

Brain Activation During Emotional Memory Processing Associated with Subsequent Course of Depression

Hui Ai^{*1}, Esther M Opmeer¹, Dick J Veltman^{2,3}, Nic JA van der Wee^{4,5}, Mark A van Buchem^{5,6}, André Aleman^{1,7} and Marie-José van Tol¹

¹Department of Neuroscience, Neuroimaging Center, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; ²Department of Psychiatry, VU University Medical Center Amsterdam, Amsterdam, The Netherlands; ³Neuroscience Campus Amsterdam, VU University, Amsterdam, The Netherlands; ⁴Department of Psychiatry, Leiden University Medical Center, Leiden, The Netherlands; ⁵Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands; ⁶Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands; ⁷Department of Psychology, University of Groningen, Groningen, The Netherlands

Major depressive disorder (MDD) is characterized by a heterogeneous course and identifying patients at risk for an unfavorable course is difficult. Neuroimaging studies may identify brain predictors of clinical course and may help to further unravel the neurobiological processes underlying an unfavorable course. We investigated whether brain activation during an emotional memory paradigm is associated with depressive course. To this end, we followed 74 MDD patients and 45 healthy controls (HCs) for 2 years. At baseline, participants performed an emotional word-encoding and -recognition task during functional magnetic resonance imaging. Activation patterns were compared between patients with fast remission ($n=22$), remission with recurrence ($n=23$), non-remission ($n=29$), and HCs. Additionally, linear relations of brain activation and time to remission during the follow-up period were investigated across patients. We observed that during encoding of negative words, non-remitters showed higher activation of the left insula than HCs. Groups also differed in activation of the right hippocampus and left amygdala during negative encoding, with a trend for higher activation in non-remitters compared with HCs. Furthermore, hippocampal activation during negative word encoding was significantly and positively correlated with time to remission, irrespective of illness severity. Our findings suggest that higher activation in the left insula could serve as a neural marker of a naturalistic non-remitting course, whereas higher hippocampal activation is associated with delayed remission. Longitudinal analyses should clarify whether abnormal activation progresses further as a function of time with depression or may serve as load-independent markers of MDD course.

Neuropsychopharmacology (2015) **40**, 2454–2463; doi:10.1038/npp.2015.96; published online 29 April 2015

INTRODUCTION

Major depressive disorder (MDD) is the leading cause of disability worldwide and a substantial percentage of patients with MDD will not fully remit or will experience new episodes after remission (Simon *et al*, 2004). Such non-remitted course may predict an unfavorable chronic course of MDD (Rush *et al*, 2006; Trivedi *et al*, 2006). However, it is currently unclear which patients will show a favorable or unfavorable course. Studies on treatment response have indicated the importance of structure and functioning of the lateral prefrontal cortex (PFC) (Heller *et al*, 2013; Langenecker *et al*, 2007; Ritchey *et al*, 2011), medial PFC (Ritchey *et al*, 2011), anterior cingulate cortex (ACC) (Fu *et al*, 2013; Kemp *et al*, 2008; Pizzagalli, 2010), insula (Fu *et al*, 2013;

Langenecker *et al*, 2007; McGrath *et al*, 2013), amygdala (Langenecker *et al*, 2007), and hippocampus (Frodl *et al*, 2004; Fu *et al*, 2013) for predicting remission following treatment. However, treatment-response markers are not necessarily markers of a naturalistic course of depression. A naturalistic prospective investigation could elucidate the neurocognitive mechanisms mediating long-term episodic stages of depression (Frodl *et al*, 2008). Such markers might facilitate early detection and intervention to minimize cumulative effects (Clark *et al*, 2009).

Memory biases are an important aspect of the cognitive symptoms in MDD (Airaksinen *et al*, 2007; Ebmeier *et al*, 2006). Studies consistently point to enhanced memory for negative emotional information and diminished memory for positive emotional information (Bower, 1981; Gotlib *et al*, 2004; Leppänen, 2006; Rinck and Becker, 2005). It has been proposed that altered emotional memory formation may predict remission or the risk of recurrence (Pringle *et al*, 2011). Notably, it has been shown that patients who were going to improve 9 months later had an enhanced memory for positive stimuli at the start of the study than those who

*Correspondence: H Ai, Department of Neuroscience, Neuroimaging Center, University Medical Center Groningen, University of Groningen, Antonius Deusinglaan 2, Groningen 9713AW, The Netherlands, Tel: +31(0) 50 363 8792, Fax: +31(0)50 363 8875, E-mail: h.ai@umcg.nl
Received 3 September 2014; revised 10 February 2015; accepted 1 April 2015; accepted article preview online 10 April 2015

did not improve (Johnson *et al*, 2007). Moreover, activation of cortical midline structures during encoding of negative pictures has been found predictive of worsening of depressive symptoms in a small sample (Foland-Ross *et al*, 2014). However, the neural mechanisms associated with depressive course trajectories that relate to memory processing of both positive and negative materials have not been explored so far.

Previously, we observed a blunted response of the hippocampus (a key structure for memory processing) during positive memory encoding and an elevated insular response during negative memory encoding in depressed patients (van Tol *et al*, 2012). Because hippocampal and insular effects were present irrespective of current clinical status and therefore may represent trait markers associated with the course of MDD, we hypothesized that hippocampal and insular activation differentiates patients with a favorable course trajectory from patients with an unfavorable course trajectory.

The aim of this study was to investigate whether the neural correlates of memory processing are associated with the subsequent course of depression. Patients in a current depressive episode and healthy controls (HCs) were scanned using functional magnetic resonance imaging (fMRI). We hypothesized that patients who would not remit within 2 years would show better memory performance for negative words and worse memory performance for positive words at baseline compared with patients who would remit and HCs. On a neural level, we hypothesized that blunted hippocampal activation during positive word encoding and higher hippocampal, amygdalar, and insular activation during negative word encoding is associated with non-remitting course. Second, we aimed to more broadly explore whether the activation in regions previously related to treatment response (ie, ACC, lateral PFC, medial PFC, hippocampus, insula, and amygdala) during positive and negative word processing is associated with subsequent course.

MATERIALS AND METHODS

Participants

Participants were recruited from the ongoing longitudinal naturalistic Netherlands Study of Depression and Anxiety (NESDA), involving the University Medical Center Groningen (UMCG), Academic Medical Center (AMC), and Leiden University Medical Center (LUMC). The ethical review boards of each center approved this study and all participants gave written informed consent. Complete data at baseline (S1) was available for 215 participants, of which 110 patients had a half-year diagnosis of MDD based on the Composite International Diagnostic Interview (CIDI life time-version 2). We included only patients who showed depressive symptoms indicative of at least a mild depressive episode according to the Inventory of Depressive Symptomatology ($IDS > 13$; Rush *et al*, 1996) on the day of scanning and of whom longitudinal 2-year follow-up data (S2) on course was available ($n = 74$). During the follow-up period, patients received treatment as usual (or no treatment if wished by the patient). In addition, 45 HCs who did not have a current or lifetime diagnosis of a DSM-IV disorder were included. Detailed sample characteristics and exclusion criteria are described in Table 1 and Supplementary Materials and Methods.

Diagnostic status and depressive state were assessed at S1 and S2 (Table 1) with the CIDI and life chart method (LCM) (Lyketsos *et al*, 1994). The methodology of the LCM has shown to have high validity and reliability (Warshaw *et al*, 2001). Based on the LCM at S2, patients were divided into four trajectory groups. Patients who remitted within 1 year after S1 without recurrence in the follow-up period were defined as remitters (REM, $n = 22$). We defined remission as a period of 3 months without symptoms or symptoms without burden. Recurrence (REC, $n = 23$) was defined as recurrence of symptoms for at least 1 month with burden after firstly obtaining remission. Non-remission (NONREM, $n = 25$) was defined as experiencing symptoms with burden in every month following S1 during the entire follow-up period. Patients who obtained remission later than 1 year after S1 were defined as slow remitters ($n = 4$). We included slow remitters into NONREM for power reasons and followed up the results in NONREM by performing a sensitivity analysis leaving out these subjects. In total, we included 119 participants: 22 REM, 23 REC, 29 NONREM, and 45 HCs (Table 1). Furthermore, time to remission (months) was calculated for each patient and used as an index of course trajectory in the regression analyses.

Task Paradigm

All participants performed an event-related, subject-paced, emotional word-encoding and -recognition task (Daselaar *et al*, 2003; van Tol *et al*, 2012) that was presented using E-prime software (Psychological Software Tools, Pittsburgh, PA) during fMRI scanning. During encoding, participants classified 40 negative, 40 positive, and 40 neutral words according to their valence. Words were presented in pseudorandomized order. After a 10-min interval, these 120 words were presented again, mixed with 120 new emotional words, and participants had to indicate whether the word was seen during the encoding phase. A detailed description is in the Supplementary Materials and Methods.

fMRI Data Acquisition

fMRI data were collected with 3 T Philips MR scanners located at the three sites. A SENSE-8 channel head-coil was used in Groningen and Leiden and a SENSE-6 channel head-coil in Amsterdam. In Groningen, echo planar imaging volumes of 39 axial slices were acquired in interleaved ascending order (no gap) using a T2*-weighted gradient echo sequence (TR = 2300 ms, TE = 28 ms, matrix size = 64×64 , plane resolution = $3 \times 3 \text{ mm}^2$, slice thickness = 3 mm). Settings for Leiden and Amsterdam were slightly different: 35 slices, TR = 2300 ms, TE = 30 ms, matrix size = 96×96 , plane resolution = $2.29 \times 2.29 \text{ mm}^2$, slice thickness = 3 mm. Additionally, an anatomical MRI was obtained with a 3D gradient-echo T1-weighted sequence (TR = 9 ms, TE = 3.5 ms, matrix size = 256×256 , voxel size = $1 \times 1 \times 1 \text{ mm}^3$, 170 slices).

Data Analysis

Clinical variables and behavioral data. Demographic, psychometric assessment, and behavioral data were analyzed in SPSS v.16.0 (SPSS, Chicago, IL). For the demographic and

Table 1 Demographic Description

	HC	REM	REC	NONREM	F	χ^2	P-value
N	45	22	23	29			
Scan site (AMC/LUMC/UMCG) (N)	18/18/9	9/9/4	10/6/7	6/15/7	—	4.8	0.57
Sex (male/female) (N)	17/28	9/13	4/19	12/17	—	4.12	0.25
Age, mean (SD)	38.93 (9.69)	35.09 (8.94)	34.17 (11.65)	39.48 (10.26)	1.96	—	0.13 ^a
Years of education, mean (SD)	13.78 (2.60)	11.86 (2.29)	13.30 (2.89)	12.97 (2.82)	2.6	—	0.06 ^a
Interval months, mean (SD)	21.86 (1.4)	22.64 (1.5)	22.04 (1.8)	22.41 (1.24)	1.7	—	0.17 ^a
IDS_S1, mean (SD)	3.4 (3.1)	27.32 (8.94)	26.43 (5.74)	28.21 (7.74)	.35	—	0.70 ^b
IDS_S2, mean (SD)	4.93 (3.66)	14.32 (10.29)	21.22 (13.17)	25.17 (10.31)	5.84	—	0.004 ^{b,c}
BAI ^d _S1, mean (SD)	1.91 (2.29)	14.86 (9.49)	14.7 (8.78)	15.14 (9.15)	0.02	—	0.99 ^b
BAI_S2, mean (SD)	1.95 (2.28)	7.86 (7.88)	9.96 (8.32)	11.97 (7.76)	1.66	—	0.20 ^b
SSRI use_S1 (N)	—	9	7	10	—	0.55	0.76
SSRI use_S2 (N)	—	5	5	6	—	2.78	0.59
Psychotherapy use ^e _S1 (N)	—	17	17	25	—	1.31	0.52
Psychotherapy use between S1 and S2 (N)	—	9	6	18	—	7.52	0.023 ^c
Childhood trauma, ^f (N)	—	11	17	21	—	11.7	0.17
Comorbidity with anxiety (yes/no)_S1, (N)	—	13/9	13/10	21/8	—	1.66	0.44
Time to remission (months), mean (SD)	—	5.09 (2.04)	6.26 (3.88)	21.97 (2.21)	299.7	—	<0.001 ^c
Months with depression before S1, mean (SD)	—	21.07 (16.5)	14.29 (13.6)	21.35 (16.6)	1.15	—	0.32
Months with depression between S1 and S2 (%)	—	0.04 (0.08)	0.34 (0.18)	0.91 (0.15)	240.0	—	<0.001 ^c
Years since onset of depression_S1, mean (SD)	—	7.00 (6.8)	11.43 (11.5)	13.78 (11.1)	2.14	—	0.13

Abbreviations: AMC, Academic Medical Center; BAI, Beck Anxiety Inventory; HC, healthy controls; IDS, Inventory of Depressive Symptomatology; LUMC, Leiden University Medical Center; NONREM, non-remitted patients; REC, remitters with recurrence; REM, remitters; SSRI, selective serotonin reuptake inhibitor; S1, baseline measurement; S2, 2 years follow-up measurement; UMCG, University Medical Center Groningen.

^aComparisons of all four groups (include HCs).

^bComparisons of patients groups.

^cSignificant at $p < 0.05$.

^dBAI was used to assess the severity of anxiety (Beck et al, 1988).

^eIncludes mental health care, independent psychiatrist or psychotherapist, first-line psychologist, a social worker, or a social psychiatric nurse.

^fIncludes emotional neglect, psychological abuse, physical abuse, and sexual abuse before age of sixteen (measured by semistructured childhood trauma interview previously used in the Netherlands Mental Health Survey and Incidence Study, NEMESIS (De Graaf et al, 2004).

psychometric data, we used analyses of variance (ANOVA), χ^2 tests, and t -tests where appropriate with a significance level of $p < 0.05$ (after Bonferroni correction if appropriate).

For the behavioral data, reaction times (hits and false alarms), number of words classified according to valence and recognition accuracy (proportion hits and false alarms) (Tulving, 1985) were calculated. Main effects of group (4; HC, REC, REM, NONREM) and valence (3; positive, neutral, negative) and the interaction of group and valence were investigated with a repeated-measures ANOVA. In case a significant main effect or interaction effect was detected ($p < 0.05$), *post hoc t*-tests were conducted at a significant level of $p < 0.05$ (two-tailed) after Bonferroni correction.

Imaging data. A full description of preprocessing and modeling using Statistical Parametric Mapping software version 5 (<http://www.fil.ion.ucl.ac.uk/spm>) can be found in van Tol et al (2012) and in the Supplementary Materials and Methods. For each participant, we defined the following contrasts: [successfully encoded positive words > successfully encoded neutral words], [successfully encoded negative words > successfully encoded neutral words], [correctly recognized positive words > correctly recognized neutral

words], and [correctly recognized negative words > correctly recognized neutral words]. These contrasts were chosen to focus on valence specificity of both encoding and recognition phases. The number of error trials was too low (reported in van Tol et al, 2012) to test for proper memory effects (ie, [successfully encoded words > missed words], [successfully recognized words > missed words]).

Group Analyses. At second level, two four (group; HC, REM, REC, NONREM) by two (valence; positive > neutral and negative > neutral) flexible factorial models were set up separately for the encoding and recognition phase. In the first model, the main effect of group was assessed (across valence; model 1), and in the second model, the interaction between group and valence (model 2) was modeled. We set up these two models, because the sensitivity to detect interaction effects increases by leaving out the general differences between groups, which is not specific for the task (Gläscher and Gitelman, 2008), and thus main effects and interactions cannot be estimated in the same model. To test for valence-specific effects of group, we subsequently set up a full-factorial model, with group as between-subject factor and valence as within-subject factor. This model was

only used to test for group effects during either positive or negative encoding/recognition. Site was added by means of two dummy variables to each model, in addition to age and years of education. Analyses were repeated after excluding slow remitters ($n = 4$).

Consistent with our previous report (van Tol *et al*, 2012), main effects and interactions (F-tests) were explored at $p < 0.005$ uncorrected. *Post hoc t*-tests had to meet $p < 0.05$ familywise error (FWE) corrected at the voxel level for the spatial extent of *a priori* defined regions of interests (ROIs) (with an initial threshold of $p < 0.005$, uncorrected). To follow-up our previous results (van Tol *et al*, 2012), we created two sets of masks for testing main effects of trajectory group: the right hippocampus and left ACC for positive word encoding, and the right hippocampus, left amygdala, and left insula for negative word encoding. The regions were defined by the Anatomical Automatic Labeling (AAL) system (Tzourio-Mazoyer *et al*, 2002) implemented in the Wake Forest University Pick Atlas toolbox (Wake Forest University, Winston Salem, NC). We set significance for effects occurring in these regions at $p_{FWE} < 0.05$, corrected for the number of ROIs (ie, 2 for positive encoding; 3 for negative encoding). As signal in these regions are non-independent, we took their interdependency into account for calculating α -level by using the Simple Interactive Statistical Analysis Bonferroni tool (SISA Bonferroni; <http://www.quantitativeskills.com/sisa/calculations/bonfer.htm>). Correlation analysis showed a mean correlation of $r = 0.64$ between the signal in the left insula, left amygdala, and right hippocampus during negative word encoding as defined by the AAL label. Therefore, the threshold was set to $\alpha = 0.034$ to hold FWE-control during negative word encoding after SISA-Bonferroni adjustment. A mean correlation between the right hippocampus and left ACC was $r = 0.51$ in the corresponding AAL labels and the SISA-Bonferroni adjustment was set as $\alpha = 0.029$ to hold FWE-correction during positive encoding.

For exploring the effect in a broader set of regions previously associated with depression and treatment response (ie, bilateral ACC, lateral PFC, medial PFC, hippocampus, insula, and amygdala), we defined a separate set of *a priori* AAL masks. SISA-Bonferroni adjustment values was set for positive ($\alpha = 0.027$; mean $r = 0.65$) and negative encoding ($\alpha = 0.029$; mean $r = 0.69$) for these ROIs ($n = 6$). Effects occurring outside our ROIs had to meet $p < 0.05$ FWE, whole brain corrected.

Regression Analyses. To test the effect of illness trajectory within patients, we built two full factorial models for encoding and recognition with valence as factor (2; positive > neutral and negative > neutral) and time to remission as an interacting covariate with valence. Age, years of education, and site (two dummy variables) were added as covariates. We used the same set of *a priori* AAL masks as used for the explorative group analyses. SISA-Bonferroni adjustment values were set for positive ($\alpha = 0.028$; mean $r = 0.68$) and negative encoding ($\alpha = 0.031$; mean $r = 0.74$) for these ROIs ($n = 6$). Effects occurring outside our ROIs had to meet $p < 0.05$ FWE, whole brain corrected.

Next, we added illness severity at S1 (IDS scores) as a covariate to each model to test for its possible confounding

effect. Finally, to control for medication use, we repeated these analyses after excluding the selective serotonin reuptake inhibitor (SSRI) users. In addition, to control for a possible effect of psychotherapy use at S1, we repeated the analyses by adding it as a dummy covariate (yes/no).

RESULTS

Sample Characteristics

Sample characteristics are listed in Table 1. The four groups were comparable on age, sex, and years of education. The patient groups were comparable on SSRI use at S1 and S2, psychotherapy use at S1, childhood trauma, IDS scores at S1, Beck Anxiety Inventory scores at S1 and S2, comorbidity with anxiety at S1, years since onset of depression at S1, and months with depression in the 5 years before S1. In the follow-up period, NONREM received more psychotherapy than other patient groups, but pharmacological treatment was taken in similar extents. IDS scores at S2 in REM were significantly lower than in NONREM.

There was a significant group effect in time to remission: time to remission in REC and REM was shorter compared with NONREM. The REC and REM group did not differ in this.

No effect of trajectory group was observed on memory performance. Behavioral results are summarized in Figure 1 and Supplementary Table S1.

fMRI Results

Encoding. Encoding results are listed in Table 2A and Figure 2a (group analysis) and Table 3 and Figure 2b (regression analysis). During successful encoding, a main effect of group was observed in the right insula, right dorsolateral PFC (DLPFC), right parahippocampal gyrus, and right posterior cingulate cortex extending to the fusiform gyrus (model 1), indicating higher activation in NONREM compared with HCs. However, *post hoc t*-tests across valences did not survive correction for multiple comparisons in these areas.

Interactions of group and valence were observed in the left insula, right DLPFC, and right hippocampus (model 2), indicating higher activation in NONREM compared with HCs and REM during negative word encoding, but not during positive word encoding. We followed up these interactions with valence-specific group comparisons (within the full-factorial models). These confirmed that during negative word encoding, a main effect of group was present in the left insula (Figure 2A). *Post hoc t*-tests revealed that NONREM showed higher activation in this region than HCs and subthreshold higher activation than REM ($p_{FWE} = 0.24$), during negative word encoding. A similar pattern of insula activation was also found after excluding SSRI users (Supplementary Table S3a) and excluding the slow remitting patients from NONREM (Supplementary Table S4a). Adding illness severity at S1 (demeaned within group) or psychotherapy use at S1 as covariate did not affect the results ($Z = 3.85$, $p_{FWE} = 0.014$ for IDS; $Z = 3.66$, $p_{FWE} = 0.027$ for psychotherapy use).

Finally, trends of higher activation in the right hippocampus were found when comparing NONREM with REM

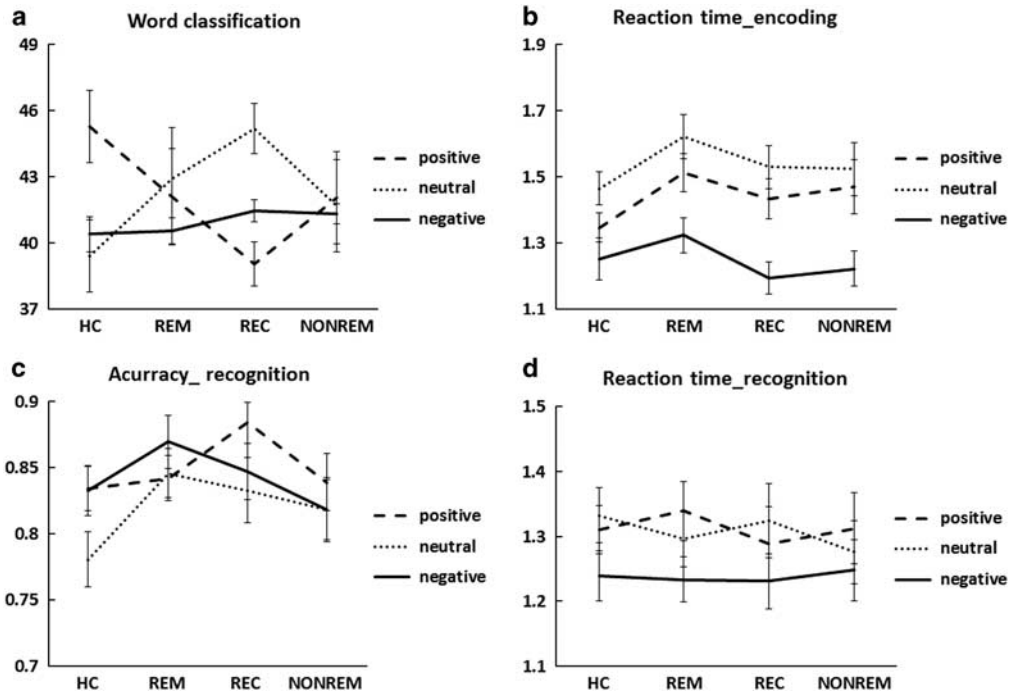


Figure 1 Behavioral data. (a) Word classification during encoding, Y-axis: number of words (b) reaction time (Encoding phase), Y-axis, seconds (c) recognition accuracy (Hits), Y-axis, proportion and (d) reaction time (Recognition phase) Y-axis, seconds. HC, healthy controls; NONREM, non-remitted patients; REC, remitters with recurrence; REM, fast remitters.

($p_{FWE}=0.07$) and HCs ($p_{FWE}=0.079$), and in the right amygdala comparing NONREM and HCs ($p_{FWE}=0.049$).

Multiple regression analysis revealed that activation of the right hippocampus was positively related to time to remission during negative word encoding. Adding IDS scores or psychotherapy use at S1 to the model did not change the result ($Z=3.97$, $p_{FWE}=0.01$ for IDS; $Z=3.94$, $p_{FWE}=0.01$ for psychotherapy use). After excluding the SSRI users ($n=24$), the effect was observed subthreshold ($Z=2.62$, $p_{uncorrected}<0.005$, $p_{FWE}=0.43$). To check *post hoc* whether SSRI use had a direct effect on hippocampal activation during negative encoding, we directly compared SSRI users with non-users. However, no effect was observed ($t=0.65$, $p=0.52$).

Recognition. Recognition effects are listed in Table 2B. During successful recognition, a main effect of group was observed in the left ventrolateral PFC (VLPFC) and interactions between group and valence were found in the left insula, left amygdala, and right ACC. Planned valence-specific comparisons revealed a main effect of group during negative word recognition in the left insula, left VLPFC, and right DLPFC. *Post hoc t*-tests did, however, not survive corrections for multiple comparisons in these areas.

No association was found between brain activation during recognition and time to remission.

DISCUSSION

In this study, we investigated whether regional brain activation during emotional encoding and recognition in depressive patients was associated with subsequent course

trajectory. We found that higher activation of the left insula during negative word encoding related to a non-remitting course. Groups also differed in activation of the right hippocampus and left amygdala during negative encoding, with a trend for higher activation in non-remitters compared with HCs. Moreover, higher hippocampal activation during negative word encoding was associated with delayed remission. Effects were unrelated to illness severity at baseline, although the association between time to remission and hippocampal activation was subthreshold only after exclusion of SSRI users. Memory performance or encoding behavior was not related to course. Taken together, these results indicate that insular and hippocampal activation during negative information processing may serve as neural markers of an unfavorable course in MDD.

Previously, in our study regarding the baseline fMRI measurement, we observed that elevated insula activation was associated with MDD, irrespective of current symptomatic state (van Tol *et al*, 2012). By including longitudinal data on the course trajectory, we could now add to this finding that insular function is indeed a neural marker of subsequent course of MDD that is not related to severity of depressive symptomatology. The insula has been associated with subjective awareness of negative feelings and the experience of visceral states following emotional events (Menon and Uddin, 2010; Singer *et al*, 2009). Higher activation of the left insula in non-remitters during negative word encoding might reflect increased sensitivity to negative stimuli (Surguladze *et al*, 2010). Such potentiation of the insula might contribute to cognitive symptoms such as difficulty in disengaging from negative information (Fu *et al*, 2013) and enhanced biased processing of negative information (Herwig *et al*, 2007), thereby hampering remission. Treatment studies have

Table 2A Main Effect of Group and Interaction Between Group and Valence During Successful Encoding

Regions	MNI coordinates								
	K ^a	Side	BA	x	y	z	F	Z	P-value
<i>Main effect of group</i>									
Insula	3	R	13	36	0	12	5.03	2.85	0.002
Middle prefrontal cortex	1	R	8	21	6	45	4.65	2.69	0.004
Posterior cingulate cortex	9	R	31	18	-39	36	6.84	3.54	<0.001
Parahippocampal gyrus, extends to fusiform gyrus	23	R	37	33	-39	-15	4.88	2.79	0.003
<i>Interaction</i>									
Insula	5	L	47	-42	15	-15	5.56	2.98	0.001
Middle prefrontal cortex	25	R	8	24	31	33	8.37	3.92	<0.001
Hippocampus	1	R	—	36	-27	-9	4.52	2.58	0.005
<i>Main effect of group_positive</i>									
Insula	13	R	13	36	-3	12	6.25	3.33	<0.001
Posterior cingulate cortex	1	L	31	-9	-15	45	4.54	2.64	0.004
<i>Main effect of group_negative</i>									
Insula	10	L	47	-42	12	-12	5.22	2.93	0.002
	K ^b	Side	BA	x	y	z	T	Z	P _{FWE}
<i>Post hoc t-tests</i>									
NONREM > HC (negative)									
Insula	29	L	47	-42	12	-12	3.87	3.80	0.017 ^c
Hippocampus	16	R	—	33	-12	-21	3.13	3.09	0.079
Amygdala	4	L	—	-18	-6	-15	2.80	2.77	0.049
NONREM > REM (negative)									
Hippocampus	4	R	—	33	-6	-18	3.16	3.12	0.07
Insula	4	L	13	-33	12	-12	2.87	2.84	0.24

Abbreviations: HC, healthy controls; NONREM, non-remitted patients; REC, remitters with recurrence; REM, remitters.

^aNumber of voxels in whole brain.

^bNumber of voxels in regions of interest.

^cSignificant at $p_{FWE} < 0.034$. SISA-Bonferroni adjustment values was set for positive ($\alpha = 0.029$) and negative encoding ($\alpha = 0.034$) for three ROIs across all participants (see Materials and Methods section).

suggested that higher insular activation is predictive of poor response to treatment (Fu *et al*, 2013) and insular metabolism might differentiate between responders to cognitive behavioral therapy or citalopram treatment (McGrath *et al*, 2013). Our study suggests that higher activation of the left insula during negative information processing may also relate to the naturalistic course of MDD. Thus, insular function may affect the odds of remission, but may also affect the odds that a patient responds to a specific treatment strategy. Future studies should therefore investigate putative mechanisms that can help explain how insular activation and metabolism may contribute to an unfavorable course.

In contrast to our hypothesis, course was not associated with posterior hippocampal activation during positive encoding, but to anterior hippocampal activation during negative encoding. Given the important role of the hippocampus in memory of emotional items (Dolcos *et al*, 2004) and relational

processing in memory encoding (Poppenk *et al*, 2013; Schacter and Wagner, 1999), our results suggest that MDD patients, who are likely to suffer from depressive symptom for longer time, tend to allocate more resources to encoding negative stimuli. This might suggest a 'potentiation' of negative information, which in turn may perpetuate course of depression and postpone remission. Moreover, the anterior hippocampus is richly interconnected with the amygdala, which has been found to facilitate memory processing of emotional stimuli by modulating hippocampal activation in healthy individuals (Disner *et al*, 2011). Adding to the finding of an association between small hippocampal volume and poor clinical response (Fu *et al*, 2013), our finding indicates that hippocampal activation during negative encoding might be a neural marker of prolonged course of depression. In addition, amygdalar activation has been found to correlate with hippocampal activation during mood-congruent

Table 2B Main Effect of Group and Interaction Between Group and Valence During Correct Recognition

Regions	MNI coordinates								
	K ^a	Side	BA	x	y	z	F	Z	P-value
<i>Main effect of group</i>									
Inferior frontal gyrus	8	L	10	-33	48	6	5.29	2.96	0.002
<i>Interaction</i>									
Insula	9	L	13	-36	-12	6	7.13	3.55	<0.001
Amygdala	2	L	-	-18	-6	-21	5.71	3.06	0.001
Medial frontal gyrus	3	R	10	15	57	0	5.55	3.00	0.001
<i>Main effect of group_positive</i>									
Middle frontal cortex	9	R	9	36	30	39	5.91	3.21	0.001
Medial frontal cortex	2	R	8	15	39	45	4.45	2.60	0.005
	K ^b	Side	BA	x	y	z	T	Z	P _{FWE}
<i>Post hoc t-tests</i>									
<i>NONREM > HC (positive)</i>									
Insula	42	R	13	48	3	0	3.15	3.11	0.19
	K ^a	Side	BA	x	y	z	F	Z	P-value
<i>Main effect of group_negative</i>									
Inferior frontal gyrus	2	L	10	-33	48	6	4.93	2.81	0.002
Middle frontal gyrus	6	R	9	24	39	39	5.28	2.96	0.002
Insula	2	L	13	-39	-15	18	4.58	2.66	0.004
	K ^b	Side	BA	x	y	z	T	Z	P _{FWE}
<i>Post hoc t-tests</i>									
<i>REM > NONREM (negative)</i>									
Insula	16	L	13	-36	-18	18	3.61	3.55	0.062
<i>HC > NONREM (negative)</i>									
Insula	7	L	13	-36	-12	6	2.94	2.91	0.34

Abbreviations: HC, healthy controls; NONREM, non-remitted patients; REC, remitters with recurrence; REM, remitters.

^aNumber of voxels—whole brain.

^bNumber of voxels in regions of interest.

SISA-Bonferroni adjustment values was set for positive ($\alpha=0.028$) and negative encoding ($\alpha=0.031$) for six ROIs (see Materials and Methods section).

memory encoding in MDD patients (Hamilton and Gotlib, 2008). Although we only found a trend of higher amygdalar activation, our results imply that hyperactivity in the anterior hippocampus might reflect an additional modulation of the amygdala during mood-congruent memory. We observed that hippocampal activation was unrelated to depressive severity, which implies that it could serve as a neural marker of depressive course independent of illness severity. Of note, associations were less strong after excluding the SSRI users from the analysis, which most likely reflects a substantial drop in power in this sensitivity analysis, as no significant effect of SSRI use on hippocampal activation was observed. We therefore conclude that the drop in significance is not a confounder of SSRI use on hippocampal activation *per se*, but

is due to the relative drop in load scores after excluding SSRI users.

Contrary to our expectation, no differences were observed in the PFC or ACC during emotional encoding or recognition, indicating that activation of these regions during this task may be less related to clinical course. Because the lateral PFC has been found to be hyperactivated during retrieval of mood-incongruent stimuli in symptomatic phases of MDD and after recovery (Van Wingen *et al*, 2010), our results suggest a trait-related rather than a course-predictive role of this region in memory processing. Activation of the medial PFC and posterior cingulate cortex during encoding of negative pictures has been associated with worsening of symptoms over an 18-month period (Foland-Ross *et al*, 2014). However, we focused on the

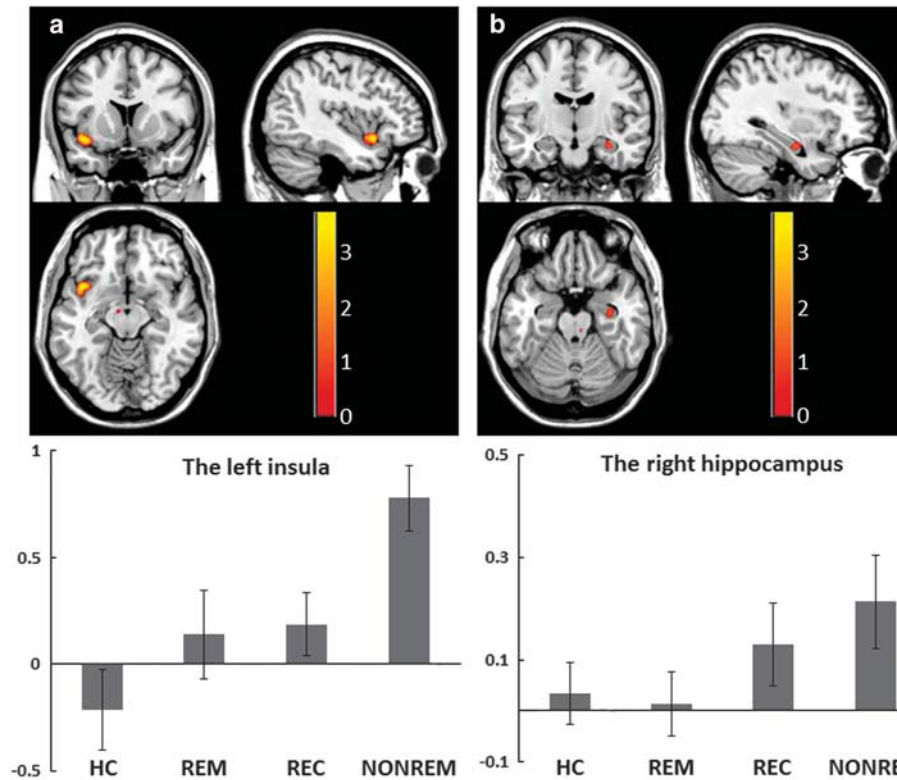


Figure 2 Brain activation during emotional memory task. (a) Higher activation of the left insula in patients with non-remission during negative word encoding in group analysis. (Contrast: NONREM>HC; effects are displayed at $T>2.60$, $p<0.005$ uncorrected). (b) Higher activation of the right hippocampus in patients with non-remission during negative encoding in regression analysis. (Contrast: NONREM>HC, effects are displayed at $T>2.60$, $p<0.005$ uncorrected). NONREM: non-remitted patients; REM: fast remitters; REC: remitters with recurrence; HC: healthy controls.

Table 3 Multiple Regression Analyses with Time to Remission as Predictor Across Patient Groups^a

Regions	MNI coordinates							β_{FWE}	
	K^b	Side	BA	x	y	z	T		Z
Hippocampus	27	R	—	33	-27	-12	4.03	3.91	0.012 ^c

^aCorrected for spatial extent of the region as defined by the AAL labels ROI (after SISA-Bonferroni correction for the number of regions considered (medial frontal gyrus, inferior frontal gyrus, amygdala, hippocampus, ACC, insula)).

^bNumber of voxels in regions of interest. ^cSignificant at $p_{FWE}<0.027$. SISA-Bonferroni-adjusted adjustment values was set for negative encoding ($\alpha=0.027$) for six ROIs after excluding all HCs (see Materials and Methods section).

trajectories describing the course of MDD over 2 years and did not take change in symptom severity at a certain timepoint as our end point, which might explain the differential association with cortical midline activation. Pretreatment pregenual-ACC activation has also been identified as a marker of treatment response in depression (Fu *et al*, 2013; Kemp *et al*, 2008; Pizzagalli 2010), although most studies focused on comparisons between responders and non-responders to short-term treatment during the processing of relatively simple tasks of emotional processing or during rest (reviewed by Pizzagalli 2010). The current results, however, do not support pregenual-ACC activation

during emotional memory as a potentially neural marker of naturalistic remission of depression.

Some limitations of our study should be noted. First, we calculated our results in the context of emotional word evaluation for the purpose of later recognition and could not contrast remembered vs forgotten words. Therefore, inferences on memory success with respect to forgotten words could not be drawn because of the relative lack of forgotten words (possibly caused by the short retention interval of only 10 min). Second, despite that we controlled for site effects by adding it as a covariate, different settings of data acquisition in the three sites could still confound the results, albeit a minor one, as no bias of group by site was present. Third, although patient groups did not differ in medication and psychotherapy use at the time of scanning, there could still be an effect of medication dose and frequency of treatment. Fourth, we only focused on the depressive course but did not control for the course of comorbid anxiety symptoms. Anxiety severity in patient groups, however, did not differ at baseline and follow-up measurement, suggesting this could not explain our findings. Fifth, the age-matched patient groups showed variations in their course of depression in the follow-up period, which were however not defined by the months with depression in the 5 years before S1. This could be explained by the comparable young and relative mild groups we included. Sixth, although the LCM has been reported to have high reliability and validity (Warshaw *et al*, 2001), this retrospective method could possibly be biased by current mood state. Finally, although we have over 20

subjects in each subgroup to get sufficient reliability (Thirion *et al*, 2007), subgroups were rather small and we did not correct the omnibus test for multiple comparisons. Therefore, replication of our naturalistic study is needed.

In conclusion, our results suggest that a prolonged course of depression is associated with higher activation in the left insula during negative memory processing, whereas a course with delayed remission is associated with higher hippocampal activation. Further longitudinal studies are necessary to clarify whether abnormal insular and hippocampal function change as a function of time with depression or may serve as 'load'-independent markers of MDD course.

FUNDING AND DISCLOSURE

The infrastructure for the NESDA study (www.nesda.nl) is funded through the Geestkracht program of the Netherlands Organization for Health Research and Development (Zon-Mw, grant no. 10-000-1002). This study is supported by the participating universities and mental health-care organizations: VU University Medical Center Amsterdam, University Medical Center Groningen, Leiden University Medical Center, GGZ inGeest, Arkin, GGZ Rivierduinen, Lentis, GGZ Friesland, GGZ Drenthe, Scientific Institute for Quality of Healthcare (IQ healthcare), Netherlands Institute for Health Services Research (NIVEL) and Netherlands Institute of Mental Health and Addiction (Trimbos Institute). Hui Ai, Esther Opmeer, Mark van Buchem, Dick Veltman, and Marie-José van Tol declare no conflict of interest. Nic van der Wee received speaking fees from Eli Lilly and Wyeth; and served on advisory panels of Eli Lilly, Pfizer, Wyeth and Servier. André Aleman received speaker fees from Lundbeck. All of these activities are not directly related to the present study and, therefore, do not form a conflict of interest. (don't insert this; take from acknowledgement and move to funding).

REFERENCES

- Airaksinen E, Wahlin Å, Forsell Y, Larsson M (2007). Low episodic memory performance as a premorbid marker of depression: evidence from a 3-year follow-up. *Acta Psychiatr Scand* **115**: 458–465.
- Beck AT, Steer RA, Carbin MG (1988). Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clin Psychol Rev* **8**: 77–100.
- Bower GH (1981). Mood and memory. *Am Psychologist* **36**: 129–148.
- Clark L, Chamberlain SR, Sahakian BJ (2009). Neurocognitive mechanisms in depression: implications for treatment. *Annu Rev Neurosci* **32**: 57–74.
- Daselaar S, Veltman D, Rombouts S, Raaijmakers J, Jonker C (2003). Neuroanatomical correlates of episodic encoding and retrieval in young and elderly subjects. *Brain* **126**: 43–56.
- De Graaf R, Bijl R, Ten Have M, Beekman A, Vollebergh W (2004). Rapid onset of comorbidity of common mental disorders: findings from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Acta Psychiatr Scand* **109**: 55–63.
- Disner SG, Beevers CG, Haigh EAP, Beck AT (2011). Neural mechanisms of the cognitive model of depression. *Nat Rev Neurosci* **12**: 467–477.
- Dolcos F, LaBar KS, Cabeza R (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron* **42**: 855–864.
- Ebmeier KP, Donaghey C, Steele JD (2006). Recent developments and current controversies in depression. *Lancet* **367**: 153–167.
- Foland-Ross LC, Hamilton P, Sacchet MD, Furman DJ, Sherdell L, Gotlib IH (2014). Activation of the medial prefrontal and posterior cingulate cortex during encoding of negative material predicts symptom worsening in major depression. *NeuroReport* **25**: 324–329.
- Frodl TS, Koutsouleris N, Bottlender R, Born C, Jäger M, Scupin I *et al* (2008). Depression-related variation in brain morphology over 3 years effects of stress? *Arch Gen Psychiatry* **65**: 1156–1165.
- Frodl T, Meisenzahl EM, Zetsche T, Höhne T, Banac S, Schorr C *et al* (2004). Hippocampal and amygdala changes in patients with major depressive disorder and healthy controls during a 1-year follow-up. *J Clin Psychiatry* **65**: 492–499.
- Fu CH, Steiner H, Costafreda SG (2013). Predictive neural biomarkers of clinical response in depression: a meta-analysis of functional and structural neuroimaging studies of pharmacological and psychological therapies. *Neurobiol Dis* **52**: 75–83.
- Gläscher J, Gitelman D (2008). Contrast weights in flexible factorial design with multiple groups of subjects. Unpublished tutorial. (<http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0803&L=SPM&P=R16629>).
- Gotlib IH, Kasch KL, Traill S, Joormann J, Arnow BA, Johnson SL (2004). Coherence and specificity of information-processing biases in depression and social phobia. *J Abnorm Psychol* **113**: 386–398.
- Hamilton JP, Gotlib IH (2008). Neural substrates of increased memory sensitivity for negative stimuli in major depression. *Biol Psychiatry* **63**: 1155–1162.
- Heller AS, Johnstone T, Peterson MJ, Kolden GG, Kalin NH, Davidson RJ (2013). Increased prefrontal cortex activity during negative emotion regulation as a predictor of depression symptom severity trajectory over 6 months. *JAMA Psychiatry* **70**: 1181–1189.
- Herwig U, Kaffenberger T, Baumgartner T, Jäncke L (2007). Neural correlates of a 'pessimistic' attitude when anticipating events of unknown emotional valence. *NeuroImage* **34**: 848–858.
- Johnson SL, Joormann J, Gotlib IH (2007). Does processing of emotional stimuli predict symptomatic improvement and diagnostic recovery from major depression? *Emotion* **7**: 201–206.
- Kemp AH, Gordon E, Rush AJ, Williams LM (2008). Improving the prediction of treatment response in depression: integration of clinical, cognitive, psychophysiological, neuroimaging, and genetic measures. *CNS Spectr* **13**: 1066–1086.
- Langenecker SA, Kennedy SE, Guidotti LM, Briceno EM, Own LS, Hooven T *et al* (2007). Frontal and limbic activation during inhibitory control predicts treatment response in major depressive disorder. *Biol Psychiatry* **62**: 1272–1280.
- Leppänen JM (2006). Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. *Curr Opin Psychiatry* **19**: 34–39.
- Lyketsois CG, Nestadt G, Cwi J, Heithoff K (1994). The Life Chart Interview: a standardized method to describe the course of psychopathology. *Int J Methods Psychiatr Res* **4**: 143–155.
- McGrath CL, Kelley ME, Holtzheimer PE, Dunlop BW, Craighead WE, Franco AR *et al* (2013). Toward a neuroimaging treatment selection biomarker for major depressive disorder: treatment-specific biomarker for major depression. *JAMA Psychiatry* **70**: 821–829.
- Menon V, Uddin LQ (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct* **214**: 655–667.
- Pizzagalli DA (2010). Frontocingulate dysfunction in depression: toward biomarkers of treatment response. *Neuropsychopharmacology* **36**: 183–206.

- Poppenk J, Evensmoen HR, Moscovitch M, Nadel L (2013). Long-axis specialization of the human hippocampus. *Trends Cogn Sci (Regul Ed)* **17**: 230–240.
- Pringle A, Browning M, Cowen PJ, Harmer CJ (2011). A cognitive neuropsychological model of antidepressant drug action. *Prog Neuro-Psychopharmacol Biol Psychiatry* **35**: 1586–1592.
- Rinck M, Becker ES (2005). A comparison of attentional biases and memory biases in women with social phobia and major depression. *J Abnorm Psychol* **114**: 62–74.
- Ritchev M, Dolcos F, Eddington KM, Strauman TJ, Cabeza R (2011). Neural correlates of emotional processing in depression: changes with cognitive behavioral therapy and predictors of treatment response. *J Psychiatr Res* **45**: 577–587.
- Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH (1996). The inventory of depressive symptomatology (IDS): psychometric properties. *Psychol Med* **26**: 477–486.
- Rush A, Trivedi M, Wisniewski S, Nierenberg A, Stewart J, Warden D et al (2006). Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry* **163**: 1905–1917.
- Schacter DL, Wagner AD (1999). Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus* **9**: 7–24.
- Simon GE, Fleck M, Lucas R, Bushnell DM (2004). Prevalence and predictors of depression treatment in an international primary care study. *Am J Psychiatry* **161**: 1626–1634.
- Singer T, Critchley HD, Preuschoff K (2009). A common role of insula in feelings, empathy and uncertainty. *Trends Cogn Sci (Regul Ed)* **13**: 334–340.
- Surguladze SA, El-Hage W, Dalgleish T, Radua J, Gohier B, Phillips ML (2010). Depression is associated with increased sensitivity to signals of disgust: a functional magnetic resonance imaging study. *J Psychiatr Res* **44**: 894–902.
- Thirion B, Pinel P, Mériaux S, Roche A, Dehaene S, Poline J (2007). Analysis of a large fMRI cohort: statistical and methodological issues for group analyses. *NeuroImage* **35**: 105–120.
- Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L et al (2006). Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry* **163**: 28–40.
- Tulving E (1985). Memory and consciousness. *Can Psychol* **26**: 1–12.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N et al (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* **15**: 273–289.
- van Tol M, Demenescu LR, van der Wee Nic JA, Kortekaas R, Marjan MAN et al (2012). Functional magnetic resonance imaging correlates of emotional word encoding and recognition in depression and anxiety disorders. *Biol Psychiatry* **71**: 593–602.
- Van Wingen GA, Van Eijndhoven P, Cremers HR, Tendolkar I, Verkes RJ, Buitelaar JK et al (2010). Neural state and trait bases of mood-incongruent memory formation and retrieval in first-episode major depression. *J Psychiatr Res* **44**: 527–534.
- Warshaw MG, Dyck I, Allsworth J, Stout RL, Keller MB (2001). Maintaining reliability in a long-term psychiatric study: an ongoing inter-rater reliability monitoring program using the longitudinal interval follow-up evaluation. *J Psychiatr Res* **35**: 297–305.

Supplementary Information accompanies the paper on the Neuropsychopharmacology website (<http://www.nature.com/npp>)