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

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

January 26, 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

Word count
Abstract:
Total:
Figures & Tables:
References:

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 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 were randomized to task-based fMRI measurement of amygdala reactivity either twice before, or after discontinuing antidepressants. 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, 80 (mean (SD) age 35.5 (11.4) years; 60 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 did not differ from controls (left: $t = 0.77, p = 0.44$, right: $t = 0.88, p = 0.37$). 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; $\beta = 25.1$, 95%-CI $(4.4, 45.8), p = 0.018$). Amygdala reactivity change was associated with shorter times to relapse (hazard ratio 1.05, 95%-CI $(1.016, 1.085)$), and predictive of relapse (LOOCV balanced accuracy 71%, 95%-PPI $(57\%, 84\%)$).

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 cannot be continued indefinitely. 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 yet 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 in order 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 (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 assessment 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 a face depicted at the top of the screen to one of two other faces at the bottom of the screen; the target faces alternatively show angry and sad 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.

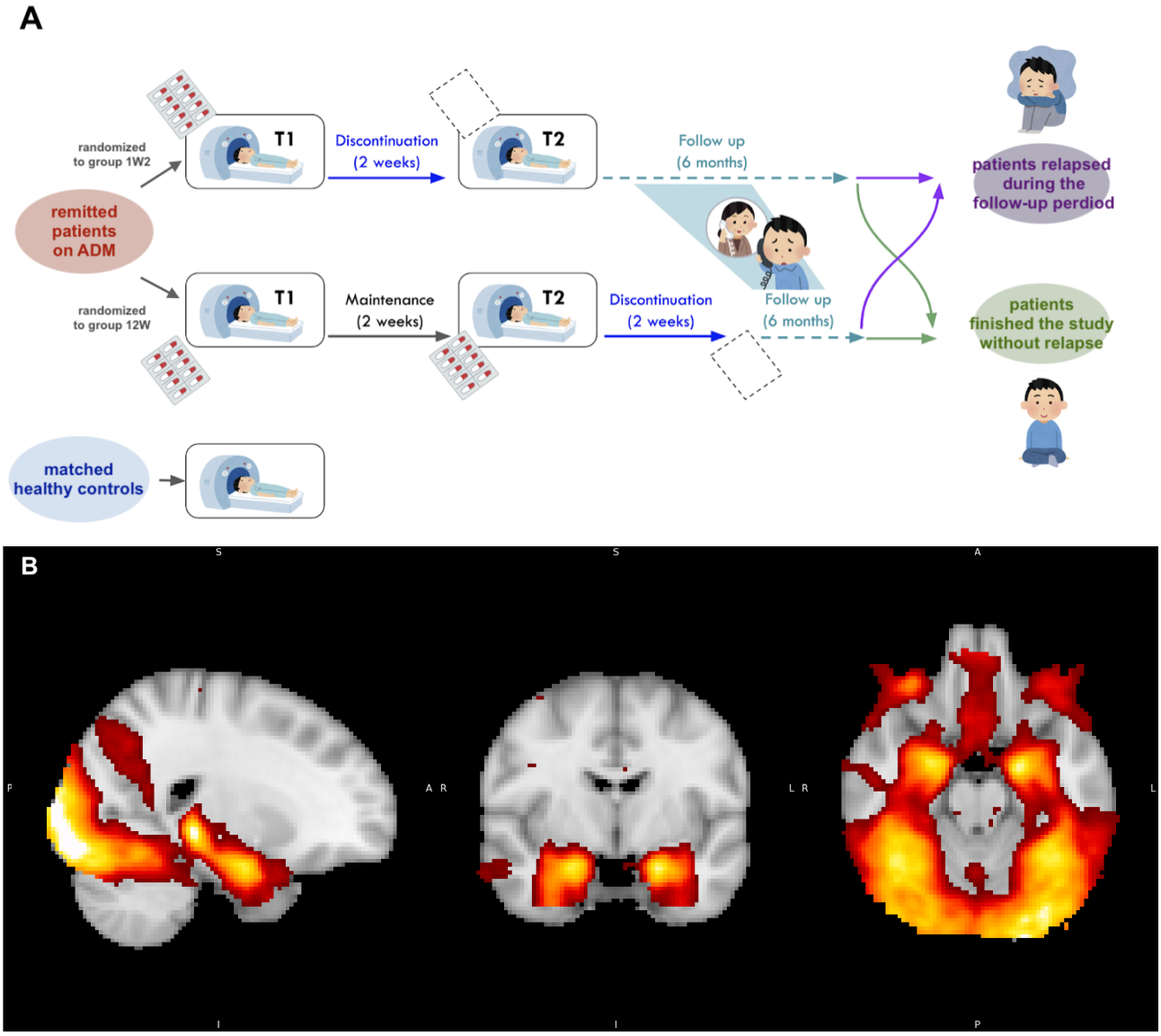

Figure 1: Study design (top) and whole brain fMRI results (bottom). Panel A) Study design: patients with remitted Major Depressive Disorder were randomized to either undergo fMRI before and after ADM discontinuation (top), or to undergo fMRI twice before ADM discontinuation. After discontinuation, all 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-level FWE correction at $p < 0.05$ and a cluster defining threshold of $p < 0.001$.

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.

Functional MRI acquisition Images were acquired at using a Phillips 3T Ingenia in Zurich and a Siemens 3T Trio in Berlin. In Zurich, a 32-channel coil was used to acquire echo-planar images (EPIs; 108 volumes; 40 axial slices; 2.5mm slice sickness; descending sequential acquisition, repetition time: 2560 ms; echo time: 27 ms, field of view: 210 x 210 x 119.5, acquisition matrix: 84 x 82, reconstructed voxel size: 2.19 x 2.19 x 2.50 mm, flip angle: 90°). Additionally, we acquired T1-weighted magnetization-prepared rapid-acquisition gradient-echo (MPRAGE) structural images (301 axial slices; slice thickness: 1; repetition time: 7.9 ms; echo time: 3.7 ms, field of view: 250 x 250 x 180.6, acquisition matrix: 252 x 251, reconstructed voxel size: 0.98 x 0.98 x 0.60 mm, flip angle: 8°). In Berlin, a 32-channel coil was used for EPIs (108 volumes; 40 axial slices; 3 mm slice thickness including a gap of 0.5 mm; descending sequential acquisition, repetition time: 2560 ms; echo time: 27 ms, field of view: 210 x 210 x 120, acquisition matrix: 84 x 84, voxel size: 2.50 x 2.50 x 2.50 mm, flip angle: 90°). T1-weighted MPRAGE structural images (192 axial slices; slice thickness: 1 mm; repetition time: 1900 ms; echo time: 2.52 ms, field of view: 256 x 256 x 192, acquisition matrix: 256 x 256, reconstructed voxel size: 0.98 x 0.98 x 0.60 mm, flip angle: 9°) were also acquired.

fMRI data preprocessing Imaging data were preprocessing using standard settings of the FSL (FMRIB Software Library v6.0) FEAT software. The standard pipeline includes brain extraction from the T1-weighted images (using the BET program with initial threshold of 0.2); motion correction using rigid-body registration (via MCFLIRT program) to the central volume of the functional time series; and high-pass filtering with a cutoff of 1/100Hz. EPI images were registered to the high-resolution structural images and normalized into standard (MNI152) space using affine (boundary-based), and nonlinear registration, respectively (via FLIRT and FNIRT programs). Functional data were spatially smoothed using a Gaussian filter with 6 mm full-width at half maximum (via SUSAN program). We did not perform slice timing correction, but instead included temporal derivatives in the first-level statistical model as per the FEAT default. All summary images for the registration procedure were visually inspected, and for subjects with poor registration results the procedure was repeated with the following changes. Firstly, we changed the threshold parameter for the brain extraction program, and, secondly, we changed the functional image that was used as template in the registration (the middle volume) in case that image was affected by movement. 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 and six motion regressors obtained from the motion correction step during preprocessing. 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. 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) and for each side (left and right amygdala).

To assess the difference in left/right 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 and slopes for the time effect.

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. In order to examine whether amygdala reactivity might contain predictive information, we took a machine learning approach. The features included in the models consisted of the face vs. form contrast estimates from all voxels in the amygdala ROI (366 for right amygdala and 306 for left) returned from the first-level analysis of the patient sample. 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 at the first assessment, before any of the patients had discontinued their ADM. The second model used only the amygdala activity from the second time point and the third model used the difference (patient-wise) of amygdala activity in each voxel of the amygdala mask.

Results

Of the 90 patients and 59 healthy controls who completed the study, 80 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. The fMRI task was effective, resulting in an overall activation pattern similar to that reported in the literature, with prominent bilateral amygdala activation (Fig 1). The analyses reported here were limited to the amygdala ROI (Fig 2). In the following, we will denote ROI-averaged face-vs-form contrast estimates as *amygdala reactivity*.

Remitted patients vs. healthy controls

Fig. 2B shows a comparison of amygdala reactivity in patients vs controls at T1. The ROI-averaged contrast estimates of both patient and control groups were significantly greater than zero (left side: $t = 6.82, p < 10^{-9}$, right side: $t = 7.29, p < 10^{-10}$), but did not differ between groups (left side: $t = 0.77, p = 0.44$, right side: $t = 0.88, p = 0.37$).

	Controls (N=53)	Patients (N=80)	p-value	Non Relapsers (N=54)	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)) [41]	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	0.55
Neuropsychological scores						
Working Memory (Digit Span backwards of the Wechsler Adult Intelligence Scale [42])	8.17 ± 3.38	6.93 ± 2.0	0.018	7.07 ± 2.16	6.62 ± 1.58	0.339
Intelligence (MWTB; Mehrfachwahl-Wortschatz-Intelligenztest) [43]	28.28 ± 3.92	28.69 ± 4.1	0.598	28.3 ± 4.13	29.54 ± 3.98	0.203
Executive function (Trial making test A; TMT A) [44]	22.46 ± 5.55	24.88 ± 8.25	0.097	24.87 ± 8.13	24.9 ± 8.68	0.986
Executive function (TMT B) [44]	54.42 ± 21.15	56.5 ± 17.1	0.56	55.56 ± 17.06	58.58 ± 17.32	0.458

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.

Association between amygdala reactivity, discontinuation and relapse

The results of the linear mixed model for the amygdala reactivity are depicted in Table 2.l 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 4.4 to 45.8, $p = 0.018$) and as depicted in Fig. 2C. There were no main effects of group ($z = 0.31, p = 0.76$), time ($z = 0.287, p = 0.77$) or relapse ($z = -0.28, p = 0.78$). A post-hoc paired t-test indicated that this was driven by an increase in right amygdala reactivity in those patients who discontinued before T2 and who later went on to relapse (point-estimate 26.2, $n = 12$, $p = 0.0033$).

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 ($\beta = 0.05, z = 2.88, p = 0.004$).

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 right ROI) yielded a predictive (balanced) accuracy of 71% (95%-PPI (56%, 84%); left side: 49%, (36%, 63%)), with a posterior probability of 0.002% for

Name	Coef.	Std. Error	z	p-value	95%-CI
(Intercept)	19.287	4.996	3.861	0.0	(9.496, 29.079)
T: 2	1.222	4.257	0.287	0.774	(−7.122, 9.566)
Discontinuation Group: 1W2	1.728	5.613	0.308	0.758	(−9.274, 12.731)
Relapse	−1.861	6.739	−0.276	0.782	(−15.069, 11.347)
Site: Zurich	3.651	3.646	1.001	0.317	(−3.495, 10.797)
T: 2 & Disc. Group: 1W2	−3.456	6.021	−0.574	0.566	(−15.256, 8.344)
T: 2 & Relapse	3.35	7.285	0.46	0.646	(−10.929, 17.629)
Discontinuation Group: 1W2 & Relapse	−6.066	9.78	−0.62	0.535	(−25.235, 13.102)
T: 2 & Disc. Group: 1W2 & Relapse	25.099	10.582	2.372	0.018	(4.358, 45.839)

Table 2: Coefficient table of mixed-model for the right-side 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).

the accuracy being less than 50%.

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 any change in symptoms have occurred, and an increase in amygdala reactivity was only observed after discontinuation, and in those individuals who go on to relapse. 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, 45]. 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 that found that amygdala hyperactivity was increased in patients with depression, but decreased with treatment [21, 26, 31, 46].

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, 47] in this sample, several neurocognitive

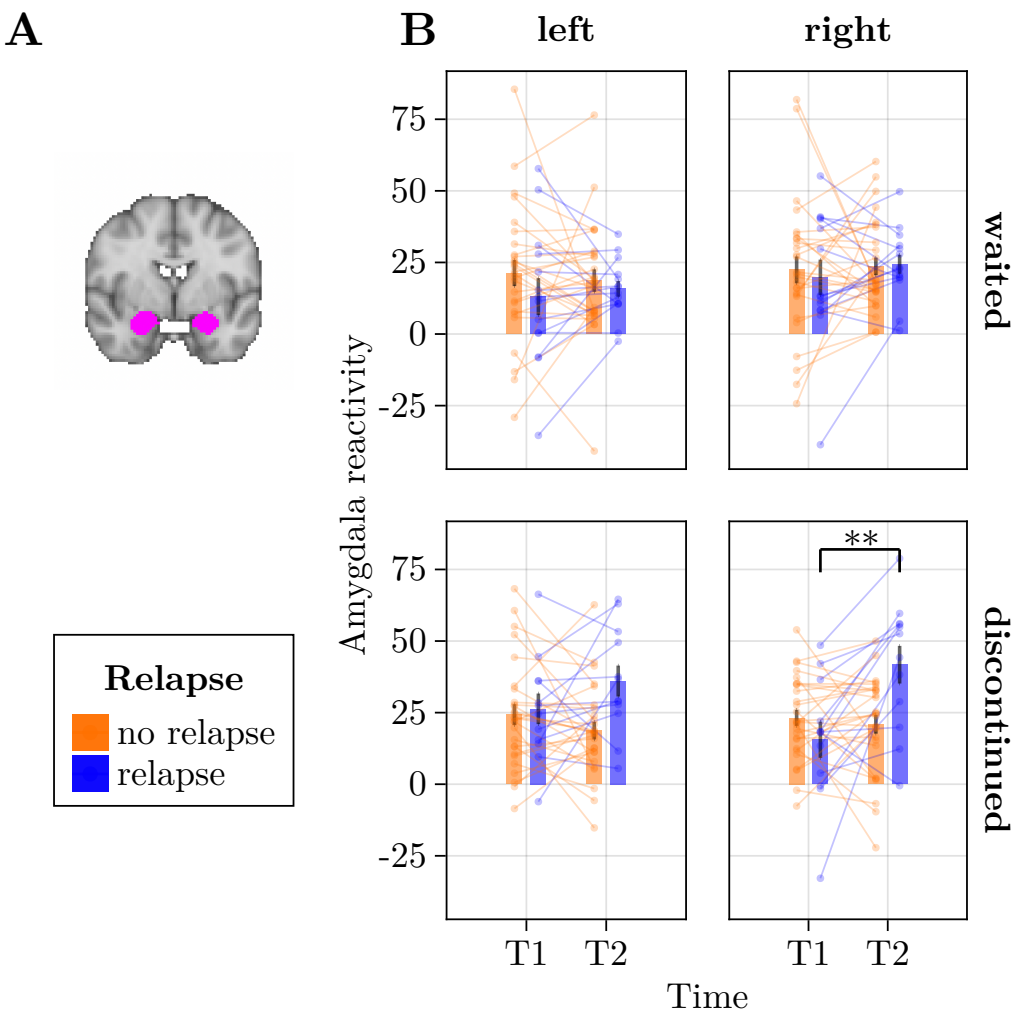

Figure 2: ROI-based analysis: Panel A) shows the two selected ROIs corresponding to left and right amygdala from the Harvard-Oxford atlas. Panel B) shows the same ROI-averaged contrast values for patients for both time points and split by discontinuation (if that patient discontinued ADM before time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

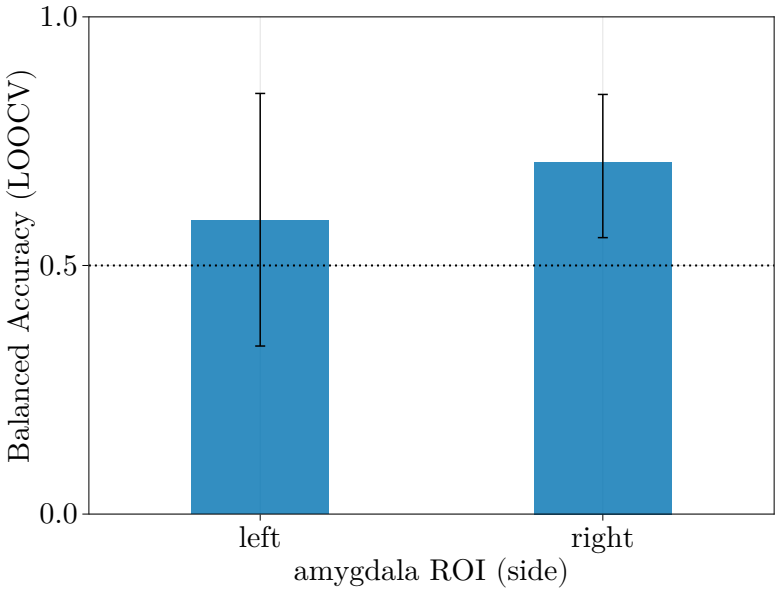

Figure 3: Relapse prediction: depicted is the predictive accuracy of the relapse classifiers based on right (blue) and left (orange) amygdala ROI (precisely: the modes of the posterior over the balanced accuracy inferred from the confusion matrix resulting from an LOOCV procedure). The classifiers were 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.

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 [48].

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 [49], emotional reactivity [50–52] and hyperconnectivity between anterior temporal and subgenual cortices while experiencing self-blaming emotions [53]. 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 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, it may be that similar effects could 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 and should be replicated in a larger-scale study. The costs of scaling neuroimaging studies in this setting is substantial, and requires smaller-scale studies such as this one. 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]. 2872

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. 2999

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-pharmacologic-treatment-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 Meta-Analysis. *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-in-depressed-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 DSM-IV-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. 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).

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

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

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

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

46. 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).

47. 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).

48. 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.

49. 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).

50. 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).

51. 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).

52. 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.

53. 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).

Supplementary Material

S1 CONSORT diagrams

The CONSORT diagram show reasons for exclusion and dropouts for patients, depicted in FigS1, and controls, depicted in Fig S2. Of the 123 patients, 84 finished the study according to protocol and 80 could be analyzed, of which 26 had a relapse. Of the 66 healthy controls (HC) who were included in the study, 8 dropped out and 2 had to be excluded for corrupted or missing data files, leading to a sample of 53 HC included in the analysis.

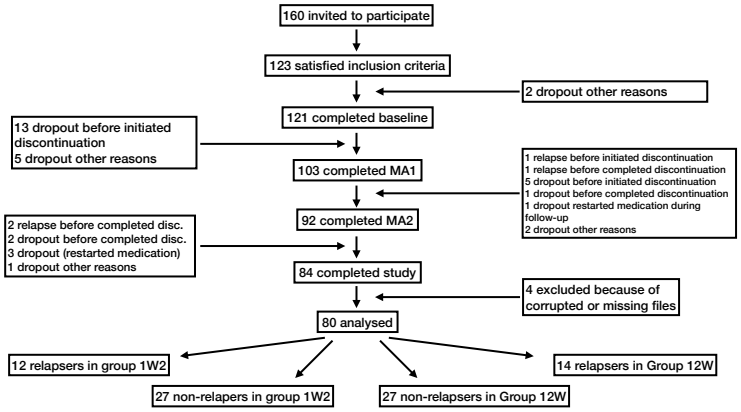

Figure S1: Reasons for dropout and exclusions for patients. Group 1W2 discontinued their ADM between main assessments 1 (MA1) and 2 (MA2). BA = baseline assessment. The respective numbers from the Zurich and Berlin sites are shown in parentheses.

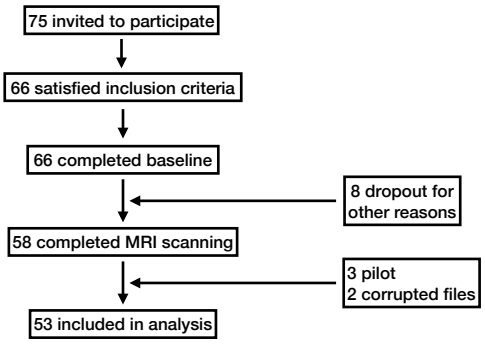

Figure S2: Reasons for dropout and exclusions for controls.

S2 Inclusion and exclusion criteria

Participants participates were deemed eligible for participation if fulfilling the following inclusion criteria:

1. age 18-55 years
2. ability to consent and adhere to the study protocol
3. written informed consent
4. fluent in written and spoken German.

Patients had to additionally fulfil the following criteria:

1. currently under medical care with a psychiatrist or general practitioner for remitted Major Depressive Disorder and willing to remain in care for the duration of the study (approx. 9 months)
2. informed choice to discontinue medication (including willingness to taper the medication over at most 12 weeks) that was independent of study participation
3. clinical remission (Hamilton Depression Score of less than 7) had been achieved under therapy with antidepressants without having undergone manualized psychotherapy; with no other concurrent psychotropic medication and had been maintained for a minimum of 30 days
4. consent to information exchange between treating physician and study team members regarding inclusion/exclusion criteria and past medical history.

Any of the following exclusion criteria led to exclusion of participants. This included the following general criteria:

1. any disease of type and severity sufficient to influence the planned measurement or to interfere with the parameters of interest (this includes neurological, endocrinological, oncological comorbidities, a history of traumatic or other brain injury, neurosurgery or longer loss of consciousness.)
2. premenstrual syndrome (ICD-10 N94.3).

and criteria related to magnetic resonance imaging (MRI):

1. MRI-incompatible metal parts in the body
2. inability to sit or lie still for a longer period
3. possibility of presence of any metal fragments in the body
4. pregnancy
5. pacemaker, neurostimulator or any other head or heart implants
6. claustrophobia

7. dependence on hearing aid.

For patients the following additional criteria would led to exclusion:

1. current psychotropic medication other than antidepressants
2. questionable history of major depressive episodes without complicating factors
3. current acute suicidality
4. lifetime or current axis II diagnosis of borderline or antisocial personality disorder
5. lifetime or current psychotic disorder of any kind, bipolar disorder
6. current posttraumatic stress disorder, obsessive compulsive disorder, or eating disorder
7. current drug use disorder (with the exception of nicotine) or within the past 5 years.

Healthy controls were excluded if there was a lifetime history of DSM-IV TR axis I or axis II disorder with the exception of nicotine dependence.

S3 Supplementary tables

Name	Coef.	exp(Coef)	Std. Error	z	Pr($>|z|$)
incr. in amygdala reactivity	−0.002	0.998	0.010	−0.205	0.837
Discontinuation group: 1W2	−0.941	0.390	0.511	−1.843	0.065
Age	0.0032	1.003	0.019	0.170	0.865
Gender	0.164	1.176	0.478	0.340	0.734
Site: Zurich	−0.660	0.517	0.455	−1.449	0.147
incr. in amygdala reactivity & Group: 1W2	0.050	1.051	0.017	2.945	**0.003**

Table S1: **Cox proportional hazards model** Coefficient table from the proportional hazards model for the relapse times, based on data from patients only. Continuous variables were centered, so that the effect of Discontinuation group (compares 1W2 vs. 12W), can be interpreted by assuming other covariates are set to zero (i.e. assuming average age and average increase amygdala reactivity).

model	side	bacc mean	bacc mode	bacc 95%-PPI	p(bacc < 0.5)
T1	r	0.549	0.548	[0.448, 0.65]	0.171
T1	l	0.535	0.533	[0.432, 0.643]	0.258
T1	bil	0.561	0.56	[0.457, 0.669]	0.128
T2	r	0.576	0.574	[0.431, 0.726]	0.158
T2	l	0.506	0.5	[0.368, 0.654]	0.474
T2	bil	0.455	0.447	[0.323, 0.601]	0.741
T2-T1	r	0.711	0.719	[0.567, 0.841]	**0.002**
T2-T1	l	0.491	0.489	[0.356, 0.631]	0.551
T2-T1	bil	0.65	0.653	[0.512, 0.783]	**0.017**

Table S2: **Relapse prediction results**. *model* denotes the time points that were included, and *side* denotes the ROI from which the features were taken. *bacc* denotes the balanced accuracy, *PPI* the 95% posterior predictive interval and $p(bacc < 0.5)$ column contains the respective posterior probability mass below the threshold of 50%.