Framing depression in a SN(a)Pshot

Woudstra, S.

2013

document version
Publisher's PDF, also known as Version of record

Chapter 5

Executive Function Related Prefrontal Recruitment as Predictor for Non-Remitted Outcome in Major Depression

Functional neuroimaging changes in MDD over 2 years

Saskia Woudstra
Marie-José van Tol
Nic J.A. van der Wee
Mark A. van Buchem
Esther M. Opmeer
André Aleman
Witte J.G. Hoogendijk
Dick J. Veltman

Under review

Abstract

Deficits in executive functioning may hamper full rehabilitation following a depressive episode in Major Depressive Disorder (MDD) and increase the risk for recurrence. However, whether executive abnormalities can predict unfavourable course or are a result of prolonged disease course is unclear.

Using longitudinal neuroimaging data acquired during a visuospatial planning task from 68 MDD patients (39 remitted and 29 non-remitted at two-year follow-up) and 45 healthy controls, we investigated whether abnormalities in executive function (EF) related brain activity in MDD 1) are a predictor for remission or non-remission of depression, 2) persist with non-remitting course or normalizes with stable remission at follow-up (2-years).

We found increased right dlPFC activity in MDD to be predictive of non-remission at two years follow-up. EF-related abnormal brain activity in MDD at baseline did not progress over time as a function of chronic course and was not associated with performance differences. Results were unaffected by presence of anxiety comorbidity, age of onset, duration and number of episodes prior to baseline, or depression severity.

Our results indicate that increased right dlPFC activity may distinguish MDD patients at risk for non-remission from those with a more favourable course. Since our study has a naturalistic design, future long-term follow-up functional neuroimaging treatment studies are necessary to investigate the role of increased dlPFC activity on treatment response prediction for improvement of individual-based treatment programs.

Keywords Major depressive disorder – functional neuroimaging – executive function – longitudinal data – recurrence – remission.

Introduction

Major depressive disorder (MDD) is a common, multifactorial disorder associated with significant morbidity, mortality, and public health costs. Next to experiencing depressed mood and/or loss of interest/pleasure, impaired ability to think and concentrate is one of the additional symptoms according to the Diagnostic and Statistical Manual of Mental Disorder (DSM-IV, TR), and relates to deficits in memory, attention, and executive functioning (EF; (Frodl *et al.*, 2006)). EF impairments are key factors affecting full social rehabilitation, increasing the risk of relapse (Gotlib & Hamilton, 2010).

A major feature of executive function is planning, the process of thinking about and organizing the steps required to achieve a desired goal. Studies in MDD have shown abnormal task performance and increased activity in task-related brain areas (Fitzgerald, Sr., 2008; Matsuo *et al.*, 2006; McClintock *et al.*, 2010; Rose *et al.*, 2006; van Tol *et al.*, 2011; Wagner *et al.*, 2006), using the Tower of London visuospatial planning task (ToL) and the n-back working memory task, although reduced activity has also been reported (Elliott *et al.*, 1997; Goethals *et al.*, 2005). These inconsistencies may be due to limited power, or differences in scanning modalities and clinical characteristics of patient groups, including medication use. In healthy controls, visuospatial planning has been consistently associated with activation of a dorsal prefrontal-parietal-striatal circuit (van den Heuvel *et al.*, 2003; van Tol *et al.*, 2011).

It is thought that executive dysfunction in MDD is a stable neuropsychological (i.e., trait) characteristic, present in both recurrent depression (Bhardwaj *et al.*, 2010; Reppermund *et al.*, 2009; Westheide *et al.*, 2007) and remission. In our previous work, however, we observed that task performance and increased recruitment of the left dlPFC during visuospatial planning was characteristic of current moderate to severe depression, and was not observed in the remitted state, consistent with recent work by (Maalouf *et al.*, 2011). However, these observations were made in a cross-sectional design, and it is unclear whether abnormal dlPFC recruitment in patients normalizes with remission, or if this abnormal dlPFC recruitment in the active phase may predict disease course, as has recently been suggested (Samson *et al.*, 2011). Support for this hypothesis is also found in studies investigating genetic or familiar risk for depression in association with abnormal dlPFC recruitment, which indicates that these abnormalities represent a premorbid vulnerability factor (Holmes *et al.*, 2010).

The aim of our study is two-fold. We aimed to investigate in a longitudinal design whether (1) abnormal EF-related brain activity in MDD is a state or trait related phenomenon and (2) constitutes a predictor for remission or non-remission of depression.

Methods

Participants

This study is part of the Netherlands Study of Depression and Anxiety (NESDA), a multisite longitudinal cohort study aimed to provide insight into the long-term course of depression and anxiety disorders in patients selected from primary care and mental health organizations (Penninx *et al.*, 2008). Participants in this study underwent extensive clinical assessment and a subgroup was additionally included in a Magnetic Resonance Imaging (MRI) study, on average within two months following clinical assessment in one of the three imaging centres

(Leiden University Medical Center (LUMC); Academic Medical Center (AMC) Amsterdam, or University Medical Center Groningen (UMCG). Participants in the present study underwent a neuroimaging protocol at two time points; at baseline (S1; described by (van Tol *et al.*, 2011)) and after two years (S2). At baseline, 301 participants were included in the NESDA neuroimaging study and 199 participants were scanned again at S2 (66.11%). 'Drop-outs' did not differ from 'included participants' regarding depression severity at baseline ($F(1,299)=1.15$; $P=0.29$). Reasons for dropping out included 'too busy', pregnancy, and MRI-contraindications including metal in body. To assess depressive symptom characteristics and severity scores, and depression severity at the day of scanning, the Inventory of Depressive Symptomatology (IDS; (Rush *et al.*, 1986)) and the Montgomery–Åsberg Depression Rating Scale (MADRS; (Montgomery & Asberg, 1979)) were used. Anxiety severity was assessed using the Beck Anxiety Inventory (BAI; (Beck *et al.*, 1988)). To establish current diagnosis, the Composite International Diagnostic Interview (CIDI; (Robins LN, 1988)) was administered followed within on average two months by the scanning sessions at both S1 and S2. Percentage of time with depressive symptoms between S1 and S2 were assessed using the life chart interview (Penninx *et al.*, 2011). Number and duration of previous episodes was determined by the number of depressive episodes prior to S1 and average months of these episodes prior to S1, respectively (self-report).

For the present report, we included participants with high-quality MR-scans at both S1 and S2. Patients had to meet criteria for current MDD (six-month recency) at S1, and HC were without lifetime DSM-IV diagnosis at S1 and S2, and not using any antidepressant or other psychotropic medication. Patients with only a diagnosis of anxiety disorder ($n=43$) at S1 or S2 were excluded from the current analysis. Exclusion criteria for patients were as previously reported (van Tol *et al.*, 2011): presence of MRI contraindications, DSM-IV axis I disorder other than MDD or common comorbid anxiety disorders (i.e. Panic Disorder, Social Anxiety Disorder, and Generalized Anxiety Disorder), dependence or recent abuse of alcohol and/or drugs, hypertension, major internal and/or neurological disorders and use of psychotropic medication (other than stable use of a selective serotonin reuptake inhibitor (SSRI) or incidental use of benzodiazepines, not within 48 hours before scanning). Exclusion criteria for healthy controls were: any DSM-IV disorder, presence of MRI contraindications, dependence or recent abuse of alcohol and/or drugs, hypertension, major internal and/or neurological disorders and use of psychotropic medication. Data from 30 participants were excluded because of incomplete data or poor data quality. In addition, similar to our cross-sectional baseline paper (van Tol *et al.*, 2011), task involvement was considered insufficiently reliable when overall performance was below 75% (on both measurements). Distribution of sufficient/insufficient performance was similar over groups. Data from bad performers were excluded from analysis ($n=7$). To obtain a good match on age, data of additional four HC and four patients were removed from further analysis. This resulted in a final sample of 113 participants, which were divided into three groups with the following criteria: non-remitted MDD patients (non-remMDD; $n=29$) with a MADRS score >10 at both time points; remitted MDD patients (remMDD; $n=39$) with a MADRS score >10 at S1 and <10 at S2; and HC ($n=45$) with a MADRS score <10 and a 'normal' BAI score (based on cut-off score of <10 (Beck *et al.*, 1988)), measured at S1 and S2 (table 5.1). The study was approved by the Ethics Committees

at the VU University Medical Center, AMC, LUMC, and UMCG and all participants provided written informed consent.

Task paradigm

We employed an event-related parametric version of the Tower of London planning task, described in detail elsewhere (van den Heuvel *et al.*, 2003; van Tol *et al.*, 2011). Briefly, while functional MRI data were collected, participants were asked to work out the minimum number of steps needed – ranging from one to five moves – to get from the start position to the goal position, which were both presented on a screen simultaneously. During baseline condition participants were asked to count the number of yellow and blue beads. Reaction times and proportion of correct answers were used as performance measures. This protocol was repeated at follow-up measurement, 2 years after baseline measurement.

Magnetic Resonance Image data acquisition

Functional neuroimaging data were acquired using Philips 3-Tesla magnetic resonance systems (Best, The Netherlands) located at the AMC, LUMC, and UMCG. In the AMC a SENSE-6 channel head coil was used at baseline measurement, and a SENSE-8 channel head coil was used at follow-up measurement. A SENSE-8 channel head coil was used for radio frequency transmission and reception in the LUMC and UMCG at both baseline and follow-up measurement. For each subject, echo-planar images (EPI) were obtained using a T2*-weighted gradient echo sequence with a repetition time (TR) of 2300ms, echo time (TE) of 28ms at UMCG and 30ms at AMC and LUMC, a flip angle of 90°, matrix size: 96x96 (64x64 voxels in UMCG at baseline measurement), 35 axial slices (39 slices in UMCG at baseline measurement), interleaved acquisition, 2.29x2.29mm in-plane resolution (3x3mm in Groningen at baseline measurement), and 3mm slice thickness. EPI images were scanned parallel to the anterior-posterior commissure plane. A T1-weighted anatomical MRI was also acquired for each subject (TR=9 ms, TE=3.5 ms, matrix size 256x256, voxel size 1x1x1 mm).

Data analysis

Task performance and clinical characteristics

Performance data (accuracy and reaction times) were analysed with IBM SPSS Statistics 20 (IBM Corporation, Armonk, New York, USA) by means of a repeated-measures analysis of variance (ANOVA), using the proportion correct scores and mean response times per trial type as dependent factors, group (non-remitted, remitted, and healthy controls) as between subject factor, and time as within-subject factor. Significance for behavioural analyses was set at $\alpha=.05$, and *post-hoc* paired tests were Bonferroni-corrected for multiple comparisons. For the clinical and demographic data we used analysis of variance (ANOVA), Chi-square and non-parametric tests as appropriate.

Image processing

Functional neuroimaging data were preprocessed and analysed using Statistical Parametric Mapping software (SPM5, http://www.fil.ion.ucl.ac.uk/spm/) implemented in Matlab

7.1.5 (MathWorks, Natick, MA, USA). Preprocessing included slice time correction, image realignment, registration of the T1-scan to the mean EPI, warping to MNI-space as defined by the SPM5 T1-template, reslicing to 3x3x3mm voxels and spatial smoothing using an 8mm full width at half maximum Gaussian kernel. Subject movement greater than 3mm in any direction resulted in exclusion of the data from further analysis. Following spatial preprocessing, data were analysed within the framework of the general linear model. Following the summary statistics approach, parametric contrast images for task load (with trial types 1-5 having weights [-1.5 -1 -0.5 1 2], see (van den Heuvel *et al.*, 2003; van Tol *et al.*, 2011)) were calculated per subject on a voxel-by-voxel basis and entered into second-level analyses. A 3 (group) by 2 (time) mixed ANCOVA was designed, with group as between-subject variable and time as within-subject variable. Scan centre, and head coil (by means of three dummy variables) were included as covariates.

The main effect of task is reported at a threshold of $P<.05$ whole-brain corrected for Family Wise Error (FWE; initial voxel wise threshold ($Z>2.97$). Based on our previous study, we chose the dorsolateral prefrontal cortex (dlPFC) as region of interest (ROI). To restrict the search for interaction effects to voxels which were identified in the main effect, we constructed a mask based on the main effects of the task from our previous study (van Tol *et al.*, 2011). The peak voxels from this main effect of increasing planning load calculated over 274 participants (left dlPFC, right dlPFC: MNI [-36,30,39] [33,24,48] respectively (van Tol *et al.*, 2011)) were used as the centres of 16-mm radius spheres to create a bilateral mask. This mask was used for Small Volume Correction (SVC) in the main comparisons (effect of diagnosis, and longitudinal effects) to protect against Type-I error. Effects in the resulting volumes of interest had to meet $P<.05$ FWE corrected for the spatial extent of the ROI, using an initial voxel-wise threshold of $Z=2.97$. Non-ROIs had to meet $P<.05$, whole-brain FWE corrected. All analyses were repeated after excluding SSRI using patients.

Results

Sample characteristics

Table 5.1 lists sample characteristics. Groups were matched on age, gender, handedness, comorbidity, and distribution of diagnosis over scan-centre. An effect of group * education was found ($F(2,110)=3.3$, $P=0.04$). *Post-hoc* t-tests showed that healthy controls were higher educated than non-remitted and remitted MDD patients ($T(111)=2.4$, $P=.016$), and patient groups were similarly educated ($t(66=-.76$, $P=.45$). SSRI use was not significantly different between patient groups at S1 or at S2 (S1: $c^2(1)=.02$, $P=.56$; S2: $c^2(1)=.08$, $P=.51$). No difference between patients was found for number of previous MDD episodes, duration of previous episodes (i.e. episodes prior to S1; # episodes: $F(1,66)=.69$, $P=.41$; duration: $F(1,65)=1.17$, $P=.28$) or age of onset ($F(1,63)=.13$, $P=.71$). Non-remitted patients showed a higher percentage of time with depressive symptoms between S1 and S2 than remitted patients ($F(1,66)=22.44$, $P<.001$). Patients showed higher MADRS, IDS, and BAI scores than healthy controls at S1 ($F\_all>8$, $P<.001$) and non-remitted patients showed higher MADRS score than remitted patients at S1 ($F(1,66)=9.2$, $P=.003$). Of note, no difference between patient groups was found for IDS and BAI scores at S1. As expected, at S2, differences in MADRS, IDS, or BAI score were found between patients and healthy controls ($F\_all>8$,

$P<.001$). In the patient group we found that non-remitted patients had higher MADRS and IDS, but not BAI scores than remitted patients (table 5.1).

Task Performance

Table 5.2 lists behavioural characteristics. All groups showed higher performance speed over time (table 5.2), but no group x time interaction effects were found ($F(2,110)=1.09$, $P=.342$). Groups performed similarly over time for planning accuracy ($F(2,110)=.57$, $P=.56$).

Imaging results

At S1, we found increased right dlPFC activity in non-remitted patients relative to remitted patients (MNI [27,24,36], $Z=3.64$, $P_{FWE}=0.03$) but the difference relative to HC failed to reach significance (MNI [27,21,39], $Z=3.11$, $P_{FWE}=0.15$; figure 5.1). No effects in the left dlPFC were observed, and no differences between groups were observed at S2 (table 5.1).

We found no correlation of depression severity with dlPFC EF-related brain activity at S2.

To ensure that the increased right dlPFC activity at baseline was not due to differences in depression severity, we added MADRS scores to the model, but this did not change the result (MNI [27,24,36], $Z=3.89$, $P_{FWE}=0.01$).

Longitudinal association

Group x time interaction analyses showed a significant decrease of MADRS, and IDS score over time in remitted patients relative to non-remitted patients (MADRS: $F(1,65)=19.19$; $P<.001$, IDS: $F(1,64)=7.6$; $P=.008$): Non-remitted patients showed a steeper decline of MADRS and IDS score than remitted patients (non-remitted patients: MADRS: $t(28)=2.6$, $P=.014$; IDS: $t(27)=3.2$, $P=.003$; remitted patients: MADRS: $t(37)=5.6$, $P<.001$; IDS: $t(37)=8.5$, $P<.001$; BAI: $t(37)=6.4$, $P<.001$). No effect of time * group was observed on BAI scores in the patient groups ($F(1,64)=3.17$; $P=.08$). SSRI use did not significantly decline over time (Ratio SSRI S1: SSRI S2: non-remMDD 7:6; remMDD 10:9).

All groups showed altered dlPFC activity over time (table 5.3), but no significant group x diagnosis x time interaction effects was shown. *Post-hoc* analyses showed no significant difference over time within groups (table 5.3).

Table 5.1: Demographic and clinical characteristics

		group (mean and SD)						
			Patients			MDD - HC	Rem - NR	between group
Variable		Total, n=68	Not remitted, n=29	In remission, n=39	Healthy controls, n=45	P-value [a]	P-value [b]	P-value [c]
Male:Female ratio		30:38	12:17	18:21	16:29	.24	.44	.61
Age, mean (SD) in years		39.1 (10.1)	41.3 (10.2)	37.4 (9.8)	37.0 (8.1)	.25	.11	.12
Right handed		61	28	33	42	.80	.11	.18
Scan site: A/L/G	S1	25/25/18	11/9/9	14/16/9	20/15/10	.71	.65	.81
	S2	27/24/17	12/9/8	15/15/9	17/19/9	.72	.81	.89
Education, mean (SD), y		12.6 (3.1)	12.2 (3.3)	12.8 (2.9)	14.0 (3.0)	**.020**	.45	**.041**
SSRI use	S1	17	7	10	N/A	N/A	.56	N/A
	S2	15	6	9	N/A	N/A	.51	N/A
Comorbidity at S1		39	20	19	N/A	N/A	.08	N/A
# previous episode		20.0 (120.7)	6.1 (7.5)	31.3 (161.4)	N/A	N/A	.41	N/A
duration of prev. episodes in months		19.1 (14.1)	21.3 (11.7)	17.5 (15.7)	N/A	N/A	.28	N/A
% depressive symptoms between S1 and S2		.35 (.30)	.52 (.27)	.22 (.25)	N/A	N/A	<.001	N/A
Age of onset		26.1 (11.8)	25.4 (12.6)	26.5 (11.2)	N/A	N/A	.71	N/A
MADRS, mean (SD)	S1	17.9 (9.3)	21.6 (9.1)	15.1 (8.5)	2.3 (3.7)	**<.001**	**.004**	**<.001**
	S2	11.04 (8.9)	15.6 (8.8)	7.6 (7.5)	1.5 (2.6)	**<.001**	**<.001**	**<.001**
IDS, mean (SD)	S1	30.6 (11.7)	33.2 (10.9)	29.6 (11.1)	8.5 (8.9)	**<.001**	.41	**<.001**
	S2	20.8 (10.9)	26.4 (10.4)	16.5 (9.4)	5.7 (4.2)	**<.001**	**<.001**	**<.001**
BAI, mean (SD)	S1	14.8 (9.4)	15.8 (9.8)	14.6 (8.9)	3.8 (6.1)	**<.001**	.80	**<.001**
	S2	9.6 (7.8)	12.4 (9.9)	7.6 (5.2)	1.9 (2.1)	**<.001**	**.012**	**<.001**

p-value [a]: MDD vs. HC; p-value [b]: remitted vs. nonremitted; p-value [c]: remitted vs. nonremitted vs. HC.
SD: standard deviation; MDD: major depressive disorder; HC: healthy controls; Rem: remitted patients; NR: non-remitted patients; n: number; y:years; m:months; scan site A/L/G: A: Amsterdam; L: Leiden; G: Groningen; MADRS score= total scale score all items ; IDS score = total scale score all items; BAI score = total scale score all items.

Table 5.2: Task performance

		group (mean and SD)						
			Patients			MDD - HC	Rem - NR	between group
Variable		Total, n=68	Not remitted, n=29	In remission, n=39	Healthy controls, n=45	P-value [a]	P-value [b]	P-value [c]
propBL	S1	.982 (.023)	.9768 (.02991)	.9852 (.02475)	.9843 (.02194)	.60	.21	.35
	S2		.9619 (.04579)	.9779 (.03429)	.9644 (.04615)		.10	.24
Prop step1	S1	.965 (.046)	.9577 (.04074)	.9697 (.04918)	.9706 (.04633)	.43	.29	.42
	S2		.9593 (.03626)	.9648 (.03994)	.9675 (.03234)		.56	.60
Prop step 2	S1	.944 (.060)	.9507 (.06475)	.9381 (.05703)	.9391 (.07585)	.79	.40	.71
	S2		.9342 (.05895)	.9511 (.06932)	.9408 (.06185)		.29	.55
Prop step 3	S1	.960 (.052)	.9529 (.06039)	.9653 (.04540)	.9370 (.08731)	.10	.34	.20
	S2		.9598 (.06053)	.9502 (.06120)	.9456 (.06410)		.53	.62
Prop step 4	S1	.844 (.147)	.8114 (.16691)	.8679 (.12737)	.8622 (.13003)	.34	.12	.15
	S2		.8161 (.14789)	.8461 (.11678)	.8263 (.12016)		.35	.62
Prop step 5	S1	.828 (.189)	.8194 (.21130)	.8335 (.17377)	.8022 (.15201)	.60	.76	.82
	S2		.8011 (.19229)	.8153 (.17017)	.7892 (.16220)		.75	.78
RT_BL	S1	3.91 (1.18)	3.87 (1.125)	3.93 (1.240)	3.47 (1.120)	.028	.82	.09
	S2		3.14 (1.011)	3.14 (.883)	2.79 (.704)		.96	.08
RT_step1	S1	5.02 (1.35)	5.02 (1.476)	5.02 (1.273)	4.49 (1.142)	.014	.99	.051
	S2		4.2526 (1.311)	4.301 (.916)	3.86 (.803)		.86	.055
RT_step2	S1	6.22 (1.56)	6.12 (1.518)	6.29 (1.604)	5.91 (1.940)	.17	.65	.35
	S2		5.46 (1.553)	5.50 (1.355)	5.00 (1.04)		.89	.074
RT_step3	S1	8.33 (2.23)	8.36 (2.445)	8.31 (2.095)	8.12 (3.212)	.48	.92	.77
	S2		7.433 (2.347)	7.36 (1.633)	6.77 (1.54)		.89	.066
RT_step4	S1	12.10 (3.83)	11.81 (4.228)	12.32 (3.548)	11.74 (4.309)	.50	.59	.70
	S2		9.43 (3.218)	10.15 (2.79)	9.40 (2.71)		.32	.25
RT_step5	S1	17.48 (5.38)	17.10 (5.493)	17.76 (5.342)	16.61 (4.438)	.40	.62	.61
	S2		14.263 (5.782)	12.50 (18.77)	13.35 (3.940)		.63	.83

p-value [a]: MDD vs. HC; p-value [b]: remitted vs. nonremitted; p-value [c]: remitted vs. nonremitted vs. HC. SD: standard deviation; MDD: major depressive disorder; HC: healthy controls; Rem: remitted patients; NR: non-remitted patients; n: number; prop: proportion good trials; RT: average reaction time (in seconds); BL: baseline condition; S1: time point 1; S2: time point 2. To account for the number of tests, we corrected the p-values using Bonferonni correction, resulting in a critical corrected p-value of P< .017.

Table 5.3: Longitudinal Characteristics

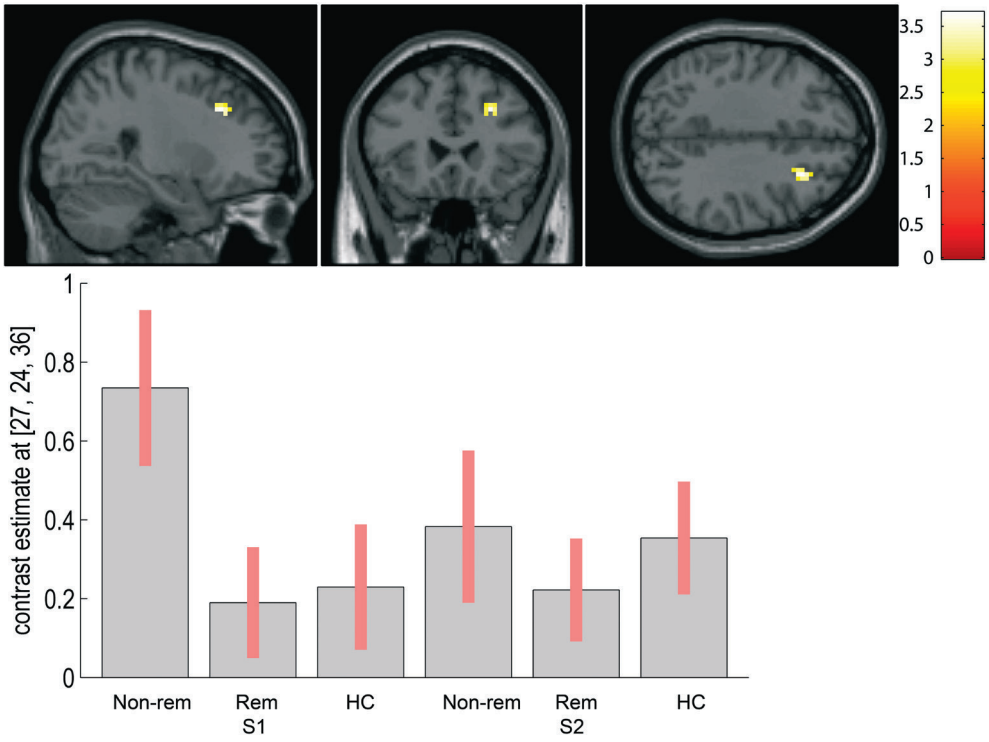
	group (mean and SD)							Within group		
		Patients			MDD - HC	Rem - NR	Between group	NR	Rem	HC
Variable	Total, n=68	Not remitted, n=29	In remission, n=39	Healthy controls, n=45	P-value [a]	P-value [b]	P-value [c]			
Δ Left dlPFC activity	0.008 (.47)	1.62 (5.06)	0.65 (14.34)	0.92 (6.61)	.33	.65	.92	.82	.66	.19
Δ Right dlPFC activity	0.051 (.49)	-3.85 (20.9)	-3.52 (15.16)	-.34 (3.31)	.37	.73	.47	.42	.72	.71
MADRS, mean (SD)	6.1 (11.8)	6.0 (12.4)	7.6 (8.2)	.97 (3.1)	**<.001**	**<.001**	**<.001**	**.014**	**<.001**	.066
IDS, mean (SD)	9.9 (10.1)	5.6 (10.3)	12.7 (9.2)	2.8 (7.6)	**<.001**	**.011**	**<.001**	**.003**	**<.001**	.019
BAI, mean (SD)	4.9 (7.8)	2.8 (8.9)	6.6 (6.4)	1.8 (5.1)	**<.001**	.13	**<.001**	.077	**<.001**	.025
Pr baseline	.01 (.04)	0.015 (.06)	0.007 (.04)	0.019 (.05)	.70	**.039**	.14	.16	.29	**.009**
Pr step 1	.002 (.06)	0.002 (.05)	0.005 (.07)	0.004 (.04)	.31	.24	.32	.86	.65	.55
Pr step 2	.0004 (.06)	0.017 (.07)	0.013 (.06)	0.003 (.06)	.85	.87	.97	.22	.19	.75
Pr step 3	.006 (.07)	0.007 (.09)	0.015 (.06)	0.008 (.09)	.12	.89	.29	.69	.14	.59
Pr step 4	.01 (.14)	0.005 (.17)	0.021 (.12)	0.038 (.12)	.61	.14	.26	.88	.25	**.034**
Pr step 5	.13 (.18)	0.018 (.13)	0.018 (.18)	0.15 (.14)	.52	.73	.76	.47	.53	.33
Rt baseline	.76 (89)	0.733 (.85)	0.781 (.93)	0.647 (.81)	.025	.87	.053	**<.001**	**<.001**	**<.001**
Rt step 1	.74 (.96)	0.769 (1.03)	0.722 (.93)	0.597 (.85)	**.008**	.93	.032	**<.001**	**<.001**	**<.001**
Rt step 2	.73 (1.32)	0.657 (1.32)	0.788 (1.34)	0.856 (1.31)	.049	.74	.14	**.012**	**.001**	**<.001**
Rt step 3	.93 (1.52)	0.931 (1.78)	0.944 (1.32)	1.375 (2.92)	.13	.90	.31	**.009**	**<.001**	**.003**
Rt step 4	2.25 (2.76)	2.383 (2.74)	2.16 (2.81)	2.421 (3.72)	.31	.42	.43	**<.001**	**<.001**	**<.001**
Rt step 5	4.22 (15.5)	2.84 (6.96)	5.26 (19.64)	3.509 (4.17)	.71	.78	.880	.036	.103	**<.001**

p-value [a]: MDD vs HC; p-value [b]: remitted vs nonremitted; p-value [c]: remitted vs nonremitted vs HC. SD: standard deviation; MDD: major depressive disorder; HC: healthy controls; Rem: remitted patients; NR: non-remitted patients; n: number; prop: proportion good trials; RT: average reaction time (in seconds); BL: baseline condition; S1: time point 1; S2: time point 2; activity: (S1-S2/S1). To account for the number of tests, we corrected the p-values using Bonferonni correction, resulting in a critical corrected p-value of P< .017.

Effects of SSRI use

To exclude possible confounding effects of SSRI treatment, we repeated our analyses with 90 SSRI-free participants. Groups were similar to the overall groups with respect to demographic and clinical characteristics. Omitting the SSRI users at both time-points did not change behavioural or performance results. The direction of the neuroimaging effects were similar, although the difference between remitted and non-remitted patients at S1 was observed subthreshold only (right dlPFC: MNI [27,24,36], Z=3.26, P_{FWE}=0.09).

Figure 5.1: Overview of group effect at S1 during increasing task load: Non-remitted MDD patients show increased activity in the right dlPFC relative to remitted MDD patients. The upper panel shows a sagittal, coronal, and axial slice. The lower panel shows the parameter estimate plot (arbitrary units). Non-rem: Non-remitted MDD patients; rem: remitted MDD patients; HC: healthy controls; S1: Time point 1; S2: Time point 2.



Discussion

In this longitudinal study, we tested whether abnormal EF-related brain activity in MDD constitutes a state or trait related phenomenon and whether it may predict recurrence or remission. Taken together, our results indicate that abnormal EF-related brain activity in MDD does not persist or become worse over time in recurrent or chronic MDD. Rather, our results indicate that EF abnormalities (right dlPFC activity) represent a state abnormality

of MDD. These specific EF abnormalities are predictive of a non-remitting course of MDD. EF abnormalities were not associated with performance differences and the observed longitudinal performance improvement within groups likely indicates a general training effect.

Our findings of increased right dlPFC activity in non-remitted MDD patients at baseline is in line with our previous study, in which we found increased dlPFC activity in MDD, although in the contralateral hemisphere (van Tol *et al.*, 2011). Increased dlPFC activity in MDD has been found before and may represent neural over-recruitment in order to perform at a similar level as healthy controls, a suggestion in line with well-known models of depression (Clark *et al.*, 2009; Drevets *et al.*, 2008; Mayberg, 1997; Phillips *et al.*, 2003). Our results indicate that over-recruitment of the *right* dlPFC has meaning for course prediction as it might potentially distinguish those MDD patients with an unfavourable outcome. Recruitment of the *left* dlPFC, as found in our previous study (van Tol *et al.*, 2011), may relate to current disease severity.

To our knowledge, this is the first longitudinal study that investigated neural correlates of executive function over the course of depression. As no evidence for persistence or deterioration of abnormal EF-related brain activity was observed, our findings may be reflective of neurodevelopmental abnormalities that predispose to the occurrence of depression, as suggested previously for neuroanatomical abnormalities (Woudstra et al. 2013 *submitted*, and (Amico *et al.*, 2011), and supported by other studies (Halari *et al.*, 2009; Jonassen *et al.*, 2012). Another hypothesis that might explain our findings is that EF abnormalities are associated with depression severity (McClintock *et al.*, 2010; Snyder, 2013; van Tol *et al.*, 2011), rather than predictive of non-remitted course of depression. Indeed, a difference in MADRS score was observed between non-remitted and remitted patients, as well as a difference in right dlPFC activity between non-remitted and remitted patients. This suggests a correlation between EF abnormalities and clinical severity at S1. Moreover, in the study of van Tol et al. (van Tol *et al.*, 2011) a correlation between planning activity and depression severity was found in the left dlPFC. This correlation was observed in MDD, but not in comorbid depression-anxiety patients, and was found to be specific for moderately to severely depressed MDD patients. Depression severity was determined by creating three clinically relevant subgroups based on MADRS score (remitted: MADRS 0-8; moderate: MADRS 9-18; severe: MADRS >18; (van Tol *et al.*, 2011)). Unfortunately, replication of this analysis in our sample was not possible due to the low number of MDD patients without comorbid anxiety per group (MDD-remitted: 8; MDD-mild: 9; MDD-severe: 8). An additional regression analyses of MADRS scores on EF related brain activity did not reveal any significant results. Moreover, repeating our analyses with depression severity as covariate did not change our results. Together these results indicate that the EF abnormalities we found are indeed predictive for non-remitting course of depression and not associated with depression severity.

Of note, we found non-remitted patients to show higher MADRS score than remitted patients at S1, whereas no difference was found for IDS and BAI. The MADRS and IDS are both questionnaires for measurement of depression severity. However, whereas the IDS is a self-report questionnaire, the MADRS is a structured clinical interview. Moreover, the IDS contains more anxiety items, which may explain the similarity with BAI scores.

The present study has a number of strengths. We obtained longitudinal functional neuroimaging data from a large sample (n=111). We recruited a large and representative

outpatients sample through the NESDA study, which made it possible to investigate the effects of depression at the neural level during a naturalistic course. Furthermore, by using an event-related fMRI design, we were able to decrease expectancy effects and to allow *post-hoc* selection of correct trials. Some limitations should also be noted. First, since the epidemiological NESDA cohort was recruited through general practitioners and outpatient clinics, we may not have been fully able to capture the most severe MDD patients. In addition, the NESDA study is an ongoing cohort study designed to investigate the naturalistic course and consequences of depressive and anxiety disorders and treatment intervention is not tested. Although we repeated our analyses omitting SSRI users (and found EF abnormalities at a subthreshold level, which likely reflects an effect of power), we were not able to investigate the role of SSRIs response and non-response in greater detail due to the study design. Future long-term follow-up functional neuroimaging treatment studies with sufficient power are needed to investigate the role of increased dlPFC activity in treatment response prediction (and comparison between SSRIs and cognitive behavioural therapy).

Second, remitted and non-remitted patients differed trend-wise from HC in years of education, although neither performance nor task-associated brain activity was significantly correlated with education level (data not shown). Third, due to the low number of HC who became depressed in the two-year interval (n=2), we were not able to study whether EF abnormalities were present before the onset of the disease.

Conclusion

In conclusion, our results indicate that EF abnormalities (right dlPFC activity) represent a state abnormality of MDD, and that such abnormalities are prognostic for a non-favourable naturalistic course of MDD. We found the EF abnormalities at S1 to be independent of anxiety comorbidity, age of onset, number of previous episodes, duration of previous episodes, and depression severity. Furthermore, our data suggest that over-recruitment of right dlPFC may distinguish those MDD patients at risk for an unfavourable outcome, which may be relevant for relapse prevention programs. This implication however, requires further investigation. As no evidence for persistence or deterioration of abnormal EF-related brain activity was observed, our findings may be reflective of abnormalities that predispose to the development of depression. Future studies using a longitudinal design that specifically investigates treatment response should elucidate the predictive value of increased dlPFC activity for treatment response in MDD to improve individual-based treatment programs.

Supplemental information

Supplementary table S5.1: Behavioural and clinical characteristics SSRI free patients

		group (mean and SD)						
		Patients				MDD - HC	Rem - NR	Between group
Variable		Total, n=45	Not remitted, n=18	In remission, n=27	Healthy controls, n=45	P-value [a]	P-value [b]	P-value [c]
Male:Female ratio		22:23	7:11	15:12	16:29	.20	.21	.24
Age, mean (SD), y		39.22 (9.63)	42.1 (9.9)	37.3 (9.1)	37.0 (8.1)	.238	.108	.11
Right handed		39	17	22	42	.29	.21	.21
Scan site: A/L/G	S1	16/15/14	6/5/7	10/10/7	20/15/10	.57	.64	.72
	S2	18/16/13	7/5/6	11/9/7	17/19/9	.47	.85	.76
Education, mean (SD), y		12.76 (3.21)	12,28(3,659)	13,07(2,895)	14.0 (3.0)	.058	.42	.12
Comorbidity at S1		22	10	12	N/A	N/A	.34	N/A
# previous episode		5.23 (7.37)	6.16 (7.97)	4.57 (7.01)	N/A	N/A	.49	N/A
duration of prev. episodes (m)		17.18 (13.25)	18,78 (10.51)	16.08 (14.95)	N/A	N/A	.51	N/A
% time with depressive symptoms S1 – S2		.33 (.31)	.51 (.30)	.21 (.25)	N/A	N/A	**.001**	N/A
age of onset		25.0 (10.6)	22.9 (11.5)	26.5 (9.9)	N/A	N/A	.28	N/A
MADRS, mean (SD)	S1	17.14 (8.41)	20.56 (6.9)	14.77 (8.6)	2.3 (3.7)	**<.001**	**.023**	**<.001**
	S2	9.89 (7.74)	14.0 (7.69)	7.15 (6.57)	1.5 (2.6)	**<.001**	**.003**	**<.001**
IDS, mean (SD)	S1	29.49 (11.62)	30.9 (9.87)	29.74 (11.36)	8.5 (8.9)	**<.001**	.72	**<.001**
	S2	19.64 (11.46)	25.83 (10.96)	16.19 (9.52)	5.7 (4.2)	**<.001**	**.003**	**<.001**
BAI, mean (SD)	S1	13.4 (8.55)	12.58 (6.21)	14.48 (9.44)	3.8 (6.1)	**<.001**	.47	**<.001**
	S2	8.82 (5.96)	10.22 (6.22)	7.85 (5.69)	1.9 (2.1)	**<.001**	.20	**<.001**

p-value a: MDD vs HC; p-value b: rem vs. nonremitted; p-value c: rem vs nonremitted vs HC. SD: standard deviation; MDD: major depressive disorder; HC: healthy controls; Rem: remitted patients; NR: non-remitted patients; n: number; y:years; m:months; scan site A/L/G: A: Amsterdam; L: Leiden; G: Groningen; MADRS score= total scale score all items ; IDS score = total scale score all items; BAI score = total scale score all items.