Journal of Affective Disorders 150 (2013) 1213–1216

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Brief report

Predicting obsessive–compulsive disorder severity combining neuroimaging and machine learning methods

Marcelo Q. Hoexter [a,*], Euripedes C. Miguel [a], Juliana B. Diniz [a], Roseli G. Shavitt [a], Geraldo F. Busatto [a], João R. Sato [b]

[a] Department & Institute of Psychiatry, University of São Paulo Medical School, Rua Ovídio Pires de Campos 785, 3° Andar Ala Norte, Sala 9 (PROTOC), CEP 05403-010, São Paulo, Brazil
[b] Center of Mathematics, Computation and Cognition, Universidade Federal do ABC, Santo Andre, Brazil

ARTICLE INFO

Article history:
Received 22 April 2013
Accepted 17 May 2013
Available online 14 June 2013

Keywords:
Machine learning
Magnetic resonance imaging
Neuroimaging
Obsessive–compulsive disorder
Support vector regression
Symptom severity

ABSTRACT

Background: Recently, machine learning methods have been used to discriminate, on an individual basis, patients from healthy controls through brain structural magnetic resonance imaging (MRI). However, the application of these methods to predict the severity of psychiatric symptoms is less common.
Methods: Herein, support vector regression (SVR) was employed to evaluate whether gray matter volumes encompassing cortical–subcortical loops contain discriminative information to predict obsessive–compulsive disorder (OCD) symptom severity in 37 treatment-naïve adult OCD patients.
Results: The Pearson correlation coefficient between predicted and observed symptom severity scores was 0.49 ($p=0.002$) for total Dimensional Yale-Brown Obsessive–Compulsive Scale (DY-BOCS) and 0.44 ($p=0.006$) for total Yale-Brown Obsessive–Compulsive Scale (Y-BOCS). The regions that contained the most discriminative information were the left medial orbitofrontal cortex and the left putamen for both scales.
Limitations: Our sample is relatively small and our results must be replicated with independent and larger samples.
Conclusions: These results indicate that machine learning methods such as SVR analysis may identify neurobiological markers to predict OCD symptom severity based on individual structural MRI datasets.

1. Introduction

Several analytical approaches have been used to classify or discriminate patients with psychiatric disorders from healthy controls through brain magnetic resonance imaging (MRI) as an attempt to understand the neurobiological underpinnings of mental disorders (Pereira et al., 2009).

Structural MRI has been previously used to assess the feasibility of differentiating individual subjects with obsessive–compulsive disorder (OCD) from healthy controls. The overall classification accuracy to discriminate a group of OCD patients (not used for training the classifier) from control subjects was 76.6% (Soriano-Mas et al., 2007). More recently, using a functional MRI protocol involving the processing of fear and disgust stimuli in OCD, a multivariate pattern recognition approach revealed that the orbitofrontal cortex (OFC) and caudate nucleus encoded diagnostic information that differentiated OCD patients from healthy controls with an accuracy of 100% (Weygandt et al., 2012). However, most of patients included in previous studies have been exposed to medications, which most probably have influenced the results (Hoexter et al., 2012).

Given the limitations in the diagnostic constructs currently under use in psychiatry (Insel et al., 2010), it seems mandatory to evaluate if additional characteristics correlate with brain imaging patterns beyond the presence of a specific disorder. One of the such characteristics could be the global symptom severity that varies across sufferers from a given disorder. The combination of computational neuroanatomy based on neuroimaging and machine learning methods is a suitable approach to address this issue (Sato et al., 2012a, 2012b) and may allow the identification of predictive neurobiological markers of future outcomes in individual subjects (de Almeida and Phillips, 2013).

Numerous neuroimaging studies have reported associations between OCD severity and gray matter (GM) volume abnormalities in several brain regions (Radua and Mataix-Cols, 2009; Rotge et al., 2010). Moreover, OCD severity is also a consistent variable for predicting outcome in OCD (Denys et al., 2003; Mataix-Cols et al., 1999).

Herein, based on structural MRI data from a sample described previously (Hoexter et al., 2012), we combine support vector

* Corresponding author. Tel.: +55 11 2661 7896.
E-mail address: mqhoexter@gmail.com (M.Q. Hoexter).

0165-0327/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jad.2013.05.041

regression to evaluate whether specific GM volumes encompassing cortical–subcortical loops contain relevant information to predict OCD severity in a sample of treatment-naïve patients. This approach has not been applied previously and may provide objective measures of the severity of the disorder, and address specific dimensions that may characterize subgroups of patients.

The enrolling of only treatment-naive patients in such a study design is essential to rule out the potential influence of previous treatments on brain morphometry.

2. Methods

2.1. Subjects

Thirty-seven adult OCD patients participated in this study. Patients were recruited from the Obsessive–Compulsive Spectrum Disorders Program at the Institute of Psychiatry, University of São Paulo Medical School, Brazil. All subjects provided written informed consent to participate in this study, approved by the local Ethics Committee (Hoexter et al., 2009).

Participants were treatment-naïve, aged between 18 and 65 years, had OCD as the primary diagnosis (DSM-IV) and presented a Yale-Brown Obsessive–Compulsive Scale (Y-BOCS) (Goodman et al., 1989) score≥16 or >10 for only obsessions or compulsions. Exclusion criteria were: previous treatment with psychotropic medication (benzodiazepines, antipsychotics, antidepressants, stimulants, mood stabilizers), at least 12 sessions of cognitive-behavior therapy, substance abuse/dependence, any medical disorder possibly affecting the central nervous system. Further details can be found in (Hoexter et al., 2009, 2012).

2.1.1. Clinical assessment

Psychopathological assessment included the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) (First et al., 1997), Y-BOCS, and Dimensional Yale-Brown Obsessive–Compulsive Scale (Rosario-Campos et al., 2006). The Y-BOCS is a 10-item, clinician-administered instrument, in which each item is scored on a four-point scale from 0="no symptoms" to 4="extreme symptoms" (total range from 0 to 40). The DY-BOCS assesses obsessive–compulsive symptoms in accordance to specific dimensions ("contamination/washing", "aggression/checking", "sexual/religious", "symmetry/ordering" and "hoarding"), which include obsessions and related compulsions. The severity of each dimension can be independently quantified. This instrument also investigates the time spent with obsessive–compulsive symptoms, the level of anxiety and interference, with scores ranging from 0–5 (minimum score=0 and maximum score=15 for each dimension). The total negative impact of obsessive–compulsive symptoms is also measured (total DY-BOCS ranges from 0–30).

2.2. Image acquisition

MRI data were acquired in a 1.5 T GE Signa scanner (General Electric, Milwaukee WI, USA) using a T1-3D SPGR sequence (axial, TE=4.20 ms, TR=10.5 ms, flip angle=15°, 256×192 matrix, 248 slices with thickness=1.6 mm). The images were then reconstructed based on a voxel size of $0.94 \times 0.94 \times 0.80$ mm^3.

2.3. Image processing and support vector regression

Structural MRI datasets from all patients were processed using the automated cortical and subcortical segmentation methods (aseg.volume.stats and bilateral aparc.volume.stats output files) in Freesurfer package (http://surfer.nmr.mgh.harvard.edu), by using the "recon-all" routine. This pipeline is based on the following steps: cortical surface modeling, spherical coordinate transformation, curvature (sulci and gyri) registration to standard space, and automated parcellation and labeling of cortical and subcortical structures. The volume of labeled regions was used as the explanatory variables to predict the behavioral scales. Further details about recon-all pipeline can be found at Fischl et al. (1999).

Support vector regression (SVR) with linear kernel (Smola and Scholkopf, 1998) was applied to predict OCD severity based on the volumes of labeled brain regions. The areas set as predictors were: lateral and medial OFC, anterior cingulate cortex (ACC), caudate, putamen, pallidum, accumbens, and thalamus at each hemisphere, resulting in 16 regressors. These regions were chosen based on the regions of the cortical–subcortical pathways classically associated to OCD (Saxena and Rauch, 2000). A feature selection algorithm was applied through recursive feature elimination (Guyon et al., 2002). The relevance of each region was quantified by using the proportion of feature selection in leave-one-out procedure (Sato et al., 2012a, 2012b).

3. Results

The mean age ± standard deviation (SD) of our sample was 31.9 ± 10.1 years. Twenty-two of 37 patients were female (59%) and 36 were right-handed (97%). Clinical characteristics, Y-BOCS and DY-BOCS scores of the sample are presented in Table 1.

Fig. 1 shows a scatter-plot of predicted and observed values and the relevance (contribution) of each brain region for DY-BOCS and Y-BOCS scores. Of note, the regions that contained the most discriminative information were the left medial OFC and the left putamen for both scales. The Pearson correlation coefficient between predicted and observed scores was 0.49 ($p=0.002$) for total DY-BOCS and 0.44 ($p=0.006$) for Y-BOCS. Moreover, the Pearson correlation coefficient between predicted and observed scores for the "sexual/religious" dimension was 0.37 ($p=0.02$). The best predictors of the "sexual/religious" scores were the left medial OFC, right lateral OFC and left ACC. No statistically significant correlations were found between predicted and observed values for other dimensions.

4. Discussion

Herein we have applied a SVR analysis to predict OCD severity based on individual structural MRI scans obtained from treatment-naïve patients. The enrollment of only treatment-naive patients was essential to rule out the potential influence of previous treatments on brain morphometry. While previous MRI studies in OCD have used classification models to discriminate patients from controls (categorical outcome) (Soriano-Mas et al., 2007;

Table 1
Clinical characteristics of treatment-naïve OCD patients ($n=37$).

Variable	Mean ± SD (range)
Age, years	31.9 ± 10.1 (18–60)
Age of onset of OCS, years	13.4 ± 7.5 (5–35)
Illness duration, years	18.3 ± 10.5 (0–42)
Y-BOCS scores (total)	25.0 ± 5.2 (15–36)
DY-BOCS (total)	20.5 ± 4.5 (8–30)
Aggression/checking	5.6 ± 4.7 (0–12)
Sexual/religious	3.2 ± 4.7 (0–14)
Symmetry/ordering	7.6 ± 3.8 (0–14)
Contamination/washing	5.7 ± 5.1 (0–13)
Hoarding	2.9 ± 3.3 (0–10)

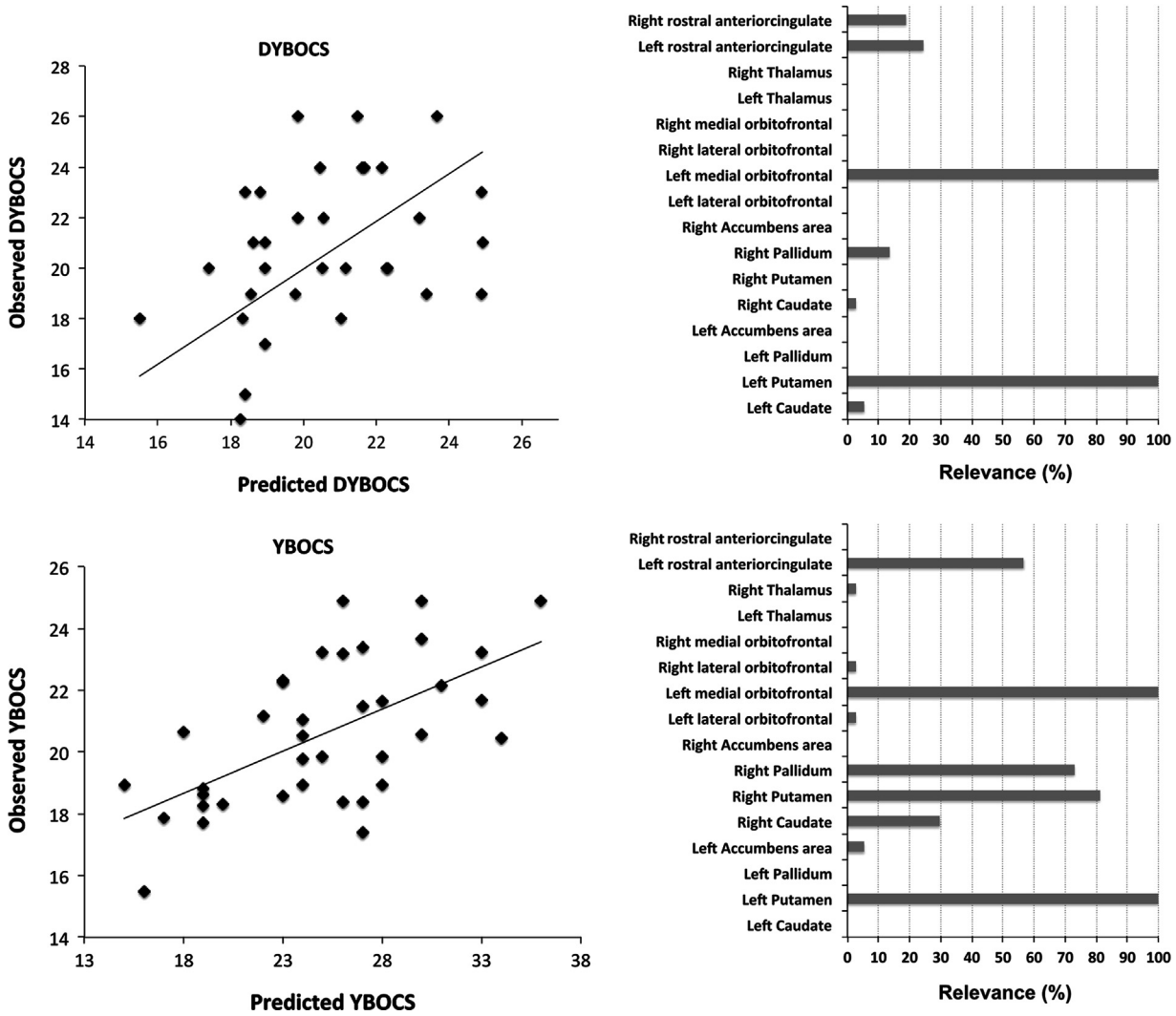

Fig. 1. Left: observed and predicted total-DYBOCS (Dimensional Yale-Brown Obsessive–Compulsive Scale) and total-YBOCS (Yale-Brown Obsessive–Compulsive Scale) based on support vector regression from volumetric data of orbitofrontal cortex, anterior cingulate cortex caudate, putamen, pallidum, accumbens, and thalamus at each hemisphere ($r=0.49$, p-value$=0.002$ for DY-BOCS and $r=0.44$, p-value$=0.006$ for Y-BOCS). Right: the relevance of each brain region to obtain the predicted total-DYBOCS and total-YBOCS values is shown in the bar charts. The total DY-BOCS scores vary from 0 to 30 and the total Y-BOCS scores from 0 to 40.

Weygandt et al., 2012), a methodological advantage of applying a regression method for this purpose is the possibility to identify neurobiological markers that predict continuous variables (Smola and Scholkopf, 1998), such as symptom severity. This is relevant, given that disorders such as OCD are clearly heterogeneous and may be understood as a spectrum of potentially overlapping vulnerabilities that do not establish an absolute, qualitative break between 'normality' and 'abnormality' (Miguel et al., 2005). Moreover, as opposed to non-linear methods, a linear kernel approach allows a straightforward quantification of the contribution of each brain region for obtaining the predicted values.

Our findings indicate that GM volumes within cortical–subcortical circuits, such as the OFC, ACC and striatum, contain discriminative information to predict OCD severity. Despite the contribution of several brain areas to the prediction of symptom severity, the regions that contained the most relevant information were the left medial OFC and left putamen. As expected, this was true for both total DY-BOCS and Y-BOCS (Rosario-Campos et al., 2006). Voxel-based morphometry (VBM) studies have consistently reported GM alterations within these structures in OCD at a group level (Radua and Mataix-Cols, 2009; Rotge et al., 2010). In a previous VBM study, we observed decreased GM volumes in bilateral medial OFC, left putamen and right ACC in treatment-naive OCD patients (same sample included herein, with the exception of one patient) compared with controls (Hoexter et al., 2012). The structures that contained the most relevant information to predict OCD symptom severity were the left medial OFC and left putamen.

GM volumes within cortical–subcortical circuits, mainly the left medial OFC, right lateral OFC and left ACC, also allowed the prediction of scores in the "sexual/religious" dimension. The presence of "sexual/religious" symptoms has been associated with refractoriness in OCD (Ferrao et al., 2006). Moreover, direct linear correlations between scores on the "sexual/religious" dimension and GM volumes within the right middle lateral OFC and bilateral ACC were found in our previous VBM study (Alvarenga et al., 2012). The scores of the "sexual/religious" dimension and total DY-BOCS were not correlated. Therefore, our results are specific to that dimension and not an artifact of the dimension impact on global severity. Notwithstanding, given that a considerable number of patients scored zero in several dimensions, this finding must be interpreted cautiously due to zero-weighting issues (Delucchi and Bostrom, 2004). This bias might also explain the lack of findings involving the severity prediction of other symptom dimensions.

Caution should be taken when interpreting our results due to other study limitations. Firstly, while the application of SVR

analysis methods to predict symptom severity based on individual MRI datasets aims to provide additional information about the neurobiological underpinnings related to the disorder and its varying clinical manifestations, the evaluation of symptom severity continues to be clinically assessed. Secondly, our sample is relatively small and our results must be replicated with independent and larger samples. Lastly, whole-brain analysis (rather than the region-of-interest approach) may provide important neurobiological clues about the involvement of different neural substrates not chosen a priori.

In conclusion, the use of a SVR analysis method has shown robust results in predicting OCD severity based on structural MRI of individual subjects. Reliable objective measures, obtained through neuroimaging, can help to identify neurobiological markers that quantify the severity of a disorder, as well as specific dimensions that are important to the characterization of subgroups within samples of sufferers from a specific mental disorder.

Role of funding source

This study received financial support in the form of Grants provided by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Foundation for the Support of Research in the State of Sao Paulo) to Dr. Miguel (2005/55628-8), to Dr. Shavitt (06/61459-7) and to Dr. Diniz (06/50273-0). Dr Hoexter was supported by a Ph.D. Grand from FAPESP (2005/04206-6) and by a doctorate "sandwich" scholarship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Agency for Support and Evaluation of Graduate Education) (4375/08-4).

Conflict of interest

Dr. Hoexter, Dr. Miguel, Dr. Diniz, Dr. Shavitt, Dr. Busatto and Dr. Sato have declared no conflict of interest.

Acknowledgments

The authors thank Fabio L.S. Duran, Marcelo C. Batistuzzo and Antonio C. Lopes for previous contribution. We also wish to thank all patients who participated in this investigation.

References

Alvarenga, P.G, do Rosário, M.C., Batistuzzo, M.C., Diniz, J.B., Shavitt, R.G., Duran, F.L., Dougherty, D.D., Bressan, R.A., Miguel, E.C., Hoexter, M.Q., 2012. Obsessive–compulsive symptom dimensions correlate to specific gray matter volumes in treatment-naive patients. Journal of Psychiatric Research 46, 1635–1642.

Delucchi, K.L., Bostrom, A., 2004. Methods for analysis of skewed data distributions in psychiatric clinical studies: working with many zero values. American Journal of Psychiatry 161, 1159–1168.

Denys, D., Burger, H., van Megen, H., de Geus, F., Westenberg, H., 2003. A score for predicting response to pharmacotherapy in obsessive–compulsive disorder. International Clinical Psychopharmacology 18, 315–322.

Ferrao, Y.A., Shavitt, R.G., Bedin, N.R., de Mathis, M.E., Carlos Lopes, A., Fontenelle, L. F., Torres, A.R., Miguel, E.C., 2006. Clinical features associated to refractory obsessive–compulsive disorder. Journal of Affective Disorders 94, 199–209.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B., 1997. Structured Clinical Interview for DSM-IV Axis I Disorders: Clinical Version (SCID CV). American Psychiatric Press, Washington, DC.

Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis.II: inflation, flattening, and a surface-based coordinate system. NeuroImage 9, 195–207.

Goodman, W.K., Price, L.H., Rasmussen, S.A., Mazure, C., Fleischmann, R.L., Hill, C.L., Heninger, G.R., Charney, D.S., 1989. The Yale-Brown obsessive compulsive scale. I. Development, use, and reliability. Archives of General Psychiatry 46, 1006–1011.

Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using support vector machines. Machine Learning 46, 389–422.

Hoexter, M.Q., Shavitt, R.G., D'Alcante, C.C., Cecconi, J.P., Diniz, J.B., Belotto-Silva, C., Hounie, A.G., Borcato, S., Moraes, I., Joaquim, M.A., Cappi, C., Sampaio, A.S., Mathis, M.A., Batistuzzo, M.C., Lopes, A.C., Rosa, A.C., Muniz, R.K., Marques, A.H., Santos, L.C., Taub, A., Duran, F.L., Dougherty, D.D., Busatto, G.F., Bressan, R.A., Miguel, E.C., 2009. The drug-naive OCD patients imaging genetics, cognitive and treatment response study: methods and sample description. Revista Brasileira de Psiquiatria 31, 349–353.

Hoexter, M.Q., de Souza Duran, F.L., D'Alcante, C.C., Dougherty, D.D., Shavitt, R.G., Lopes, A.C., Diniz, J.B., Deckersbach, T., Batistuzzo, M.C., Bressan, R.A., Miguel, E. C., Busatto, G.F., 2012. Gray matter volumes in obsessive–compulsive disorder before and after fluoxetine or cognitive-behavior therapy: a randomized clinical trial. Neuropsychopharmacology 37, 734–745.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D.S., Quinn, K., Sanislow, C., Wang, P., 2010. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. American Journal of Psychiatry 167, 748–751.

Mataix-Cols, D., Rauch, S.L., Manzo, P.A., Jenike, M.A., Baer, L., 1999. Use of factor-analyzed symptom dimensions to predict outcome with serotonin reuptake inhibitors and placebo in the treatment of obsessive–compulsive disorder. American Journal of Psychiatry 156, 1409–1416.

Miguel, E.C., Leckman, J.F., Rauch, S., do Rosario-Campos, M.C., Hounie, A.G., Mercadante, M.T., Chacon, P., Pauls, D.L., 2005. Obsessive–compulsive disorder phenotypes: implications for genetic studies. Molecular Psychiatry 10, 258–275.

Pereira, F., Mitchell, T., Botvinick, M., 2009. Machine learning classifiers and fMRI: a tutorial overview. NeuroImage 45, S199–S209.

Radua, J., Mataix-Cols, D., 2009. Voxel-wise meta-analysis of grey matter changes in obsessive–compulsive disorder. British Journal of Psychiatry 195, 393–402.

Rosario-Campos, M.C., Miguel, E.C., Quatrano, S., Chacon, P., Ferrao, Y., Findley, D., Katsovich, L., Scahill, L., King, R.A., Woody, S.R., Tolin, D., Hollander, E., Kano, Y., Leckman, J.F., 2006. The dimensional Yale-Brown obsessive–compulsive Scale (DY-BOCS): an instrument for assessing obsessive–compulsive symptom dimensions. Molecular Psychiatry 11, 495–504.

Rotge, J.Y., Langbour, N., Guehl, D., Bioulac, B., Jaafari, N., Allard, M., Aouizerate, B., Burbaud, P., 2010. Gray matter alterations in obsessive–compulsive disorder: an anatomic likelihood estimation meta-analysis. Neuropsychopharmacology 35, 686–691.

Sato, J.R., de Araujo Filho, G.M., de Araujo, T.B., Bressan, R.A, de Oliveira, P.P., Jackowski, A.P., 2012a. Can neuroimaging be used as a support to diagnosis of borderline personality disorder? An approach based on computational neuroanatomy and machine learning. Journal of Psychiatric Research 46, 1126–1132.

Sato, J.R., Kozasa, E.H., Russell, T.A., Radvany, J., Mello, L.E., Lacerda, S.S., Amaro Jr., E, 2012b. Brain imaging analysis can identify participants under regular mental training. PLoS One 7, e39832.

Saxena, S., Rauch, S.L., 2000. Functional neuroimaging and the neuroanatomy of obsessive–compulsive disorder. Psychiatric Clinics of North America 23, 563–586.

Smola, A.J., Scholkopf, B., 1998. On a kernel-based method for pattern recognition, regression, approximation and operator inversion. Algorithmica 22, 211–231.

Soriano-Mas, C., Pujol, J., Alonso, P., Cardoner, N., Menchón, J.M., Harrison, B.J., Deus, J., Vallejo, J., Gaser, C., 2007. Identifying patients with obsessive–compulsive disorder using whole-brain anatomy. NeuroImage 35, 1028–1037.

Weygandt, M., Blecker, C.R., Schafer, A., Hackmack, K., Haynes, J.D., Vaitl, D., Stark, R., Schienle, A., 2012. fMRI pattern recognition in obsessive–compulsive disorder. NeuroImage 60, 1186–1193.

de Almeida, J.R., Phillips, M.L., 2013. Distinguishing between unipolar depression and bipolar depression: current and future clinical and neuroimaging perspectives. Biological Psychiatry 73, 111–118.