of these circuits between patients with autism and controls were tested with Enrichment Analysis.

The main aims of the paper are to describe this new approach to identifying resting state networks that may be related to particular functions, and then to show how this approach can be used in one example of a psychiatric disorder, autism, to identify whole networks, in contrast to single connectivity links, that may be different in a disorder or particular group of individuals.

METHODS AND MATERIALS

Participants and data preprocessing

Participants’ demographic and clinical characteristics are summarized in Table S1. All data are fully anonymized in accordance with HIPAA (Health Insurance Portability and Accountability) guidelines, and research procedures and ethical guidelines were followed in accordance with the Institutional Review Boards (IRB) of each participating institution. More details of subjects are provided in the Supplementary Material and the collection and preprocessing of the data are provided at the ABIDE website (http://fcon_1000.projects.nitrc.org/indi/abide).

Neuroimaging pre-processing included dropping the first 4 volumes, slice-timing correction, motion realignment, spatial normalization to the stereotactic space of the Montreal Neurological Institute (MNI) with voxel size of 3×3×3 mm³, band-pass filtering (0.01–0.1 Hz) and smoothing (FWHM 6 mm). Nuisance covariates including head motion parameters (Friston 24 head motion parameters), linear and quadratic trends, white matter signals and cerebrospinal fluid signals (standard masks warped back form MNI space to native space)
were regressed out from the BOLD signals. The global signal was not removed for reasons described elsewhere (Cheng et al., 2016). All the data preprocessing procedures were done by the Preprocessed Connectomes Project (http://preprocessed-connectomes-project.github.io/abide/index.html). All the details of preprocessing are available on the official website provided above for these valuable datasets made available by Autism Brain Imaging Data Exchange (ABIDE, http://fcon_1000.projects.nitrc.org/indi/abide/). The exclusion criteria for sample selection included: 1) Mean framewise displacement exceeds 0.5 mm. 2) the percentage of ‘bad’ points (framewise displacement >0.5 mm) exceeds 35% in volume censoring (scrubbing). 3) fMRI data failed to normalize to MNI space which is visually inspected by members of the ABIDE project. After quality control, a total of 867 subjects (473 controls and 394 patients with autism) were left for the following analysis.

Task or function-related networks

The construction of task or function-related networks (sometimes collectively referred to as task networks for brevity) was performed using the Neurosynth database (http://neurosynth.org), which contained the results of more than 11,000 journal articles on the location of task-related activations at the time of our analyses (May 2015). The search system provides more than 3,000 search terms on the basis of which coordinates of activations are extracted from each paper. (For convenience, we call each such set of coordinates produced by a search term, such as “inhibition”, an activation map for that search term). However, many of the automatically generated search terms were not useful in identifying tasks (e.g., ‘able’, ‘abstract’ etc.), and therefore for our analysis we selected 219 search terms which have clear biological significance, and provide these search terms in Table S2. It should be mentioned that in our method the selection of terms will not affect the results except for increasing the burden of multiple comparison correction. The statistical significance level corresponding to each search term is independent of the selection of the search terms. The details of creating search term-to-coordinate mappings are described by
Yarkoni et al. (2011). In this paper, we used a forward inference map which indicates the degree to which each voxel is consistently activated in studies that use a given term with a FDR corrected P < 0.01 (Yarkoni et al., 2011). Although reverse inference maps indicate how much a given region is selectively related to a particular term, they do not work very well with some search terms such as ‘visual information’, ‘passive viewing’, and ‘monitor’, which identify only a small number of activated areas, while other search terms such as ‘face’, ‘working memory’, and ‘language’ identify a large number of activated areas. For better and fairer comparison of the networks identified with different search terms, we used forward inference map in our analysis. However, we also provide the results based on reverse inference maps for comparison in the Supplementary Material (Fig. S3).

Given the huge number of voxels of the whole brain map with a resolution of 2×2×2 mm, it is impossible to perform functional connectivity analysis at the voxel level, especially in our case, since a permutation test is needed to obtain the significance of the effects. However, the number of regions defined in atlases such as the Brodmann and AAL2 may be too small to capture sufficient information to be useful in the type of analysis described here. We therefore increased the number of brain regions that were entered into the analysis to 2048 as follows. We first used the Brodmann areas (https://en.wikipedia.org/wiki/Brodmann_area) for the parcellation of the cerebral cortex and the AAL2 atlas (Rolls et al., 2015; Tzourio-Mazoyer et al., 2002) for subcortical regions to construct the regions of interest. Considering the low resolution of the Brodmann and AAL atlases, the native Brodmann and AAL segmentation was then parcellated into 2048 micro regions of interest (ROI) using the method provided by Zalesky et al (Zalesky et al., 2010) which has widely used in network analysis (Crossley et al., 2013; Wang et al., 2013; Zhang et al., 2011). Specifically, to construct parcellation templates at a high resolution scale, each node comprising the native Brodmann areas (cerebral cortex) and AAL2 template (subcortical regions) was subdivided into micro-nodes while constraining the ROI volumes to be as uniform as possible. Each micro-node was constrained to lie within the volume encapsulated by its parent low-resolution AAL2 or Brodmann node. The number of micronodes within each node was
set to be the number of (3×3×3 mm) voxels in that node divided by 25. This process produced 2048 micronodes. If 40% of voxels in the micronode were activated in the activation map, then the micronode was defined as one of the micronodes related to the search term corresponding to the activation map. A variety of thresholds other than 40% were tested to confirm the robustness of the main findings (see supplementary material, Table S6). After the above processing, we obtained many task networks, each of which contained 7-510 micronodes. Each micronode may be referred to as an ROI in the rest of this paper.

Knowledge based functional connectivity Enrichment Analysis (KEA)

Given the atlas obtained from all task search terms using the Neurosynth database, the goal of the knowledge-based association study was to determine how the functional connectivities between the ROIs defined by the task networks were different between healthy controls and patient groups. The functional connectivities between the ROIs were performed using resting-state fMRIs from individuals not in the Neurosynth database. The analysis includes the following steps (see Fig. 1 step 2):

Step 1. Meta-analysis to assess the difference of functional connectivity between two groups.

We first performed meta-analysis of functional connectivities between ROIs of a specific task network \(i\) incorporating multi-site resting-state fMRI data. For each pair of ROIs in each network (say \(n_i\) ROIs), the time series were extracted and their Pearson correlation was calculated for each subject followed by \(z\)-transformation. Given that we did not remove the global signal in our data preprocessing, most of the Pearson correlations were positive and the sign of the Pearson correlation was kept in the following analysis. Two-sample t-tests were performed on the Fisher’s \(z\)-transformed correlation coefficients to identify significantly altered functional connectivity links in patients compared to controls within each imaging centre that provided resting-state fMRI data. The effects of age, gender ratios, head motion and full IQ were regressed within each dataset in this step by generalized linear models (Barnes et al., 2010; Di Martino et al., 2013). After obtaining the t-test results (p value for
each functional connectivity, denote as \( p_k \) for each centre, the Liptak-Stouffer \( z \) score method (Liptak, 1958) which has been described in detail in our previous studies (Cheng et al., 2015a; Cheng et al., 2015b; Cheng et al., 2016) was then used to combine the results from the individual datasets. Specifically, the p-value of each functional connectivity result from the two-sample t-test in the previous step was converted to its corresponding \( z \) score. This was calculated firstly as in equation: \( z_k = \Phi^{-1}(1 - p_k) \), where \( \Phi \) is the standard normal cumulative distribution function and \( k \) represent the \( k \) centre. Next, a combined \( z \) score for a functional connectivity was calculated using the Liptak-Stouffer formula:

\[
z = \sum w_k z_k' / \sum w_k^2,
\]

where \( w_k = \sqrt{\text{sample size}} \) is the weight of the \( k \) dataset. Finally, we obtained \( n_i(n_i - 1)/2 \) \( z \) values for the resting state functional connectivities, which represent the different connectivity strength in the patient group compared to the control group for each set of ROIs within one given task network. Suppose that there are \( M \) task networks; we define in this way a set of \( z \) values for task \( i \) as \( S_i = (z_1^i, z_2^i, \ldots, z_{n_i(n_i-1)/2}^i) \) (\( i=1, 2, \ldots, M \)).

**Step 2. Calculation of a modified Enrichment Score (MES).**

The Knowledge based enrichment analysis is in line with gene set enrichment analysis (GSEA) (Subramanian et al., 2005), which is one of the most widely used methods for detecting enrichment in gene data (Baur et al., 2006; Gupta et al., 2009; Huang et al., 2008). This is based on a maximum deviation statistic of two empirical distribution functions. First, the \( z \) values of functional connectivities (denoted by \( f_1, f_2, \ldots, f_N \), where \( N = 2048 \times 2047/2 \) of the whole brain are ranked in ascending order according to their absolute value, \( z(f_i) = z_i \), denoted by \( L = \{z_1, z_2, \ldots, z_N\} \). Then the two empirical cumulative distribution functions for the FCs in the set \( S_i \) (the FC set that is to be tested) and the complementary set \( \overline{S_i} \) (all FCs except the set \( S_i \)) are defined as:

\[
P_{S_i}(l) = \frac{\sum_{j \in S_i} |z_j|^p}{\sum_{j \in \overline{S_i}} |z_j|^p}
\]
where $N_S$ is the number of FCs in the set $S_i$; $N$ is the total number of FCs; $l$ is a given position of FC list $L$, $p$ is a parameter that gives higher weight to FCs with extreme $z$ values. The enrichment score ($ES$) is the maximum deviation from zero of $|P_{S_i} - P_{\bar{S_i}}|$. The $ES$ statistic reduces to the standard Kolmogorov–Smirnov statistic in the case $p=0$.

Although GSEA is the most commonly used tool to determine enrichment for microarray analysis, the problem of low power is not well addressed, as can be seen from the recommended FDR threshold (FDR 0.25) in the software manual (Goeman and Bühlmann, 2007). Therefore, in the present study, we made some modifications in the definition of $ES$ to avoid the relatively lower power of the original GSEA. Specifically, we used the ratio of significant FCs under a given threshold $\alpha$ (in the present study $\alpha = 0.01$) as the weight for each set rather than the parameter $p$ in the calculation of $P_{S_i}$. The modified definition of $ES$ for task $i$ is given by:

$$MES_i = \max_i \left| \frac{\sum_{j \in S_i} \frac{1}{N_S} - \sum_{j \in \bar{S_i}} \frac{1}{N - N_S}}{N_S} \right| \times \frac{\# FCs of task i with p value < \alpha}{N_S}.$$

The modified $ES$ is similar to the weighted version ($p>0$) of GSEA, and both of them aim to highlight the FCs at the extremes (those with a high or low $z$ value) of the entire ranked list $L$. However, the modified definition is much more powerful which we show in the comparison between them in the Supplementary Material (Fig. S1).

It should be noted that the whole brain has been divided into 2048 regions of interest as described in section ‘task or function-related networks’. The complementary set $\bar{S_i}$ is only related with task $i$ and is independent of other tasks. This is the reason that the selection of terms in our method will not affect the results for each task. Apart from increasing the burden of correction for multiple comparisons.
Step 3. Estimation of the significance level of the modified Enrichment Score, \( MES \).

The statistical significance (nominal p value) of the \( MES \) was estimated by a ‘group-based’ permutation test, where the groups in this case are patients vs controls. Specifically, first we permuted the disease labels of all samples, then we reordered the FCs and recomputed the enrichment score as described above. We repeated the above step for 10,000 permutations to generate a null distribution for the \( MES \). Finally, the empirical, nominal p value of the observed \( MES \) was then calculated relative to this null distribution. As there are 219 search terms, correction for multiple comparisons is needed, and FDR correction was used to correct for these multiple comparisons. We also highlight in places some significant results that can also pass Bonferroni correction.

Correlation with symptoms

We investigated whether the dysconnectivity within task networks identified by the knowledge-based association study correlated with clinical variables. The severity of autism was quantified with the Autism Diagnostic Observation Schedule (ADOS) (Lord et al., 2012) total, communication and social scores. We first divided the FCs of each significant task network into increased and decreased FC groups compared with healthy controls. For each group, we then calculated the mean value as an integration of the high-dimensional FCs, which is used to calculate the partial correlation between the task network and ADOS scores after removing the effects of head motion, sex and age. Since most FC differences that we identified were decreases in autism, we only considered decreased FCs in this analysis, and used p values for one tailed statistics. It should be noted that the correlation analysis between the symptoms and FCs with significant task networks are performed within the autism group.

RESULTS

Task or function-related networks

Fig. 2 shows examples of the activations identified using some of the individual task search terms when using the automated Neurosynth meta-analyses (Yarkoni et al., 2011). The
right of each activation diagram in each panel of Fig. 2 shows the corresponding parcellated atlas. Activations produced by the individual search terms of ‘self referential’, ‘autobiographical memory’, ‘face recognition’, ‘object recognition’, ‘percept’, ‘inhibition’, ‘autism’, and ‘finger movements’ are illustrated. Each set of activations can be described as a circuit or set of regions that are functionally related in being selected by a search term.

**Significant differences in autism identified by KEA**

The results of applying the Knowledge based functional connectivity Enrichment Analysis (KEA) to the resting state fMRI data of the autism vs the control group are shown in Fig. 3. The FC networks between the ROIs identified by the search term ‘self referential’ was significant after Bonferroni correction (p<0.05). In addition, networks identified by the search terms related to ‘face’, ‘visual’ and ‘percept’ were also significantly different between people with autism and controls using FDR correction (p<0.05) (Fig. 3). Importantly, the term ‘autism’ was also significant after FDR correction (p = 0.0018). Fig. 4 shows the results for the top four terms identified by KEA. For each term, the left part shows the distribution of t values (top left) and the Kolmogorov-Smirnov (KS) statistic (bottom left) for these search term-related FC differences between patients and controls for these three significantly different search terms. The right part shows the significant FC differences after FDR correction (p<0.05). For all these task search terms, there were significant decreases in functional connectivities in autism in the networks defined by these task search terms as significantly different in patients with autism. The FCs for ‘self-referential’ were most significantly different in autism with enrichment analysis (modified Enrichment Score, $MES = 0.1104$, p < 0.0001), followed by ‘autobiographical memory’ with $MES = 0.0427$ (p < 0.001), and then by ‘object recognition’ with $MES = 0.0393$ (p < 0.001). It is noted that the activation networks for ‘autobiographical memory’ and ‘self-referential’ are very similar and have some overlap with the default mode network. The results for other significant task search terms are summarized in Table 1.
In more detail, the resting state functional connectivities for the task search term related ROIs that were different from controls after Bonferroni correction were for the networks identified by ‘self referential’, with the networks shown in Fig. 2. The functional connectivity networks that were different in autism for this search term is shown in Fig. 4A. Orbitofrontal (ORB), anterior cingulate (ACC), middle temporal gyrus (MTG), and precuneus (PCUN) regions were notable in showing decreased functional connectivity in autism, consistent with a recent study in which these functional connectivity differences were related to the altered face processing, sense of self, and emotion in autism (Cheng et al., 2015b). For the search term ‘autobiographical memory’ (Fig. 4B), reduced connectivity in autism was found for links that included the precuneus, parahippocampal gyrus, fusiform gyrus, middle temporal gyrus, and gyrus rectus (medial orbitofrontal cortex). For the search term ‘object recognition’, reduced functional connectivity in autism was found for links that included the inferior and middle temporal gyrus and fusiform gyrus and some occipital visual areas, and the inferior frontal gyrus (Fig. 4C). For the search term ‘face recognition’, reduced functional connectivity in autism was found for links that included the inferior and middle temporal gyrus and fusiform gyrus and some occipital visual areas, and the inferior frontal gyrus (Fig. 4D).

In the present investigation, the global signal was not regressed out, for reasons described elsewhere (Cheng et al., 2016), and there were few links that were more positive in the autistic than control group. In a previous investigation, we showed that some apparently more positive links involving the middle temporal gyrus were present only with global signal removal, which we now regard as difficult to interpret, as these particular links involved negative correlations (Cheng et al., 2015b), which can arise with global signal removal and which can be artefactual (Cheng et al., 2016) (see Supplementary Material).

Correlations with clinical measures of differences in functional connectivity in autism
As can be seen from Table 2 there were significant correlations (p<0.05 FDR corrected) between the altered FCs and the symptom severity scores in autism. Specifically, the
decreased FC links between the ROIs of all the significant search terms show negative correlations with the ADOS scores. Importantly, most of the correlations between the significant search term related to ‘face’ and the ADOS total, communication and social scores were significant after FDR correction.

In additional analyses, we showed that the functional connectivities that were significantly different between the autistic group and controls tended to have correlations with the ADOS symptom scores (as shown in Table S3 and 4), whereas the functional connectivities that were not significantly different between the autistic group and controls tended to have no or very low correlations with the ADOS symptom scores.

The analyses described in the main text were supplemented by analyses described in the Supplementary Material in which communities of the networks identified by the Neurosynth search terms were identified, to add to the interpretation of the findings on the functions of the networks that were identified to be related to autism (Table S4, 5, Figure S4, 5, 6).

**DISCUSSION**

This paper introduces Knowledge based functional connectivity Enrichment Analysis (KEA) to explore how resting state functional connectivity at the brain circuit level in patient groups vs controls can be related to the functions of the networks identified based on their activation of task. The resting-state approach is unbiased or completely data driven, while the task-related approach usually has a priori assumptions. Here we have provided a way to bridge between these two types of analysis, by using evidence from task-related analyses to help to give insights into the functions of the networks identified as different using resting state analyses. Moreover, the approach we describe enables data from thousands of papers and brain scans of different individuals to be taken into account when performing this bridging. Key steps in the analysis are that Regions of Interest (ROIs) of different task networks are identified in a large dataset, then functional connectivity between those ROIs is calculated from resting state fMRI in patient and control groups, and then a Kolmogorov-Smirnov-like statistic, the enrichment score (ES), is used to identify which
functional connectivity links are different between the patient (or more generally experimental) and control groups.

Another advance is that we introduced a way to reduce the number of activation maps produced by the Neurosynth approach (Yarkoni et al., 2011), one for each of up to 3000 search terms, to a much smaller set of maps that reflect the similarities between the brain maps. The method for this categorisation or cluster analysis of the maps that we used was hierarchical cluster analysis. This grouped together for example all the maps produced by similar search terms from the 219 used, such as “tapping” and “finger movement”. The similarity of the search terms for each of the resulting 27 task community-based activation maps is clear from Table S2 and Fig. S2, which provide strong support for the approach. The results of this make the results produced by the Neurosynth approach much more easily interpretable. The resulting maps for each of the task-related communities are shown in Fig. S6.

With these new approaches, interesting results were obtained from the analyses of the resting-state fMRI data for differences between the group with autism spectrum disorder and the controls. The new findings here provide useful additions to previous investigations of functional connectivity in autism (Cheng et al., 2015b; Di Martino et al., 2013; Ecker et al., 2015; Maximo et al., 2014), by providing powerful and robust analyses because of the use of the new KEA method and the use of a large number of participants. For both the single search term related networks, and the cluster-based networks, the brain regions that frequently had reduced functional connectivity as shown in Figs. 4 and S6 included the Orbitofrontal cortex (ORB and PFCventmed); cingulate cortex (ACC and MCC); temporal lobe visual cortical areas including the MTG, STG, ITG, TPO, and fusiform gyrus (in which the fusiform face area is located); precuneus (PCUN), memory-related regions including the parahippocampal gyrus and hippocampus, and the insula. Given the search terms or clusters of search terms that helped to identify these areas as having reduced functional connectivity in autism, we now discuss how the reduced functional connectivity between some of these areas may be related to autism spectrum disorder.
The temporal lobe visual cortical areas are involved in face as well as object perception, and face processing is changed in autism. The fusiform gyrus contains the fusiform face area involved in face recognition. Of especial interest in relation to autism is the middle temporal gyrus (MTG) which is a region implicated in face expression and gesture representation impairments and to theory of mind impairments in autism (Cheng et al., 2015b). Indeed, Rolls and colleagues discovered face cells not only in the macaque inferior temporal visual cortex (Perrett et al., 1982; Rolls, 2011), where they provide transform-invariant representations of face identity (Rolls, 2012; Rolls and Treves, 2011), but also in the cortex in the superior temporal sulcus (Baylis et al., 1987; Hasselmo et al., 1989a), where the neurons respond to face expression (Hasselmo et al., 1989a) and to movements and gestures of the head and body and gaze direction used in social communication (Hasselmo et al., 1989b; Perrett et al., 1985a; Perrett et al., 1985b). Faces including those with expressions activate the MTG region in humans, with coordinates that include [-49 -36 2], [-49 -42 4] and [52 -42 4] (Critchley et al., 2000). The altered functional connectivity in this area in autism is of great interest, for some of the key typical symptoms of autism spectrum disorder (ASD) are face processing and especially face expression processing deficits (Lai et al., 2014), which will impair social and emotional communication (Rolls, 2014). The medial temporal gyrus (MTG) region also includes a region with activations related to theory of mind (ToM) and speech (Hein and Knight, 2008). Thus the MTG region that we identified with reduced functional cortical connectivity in ASD has functions related also to theory of mind and even speech processing. This is fascinating, given that there are major impairments of theory of mind and communication in ASD (Lai et al., 2014).

The orbitofrontal cortex, including the ventromedial prefrontal cortex (PFCventmed), is a region involved in emotion (Rolls, 2014, 2015), including neurons with firing related to face expression (Rolls et al., 2006). This is of interest, for the orbitofrontal cortex is important in social behavior communication (Rolls, 2014), when damaged in humans produces face and voice emotional expression identification impairments (Hornak et al., 2003; Hornak et al., 1996), and has connections with the amygdala which also contains face-selective neurons.
Leonard et al., 1985), and both are implicated in autism in some other approaches (Baron-Cohen et al., 1999; Lombardo et al., 2010; Nordahl et al., 2012). Further evidence for the importance of the ventromedial prefrontal cortex in autism is that has reduced functional connectivity as shown in this study and elsewhere (Cheng et al., 2015b). The reduced functional connectivity of the middle temporal gyrus with areas involved in emotion, the ventromedial prefrontal cortex, and areas involved in the sense of self (the precuneus and its connected areas), appears to be very relevant to autism spectrum disorder, in which disorders of face processing, emotional and social responses, and theory of mind (to which the sense of self contributes) are important (Cheng et al., 2015b).

Another area with reduced functional connectivity is the precuneus (PCUN) which is part of medial parietal cortex area 7. The precuneus is a region with spatial representations not only of the self, but also of the spatial environment, and it may be partly in relation to this type of representation that damage to this region impairs the sense of self and agency (Cavanna and Trimble, 2006). The reduced functional connectivity of this region is therefore of great interest in relation to the symptoms of autism that relate to not having a theory of others’ minds, for which a representation (or ‘theory’) of oneself in the world may be important (Lombardo et al., 2010). The precuneus has associated with it the adjoining paracentral lobule (PCL) which is part of the superior parietal cortex with somatosensory and perhaps visual spatial functions, and which has strong anatomical connections with the precuneus (Margulies et al., 2009). Both the paracentral lobule with its body and spatial representation, and the precuneus, operate together to produce a sense of self, in which the representation of the body and how it acts in space is likely to be an important component (Cavanna and Trimble, 2006). We therefore hypothesize that the reduced functional connectivity of these precuneus and related parietal cortex regions is related to the altered representation or disconnection of the representation of oneself in the world that may contribute to the reduction in the theory of mind in autism (Lombardo et al., 2010). In this context the reduced functional connectivity of this precuneus region with the MTG/ITG/STS areas is of interest, for theory of mind including of oneself and others, and face and voice
communication with others, would seem to be a set of functions that should normally be usefully communicating to implement social behavior, which is impaired in autism. The reduced functional connectivity of some parietal cortex areas with the somatosensory cortex (PoCG) (Fig. S6) in autism may reflect the fact that somatosensory inputs are likely to be important in the body and self-related functions of the parietal cortex.

Part of the value of the approach described here is that the resting state functional connectivity differences in mental states such as autism, schizophrenia, and ADHD can be related to the functions of the networks involved as identified by activations of different tasks in other subjects. Examples which illustrate how the functions of the different networks with reduced functional connectivity in autism can be identified are shown in Figs. 4 and S6. This places less reliance on whether, and if so how, these different tasks can be performed during neuroimaging by different patient groups.

There are several advantages of KEA. First, the changes in functional connectivity found in resting-state fMRI can be interpreted in terms of the functions of the different links. Unlike the traditional approach for functional connectivity analysis based on resting-state fMRI which focuses on identifying the most changed FCs under two different conditions, all FCs are first categorized based on their activation in tasks. Second, the KEA approach is likely to be helpful as a post hoc analysis for traditional resting-state FC analysis, by helping to identify tasks to which the altered FCs are related. Third, just as with GSEA, KEA is flexible. We created 219 FC sets, each one of them corresponding to one task or cognitive description based on the Neurosynth database. Different search terms can easily be used in future analyses to perform KEA. Fourth, analysis focusing on task-selected FCs tends to be more reproducible and less sensitive to noise, and, importantly, can also reduce the burden on correction for multiple comparisons. Fifth, significance is assessed by permuting the class labels in KEA. Therefore, the correlation between FCs is preserved, which makes the null model more accurate. Sixth, although we only applied it in the autism resting-state fMRI data to find the difference between those with autism and controls, KEA is flexible and can easily be used in other analyses. For example, we can find the significantly correlated ‘circuits’
given a behavioral score, such as IQ, amount of drinking, etc. Our method is an extension of traditional analyses to which we just add enrichment analysis to find out which cognitive circuits are those with altered functional connectivities. Almost all previous analyses involving single functional connectivity links can be easily extended to the ‘circuit’ level by our method. The approach does benefit from a large sample size, for we found that the results were not significant if we used subsets of the data with 51-172 individuals rather than the 867 individuals in the full dataset, thought the trends were in the same direction (Table S8). Finally, we note that because KEA is hypothesis driven, the set of functional connectivities being tested for significance is smaller than with a whole brain analysis, and thus KEA may reveal significant effects that might not be detected with a whole-brain analysis. Of course, correspondingly, a whole brain analysis may still be useful, because it may detect group differences that might be missed because of a hypothesis-driven approach which restricts the search to a subset of functional connectivities. The KEA approach may also produce more statistically significant results because whole circuits are being searched for, rather than individual functional connectivity links between a pair of voxels or brain regions. We do note that the Neurosynth database approach can be powerful compared to other approaches such as the Surface Management System database approach in which human intervention is used to extract the required information, because the Neurosynth approach allows very many studies to be interrogated, because it uses automatic extraction of data (Poldrack and Yarkoni, 2016; Yarkoni et al., 2011).

Code availability: the code of the KEA is available at the following website: http://www.dcs.warwick.ac.uk/~feng/.
ACKNOWLEDGEMENTS

The use of the Neurosynth database (http://neurosynth.org) (Yarkoni et al., 2011) is acknowledged. J.Feng is a Royal Society Wolfson Research Merit Award holder. J.Feng is also partially supported by the National High Technology Research and Development Program of China (No. 2015AA020507) and the key project of Shanghai Science & Technology Innovation Plan (No. 15JC1400101). The research was partially supported by the National Centre for Mathematics and Interdisciplinary Sciences (NCMIS) of the Chinese Academy of Sciences (CAS), Key Program of National Natural Science Foundation of China (No. 91230201), and the Shanghai Soft Science Research Program (No. 15692106604). W.C is partly supported by grants from the National Natural Sciences Foundation of China (No.11471081). J.Zhang is supported by National Science Foundation of China (NSFC 61104143 and 61573107), and special Funds for Major State Basic Research Projects of China (2015CB856003). LW is supported by the NSFC grants (No. 11571349 and No. 11201460), the NCMIS of the CAS, the Youth Innovation Promotion Association of the CAS, and the Strategic Priority Research Program of the CAS (XDB13040600).

References


---

**Table 1.** The enrichment analysis results on the functional connectivities of the single search term related networks (FDR<0.05). MES - modified Enrichment Score (see Methods).
<table>
<thead>
<tr>
<th>Terms</th>
<th>MES</th>
<th>p value</th>
<th>FDR</th>
<th>Terms</th>
<th>MES</th>
<th>p value</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>self referential</td>
<td>0.1104</td>
<td>0.0000</td>
<td>0.0000</td>
<td>image</td>
<td>0.0295</td>
<td>0.0028</td>
<td>0.0451</td>
</tr>
<tr>
<td>autobiographical memory</td>
<td>0.0427</td>
<td>0.0010</td>
<td>0.0451</td>
<td>touch</td>
<td>0.0229</td>
<td>0.0029</td>
<td>0.0451</td>
</tr>
<tr>
<td>object recognition</td>
<td>0.0393</td>
<td>0.0010</td>
<td>0.0451</td>
<td>moral</td>
<td>0.0271</td>
<td>0.0030</td>
<td>0.0451</td>
</tr>
<tr>
<td>face recognition</td>
<td>0.0462</td>
<td>0.0011</td>
<td>0.0451</td>
<td>judgment</td>
<td>0.0171</td>
<td>0.0031</td>
<td>0.0451</td>
</tr>
<tr>
<td>fusiform face</td>
<td>0.0339</td>
<td>0.0012</td>
<td>0.0451</td>
<td>percept</td>
<td>0.0240</td>
<td>0.0033</td>
<td>0.0451</td>
</tr>
<tr>
<td>visual stream</td>
<td>0.0570</td>
<td>0.0014</td>
<td>0.0451</td>
<td>motion</td>
<td>0.0149</td>
<td>0.0035</td>
<td>0.0451</td>
</tr>
<tr>
<td>autism</td>
<td>0.0718</td>
<td>0.0018</td>
<td>0.0451</td>
<td>decision task</td>
<td>0.0263</td>
<td>0.0039</td>
<td>0.0469</td>
</tr>
<tr>
<td>matching task</td>
<td>0.0310</td>
<td>0.0023</td>
<td>0.0451</td>
<td>neurocognitive</td>
<td>0.0323</td>
<td>0.0041</td>
<td>0.0469</td>
</tr>
<tr>
<td>hand movements</td>
<td>0.0294</td>
<td>0.0024</td>
<td>0.0451</td>
<td>control network</td>
<td>0.0239</td>
<td>0.0044</td>
<td>0.0469</td>
</tr>
<tr>
<td>primary somatosensory</td>
<td>0.0273</td>
<td>0.0025</td>
<td>0.0451</td>
<td>face stimuli</td>
<td>0.0270</td>
<td>0.0045</td>
<td>0.0469</td>
</tr>
<tr>
<td>visual spatial</td>
<td>0.0404</td>
<td>0.0025</td>
<td>0.0451</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Correlations between the functional connectivities of the 21 significant networks identified with search terms and the symptom severity scores. The correlations that are significant (p<0.05) after FDR correction for multiple comparisons are indicated in bold. A negative correlation indicates that the decrease in FC was correlated with the symptom severity.

<table>
<thead>
<tr>
<th>Task search term</th>
<th>ADOS total correlation</th>
<th>p value</th>
<th>ADOS communication correlation</th>
<th>p value</th>
<th>ADOS social correlation</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>self referential</td>
<td>-0.11393</td>
<td>0.05409</td>
<td>-0.13767</td>
<td>0.02595</td>
<td>-0.08058</td>
<td>0.12833</td>
</tr>
<tr>
<td>autobiographical memory</td>
<td>-0.12735</td>
<td>0.03617</td>
<td>-0.13062</td>
<td>0.03262</td>
<td>-0.10356</td>
<td>0.07225</td>
</tr>
<tr>
<td>object recognition</td>
<td>-0.19959</td>
<td>0.00230</td>
<td>-0.18628</td>
<td>0.00413</td>
<td>-0.17301</td>
<td>0.00715</td>
</tr>
<tr>
<td>face recognition</td>
<td>-0.19100</td>
<td>0.00337</td>
<td>-0.18753</td>
<td>0.00392</td>
<td>-0.16018</td>
<td>0.01173</td>
</tr>
<tr>
<td>fusiform face</td>
<td>-0.19332</td>
<td>0.00305</td>
<td>-0.19171</td>
<td>0.00327</td>
<td>-0.16103</td>
<td>0.01137</td>
</tr>
<tr>
<td>visual stream</td>
<td>-0.21575</td>
<td>0.00108</td>
<td>-0.20840</td>
<td>0.00153</td>
<td>-0.18292</td>
<td>0.00476</td>
</tr>
<tr>
<td>autism</td>
<td>-0.12936</td>
<td>0.03395</td>
<td>-0.11600</td>
<td>0.05095</td>
<td>-0.11488</td>
<td>0.05263</td>
</tr>
<tr>
<td>matching task</td>
<td>-0.12607</td>
<td>0.03764</td>
<td>-0.10835</td>
<td>0.06335</td>
<td>-0.11468</td>
<td>0.05295</td>
</tr>
<tr>
<td>hand movements</td>
<td>-0.18018</td>
<td>0.00534</td>
<td>-0.18471</td>
<td>0.00442</td>
<td>-0.14657</td>
<td>0.01918</td>
</tr>
<tr>
<td>primary</td>
<td>-0.13971</td>
<td>0.02425</td>
<td>-0.15191</td>
<td>0.01588</td>
<td>-0.10861</td>
<td>0.06289</td>
</tr>
<tr>
<td>somatosensory</td>
<td>-0.15220</td>
<td>0.01572</td>
<td>-0.12779</td>
<td>0.03567</td>
<td>-0.14019</td>
<td>0.02385</td>
</tr>
<tr>
<td>visual spatial</td>
<td>-0.16883</td>
<td>0.00843</td>
<td>-0.16216</td>
<td>0.01089</td>
<td>-0.14368</td>
<td>0.02119</td>
</tr>
<tr>
<td>image</td>
<td>-0.12778</td>
<td>0.03569</td>
<td>-0.13641</td>
<td>0.02705</td>
<td>-0.10080</td>
<td>0.07777</td>
</tr>
<tr>
<td>touch</td>
<td>-0.09086</td>
<td>0.10035</td>
<td>-0.09415</td>
<td>0.09241</td>
<td>-0.07334</td>
<td>0.15102</td>
</tr>
<tr>
<td>moral</td>
<td>-0.11438</td>
<td>0.05340</td>
<td>-0.10519</td>
<td>0.06912</td>
<td>-0.10005</td>
<td>0.07933</td>
</tr>
<tr>
<td>judgment</td>
<td>-0.15250</td>
<td>0.01555</td>
<td>-0.15613</td>
<td>0.01363</td>
<td>-0.12418</td>
<td>0.03989</td>
</tr>
<tr>
<td>percept</td>
<td>-0.15919</td>
<td>0.01218</td>
<td>-0.15189</td>
<td>0.01589</td>
<td>-0.13606</td>
<td>0.02736</td>
</tr>
<tr>
<td>motion</td>
<td>-0.13007</td>
<td>0.03320</td>
<td>-0.12931</td>
<td>0.03401</td>
<td>-0.10815</td>
<td>0.06371</td>
</tr>
<tr>
<td>decision task</td>
<td>-0.10305</td>
<td>0.07323</td>
<td>-0.10331</td>
<td>0.07273</td>
<td>-0.08519</td>
<td>0.11518</td>
</tr>
<tr>
<td>neurocognitive</td>
<td>-0.09284</td>
<td>0.09550</td>
<td>-0.08759</td>
<td>0.10873</td>
<td>-0.07994</td>
<td>0.13026</td>
</tr>
<tr>
<td>control network</td>
<td>-0.18866</td>
<td>0.00373</td>
<td>-0.18287</td>
<td>0.00477</td>
<td>-0.15959</td>
<td>0.01199</td>
</tr>
</tbody>
</table>

Figure 1. Schematic overview of knowledge based functional connectivity enrichment analysis.
Step 1: Task or function-related networks

- Term-based search
- Working memory
- Meta analysis

Related studies
Forward inference map
Hierarchical clustering analysis

Step 2: Knowledge based Enrichment Analysis

- Task state
- Sylvian MRI data

Step 3: Clinical utility at two levels

- Task based enrichment analysis
- Significant links
- Correlation between FC set and behavior

Functional connectivities
Clinical variables
Figure 2. Task or function-related networks. Examples of the activations identified for some of the individual search terms as revealed by the automated Neurosynth meta-analyses. The activation maps (2×2×2 mm) are shown on the left of each panel, and the corresponding parcellated atlas is shown on the right of each panel. The colors in the left parts represent a heat map of the z values; and the colors in the right parts indicate the different parcels, referred to in the paper as “task networks”.
**Figure 3.** The order of significance of the differences between patients and controls for FCs identified from the task-related ROIs identified by each of the search terms indicated for autism (only show top 80 search terms). The yellow and pink bars indicate the FCs that were significant after Bonferroni correction ($p < 0.05$) and FDR correction ($p < 0.05$) respectively.
Figure 4. Functional connectivity differences in autism revealed by KEA. (Results are shown for the four search terms that produced the most significant differences in autism.) A) Result for top search term ‘self referential’. The top left shows the distribution of t values and ES values for the differences of FCs between task-related ROIs for people with autism and controls. The blue curves mark the distribution of significant tasks for the null hypothesis derived from permutation, and the red curves mark the distribution of observed data. The bottom left shows the Kolmogorov–Smirnov statistic between these two distributions. The blue and red lines each correspond to the empirical distribution function for all tasks and the task of interest respectively, and the black brace is the ES value (i.e. the two-sample Kolmogorov–Smirnov statistic). The right part shows the significantly different functional connectivity links (FDR correction p < 0.05) for the search term related networks between the autism group and controls identified by the knowledge-based enhancement association study (KEA). The links are shown in red if they are significantly stronger in the autism group, and in blue if they are significantly weaker in the autism group. The thickness of the lines indicates the degree of alteration of the functional connectivity. The anatomical abbreviations are for the areas in the automated anatomical atlas AAL2 (Rolls et al., 2015), with abbreviations shown in Table S3. The innermost circle indicates the mean z value of significantly decreased links in autism involving each brain region, and the middle circle indicates the mean z value of significantly increased links. The brain regions in the left hemisphere are in the left semicircle of the diagram. B, C, D) Results for the search terms ‘autobiographical memory’, ‘object recognition’, and ‘face recognition’, using the same conventions as A.
Highlights

- Knowledge based functional connectivity Enrichment Analysis (KEA) is described.
- KEA allows functional identification of whole networks from resting state fMRI using Neurosynth.
- Applied to 394 patients with autism and 473 controls the different networks involved in autism:
  - The orbitofrontal, anterior cingulate, middle temporal gyrus cortex, and the precuneus.
  - These networks are implicated in face processing, theory of mind, and the sense of self.