Amygdala reactivity, antidepressant discontinuation and relapse: a longitudinal, observational study with a randomized component

Tore Erdmann, PhD[1], Isabel Berwian, PhD[2,3], Klaas Enno Stephan, MD, PhD[3,4], Erich Seifritz, MD[5], Henrik Walter, MD, PhD[6,*], and Quentin JM Huys, MD, PhD[1,3,5,*]

[1] Applied Computational Psychiatry Lab, Mental Health Neuroscience Department, Division of Psychiatry and Max Planck Centre for Computational Psychiatry and Ageing Research, Queen Square Institute of Neurology, UCL

[2]Princeton Neuroscience Institute & Psychology Department, Princeton University, Princeton, USA

[3]Translational Neuromodeling Unit, University of Zurich and Swiss Federal Institute of Technology in Zurich, Zurich, Switzerland

[4]Max Planck Institute for Metabolism Research, Cologne, Germany

[5]Department of Adult Psychiatry and Psychotherapy, Hospital of Psychiatry, University of Zurich, Zurich, Switzerland

[6]Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Psychiatry and Psychotherapy [*]Equal contribution.

April 18, 2024

Corresponding Author
Quentin JM Huys
Max-Planck UCL Centre for Computational Psychiatry and Ageing Research
Russell Square House, 10-12 Russell Square, WC1B 5EH, London, UK
Email: q.huys@ucl.ac.uk

Key Points

Question Does antidepressant discontinuation increase amygdala reactivity to aversive stimuli and does this increase the risk of a depression relapse?

Findings Discontinuation of antidepressant medication increases amygdala response to negative facial expressions in individuals who go on to relapse. The increase is predictive of the risk of relapse.

Meaning The modulation of amygdala reactivity by antidepressant medications may represent a mechanism by which antidepressant medications help to maintain remission, and how antidepressant discontinuation increases is associated with relapse risk.

Abstract

Importance Antidepressant discontinuation substantially increases the risk of a depression relapse. The neurobiological mechanisms through which this happens are not known. Amygdala reactivity to negative information is a marker of negative affective processes in depression that is reduced by antidepressant medication. However, it is unknown whether amygdala reactivity is sensitive to antidepressant discontinuation, and whether any change is related to the risk of relapse after antidepressant discontinuation.

Objective To investigate whether amygdala reactivity to negative facial emotions changes with antidepressant discontinuation and relates to subsequent relapse.

Design The AIDA study was a longitudinal, observational AIDA study, where patients underwent two task-based fMRI measurements of amygdala reactivity. Patients were randomized to discontinuing antidepressants either before or after the second fMRI measurement. Relapse was monitored over a six-month follow-up period. Study recruitment took place until January 2018. Data were collected between July 1, 2015, to January 31, 2019 and statistical analyses were conducted between June 2021 and December 2023.

Setting University setting in Zurich, Switzerland, and Berlin, Germany.

Participants Patients with remitted major depressive disorder (rMDD) on antidepressants. Of 123 recruited patients, 83 (mean (SD) age 35.4 (11.4) years; 62 women (75%) were included in analyses. Of 66 recruited healthy controls matched for age, sex, and education, 53 were included in analyses (mean (SD) age 34.9 (10.7) years); 37 women (70%)).

Exposure Discontinuation of antidepressant medication.

Outcomes Task-based fMRI measurement of amygdala reactivity and MDD relapse within 6 months after discontinuation.

Results Amygdala reactivity of rMDD patients on medication initially did not differ from controls ($t = 0.33, p = 0.74$). An increase in amygdala reactivity after antidepressant discontinuation was associated with depression relapse (three-way interaction between group (continuation vs discontinuation), time point and relapse; $\beta = 18.9$, 95%-CI $(0.8, 37.1)$, $p = 0.041$). Amygdala reactivity change was associated with shorter times to relapse (hazard ratio 1.05, 95%-CI $(1.012, 1.094)$), and predictive of relapse (LOOCV balanced accuracy 67%, 95%-PPI $(53\%, 80\%)$, $p_{cor} = 0.018$).

Conclusions and Relevance An increase in amygdala reactivity is associated with risk of relapse after antidepressant discontinuation and may represent a functional neuroimaging marker that could inform clinical decisions around antidepressant discontinuation.

Introduction

Major depressive disorder (MDD) is a major cause of disability globally, affecting more than 16% of adults during their lifetime. Much of its burden arises through its high rate of recurrence [1]. More than half of patients with a first episode of depression experience a second episode and the risk of relapse increases further with every additional experienced episode [2]. Therefore, prevention of relapse is important. Indeed, relapse risk in depression has been studied and some promising mechanisms identified [3, 4], though most risk factors associated with relapse are prognostic rather than prescriptive [5].

One frequent, and clinically highly relevant, decision regarding the management of relapse risk is the decision whether to continue or discontinue antidepressant medication (ADM). ADM discontinuation confers a substantial increase in relapse risk [6, 7], but ADMs are often discontinued due to patient preference or other clinical reasons. Guidelines typically recommend 6-9 months of treatment after a first episode, and longer after more episodes, although the evidence for these recommendations is equivocal [8–10] and based on assumptions about the natural course of depressive episodes [11]. In this situation, factors—particularly mechanistically interpretable ones—that can predict which patients may be at risk of relapse and may thus be benefiting the most from continued treatment would be helpful.

However, the mechanisms leading specifically to relapse after ADM discontinuation are not well understood [12]. Previous work has shown that the predictive power of demographic and clinical variables is limited [12, 13]. Cognitive measures, such as behavioural assessments of effort-sensitivity, have recently been found to be predictive of relapse after ADM discontinuation [14], with evidence emerging also for EEG [15] and possibly resting-state fMRI measures [16].

Neurobiologically, a highly promising process is amygdala reactivity to negative affective stimuli. Theories of depression and ADM treatment effect delays have considered it a marker of negative affective bias, which denotes the tendency to allow negative experiences to have a greater effect on one's psychological state than neutral or positive ones [17–20]. Negative affective bias, as measured by the amygdala activity in response to negative emotional faces, is thought to track the course of depression, being heightened in people in the acute depressive phase ([21–24]; though note [25]). Amygdala reactivity to negative stimuli is attenuated by tryptophan deletion; by both acute and repeated SSRI administration in healthy individuals [26–28]; by emotion regulation interventions [29, 30] relevant to the treatment of depression; and by ADM treatment [21, 31]. There is good meta-analytic evidence that amygdala reactivity to negative emotions reduces or normalizes with ADM treatment [23, 32–35], and studies suggest that pre-treatment amygdala reactivity may be predictive of ADM treatment response [32]. Importantly, amygdala reactivity in the Hariri faces task has been extensively studied with evidence on its test-retest reliability [36] and inclusion in large-scale imaging datasets [25].

Here, we examine whether amygdala reactivity is affected by antidepressant discontinuation, and whether it has potential as a predictive biomarker. We report results from the AIDA (AntIDepressiva Absetzstudie) study, a longitudinal, observational study where patients were tested before and shortly after ADM discontinuation and followed up for six months to assess relapse. We employed a well-established functional magnetic resonance (fMRI) paradigm that examines the blood oxygen level dependent (BOLD) signal in the amygdala in response to facial emotion stimuli [37]. In keeping with the literature outlined above, we firstly expected that remitted patients before discontinuation would not differ from the control sample. Secondly, we expected that amygdala reactivity would increase with discontinuation, reflecting the converse of the established changes in response to ADM treatment [32, 34]. Thirdly, we expected that the increase in negative affective bias due to discontinuation would be related to the relapse risk. We conducted exploratory analyses to examine whether pre-treatment amygdala reactivity, or its change with discontinuation, might have potential as predictors for relapse risk after discontinuation.

Methods

Participants

The AIDA study recruited patients in remitted depression intent on discontinuing their antidepressant medication. Patients had experienced multiple or at least one severe [38] episode of Major Depressive Disorder [39]; had initiated antidepressant treatment during the last episodes; had reached a stable remitted state; and had reached the decision to discontinue their medication independently from and prior to study participation.

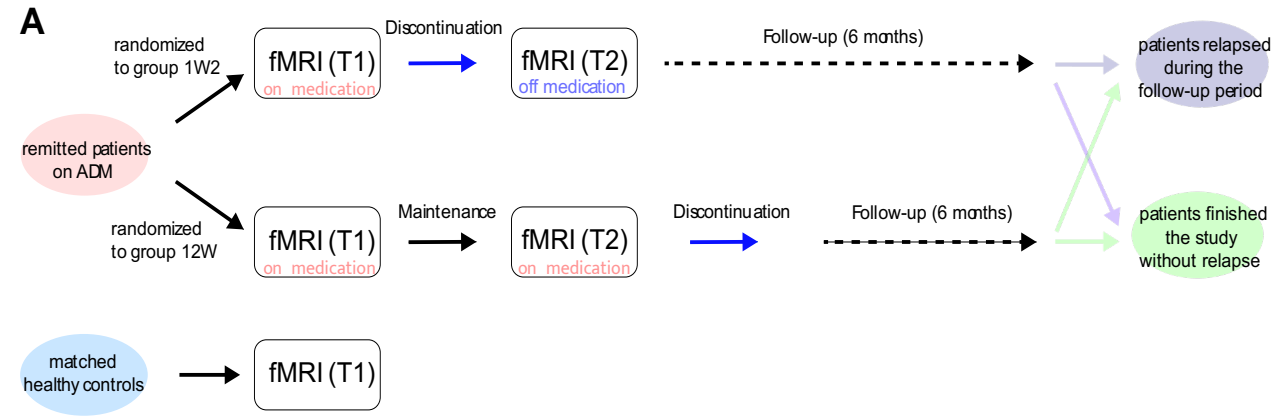
See supplement section S2 for inclusion and exclusion criteria. Healthy control participants without a history of depression were matched for age, sex, and educational level. Participants were recruited in two university settings in Zurich, Switzerland, and Berlin, Germany.

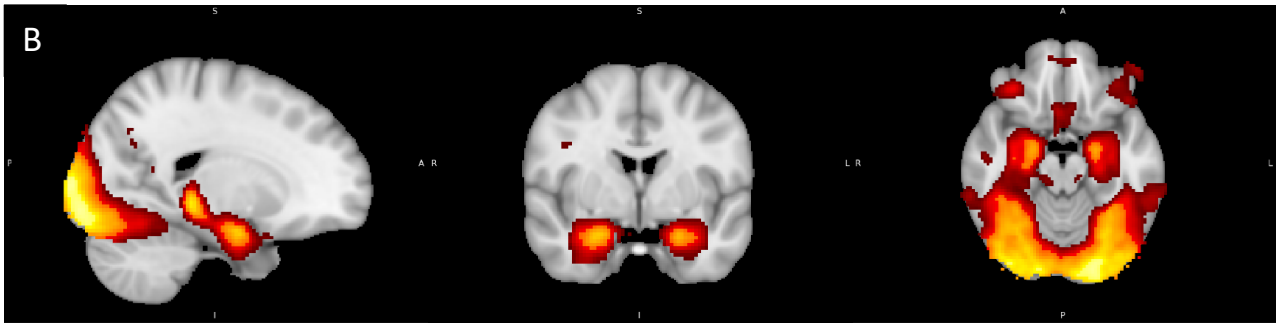
Study design

Fig. 1 shows the study design. Participants were invited after a telephone screening and underwent an in-person assessment including clinical interviews with trained staff. Participants fulfilling all inclusion criteria were randomized into one of two discontinuation groups. Participants in the discontinuation group 1W2 (withdrawal between T1 and T2) discontinued their ADM gradually (aiming for a discontinuation within 12 weeks but allowing up to 18 weeks) between assessments T1 and T2, allowing to control for repeated measurements of amygdala reactivity. Participants in the continuation group (12W; withdrawal after T1 and T2) underwent both assessments first, and then discontinued after the second assessment at T2. At each of the assessment time-points T1 and T2, participants completed a range of behavioral tasks, fMRI, electroencephalography and had blood samples taken (c.f. [13, 14, 16]). Relapse status was assessed during a six-month follow-up period. At weeks 1, 2, 4, 6, 8, 12, 16 and 21 of the follow-up period, patients were contacted for telephone assessments to determine relapse status. In case relapse was deemed probable during the telephone assessment, patients were invited to an in-person clinical interview, and, if they fulfilled diagnostic criteria [2], they underwent a final assessment. If no relapse occurred until week 26, they underwent the final assessment then. Control participants were only assessed once (at T1). Study data were collected between 1st July 2015 and 31st January 2019. Recruitment took place until January 2018. All participants provided written informed consent and received monetary compensation for participating. Ethical approval was provided by the cantonal ethics committee in Zurich and the ethics commission at the Campus Charité Mitte. All procedures were in keeping with the Declaration of Helsinki (10.1001/jama.2013.281053).

Faces task

Participants performed the Hariri faces task [37] while undergoing fMRI scanning. In the task, individuals are asked to match the face depicted at the top of the screen to one of two faces at the bottom of the screen (one of these matches the top identically); all three faces either showed angry or fearful emotions. In control trials, individuals selected which of the geometric shapes at the bottom was identical to the target shape at the top. The task consisted of 8 alternating blocks of face and form trials, with 6 trials per block.





Analysis

The samples from Berlin and Zurich were analyzed together as one group. Throughout all regression analyses site was included as a covariate of no interest. Group comparisons of symptom measures for patients versus controls and relapsers versus non-relapsers were performed via t-tests. A post-hoc test for an increase of amygdala reactivity in relapsers who discontinued before T2 was performed via a paired sample t-test.

fMRI data pre-processing Detailed information regarding imaging data acquisition is provided in section S2.1. Imaging data were pre-processed using standard settings of the FSL (FMRIB Software Library v6.0) FEAT software, with all steps described in section S2.2. We checked for excessive motion by visual inspection of the output of the motion correction program MCFLIRT but did not exclude any participant as none showed a displacement greater than half the voxel size. General linear models (GLMs) were fitted to pre-whitened data. The individual level (first level) GLM design matrix included two 5000ms-duration box-car regressors coding for the presentation of face and form stimuli, a stick function regressor for the button-presses and six motion regressors obtained from the motion correction step during pre-processing. Regressors were convolved with a hemodynamic response function (mean lag=6s, SD=3s). Each first-level GLM included three contrasts: face, form, face minus form. A group level GLM was performed using FEAT's FLAME method to obtain whole brain estimates for the face vs. form contrast.

Linear mixed effects ROI analysis Subject-wise first-level analyses were run using FEAT and resulting contrast estimates were transformed to MNI152 standard space for use in further analyses. Bilateral, left, and right amygdala ROIs were taken from the Harvard-Oxford sub-cortical atlas and used to extract the average estimates for each contrast (face, form, and face vs. form). To assess the difference in amygdala response between remitted patients and healthy controls, we calculated a t-test. We fit a linear mixed model for the

amygdala activity of only the patient group, for which we had measurements at two time points. We included time, discontinuation group (whether subjects discontinued between time points 1 and 2 (denoted by 1W2) or discontinued after T2 (denoted 12W)), relapse status, age, gender and site as predictors and subject-specific random intercepts. All models reported here converged. When reporting for left and right ROIs separately, the results are Bonferroni corrected.

Time to relapse analyses To examine the relationship between amygdala reactivity and the relapse-free interval, we entered the difference in amygdala activity (i.e. the per-person ROI-averaged contrast estimates) from time 1 to 2 as a regressor into a proportional hazards Cox model with the *time to relapse* as the right-censored dependent variable and age and gender as additional regressors.

Prediction analyses The above analyses examined the association between the average of all voxels in the selected ROIs and the intervention or clinical outcomes. To examine whether amygdala reactivity might contain predictive information, we took a machine learning approach. The features included in the models consisted of the voxel-wise face vs. form contrast estimates returned from the first-level analysis of the patient sample. We included the estimates of all voxels of the amygdala ROI (based on the Harvard-Oxford sub-cortical atlas; 366 voxels for right amygdala, 306 for left, and 672 for bilateral). These were used as predictors in a logistic regression model to predict relapse status. We used L1 regularization (i.e. variable selection) given the large number of features. Predictive performance was determined via Leave-One-Out Cross-Validation (LOOCV), with an inner (2-fold) Cross-Validation to find the optimal regularization parameter using gridsearch. We computed the posterior distribution of the balanced accuracy [40] to obtain estimates of the standard error and assess significance. We used three different models. The first model used the amygdala activity during the first scan (T1). The second model used only the amygdala activity from the second scan (T2) and the third model used the difference (patient-wise) of amygdala activity in each voxel of the amygdala mask.

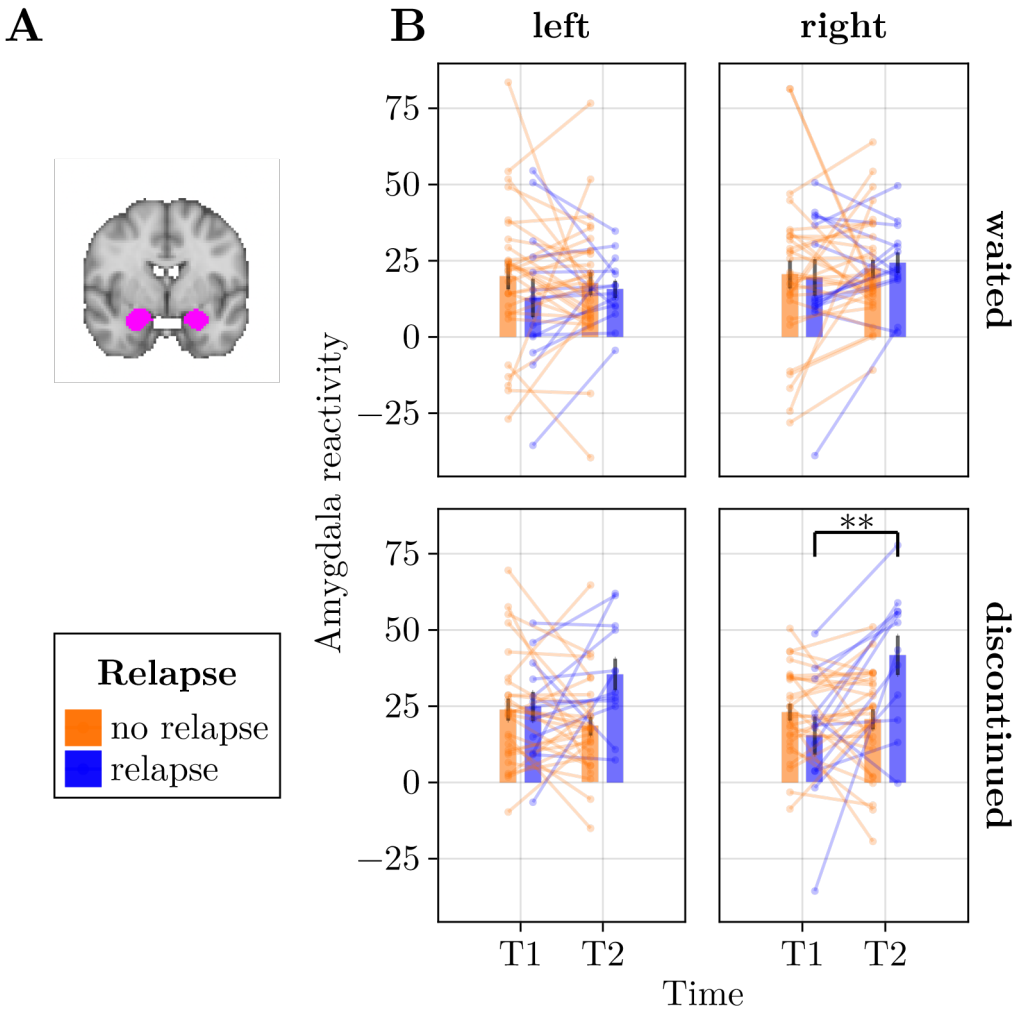
Analysis plan An analysis plan was created before data analysis commenced and is reproduced in the supplementary materials. We deviated from the analysis plan in that standard amygdala ROIs based on the Harvard-Oxford sub-cortical atlas were used rather than individual ROIs. This was done to allow for a simpler analysis pipeline entirely within FSL with fewer degrees of freedom. We added separate analyses of left/right amygdala ROIs, and the predictive analyses.

Results

Of the 84 patients and 57 healthy controls who completed the study, 83 and 53, respectively, could be included in the analyses (cf. Supplementary Material Figure S1). Table 1 shows the characteristics of the sample. The patient group was in remission, with minimal residual symptoms that were nevertheless higher than those in the never-depressed control group, and with some residual working memory impairments. At baseline, patients who went on to relapse and those who did not did not differ in any clinical or neuro-psychological variable or in terms of medication (cf. [13, 41]). The fMRI task was effective, resulting in an overall activation pattern like that reported in the literature, with prominent bilateral amygdala activation (Fig 1B). The analyses reported here were limited to the amygdala ROI (Fig 2B). In the following, we will denote ROI-averaged face-vs-form contrast estimates as *amygdala reactivity*.

Remitted patients vs. healthy controls

We compared amygdala reactivity in patients vs controls at T1. The bilateral ROI-averaged contrast estimates of both patient and control groups were significantly greater than zero (controls: $n = 55, t = 8.32, p = 10^{-10}$, patients at T1: $n = 83, t = 9.54, p = 10^{-14}$), but did not differ between groups ($n_1 = 83, n_2 = 55, t = 0.33, p = 0.74$).



[Table 1 here]

Association between amygdala reactivity, discontinuation, and relapse

The results of the linear mixed model for the amygdala reactivity are depicted in Table 2. This revealed that discontinuing ADM had a different effect for those who later did and did not relapse, as indicated by a significant three-way interaction between time (T1 before and T2 after discontinuation / wait period), discontinuation group (12W / 1W2), and relapse status (relapse / no relapse) at follow-up (95%-CI 0.8 to $37.1, p = 0.041$) and as depicted in Fig. 2B. There were no main effects of group ($z = 0.81, p = 0.42$), time ($z = -0.02, p = 0.98$) or relapse ($z = -0.49, p = 0.63$). A post-hoc paired t-test indicated that this was driven by an increase in amygdala reactivity in those patients who discontinued before T2 and who later went on to relapse (point-estimate 19.4, $n = 12$, $t = 3.03$, $p = 0.012$). At T2, amygdala reactivity in group 1W2 was higher for relapsers compared to non-relapsers (point-estimate 19.15, $n_1 = 12$, $n_2 = 27$, $t = 3.1, p = 0.007$). Examining the left and right amygdalae separately, we find a three-way interaction on the right side only (right: 95%-CI 5.6 to 45.7, $p_{cor} = 0.024$; left: 95%-CI -8.4 to 30.1, $p^{cor} = 0.61$). The increase in amygdala reactivity was significant only for the right side (point-estimate $26.4, n = 12, t = 3.7, p^{cor} = 0.007$). The group difference at T2 between those who went on to relapse versus not were significant for both sides (left point-estimate $16.97, n_1 = 12, n_2 = 27, t = 2.8, p_{cor} = 0.0218$, right point-estimate $20.98, n_1 = 12, n_2 = 27, t = 2.9, p_{cor} = 0.0204$).

Association between amygdala reactivity changes and time to relapse

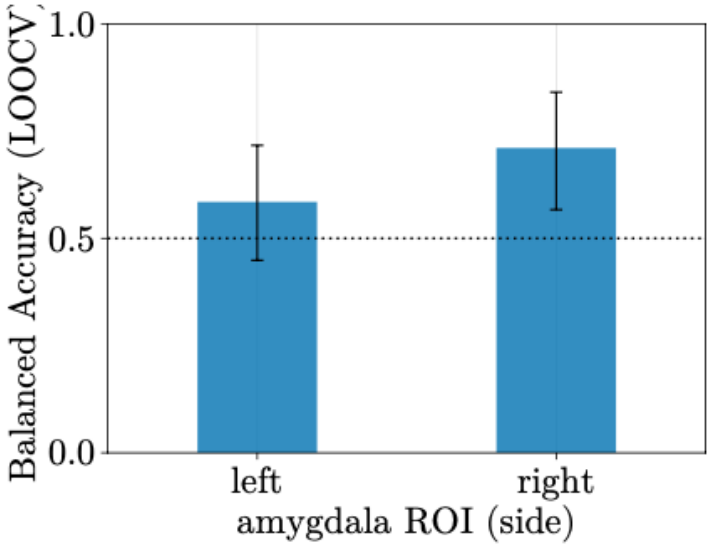
The results of the proportional hazards Cox model with the *time to relapse* as the right-censored dependent variable are shown in Table S1. In the discontinuation group, the difference in amygdala reactivity (T2 minus T1) was associated with time to relapse: patients with greater increase in amygdala tended to relapse earlier as indicated by a significant interaction of the difference in amygdala reactivity and the discontinuation group variable with a hazard ratio of 1.05 ($\beta = 0.05$, $p < 0.01$). Fitting models separately to each group yielded qualitatively the same results, with a significant effect of the change in amygdala reactivity (hazard ratio 1.05, 95%-CI $(1.015, 1.09)$ for the discontinuation group only). There were no significant effects of age, gender and site in either model.

[Table 2 here]

Prediction of relapse from amygdala reactivity

The predictive power of the change in amygdala reactivity between T1 and T2 (with all voxels in the ROI as features) is shown in Fig. 3. Results for models based only on measurements at T1 or T2 are shown in Tab. S2. We found predictive accuracies not significantly from chance for the models based on the amygdala activity of all patients at T1 before discontinuation, and amygdala reactivity at T2 in the discontinuation group.

However, the model based on the difference between T1 and T2 in amygdala reactivity (for all voxels in the bilateral ROI) yielded a predictive (balanced) accuracy of 67% (95%-PPI $(53\%,80\%)$; left side: 58%, $(45\%,72\%)$; right 71% $(57\%,84\%)$). After correcting for multiple comparisons, the posterior probability for the predictive accuracy being less than 50% is $0.18 < 0.05$ only for the model based on the voxels in the right ROI, but not the other models.



Discussion

Whether to discontinue ADM is a key clinical decision in the management of depression and brings with it a potentially substantial increase in the risk of relapse [6, 7], with few individual predictors to guide clinicians or patients in their decision-making [5, 12].

Here, we report that an increase in amygdala reactivity to negative emotional face stimuli after antidepressant discontinuation was associated with relapse during a six-month follow-up period. The findings are specific: they occur before relapse has occurred, and an increase in amygdala reactivity was only observed after discontinuation, and in those individuals who go on to relapse (Note: n=12 for this comparison). The increase in reactivity also appeared to be—to the extent this could be assessed within the study—potentially predictive of future relapse. These findings establish that there is individual variation in the impact of (mostly serotonergic) ADM discontinuation on amygdala reactivity: there was no main effect of discontinuation, meaning that amygdala reactivity only increased in those individuals who later relapsed. This raises the tantalizing possibility that amygdala reactivity was being maintained by ADM in some individuals, and by other

processes in others. Removal of ADM hence only had an adverse effect on those individuals who effectively relied on it to regulate amygdala reactivity.

Overall, the pattern of findings is consistent with the extensive literature on the relationship of amygdala reactivity, negative affective bias, and depression. The amygdala reactivity to negative emotional faces can be seen as an instance of negative affective bias, which is thought to underlie the maintenance of the depressed state [17–19, 46]. While still medicated, the remitted patients in our sample did not differ from the control group in terms of amygdala reactivity, supporting the notion that effective ADM treatment restores normal amygdala reactivity. This is in line with previous work showing that amygdala hyperactivity was increased in patients with depression, but decreased with treatment [21, 26, 31, 47].

The findings add to the existing work suggesting that neurocognitive markers may have an informative role to play in predicting relapse after antidepressant discontinuation. While clinical features and even discontinuation symptoms are not predictive of relapse [13, 48] in this sample, several neurocognitive measures have shown promise. For example, resting-state fMRI connectivity does change with discontinuation, and may be predictive of relapse [16] and a behavioral measure based on effort sensitivity, assessed at baseline, was predictive of relapse, although it was not altered by the discontinuation itself [14]. Similarly, pre-discontinuation EEG measures of affective reactivity are predictive of relapse, but we do not know about whether this changes with discontinuation [49].

Other work has identified abnormal processing of emotional stimuli that may be mediating a vulnerability to relapse after remission, such as frontotemporal connectivity during emotional face processing [50], emotional reactivity [51–53] and hyperconnectivity between anterior temporal and subgenual cortices while experiencing self-blaming emotions [54]. This does not seem to be the case for amygdala reactivity in the present study. Note that the absence of such a baseline effect strengthens the interpretation of the selective association between the discontinuation and relapse in what is an observational study, albeit with a randomized component.

The finding of stronger reactivity in the right amygdala is in line with previous work showing stronger role of the right hemisphere in processing faces and suggestions that the right amygdala plays a specific role in the processing of angry and fearful facial expressions [36, 55, 56].

The translational potential of the findings is uncertain. Whilst we found that amygdala activity changes were predictive of relapse, the analyses suggest that the measurement after discontinuation is required. This clearly substantially limits scalability. However, similar effects could potentially be observed with related pharmacological challenges, e.g. during a short-term discontinuation challenge, where patients stop medication for a couple days only. This could be more practically feasible and may potentially support further treatment decisions.

Limitations

The study has a relatively small sample size (n=41 patients in group 1W2, of which 12 relapsed) and the findings need to be treated with caution until they are replicated in a larger-scale study. The costs of scaling neuroimaging studies in this setting are substantial and thus smaller-scale studies such as this one are required. The study is unblinded: both participants and experimenters know which group participants were in, and when they discontinued. As such, it is not possible to disentangle pharmacological from psychological effects of discontinuation. To achieve this, a placebo-controlled study is required [6, 7]. Finally, the standard

version of the task employed does not allow general face processing to be disambiguated from emotion processing more specifically.

Conclusions

The AIDA study was a longitudinal, observational study with a randomized component. The design allowed four questions to be addressed, namely regarding the remitted but medicated depressed state; the effect of ADM discontinuation; the relationship between relapse and baseline features; and the relationship between the effect of ADM discontinuation and relapse. An increase in amygdala reactivity after ADM discontinuation was associated with risk of relapse. This adds to recent evidence that more specific neurobiological or behavioural measures can predict relapse and may hold promise for informing clinical treatment decisions around ADM discontinuation. Overall, the results of this study and previous results suggest that affective decision-making processes are engaged by the discontinuation and moderating relapse risk; however, the details will require further and larger-scale replication.

Author contributions

Author Contributions: TE and QJMH had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: Walter, Huys. Acquisition, analysis, or interpretation of data: Berwian, Seifritz, Stephan, Walter, Huys. Drafting of the manuscript: Erdmann, Huys. Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: Erdmann, Huys. Obtained funding: Stephan, Walter, Huys. Administrative, technical, or material support: Berwian, Seifritz, Stephan, Walter, Huys. Supervision: Stephan, Walter, Huys.

Conflict of Interest Disclosures

ES has received honoaries and educational grants from Lundbeck, OM Pharma, Janssen, Recordati, Takeda, Otsuka and Schwabe Pharma. QJMH has obtained fees and options for consultancies for Aya Technologies and Alto Neuroscience.

Funding / Support

This research was funded by research grants from the Swiss National Science Foundation (320030L_153449/1), and Wellcome Trust (221826/Z/20/Z) to QJMH, by grant WA 1539/5-1 from the Deutsche Forschungsgesellschaft to HW. Additional funds were provided by the Clinical Research Priority Program ”Molecular Imaging” at the University of Zurich. QJMH and TE acknowledge support by the UCLH NIHR BRC.

Role of Funder/Sponsor

The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contributions

Additional Contributions: Inga Schnuerer, Dipl Psych, and Daniel Renz, PhD (Translational Neuromodeling Unit, University of Zurich and ETH Zurich, Zurich, Switzerland) Julia G. Wenzel, and Leonie Kuehn and Christian Stoppel, MD, PhD (Charité Universitätsmedizin, Campus Charité Mitte, Berlin, Germany), assisted with planning, managing, and conducting the study. Ms Schnuerer and Dr Renz were compensated for their contributions. Julia Wenzel, Christian Stoppel and Leonie Kuehn were not compensated for their contributions. Yuki Yamamoto and Ryo Segawa assisted with analysis and figure generation.

References

1. Kessler, R. C. *et al.* Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* **62,** 593–602. issn: 0003-990X. https://doi.org/10.1001/archpsyc.62.6.593 (2023) (June 1, 2005).
2. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5™, 5th Ed* xliv, 947. xliv, 947. isbn: 978-0-89042-554-1 978-0-89042-555-8 (American Psychiatric Publishing, Inc., Arlington, VA, US, 2013).
3. Lythe, K. E. *et al.* Self-blame-Selective Hyperconnectivity Between Anterior Temporal and Subgenual Cortices and Prediction of Recurrent Depressive Episodes. eng. *JAMA Psychiatry* **72,** 1119–1126. http://dx.doi.org/10.1001/jamapsychiatry.2015.1813 (Nov. 2015).
4. Farb, N. A. S., Irving, J. A., Anderson, A. K. & Segal, Z. V. A two-factor model of relapse/recurrence vulnerability in unipolar depression. eng. *J Abnorm Psychol* **124,** 38–53. http://dx.doi.org/10.1037/abn0000031 (Feb. 2015).
5. Buckman, J. E. J. *et al.* Risk factors for relapse and recurrence of depression in adults and how they operate: A four-phase systematic review and meta-synthesis. *Clinical psychology review* **64,** 13–38. issn: 1873-7811 (Aug. 2018).
6. Geddes, J. R. *et al.* Relapse Prevention with Antidepressant Drug Treatment in Depressive Disorders: A Systematic Review. *The Lancet* **361,** 653–661. issn: 0140-6736. https://www.sciencedirect.com/science/article/pii/S0140673603125998 (2023) (Feb. 22, 2003).
7. Lewis, G. *et al.* Maintenance or Discontinuation of Antidepressants in Primary Care. *The New England journal of medicine* **385,** 1257–1267. issn: 1533-4406 (14 Sept. 2021). ppublish.
8. Glue, P., Donovan, M. R., Kolluri, S. & Emir, B. Meta-analysis of relapse prevention antidepressant trials in depressive disorders. eng. *Aust N Z J Psychiatry* **44,** 697–705. http://dx.doi.org/10.3109/00048671003705441 (Aug. 2010).
9. Viguera, A. C., Baldessarini, R. J. & Friedberg, J. Discontinuing antidepressant treatment in major depression. eng. *Harv Rev Psychiatry* **5,** 293–306 (1998).
10. Kaymaz, N., van Os, J., Loonen, A. J. M. & Nolen, W. A. Evidence that patients with single versus recurrent depressive episodes are differentially sensitive to treatment discontinuation: a meta-analysis of placebo-controlled randomized trials. eng. *J Clin Psychiatry* **69,** 1423–1436 (Sept. 2008).
11. Frank, E. *et al.* Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Remission, recovery, relapse, and recurrence. eng. *Arch Gen Psychiatry* **48,** 851–855 (Sept. 1991).

12. Berwian, I. M., Walter, H., Seifritz, E. & Huys, Q. J. M. Predicting Relapse after Antidepressant Withdrawal – a Systematic Review. *Psychological Medicine* **47,** 426–437. issn: 0033-2917. pmid: 27786144. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5244448/ (2023) (Feb. 2017).

13. Berwian, I. M. *et al.* Low Predictive Power of Clinical Features for Relapse Prediction after Antidepressant Discontinuation in a Naturalistic Setting. *Scientific Reports* **12,** 11171. issn: 2045-2322. https://www.nature.com/articles/s41598-022-13893-9 (2023) (1 July 1, 2022).

14. Berwian, I. M. *et al.* Computational Mechanisms of Effort and Reward Decisions in Patients With Depression and Their Association With Relapse After Antidepressant Discontinuation. *JAMA Psychiatry* **77,** 513–522. issn: 2168-622X. https://doi.org/10.1001/jamapsychiatry.2019.4971 (2023) (May 1, 2020).

15. Berwian, I. M. *et al.* Emotion-induced frontal alpha asymmetry predicts relapse after discontinuation of antidepressant medication (July 2023).

16. Berwian, I. M. *et al.* The Relationship between Resting-State Functional Connectivity, Antidepressant Discontinuation and Depression Relapse. *Scientific Reports* **10,** 22346. issn: 2045-2322. https://www.nature.com/articles/s41598-020-79170-9 (2023) (1 Dec. 18, 2020).

17. Harmer, C. J., Goodwin, G. M. & Cowen, P. J. Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. eng. *Br J Psychiatry* **195,** 102–108. http://dx.doi.org/10.1192/bjp.bp.108.051193 (Aug. 2009).

18. Disner, S. G., Beevers, C. G., Haigh, E. A. P. & Beck, A. T. Neural mechanisms of the cognitive model of depression. eng. *Nat Rev Neurosci* **12,** 467–477. http://dx.doi.org/10.1038/nrn3027 (Aug. 2011).

19. Roiser, J. P., Elliott, R. & Sahakian, B. J. Cognitive mechanisms of treatment in depression. eng. *Neuropsychopharmacology* **37,** 117–136. http://dx.doi.org/10.1038/npp.2011.183 (Jan. 2012).

20. Godlewska, B. R. & Harmer, C. J. Cognitive neuropsychological theory of antidepressant action: a modern-day approach to depression and its treatment. *Psychopharmacology* **238,** 1265–1278. issn: 1432-2072 (5 May 2021). ppublish.

21. Ruhé, H. G., Booij, J., Veltman, D. J., Michel, M. C. & Schene, A. H. Successful Pharmacologic Treatment of Major Depressive Disorder Attenuates Amygdala Activation to Negative Facial Expressions: A Functional Magnetic Resonance Imaging Study. *The Journal of Clinical Psychiatry* **72,** 20218. issn: 0160-6689. https://www.psychiatrist.com/jcp/successful-pharmacologictreatment-major-depressive/ (2023) (Aug. 9, 2011).

22. Stuhrmann, A., Suslow, T. & Dannlowski, U. Facial Emotion Processing in Major Depression: A Systematic Review of Neuroimaging Findings. *Biology of Mood & Anxiety Disorders* **1,** 10. issn: 2045-5380. https://doi.org/10.1186/2045-5380-1-10 (2023) (Nov. 7, 2011).

23. Ma, Y. Neuropsychological Mechanism Underlying Antidepressant Effect: A Systematic MetaAnalysis. *Molecular Psychiatry* **20,** 311–319. issn: 1476-5578. https://www.nature.com/articles/ mp201424 (2023) (3 Mar. 2015).

24. Mayberg, H. S. Targeted Electrode-Based Modulation of Neural Circuits for Depression. *The Journal of Clinical Investigation* **119,** 717–725. issn: 0021-9738. pmid: 19339763. https://www.jci.org/articles/view/38454 (2023) (Apr. 1, 2009).

25. Tamm, S. *et al.* No Association Between Amygdala Responses to Negative Faces and Depressive Symptoms: Cross-Sectional Data from 28,638 Individuals in the UK Biobank Cohort. *The American journal of psychiatry* **179,** 509–513. issn: 1535-7228 (7 July 2022). ppublish.

26. Godlewska, B. R., Norbury, R., Selvaraj, S., Cowen, P. J. & Harmer, C. J. Short-Term SSRI Treatment Normalises Amygdala Hyperactivity in Depressed Patients. *Psychological Medicine* **42,** 2609–2617. issn: 0033-2917, 1469-8978. https : / / www . cambridge . org / core / journals / psychological-

medicine/article/shortterm-ssri-treatment-normalises-amygdala-hyperactivity-indepressed-patients/EC7D6473AB517C5F82EFD0A4BBC8FAEC (2023) (Dec. 2012).

27. Harmer, C. J., Mackay, C. E., Reid, C. B., Cowen, P. J. & Goodwin, G. M. Antidepressant Drug Treatment Modifies the Neural Processing of Nonconscious Threat Cues. *Biological Psychiatry* **59,** 816–820. issn: 00063223. https://linkinghub.elsevier.com/retrieve/pii/S0006322305013697 (2023) (May 2006).

28. Harmer, C. J. Serotonin and Emotional Processing: Does It Help Explain Antidepressant Drug Action? *Neuropharmacology. Current Research on Serotonin: From Gene to Function to Therapy* **55,** 1023–1028. issn: 0028-3908. https://www.sciencedirect.com/science/article/pii/ S0028390808002311 (2023) (Nov. 1, 2008).

29. Aldao, A., Nolen-Hoeksema, S. & Schweizer, S. Emotion-regulation strategies across psychopathology: A meta-analytic review. eng. *Clin Psychol Rev* **30,** 217–237. http://dx.doi.org/10.1016/j. cpr.2009.11.004 (Mar. 2010).

30. Buhle, J. T. *et al.* Cognitive Reappraisal of Emotion: A Meta-Analysis of Human Neuroimaging Studies. eng. *Cereb Cortex.* http://dx.doi.org/10.1093/cercor/bht154 (June 2013).

31. Sheline, Y. I. *et al.* Increased Amygdala Response to Masked Emotional Faces in Depressed Subjects Resolves with Antidepressant Treatment: An fMRI Study. *Biological Psychiatry* **50,** 651–658. issn: 00063223. https://linkinghub.elsevier.com/retrieve/pii/S000632230101263X (2023) (Nov. 2001).

32. Gerlach, A. R. *et al.* MRI Predictors of Pharmacotherapy Response in Major Depressive Disorder. *NeuroImage: Clinical* **36,** 103157. issn: 2213-1582. https://www.sciencedirect.com/science/article/pii/S2213158222002224 (2023) (Jan. 1, 2022).

33. Delaveau, P. *et al.* Brain Effects of Antidepressants in Major Depression: A Meta-Analysis of Emotional Processing Studies. *Journal of Affective Disorders* **130,** 66–74. issn: 0165-0327. https://www.sciencedirect.com/science/article/pii/S0165032710006130 (2023) (Apr. 1, 2011).

34. Li, X. & Wang, J. Abnormal neural activities in adults and youths with major depressive disorder during emotional processing: a meta-analysis. *Brain imaging and behavior* **15,** 1134–1154. issn: 1931-7565 (2 Apr. 2021). ppublish.

35. Nord, C. L. *et al.* Neural effects of antidepressant medication and psychological treatments: a quantitative synthesis across three meta-analyses. *The British journal of psychiatry : the journal of mental science* **219,** 546–550. issn: 1472-1465 (4 Oct. 2021). ppublish.

36. Plichta, M. M. *et al.* Test–retest reliability of evoked BOLD signals from a cognitive–emotive fMRI test battery. *NeuroImage* **60,** 1746–1758 (Apr. 2012).

37. Hariri, A. R. *et al.* Serotonin transporter genetic variation and the response of the human amygdala. *Science* **297,** 400–403. http://dx.doi.org/10.1126/science.1071829 (July 2002).

38. Wakefield, J. C. & Schmitz, M. F. When Does Depression Become a Disorder? Using Recurrence Rates to Evaluate the Validity of Proposed Changes in Major Depression Diagnostic Thresholds. *World Psychiatry* **12,** 44–52. issn: 2051-5545. https://onlinelibrary.wiley.com/doi/abs/10.1002/ wps.20015 (2023) (2013).

39. First, M. B., Spitzer, R. L., Gibbon, M. & Williams, J. B. *Structured Clinical Interview for DSMIV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P)* (Biometrics Research, New York State Psychiatric Institute, 2002).

40. Brodersen, K. H., Ong, C. S., Stephan, K. E. & Buhmann, J. M. *The Balanced Accuracy and Its Posterior Distribution* in *2010 20th International Conference on Pattern Recognition* 2010 20th International Conference on Pattern Recognition (ICPR) (IEEE, Istanbul, Turkey, Aug. 2010), 3121–3124. isbn: 978-1-4244-7542-1. http://ieeexplore.ieee.org/document/5597285/ (2023).

41. Volkmann, C., Abulikemu, S., Isabel, Huys, Q. J. & Walter, H. Do discontinuation symptoms predict depression relapse after antidepressant cessation? (Feb. 2023).

42. Rush, A. J., Gullion, C. M., Basco, M. R., Jarrett, R. B. & Trivedi, M. H. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med* **26,** 477–86 (May 1996).

43. Wechsler, D. *Wechsler Adult Intelligence Scale Revised.* (The Psychological Corporation, New York, USA, 1981).

44. Lehr, S. *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B* (Spitta, Balingen, 2005).

45. Reitan, R. M. Validity of the trial making test as an indicator of organic brain damage. *Perceptual and Motor Skills* **8,** 271–276 (1958).

46. Beck, A. T. *Cognitive Therapy of Depression* isbn: 978-0-89862-919-4 (Guilford Press, Jan. 1979).

47. Williams, L. M. *et al.* Amygdala Reactivity to Emotional Faces in the Prediction of General and Medication-Specific Responses to Antidepressant Treatment in the Randomized iSPOT-D Trial. *Neuropsychopharmacology* **40,** 2398–2408. issn: 1740-634X. https://www.nature.com/articles/npp201589 (2023) (10 Sept. 2015).

48. Volkmann, C., Abulikemu, S., Berwian, I. M., Huys, Q. & Walter, H. Do Discontinuation Symptoms Predict Depression Relapse after Antidepressant Cessation? https://osf.io/fxye5 (2023) (Dec. 19, 2023).

49. Berwian, I. M. *et al. Emotion-Induced Frontal Asymmetry Predicts Relapse after Discontinuation of Antidepressant Medication* https://www.biorxiv.org/content/10.1101/2023.07.05.547831v1 (2023). preprint.

50. Goulden, N. *et al.* Reversed Frontotemporal Connectivity During Emotional Face Processing in Remitted Depression. *Biological Psychiatry. Novel Pharmacotherapies for Depression* **72,** 604– 611. issn: 0006-3223. https://www.sciencedirect.com/science/article/pii/S0006322312004155 (2023) (Oct. 1, 2012).

51. Segal, Z. V. *et al.* Cognitive reactivity to sad mood provocation and the prediction of depressive relapse. eng. *Arch Gen Psychiatry* **63,** 749–755. http://dx.doi.org/10.1001/archpsyc.63.7.749 (July 2006).

52. Farb, N. A. S., Anderson, A. K., Bloch, R. T. & Segal, Z. V. Mood-linked responses in medial prefrontal cortex predict relapse in patients with recurrent unipolar depression. eng. *Biol Psychiatry* **70,** 366–372. http://dx.doi.org/10.1016/j.biopsych.2011.03.009 (Aug. 2011).

53. Farb, N. A. S., Desormeau, P., Anderson, A. K. & Segal, Z. V. Static and treatment-responsive brain biomarkers of depression relapse vulnerability following prophylactic psychotherapy: Evidence from a randomized control trial. *NeuroImage. Clinical* **34,** 102969. issn: 2213-1582 (22). ppublish.

54. Lythe, K. E. *et al.* Self-Blame–Selective Hyperconnectivity Between Anterior Temporal and Subgenual Cortices and Prediction of Recurrent Depressive Episodes. *JAMA Psychiatry* **72,** 1119– 1126. issn: 2168-622X. https://doi.org/10.1001/jamapsychiatry.2015.1813 (2023) (Nov. 1, 2015).

55. Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F. & Weinberger, D. R. The Amygdala Response to Emotional Stimuli: A Comparison of Faces and Scenes. *NeuroImage* **17,** 317–323. issn: 10538119. https://linkinghub.elsevier.com/retrieve/pii/S1053811902911791 (2023) (Sept. 2002).

56. Hariri, A. R. The Neurobiology of Individual Differences in Complex Behavioral Traits. *Annual Review of Neuroscience* **32,** 225–247. issn: 0147-006X, 1545-4126. https://www.annualreviews.org/content/journals/10.1146/annurev.neuro.051508.135335 (2024) (Volume 32, 2009 July 21, 2009).

	Controls (N=53)	Patients (N=83)	p-value	Non Relapsers (N=57)	Relapsers (N=26)	p-value
Demographics						
Age	33.57 ± 10.7	35.42 ± 11.41	0.339	34.53 ± 11.57	37.38 ± 11.02	0.293
Male sex # (%)	16 (30.2)	21 (25.3)	0.532	14 (24.6)	7 (26.9)	0.818
BMI	23.49 ± 3.69	23.99 ± 4.3	0.481	24.17 ± 4.31	23.6 ± 4.34	0.58
Clinical measures						
Residual depression (IDS-C (Inventory of Depressive Symptomatology, Clinician-Rated) [42])	0.65 ± 1.07	3.76 ± 3.96	<0.001	3.42 ± 2.91	4.57 ± 5.71	0.248
No. prior episodes	-	2.41 ± 1.31	-	2.33 ± 1.34	2.58 ± 1.24	0.434
Medication Load[a]	-	0.76 ± 0.4	-	0.78 ± 0.4	0.72 ± 0.42	0.555
Neuropsychological scores						
Intelligence (MWTB; Mehrfachwahl-Wortschatz-Intelligenztest) [44]	28.08 ± 4.07	28.6 ± 4.23	0.47	28.18 ± 4.3	29.54 ± 3.98	0.174
Working Memory (Digit span backwards of the Wechsler Adult Intelligence Scale [43])	8.17 ± 3.38	6.93 ± 2.0	0.018	7.07 ± 2.16	6.62 ± 1.58	0.339
Executive function (Trial making test A; TMT A) [45]	23.3 ± 5.71	24.82 ± 8.3	0.207	24.78 ± 8.2	24.9 ± 8.68	0.953
Executive function (TMT B) [45]	56.52 ± 19.86	56.55 ± 17.07	0.994	55.62 ± 17.03	58.58 ± 17.32	0.466

Table 1: Table with sample characteristics and p-values for tests of group differences.

[a] defined as the dose divided by the maximal allowed dose according to the Swiss compendium (www.compendium.ch) and by the weight of the participant.

Name	Coef.	Std. Error	z value	p-value	95% CI lower bound	95% CI upper bound
(Intercept)	16.618	4.384	3.791	0.0	8.026	25.21
T: 2	-0.087	3.556	-0.024	0.981	-7.057	6.884
Discontinuation Group: 1W2	4.163	5.151	0.808	0.419	-5.933	14.259
Relapse	-3.046	6.252	-0.487	0.626	-15.299	9.206
Site: Zurich	4.481	3.275	1.368	0.171	-1.938	10.901
T: 2 & Dis. Group: 1W2	-3.669	5.167	-0.71	0.478	-13.797	6.459
T: 2 & Relapse	3.994	6.305	0.633	0.526	-8.364	16.352
Disc. group: 1W2 & Relapse	-1.058	9.165	-0.115	0.908	-19.02	16.905
T: 2 & Disc. Group: 1W2 & Relapse	18.931	9.243	2.048	0.041	0.815	37.046

Table 2: Coefficient table of mixed-model for the bilateral amygdala response (i.e. ROI-averaged voxel-wise estimates of the face vs. form contrast). Categorical variable names are binary coded, such that the coefficient for T:2 represents the difference in response for time point 2 vs. reference category (T: 1). Interactions are denoted via the & sign, that is, the coefficient for T:2 & Relapse represents the difference of the increase (from T1 to T2) for relapsers versus non-relapsers (the reference category).

Figure 1: Study design (top) and whole brain fMRI results (bottom). Panel A) Study design: patients with remitted Major Depressive Disorder were randomized to either undergo fMRI before and after ADM discontinuation (top), or to undergo fMRI twice before ADM discontinuation. After discontinuation, all patients were followed up for 6 months. A group of never-depressed control participants were assessed once only. The design enables a cross-sectional comparison of the remitted depressed state (T1 patients / controls). In the patient sample, it allows the effect of discontinuation to be related to relapse (interaction of time point (T1 / T2) with group (12W / 1W2) and relapse. Panel B) Whole-brain fMRI results for patients at both time points: Overall, the task did significantly activate the amygdala across patients and controls. Shown is the z-statistic map for the face-form contrast, with cluster-based correction with an activation threshold of $Z > 3.1$ and a cluster-extent threshold of $P < 0.001$ applied at the whole-brain level.

Figure 2: ROI-based analysis: Panel A) shows the two selected ROIs corresponding to left and right amygdala from the Harvard-Oxford atlas. Panel B) shows the same ROI-averaged contrast values for patients for both time points and split by discontinuation (if that patient discontinued ADM before time point two or after) and relapse. Bars indicate means with standard errors. ** indicates post-hoc paired-sample t-test $p < 0.01$

Figure 3: Relapse prediction: depicted is the predictive accuracy of the relapse classifiers based on right and left amygdala ROI (precisely: the modes of the posterior over the balanced accuracy inferred from the confusion matrix resulting from an LOOCV procedure). The classifiers were based-on on the voxel-wise increase in the face-form contrast estimates. Error bars indicate the 95% posterior predictive interval and the dotted line is the chance level.