Associations between white matter microstructure and cognitive decline in major depressive disorder versus controls in Germany: a prospective case-control cohort study

Kira Flinkenflügel*, Susanne Meinert*, Christopher Hirtsiefer, Dominik Grotegerd, Marius Gruber, Janik Goltermann, Nils R Winter, Frederike Stein, Katharina Brosch, Elisabeth J Leehr, Joscha Böhnlein, Katharina Dohm, Jochen Bauer, Ronny Redlich, Tim Hahn, Jonathan Repple, Nils Opel, Robert Nitsch, Hamidreza Jamalabadi, Benjamin Straube, Nina Alexander, Andreas Jansen, Igor Nenadić, Martijn P van den Heuvel, FOR2107 consortium†, Tilo Kircher, Udo Dannlowski

Summary

Background Cognitive deficits are a key source of disability in individuals with major depressive disorder (MDD) and worsen with disease progression. Despite their clinical relevance, the underlying mechanisms of cognitive deficits remain poorly elucidated, hampering effective treatment strategies. Emerging evidence suggests that alterations in white matter microstructure might contribute to cognitive dysfunction in MDD. We aimed to investigate the complex association between changes in white matter integrity, cognitive decline, and disease course in MDD in a comprehensive longitudinal dataset.

Methods In the naturalistic, observational, prospective, case-control Marburg-Münster Affective Disorders Cohort Study, individuals aged 18–65 years and of Caucasian ancestry were recruited from local psychiatric hospitals in Münster and Marburg, Germany, and newspaper advertisements. Individuals diagnosed with MDD and individuals without any history of psychiatric disorder (ie, healthy controls) were included in this subsample analysis. Participants had diffusion-weighted imaging, a battery of neuropsychological tests, and detailed clinical data collected at baseline and at 2 years of follow-up. We used linear mixed-effect models to compare changes in cognitive performance and white matter integrity between participants with MDD and healthy controls. Diffusion-weighted imaging analyses were conducted using tract-based spatial statistics. To correct for multiple comparisons, threshold free cluster enhancement (TFCE) was used to correct α-values at the family-wise error rate (FWE; $p_{tfce-FWE}$). Effect sizes were estimated by conditional, partial R^2 values (sr^2) following the Nakagawa and Schielzeth method to quantify explained variance. The association between changes in cognitive performance and changes in white matter integrity was analysed. Finally, we examined whether the depressive disease course between assessments predicted cognitive performance at follow-up and whether white matter integrity mediated this association. People with lived experience were not involved in the research and writing process.

Findings 881 participants were selected for our study, of whom 418 (47%) had MDD (mean age 36·8 years [SD 13·4], 274 [66%] were female, and 144 [34%] were male) and 463 (53%) were healthy controls (mean age 35·6 years [13·5], 295 [64%] were female, and 168 [36%] were male). Baseline assessments were done between Sept 11, 2014, and June 3, 2019, and after a mean follow-up of 2·20 years (SD 0·19), follow-up assessments were done between Oct 6, 2016, and May 31, 2021. Participants with MDD had lower cognitive performance than did healthy controls ($p<0·0001$, $sr^2=0·056$), regardless of timepoint. Analyses of diffusion-weighted imaging indicated a significant diagnosis×time interaction with a steeper decline in white matter integrity of the superior longitudinal fasciculus over time in participants with MDD than in healthy controls ($p_{tfce-FWE}=0·026$, $sr^2=0·002$). Furthermore, cognitive decline was robustly associated with the decline in white matter integrity over time across both groups ($p_{tfce-FWE}<0·0001$, $sr^2=0·004$). In participants with MDD, changes in white matter integrity ($p=0·0040$, $\beta=0·071$) and adverse depressive disease course ($p=0·0022$, $\beta=-0·073$) independently predicted lower cognitive performance at follow-up.

Interpretation Alterations of white matter integrity occurred over time to a greater extent in participants with MDD than in healthy controls, and decline in white matter integrity was associated with a decline in cognitive performance across groups. Our findings emphasise the crucial role of white matter microstructure and disease progression in depression-related cognitive dysfunction, making both priority targets for future treatment development.

Funding German Research Foundation (DFG).

Germany (Prof R Redlich, Prof N Opel); **Department of Psychiatry, Jena University Hospital/Friedrich-Schiller-University Jena, Jena, Germany** (Prof N Opel); **Core-Facility Brainimaging, Faculty of Medicine, University of Marburg, Marburg, Germany** (Prof A Jansen); **Connectome Laboratory, Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, Amsterdam, Netherlands** (Prof M P van den Heuvel PhD); **Department of Child Psychiatry, Amsterdam University Medical Center, Amsterdam Neuroscience, Amsterdam, Netherlands** (Prof M P van den Heuvel)

Correspondence to: Dr Susanne Meinert, Institute for Translational Psychiatry, University of Münster, 48149 Münster, Germany **s.meinert@uni-muenster.de**

Research in context

Evidence before this study

Cognitive deficits are the most frequent and burdensome symptoms of major depressive disorder and have been associated with worse disease outcomes. White matter microstructure, which is responsible for the rapid transmission of information in the human brain, has been linked to depressive recurrence and cognitive performance in previous cross-sectional studies. However, longitudinal studies are needed to draw conclusions about the temporal dynamics of the underlying neurobiological processes and related measures of disease course. To assess the existing evidence for these associations, we searched PubMed and Web of Knowledge databases for publications in English and German between Jan 1, 2000, and April 10, 2024, using the terms "cogniti* OR neuropsycholog*" AND "white matter* OR diffusion tensor imaging OR DTI OR diffusion weighted imaging OR DWI OR fractional anisotropy OR FA" AND "MDD OR depress*". The search identified several cross-sectional studies that reported associations between white matter integrity and cognitive deficits, but mainly in non-clinical samples or geriatric individuals with depression. When replacing "cogniti* OR neuropsycholog*" with "longitudinal OR prospective", only one study could be identified that has examined changes in white matter integrity between individuals with depression and healthy participants over a period longer than 6 months at the whole brain level. However, this study was hampered by small sample size (N=108), a cross-sectional preprocessing pipeline, and the use of averaged diffusion tensor imaging measures as the main outcome. Finally, combining all search terms, we found no study on the longitudinal associations between white matter integrity, cognitive performance, and disease course in depression.

Added value of this study

To our knowledge, this is the first study to investigate the longitudinal associations between white matter integrity, cognitive decline, and disease course in a large and deeply phenotyped clinical sample of participants including participants with depression and healthy participants. It is also the first well powered study to prospectively investigate alterations in white matter microstructure at the voxel-wise level in participants with depression across the full adult age range. By use of longitudinal and multimodal data, therefore addressing limitations of existing research, we found steeper decline in white matter integrity in participants with depression than in individuals with no history of psychiatric disorders, and robust associations between cognitive decline and decline in white matter integrity across groups. Furthermore, we report that poor depressive disease course and changes in white matter integrity both independently predicted future cognitive deficits, underscoring the clinical significance of our findings. Our results emphasise the relevance of white matter microstructure and disease trajectories in cognitive decline and add to the literature by providing evidence for a possible neural mechanism underlying cognitive deficits in major depressive disorder.

Implications of all the available evidence

Future studies with longer follow-up periods and intervals between multiple assessment points are required to support our findings and further deepen the understanding of temporal dynamics of cognitive performance alterations and white matter microstructure over the long term. These insights might be used in future research to optimise antidepressant treatment options with a stronger focus on cognitive symptoms to reduce the burden of cognitive deficits in the future.

Introduction

Major depressive disorder (MDD) is one of the leading causes of disability globally, affecting more than 220 million people worldwide as of 2021.[1] Cognitive deficits (eg, in concentration, memory, and processing speed) are among the most burdensome symptoms of MDD, with detrimental consequences for individuals' occupational success, relationship satisfaction, and health.[2,3] Approximately 94% of individuals with MDD have cognitive deficits during acute depression,[4] and these symptoms persist even after remission in approximately 73% of individuals.[5] Cognitive deficits are predictors of poor treatment response[6] and have been prospectively associated with adverse disease outcomes.[2] Moreover, cognitive impairments have been found to worsen over the course of depression,[7] as they increase with depressive episodes and correlate with treatment resistance.[5,7] Therefore, understanding the underlying mechanisms that lead to cognitive deficits as MDD progresses is of immense clinical importance.

One promising avenue to gain more insights into these mechanisms is to study white matter microstructure because fast information transmission is necessary for higher cognitive processes. Lower white matter integrity in individuals with MDD than in individuals without any history of psychiatric disorder has been found in the corpus callosum, superior longitudinal fasciculus, and corona radiata,[8,9] with alterations being most pronounced in individuals who have recurrent depressive episodes.[9,10] These structural alterations might be the result of gene–environment interactions and pathological mechanisms after heightened stress responses.[11] Similarly, associations between white matter integrity and cognitive performance have been found in the same fibre tracts in individuals without any history of psychiatric disorder and in individuals with MDD.[12,13] This finding suggests that microstructural alterations in white matter in individuals with MDD might provide a biological substrate for cognitive dysfunction. Given that cognitive deficits increase with disease progression and that alterations in white matter

integrity are most pronounced in relapsing forms of MDD,[10] deterioration in white matter integrity with each depressive episode (often described as scarring) might lead to cognitive decline over time. However, causality cannot be inferred from cross-sectional studies. Therefore, longitudinal studies are needed to draw conclusions about the temporal dynamics of the underlying neurobiological processes and related measures of the disease course.

Longitudinal diffusion-weighted imaging studies in individuals without any history of psychiatric disorder have shown age-related decreases in white matter integrity[14] that might contribute to cognitive decline in healthy ageing. In MDD, few and small longitudinal investigations have been conducted to date, yielding rather contradictory results. Although some studies have found an association between a diagnosis of MDD and loss of white matter integrity, irrespective of a patient's remission status,[15] other studies reported depression state-dependent changes (ie, changes from acute to remitted state) in white matter integrity, with an improvement in white matter integrity 2 years after study inclusion dependent on improvement in symptoms[16] or a link between worsening of depressive symptoms and greater loss of white matter integrity.[17] Despite the relevance of cognitive impairment in depression, to our knowledge, no diffusion-weighted imaging study to date has examined the longitudinal association between changes in cognitive performance and white matter integrity in individuals with MDD.

Therefore, we conducted a large-scale diffusion-weighted imaging study to investigate the longitudinal associations between white matter integrity and cognitive performance in a well powered sample of individuals with MDD and individuals with no history of psychiatric disorders (ie, healthy controls) over 2 years. We first analysed group differences in changes of cognitive performance and white matter integrity between individuals with MDD and healthy controls. Then, we tested the association between longitudinal changes in cognitive performance and white matter integrity. Furthermore, we analysed whether the depressive disease course predicts cognitive performance at follow-up and whether this association is mediated by changes in white matter integrity. We hypothesised that (1) individuals with MDD would have greater loss of cognitive performance than healthy controls, (2) changes in cognitive performance would be associated with changes in white matter integrity, and (3) an adverse depressive disease course would precede deterioration of cognitive performance at follow-up and this association would be mediated by changes in white matter integrity.

Methods

Study design and participants

The German Marburg-Münster Affective Disorders Cohort Study (MACS) is a prospective observational case-control study in which participants were recruited from local psychiatric hospitals in Münster and Marburg, Germany, and newspaper advertisements. Individuals were eligible if they were aged 18–65 years at baseline and were of Caucasian ancestry. Individuals were excluded if they had neurological abnormalities, history of seizures, severe head trauma or longer periods of unconsciousness, severe physical impairment, hypothyroidism without adequate medication, claustrophobia, colour blindness, or general MRI contraindications. Participants with a lifetime diagnosis of schizophrenia, schizoaffective disorder, bipolar disorder, or substance dependence were also excluded from the sample (appendix p 3).

At baseline, all participants were characterised with a deep-phenotyping approach[18] and underwent a psychiatric diagnosis using a structural clinical interview (SCID-I),[19] conducted by trained personnel (including KF, SM, CH, FS, and KB; details are in the appendix [p 3]). For this subsample analysis of MACS, we included individuals who were diagnosed with MDD and individuals who reported no history of any psychiatric disorder (ie, healthy controls).

Hypotheses and analyses were preregistered on May 8, 2023, on the OSF registry. The study was approved by the ethics committees of the medical faculties of the University of Marburg (07/2014) and the University of Münster (2014-422-b-S). All procedures were performed in accordance with ethical guidelines and regulations. Participants provided written informed consent before examination and received financial compensation for participation. People with lived experience were not involved in the research and writing process.

Assessment of cognitive performance and disease course

At baseline (T_0) and at 2 years of follow-up (T_1), cognitive performance was assessed with a comprehensive battery of neuropsychological tests encompassing 12 scores from seven tests covering different facets of cognitive functioning that have been shown to be altered in MDD (appendix pp 4–5).[20] These test scores were z-standardised within timepoints and with healthy controls (ie, HC) as reference points (ie MDD_{T_0}-scores were adjusted to show deviations from the mean of HC_{T_0}-scores and vice versa for MDD_{T_1}-scores to HC_{T_1}-scores) to allow classification of disorder-specific deviations in MDD from the normal variance in healthy controls while largely controlling for learning effects due to repeated administration of cognitive tests. This standardisation was necessary with respect to analyses conducted only in participants with MDD. In the absence of a (healthy) control group, this standardisation ensures the consideration of learning, and repetition effects that could otherwise bias depressive relapse-related changes in cognitive performance.

The disease course during the study interval was recorded via so-called life charting[21] (appendix p 6). Using prominent biographical anchors, participants were supported by trained raters in reconstructing the

number of depressive episodes, cumulative time in (subclinical) depressed state, number and duration of hospitalisations, number of outpatient psychiatric or psychotherapeutic contacts, and the number of days with sick leave during the follow-up interval as accurately as possible. Changes in symptom severity were measured with the 21-item Hamilton Depression Rating Scale.[22]

Diffusion-weighted imaging preprocessing

Diffusion-weighted imaging was measured using 3T MRI scanners (Marburg site: Tim Trio, Siemens, Erlangen, Germany; Münster site: Prisma, Siemens, Erlangen, Germany). Pulse sequence parameters were standardised across both scanning sites to the extent allowed by each platform to maximise the signal-to-noise ratio. Due to a body-coil and a gradient-coil exchange in the MRI scanner at the University hospital of Marburg, three dummy-coded variables, with the MRI scanner at the University of Münster as the reference category, were included as covariates in diffusion-weighted imaging analyses, following recommendations of our dedicated quality assurance protocol.[23] Preprocessing was performed with the FMRIB Software Library (version 6.0.1; FSL) using a custom longitudinal preprocessing stream, which has been reported previously (appendix p 7).[24,25] Analyses focus on fractional anisotropy as the most common diffusion metric. Other diffusion metrics that were analysed are described in the appendix (p 8).

Statistical analyses

Following the methods of Biringer and colleagues,[26] we expected cognitive performance to change in a sample of MDD over a 2-year study interval with a small Cohen's f^2 effect size of approximately 0·04. We were not aware of any study comparing changes in cognitive performance between individuals with MDD and individuals without any psychiatric disorders over time. However, on the basis of a study that examined cognitive changes in people at clinical high risk for psychosis,[27] we expected that effect sizes for the diagnosis × time interaction would not exceed small effects of approximately $f^2 = 0·03$. Effect sizes in previous smaller neuroimaging studies have probably been overestimated; therefore, we generally expected small effect sizes.[28] A conservative Cohen's f^2 effect size of 0·01 was assumed for neuroimaging analyses. A priori power analyses were conducted in G*Power (version 3.1.9.7) with a threshold of α less than 0·05 and β less than 0·20. Assuming an effect size of $f^2 = 0·03$ for the diagnosis × time analysis (analysis 1), a total sample size of 58 was required. Furthermore, assuming an effect size of $f^2 = 0·01$ for the longitudinal difference in white matter as well as for its association with cognitive changes (analysis 2), a total sample size of 200 was minimally required.

Descriptive data and factor structures were analysed using R (version 4.0.2). To generate a measure describing

general cognitive performance, we did an explorative factor analysis using HC_{T0} data and the model fit was checked via confirmatory factor analysis (CFA) using MDD_{T0}, MDD_{T1}, and HC_{T1} data (appendix pp 9–11). The analysis revealed a one-factor solution, which we designated the general cognitive performance (GCP) factor, with higher values indicating better performance in neuropsychological tests and lower scores indicating worse performance. We calculated component scores of HC_{T0} using a linear regression approach, and we applied the regression weights to the other three groups and timepoints. Other methods to quantify cognitive performance (eg individual test scores, unstandardised data) are reported in the appendix (pp 21–25).

To summarise the eight markers of disease course recorded during life charting into a comprehensive factor describing progression of depressive illness during follow-up, we performed another CFA in participants with MDD. The resulting one-factor solution (appendix p 12) was labelled course of illness (CoI), with higher values reflecting a more severe disease course and lower values reflecting less severe disease course. We chose to conduct a CFA because this method considers measurement errors and the relevance of individual variables, resulting in a better estimation of the severity of the clinical disease course than a single variable (eg relapse and time in depressed state) or the sum of all variables (disease loading; appendix p 26).

We did three key analyses. Effect sizes were estimated with conditional, partial R² values (sr^2) following the Nakagawa and Schielzeth method[29] to quantify explained variance. In analysis 1, we tested the depression-associated differences in changes in cognitive performance over time. We analysed a random intercepts, fixed slopes, linear mixed-effect model with GCP as the outcome variable, and time (T_0 vs T_1) and diagnosis (MDD vs healthy control) as regressor variables, while correcting for age. Linear mixed-effect models have the advantage that observations of the within-subjects factor can be correlated. We did not control for covariates that did not change between measurements (eg, sex) because these are already considered by the random intercepts of participants. We analysed the main effects of diagnosis and time and their interaction.

In analysis 2, we examined whether changes in white matter integrity differed between participants with MDD and healthy controls. We analysed a random intercepts, fixed slopes, linear mixed-effect model in FSL with fractional anisotropy as the outcome variable and time and diagnosis as regressor variables, while controlling for total intracranial volume (TIV), age, and coil changes. We analysed the main effect of time and the diagnosis × time interaction (analysis 2A). Next, to analyse the association between changes in cognitive performance and changes in white matter integrity across and between participants with MDD and healthy controls, we entered GCP as a regressor variable into the

previous linear mixed-effect model. We analysed the GCP×time and GCP×diagnosis×time interactions (analysis 2B). Because the linear mixed-effect model in FSL does not allow meaningful cross-sectional contrasts, we did cross-sectional group analyses and associations between GCP and white matter integrity for baseline and follow-up using separate general linear models. We did voxel-wise analyses at the whole-brain level using tract-based spatial statistics (TBSS; correction: threshold-free cluster enhancement [TFCE], 10 000 permutations for $p < 0.050$ Family wise error [FWE] correction, two-sided). Brain slice images were created with MRIcroGL (version 1.2).

For analysis 3, we did a mediation analysis in participants with MDD using a bootstrapping approach implemented in PROCESS to test whether an adverse depressive disease course precedes cognitive deficits and whether this association is mediated by white matter integrity. We entered CoI as the predictor variable, mean change in fractional anisotropy (ie, mean $\Delta FA [FA_{T1} - FA_{T0}]$) derived from significant clusters of analysis 2B as the mediator, and GCP_{T1} as the outcome variable into the model. We included GCP_{T0}, TIV_{T0}, change in TIV, age_{T0}, interscan interval, and coil changes as nuisance covariates. Moreover, to correct the effect of disease course severity before study inclusion, we also included the number of hospitalisations and episodes before baseline as nuisance covariates. Direct and indirect effects were estimated. Effects were considered to be significant if the 95% CI did not include 0.

We did extensive robustness checks, as follows: (1) realisation of general robustness checks (ie, outliers and scanner influences including an out-of-site cross-validation; appendix pp 15–18), (2) evaluation of the analysis strategy (ie, use of TIV and linear mixed-effect *vs* difference images; appendix pp 19–20), (3) testing of other approaches to define primary outcomes (ie, GCP and CoI; appendix pp 21–26), and (4) correction for sociodemographic characteristics (ie, education and sex; appendix pp 27–28) and clinical characteristics (ie, remission, medication, comorbidities, previous disease course, and symptom severity; appendix pp 29–40).

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

We included 881 participants from the MACS, of whom 418 (47%) had MDD (mean age 36·8 years [SD 13·4], 274 [66%] of 418 were female, and 144 [34%] were male) and 463 (53%) were healthy controls (mean age 35·6 years [13·5], 295 [64%] were female, and 168 [36%] were male). All participants were of Caucasian ancestry. Baseline assessments were done between Sept 11, 2014, and June 3, 2019 (T_0), and after

	Healthy controls (n=463)	Participants with MDD (n=418)	p value	Cohen's d
Age, years	35·60 (13·48)	36·78 (13·35)	0·19*	–0·088
Biological sex	..	..	0·62†	..
Male	168 (36%)	144 (34%)	..	..
Female	295 (64%)	274 (66%)	..	..
Education, years‡	14·10 (2·65)	13·39 (2·71)	<0·0001*	0·263
HDRS-21§	1·28 (1·79)	8·72 (6·94)	<0·0001*¶	–1·503

Data are mean (SD) or n (%), unless otherwise stated. HDRS-21=21-item Hamilton Depression Rating Scale. MDD=major depressive disorder. *Calculated using the unpaired two-tailed Student's t test. †Calculated using the χ^2 test. ‡Information was missing for two healthy controls and two patients with MDD. §Information was missing for one healthy control and one patient with MDD. ¶Unequal variances.

Table 1: Baseline characteristics of participants with MDD and healthy controls

	Baseline (T_0)	Follow-up (T_1)	p value	Cohen's d
Disease course*				
Number of depressive episodes†	3·81 (7·04)	4·46 (7·21)	<0·0001‡	0·744
Duration of depressive episodes, months§	42·85 (60·06)	47·55 (61·72)	<0·0001‡	0·720
Number of hospitalisations¶	1·43 (1·78)	1·73 (2·08)	<0·0001‡	0·408
Duration of hospitalisations, months||	11·06 (16·95)	11·99 (17·41)	<0·0001‡	0·393
Duration of subclinical depressive episodes during the follow-up-interval, months**	..	5·25 (5·99)	..	..
Number of outpatient psychiatric contacts during the follow-up-interval††	..	6·47 (11·59)	..	..
Number of outpatient psychotherapeutic contacts during the follow-up interval‡‡	..	27·96 (50·93)	..	..
Number of days with sick leave during the follow-up interval§§	..	2·71 (5·86)	..	..
HDRS-21¶¶	8·69 (6·91)	6·29 (6·12)	<0·0001‡	0·336
Remission status	..	..	<0·0001||||	..
Acute remission	158 (38%)	64 (15%)	..	..
Partial remission	114 (27%)	81 (19%)	..	..
Full remission	146 (35%)	273 (65%)	..	..
Medication				
Medication load index***	1·32 (1·50)	1·01 (1·35)	<0·0001‡	0·236
Antidepressant†††	247 (59%)	195 (47%)	<0·0001||||	..
SNRI	98 (23%)	84 (20%)	0·080||||	..
SSRI	115 (28%)	72 (17%)	<0·0001||||	..
NDRI	11 (3%)	15 (4%)	0·29||||	..
NaSSA	41 (10%)	15 (4%)	<0·0001||||	..
NaRI	0	1 (<1%)	<0·0001||||	..
Tricyclic antidepressants	23 (6%)	14 (3%)	0·039||||	..
MAOI	2 (<1%)	3 (1%)	0·65||||	..
Agomelatine	17 (4%)	17 (4%)	>0·99||||	..
Lithium	7 (2%)	15 (4%)	0·011||||	..
Neuroleptic	61 (15%)	48 (11%)	0·074||||	..
Anticonvulsant	11 (3%)	12 (3%)	0·80||||	..
Stimulant	5 (1%)	3 (1%)	0·32||||	..
Benzodiazepines	4 (1%)	0	<0·0001||||	..
Z-substances	5 (1%)	3 (1%)	0·41||||	..

(Table 2 continues on next page)

	Baseline (T$_0$)	Follow-up (T$_1$)	p value	Cohen's d
(Continued from previous page)				
Comorbidities				
Any comorbid psychiatric disorder	119 (28%)	99 (24%)	0·035||||	..
Generalised anxiety disorder	14 (3%)	5 (1%)	0·013||||	..
Panic disorder	22 (5%)	13 (3%)	0·020||||	..
Agoraphobia	15 (4%)	12 (3%)	0·37||||	..
Social anxiety disorder	40 (10%)	21 (5%)	0·0023||||	..
Specific phobia	21 (5%)	19 (5%)	0·65||||	..
PTSD	20 (5%)	20 (5%)	0·68||||	..
Eating disorder	22 (5%)	18 (4%)	0·39||||	..
Obsessive compulsive disorder.	14 (3%)	10 (2%)	0·046||||	..
Alcohol-use disorder	2 (<1%)	7 (2%)	0·0082||||	..
Substance-use disorder	2 (<1%)	4 (1%)	0·18||||	..

Data are mean (SD) or n (%), unless otherwise stated. HDRS-21=21-item Hamilton Depression Rating Scale. MAOI=monoamine oxidase inhibitors. MDD=major depressive disorder. NaRI=noradrenaline reuptake inhibitor. NaSSA=noradrenergic and specific serotonergic antidepressant. NDRI=norepinephrine-dopamine reuptake inhibitor. PTSD=post-traumatic stress disorder. SNRI=selective serotine-norepinephrine reuptake inhibitor. SSRI=selective serotonin reuptake inhibitor. *If the disease course value at baseline was available, T$_1$ refers to the lifetime disease course, otherwise only the disease course within the study interval is reported. ‡Calculated using the paired two-tailed Student's t test. †Information missing for 16 participants at T$_0$ and seven at T$_1$. §Information missing for 46 participants at T$_0$ and seven at T$_1$. ¶Information missing for three participants at T$_0$ and four at T$_1$. ‖Information missing for eight participants at T$_0$ and seven at T$_1$. **Information missing for nine participants at T$_1$. ††Information missing for 13 participants at T$_1$. ‡‡Information missing for 23 participants at T$_1$. §§Information missing for 20 participants at T$_1$. ¶¶Information missing for one participant at T$_0$ and one at T$_1$. ‖‖Calculated using the McNemar test for repeated measurements; for remission status the McNemar-Bowker-Test for repeated measures with more than two categories was used. ***Information missing for seven participants at T$_1$. †††Information missing for two participants at T$_1$.

Table 2: Clinical characteristics of participants with MDD (n=418) at baseline and follow-up

a mean follow-up of 2·20 years (SD 0·19), follow-up assessments were done between Oct 6, 2016, and May 31, 2021 (T$_1$). Baseline characteristics of participants are shown in table 1, and clinical characteristics of participants with MDD at baseline and at follow-up are shown in table 2. At baseline, among 418 participants with MDD, 158 (38%) had acute MDD, 114 (34%) had partial remission, and 146 (35%) were in full remission.

The analysis of the depression-associated difference in changes in cognitive performance over time (analysis 1), yielded a main effect of diagnosis for GCP (b=–0·404, SE=0·057, $t(879)$=–7·43, p<0·0001, sr^2=0·056; appendix p 41), with participants with MDD having worse scores on neuropsychological tests than healthy controls at both timepoints. The diagnosis effect was mainly driven by participants with acute depression but was still detectable in participants with partial and fully remitted MDD (appendix p 5). Furthermore, we found a main effect of time (b=0·037, SE=0·010, $t(878)$=3·79, p=0·0002, sr^2=0·001), resulting from improvement in cognitive performance over time. Although the diagnosis×time interaction was not significant (p=0·066), this improvement was driven by participants with MDD and associated with reductions in severity of depressive symptoms from baseline to follow-up (appendix p 13).

The analysis on the diagnoses-related difference in changes in white matter integrity over time (analysis 2A)

yielded no main effect of time for fractional anisotropy ($p_{tfce-FWE}$=0·18). However, we found a significant diagnosis×time interaction in the superior longitudinal fasciculus ($p_{tfce-FWE}$=0·026, sr^2=0·002, k=554 voxels in two clusters; figure 1; appendix pp 33–40). Decrease in fractional anisotropy was greater in participants with MDD than in healthy controls ($p_{tfce-FWE}$=0·049, sr^2=0·002, k=13 voxels in three clusters; appendix pp 33–40). Cross-sectional analyses yielded no significant differences between MDD and healthy controls in fractional anisotropy at baseline or follow-up (appendix p 14). However, when considering only participants with acute depression at baseline, lower fractional anisotropy was observed in MDD at both timepoints than in healthy controls (appendix p 14). These analyses were exploratory, because previous studies have highlighted the relevance of acute symptom severity in cross-sectional differences in white matter integrity.[12]

When including GCP in the model (analysis 2B), we found a GCP×time interaction in a large bilateral cluster comprising the corpus callosum, corona radiata, and superior longitudinal fasciculus, among other regions ($p_{tfce-FWE}$<0·0001, sr^2=0·003, total k=20 551 voxels in 25 clusters; figure 2; appendix pp 33–40). Across both groups, cognitive performance decline was associated with loss of fractional anisotropy over time ($p_{tfce-FWE}$<0·0001, sr^2=0·004, total k=48 834 voxels in five clusters; appendix pp 33–40). The GCP×diagnosis×time interaction was not significant ($p_{tfce-FWE}$=0·42), indicating that the association between the decline in cognitive performance and white matter integrity was similar in strength and direction across both groups. Cross-sectional analyses showed that lower GCP correlated with lower white matter integrity at both timepoints (appendix p 14). The mediation analysis (analysis 3; figure 3), yielded a negative association between CoI during the interval and GCP$_{T1}$ (total effect: β=–0·074, SE=0·028, $t(333)$=–3·11, p=0·0020; direct effect: β=–0·073, SE=0·028, $t(332)$=–3·09, p=0·0022), indicating that adverse disease course prospectively predicted cognitive deficits. Change in fractional anisotropy was also positively associated with GCP$_{T1}$ (β=0·071, SE=4·076, $t(332)$=2·90, p=0·0040), indicating that fractional anisotropy decline was associated with cognitive deficits at follow-up. Neither an association between CoI and change in fractional anisotropy (95% CI –0·0009 to 0·0006), nor an indirect (mediated) effect of CoI on GCP$_{T1}$ through change in fractional anisotropy was found (95% CI –0·012 to 0·007).

Extensive robustness checks showed that our results were not biased by outliers, scanner influences, the analysis strategy, the use of composite values, sociodemographic or clinical characteristics (appendix pp 15–32). We found a small association between current medication intake and cognitive performance decline. By contrast, no associations were observed between current medication intake and white matter integrity, nor
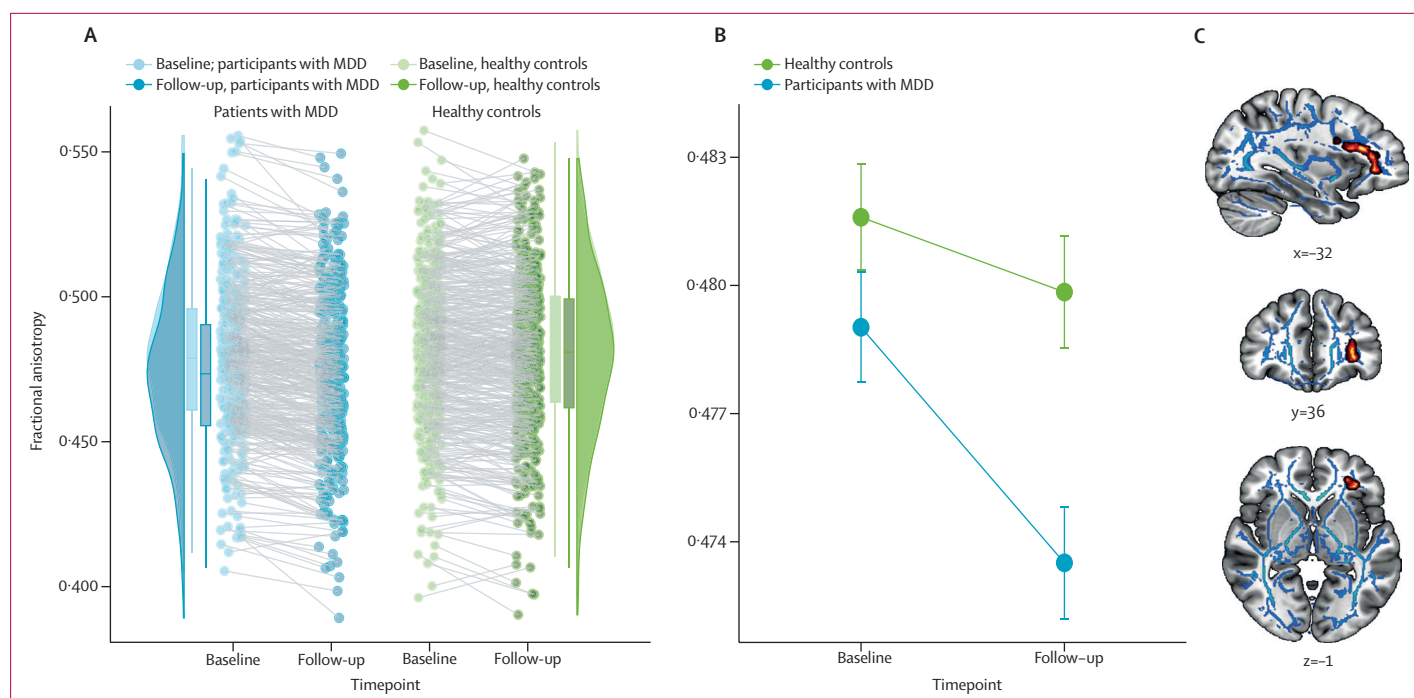
Figure 1: Longitudinal differences in white matter microstructure, measured with fractional anisotropy, between participants with MDD and healthy controls
(A) The distribution, means (indicated boxplots, with the central line being the mean, the edges of the box being the SD and the error bars being the distribution), individual measurements, and slopes of fractional anisotropy, derived from the significant *F* contrast. Each datapoint indicates one participant. (B) Lineplot depicting the longitudinal difference in fractional anisotropy between participants with MDD and healthy controls over the 2-year follow-up period. (C) Interaction effect of time and diagnosis. Statistically significant clusters from the interaction effect are displayed onto the SPM152 template in the x=–32, y=36, z=–1 plane in MNI space using MRIcroGL (version 1.2). Red-yellow colour represents voxels in which a significant interaction effect was found. The effect was mainly located in the superior longitudinal fasciculus. MDD=major depressive disorder. MNI=Montreal Neurological Institute.

between comorbidity and cognitive performance or white matter integrity. Crucially, our robustness checks showed that neither medication nor psychiatric comorbidity affected the overall pattern of results. Results for other diffusion metrics analysed are shown in the appendix (p 8). The association between changes in cognitive performance and white matter integrity was very stable, clearly pronounced, and independently replicable at both sites (Marburg and Münster). We also successfully cross-validated the models trained at one site to the other.

Discussion

To our knowledge, this study presents the most comprehensive prospective dataset to date investigating the complex longitudinal association between changes in white matter integrity, cognitive decline, and disease course in MDD. Using a battery of neurocognitive tests and diffusion-weighted imaging data taken at baseline and at 2 years, we found that participants with MDD had worse results on neurocognitive tests than did healthy controls, regardless of the timepoint. Furthermore, we observed differences in changes in white matter integrity between participants with MDD and healthy controls, with participants with MDD having a greater decrease in white matter integrity in the superior longitudinal fasciculus than healthy controls. As expected, participants with greater decline

in white matter integrity in widespread fibre tracts had a decline in cognitive performance over time, with the same effect found for all tested cognitive domains. Additionally, our analyses showed that depressive disease course during the 2-year follow-up period predicted worse cognitive performance at follow-up. This effect was mainly driven by the results of cognitive tests capturing processing speed. However, changes in white matter integrity did not mediate the link between depressive disease course and cognitive performance at follow-up.

Previous cross-sectional studies have found lower cognitive performance in individuals with MDD than in individuals without any history of psychiatric disorder,[12,20] which is replicated in the present sample at both timepoints. However, contrary to our hypothesis, we found a numerical increase in cognitive performance in participants with MDD over time that was not significant. While this result seems surprising at first, it might be explained by the fact that a large proportion of our participants were less depressed at follow-up than at baseline. This idea aligns with our previous findings that individuals in remission have higher cognitive performance than acutely depressed individuals,[12,20] highlighting the variability of cognitive performance during the clinical course of depression. Nonetheless, some cognitive deficits appear to persist in the remitted state of depression,[20]
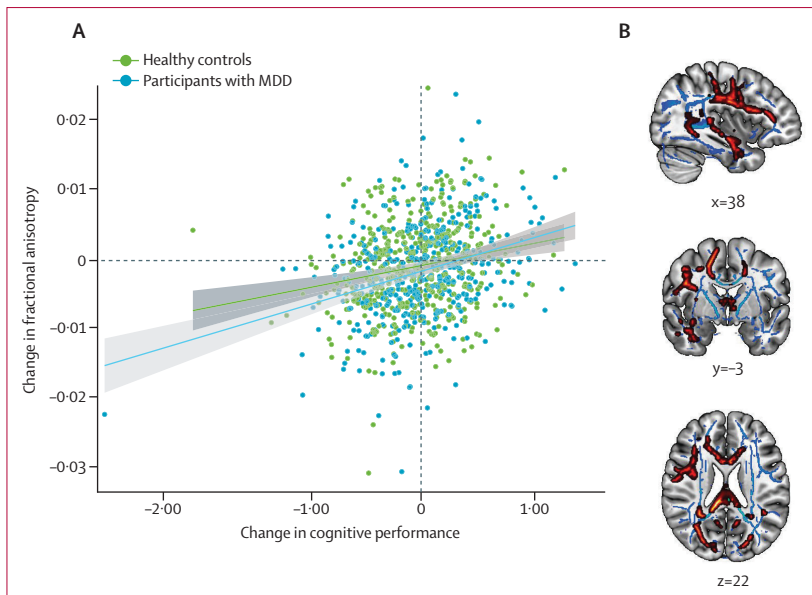
Figure 2: Longitudinal association between changes in cognitive performance and changes in white matter microstructure, measured with fractional anisotropy, in participants with MDD and healthy controls
(A) Scatterplot depicting the longitudinal association between changes in cognitive performance and changes in fractional anisotropy derived from the significant *F* contrast, across participants with MDD and healthy controls. Each datapoint represents one participant. Lines and shaded areas indicate the mean association between changes in fractional anisotropy and cognitive performance as well as the standard deviations. (B) Interaction effect of time and cognitive performance. Statistically significant clusters from the interaction effect are displayed onto the SPM152 template in the x=38, y=−3, and z=22 plane in MNI space using MRIcroGL (version 1.2). Red-yellow colour represents voxels in which a significant interaction effect was found. The effect was widespread and included the corpus callosum, superior longitudinal fasciculus, and corona radiata. MDD=major depressive disorder. MNI=Montreal Neurological Institute

which matches our finding that participants with MDD still performed worse than healthy controls at follow-up. Moreover, we found that disease progression, including the number of depressive relapses and overall time spent in depressive state during the interval, predicted poor cognitive function at follow-up, regardless of an individual's current acute symptomatology. This finding corroborates previous evidence showing an association between the duration of depressive episodes during the assessments and cognitive dysfunction at follow-up.[7] Although further studies are essential to confirm this finding, we suggest a negative effect of disease course on cognitive performance in participants with MDD.

The cross-sectional diffusion-weighted imaging literature consistently points to lower white matter integrity in individuals with MDD than in individuals without any history of psychiatric disorder;[8,9] however, the few longitudinal diffusion-weighted imaging studies to date have yielded a rather ambiguous pattern of results.[15–17] Our analyses suggest a stronger decrease in white matter integrity over a 2-year interval in participants with MDD than in healthy controls, aligning with previous studies.[15] According to established theories, chronic stress can lead to dysregulation of the hypothalamic-pituitary-adrenal axis, resulting in increased production and expression of glucocorticoids.[30] Glucocorticoids reduce growth factors,

decrease dendritic branching, and affect myelination, possibly impairing white matter fibre microstructure.[31] Notably, the effect was particularly pronounced in the superior longitudinal fasciculus, a fibre tract previously considered crucial for language and now recognised for its ipsilateral fronto-temporal and fronto-parietal connections. Given the late maturation of the superior longitudinal fasciculus and its association with healthy ageing, this fibre bundle might be particularly susceptible to neurotoxic effects of prolonged stress during disease progression.[10]

Because white matter fibres form the connections for information transfer between brain regions, impairments in white matter integrity might lead to cognitive dysfunction.[32] This hypothesis corroborates our results, which show a positive association between changes in white matter integrity and changes in cognitive performance over time. Specifically, participants whose cognitive performance decreased had a decline in white matter integrity. This association was independent of diagnosis, which is consistent with our hypothesis and previous cross-sectional findings,[12] indicating a general association between white matter integrity and cognitive performance. By contrast with the diagnosis-specific effects described here, the association between cognitive performance and white matter integrity over time was located in a large bilateral cluster, encompassing several fibre tracts, including the corpus callosum, corona radiata, and superior longitudinal fasciculus. These large association fibres have previously been shown to be most affected in MDD when compared with individuals without any history of psychiatric disorder.[8,9] These tracts, which connect distant areas of the brain, are probably key for information integration that is required for optimal cognitive function.[32] Their deterioration (eg, due to stressful life events or stress during disease progression[8,12]) could lead to cognitive impairments. In summary, these findings suggest an anatomically more global effect that might go beyond the observed diagnosis-related differences. Notably, our results show that even minor changes in cognitive performance, as observed in analysis 1, were associated with robust variations in white matter integrity. Thus, given the longitudinal nature of the current study, we might draw preliminary conclusions regarding a direct association between cognitive performance and white matter integrity across participants with MDD and healthy controls. However, we found no evidence of a mediating effect of changes in white matter integrity on the association between disease course and cognitive performance at follow-up, limiting this interpretation. Nonetheless, the absence of significant results could be due to the relatively short study interval and low variance in disease course in our MDD sample because many participants reported a rather mild disease course. Future studies with longer time intervals should further unravel the complex associations between disease progression, white matter
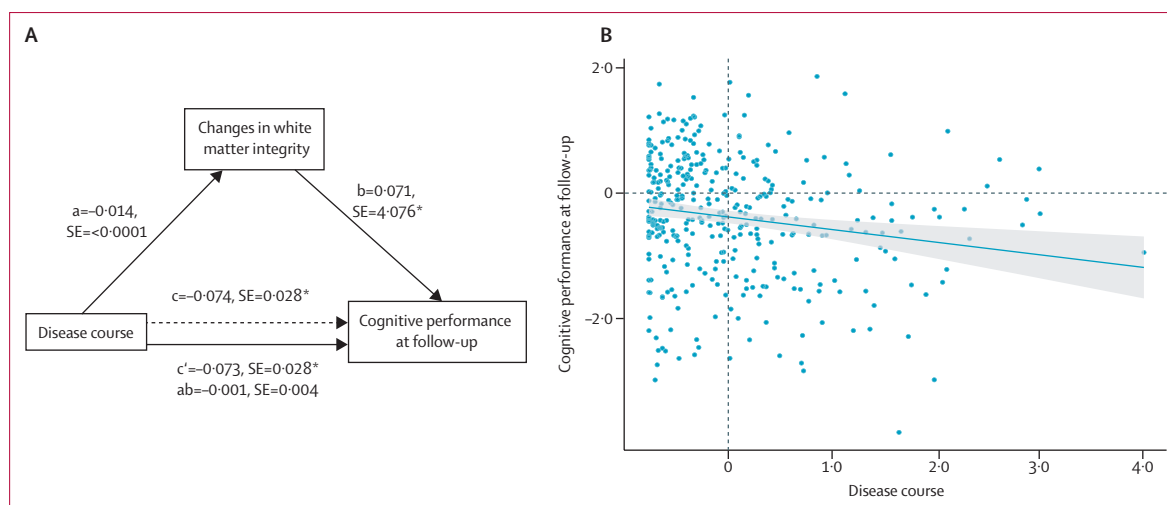
Figure 3: Mediator model with disease course as predictor variable, white matter microstructure changes as mediator, and cognitive performance at follow-up as outcome variable in participants with MDD
(A) Diagram of the mediator model with disease course during the follow-up period as the predictor variable, white matter microstructure changes (follow-up status – baseline status), measured by fractional anisotropy, as the mediator variable, and cognitive performance at follow-up as the outcome variable in participants with MDD. Standardised coefficients and standard errors for each path of the mediation model are presented. c represents the total effect, c☐ the direct effect, and ab the indirect effect. (B) Scatterplot depicting the longitudinal association between disease course in participants with MDD during the follow-up period and cognitive performance at follow-up. *Significance, at p<0·05.

microstructure, and cognitive performance in participants with MDD.

The findings of this study need to also be considered in the context of some limitations. First, the MDD disease course was captured on the basis of retrospective self-reports. Although life charting has shown reliable results elsewhere,[33] reports might be biased by current depressive symptoms. Furthermore, the influence of the MDD disease course on cognitive performance is complex. The CoI factor includes many variables that reflect the extent of impairment caused by the depressive illness. However, even with our extensive robustness checks, accounting for all possible influences known to favour a poor course of the illness, especially other comorbid diagnoses (eg, anxiety disorders), was beyond the scope of the current analysis. Second, despite successful cross-validation, any confounding effects of site or scanner cannot be ruled out completely. Third, although our findings suggest a decline in white matter integrity over time, in our cross-sectional analysis we could not detect lower white matter integrity in participants with MDD than in healthy controls, as has been found in previous diffusion-weighted imaging studies.[8,9] However, our cross-sectional results have shown that reductions in white matter integrity reflect the depressive state and symptom severity to some extent.[34] Hence, it is not surprising that reduced white matter integrity was most pronounced in participants with acute MDD versus healthy controls. These findings suggest that a complex pattern of state and scar effects might be responsible for alterations in white matter integrity in depression. Future studies with longer time intervals and multiple measurement points are needed to further pursue this idea.

Finally, cognitive performance and white matter integrity might be influenced to some degree by clinical characteristics such as medication and comorbidities. This finding might be attributed to a higher disease burden in individuals on psychiatric medication or the use of specific psychiatric agents. Furthermore, we had no information on the number of psychotherapeutic and psychiatric contacts before baseline. However, accounting for the number of previous hospitalisations before baseline did not change the results. Finally, no people with lived experience were involved in shaping the research question and study design, choosing outcome measures, planning recruitment, working as lived experience researchers, assessing the burden of interventions, writing up the study, or dissemination of its findings.

Overall, our study highlights the relevance of white matter microstructure in participants with MDD and provides evidence that changes in white matter integrity might be a possible neural mechanism associated with cognitive performance alterations across participants with MDD and healthy controls. Our results further show that worse disease course predicted poor cognitive performance, emphasising the relevance of considering cognitive alterations in disease progression and effective antidepressant treatment to prevent cognitive deterioration over time. Although the current state of research does not justify regular measurement of diffusion-weighted imaging via MRI in clinical diagnostics, our findings suggest that changes in fibre structure should be considered in future research investigating cognitive deficits in individuals with MDD. Given the negative effect of cognitive deficits on daily functioning and quality of life,[2,20] we strongly recommend greater attention be paid to

cognitive performance and related microstructural correlates in relapse prevention and personalised treatment options. Optimally, current interventions should be combined with a greater focus on enhancing cognitive performance—for example, through cognitive remediation programmes, exercise, or new psychopharmacological agents.

References
1 Institute of Health Metrics and Evaluation. Global Health Data Exchange (GHDx). https://vizhub.healthdata.org/gbd-results/ (accessed Sept 17, 2024).
2 Gonda X, Pompili M, Serafini G, Carvalho AF, Rihmer Z, Dome P. The role of cognitive dysfunction in the symptoms and remission from depression. *Ann Gen Psychiatry* 2015; **14:** 27.
3 McDermott LM, Ebmeier KP. A meta-analysis of depression severity and cognitive function. *J Affect Disord* 2009; **119:** 1–8.
4 Conradi HJ, Ormel J, de Jonge P. Presence of individual (residual) symptoms during depressive episodes and periods of remission: a 3-year prospective study. *Psychol Med* 2011; **41:** 1165–74.
5 Semkovska M, Quinlivan L, O'Grady T, et al. Cognitive function following a major depressive episode: a systematic review and meta-analysis. *Lancet Psychiatry* 2019; **6:** 851–61.
6 Roiser JP, Elliott R, Sahakian BJ. Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology* 2012; **37:** 117–36.
7 Gorwood P, Corruble E, Falissard B, Goodwin GM. Toxic effects of depression on brain function: impairment of delayed recall and the cumulative length of depressive disorder in a large sample of depressed outpatients. *Am J Psychiatry* 2008; **165:** 731–39.
8 Flinkenflügel K, Meinert S, Thiel K, et al. Negative stressful life events and social support are associated with white matter integrity in depressed patients and healthy controls: a diffusion tensor imaging study. *Biol Psychiatry* 2023; **94:** 650–60.
9 van Velzen LS, Kelly S, Isaev D, et al. White matter disturbances in major depressive disorder: a coordinated analysis across 20 international cohorts in the ENIGMA MDD working group. *Mol Psychiatry* 2020; **25:** 1511–25.
10 Meinert S, Leehr EJ, Grotegerd D, et al. White matter fiber microstructure is associated with prior hospitalizations rather than acute symptomatology in major depressive disorder. *Psychol Med* 2022; **52:** 1166–74.
11 Otte C, Gold SM, Penninx BW, et al. Major depressive disorder. *Nat Rev Dis Primers* 2016; **2:** 16065.
12 Meinert S, Nowack N, Grotegerd D, et al. Association of brain white matter microstructure with cognitive performance in major depressive disorder and healthy controls: a diffusion-tensor imaging study. *Mol Psychiatry* 2022; **27:** 1103–10.
13 Opel N, Martin S, Meinert S, et al. White matter microstructure mediates the association between physical fitness and cognition in healthy, young adults. *Sci Rep* 2019; **9:** 12885.
14 Schilling KG, Archer D, Yeh FC, et al. Aging and white matter microstructure and macrostructure: a longitudinal multi-site diffusion MRI study of 1218 participants. *Brain Struct Funct* 2022; **227:** 2111–25.
15 Doolin K, Andrews S, Carballedo A, et al. Longitudinal diffusion weighted imaging of limbic regions in patients with major depressive disorder after 6 years and partial to full remission. *Psychiatry Res Neuroimaging* 2019; **287:** 75–86.
16 Repple J, Zaremba D, Meinert S, et al. Time heals all wounds? A 2-year longitudinal diffusion tensor imaging study in major depressive disorder. *J Psychiatry Neurosci* 2019; **44:** 407–13.
17 Shen X, Adams MJ, Ritakari TE, Cox SR, McIntosh AM, Whalley HC. White matter microstructure and its relation to longitudinal measures of depressive symptoms in mid- and late life. *Biol Psychiatry* 2019; **86:** 759–68.
18 Kircher T, Wöhr M, Nenadic I, et al. Neurobiology of the major psychoses: a translational perspective on brain structure and function—the FOR2107 consortium. *Eur Arch Psychiatry Clin Neurosci* 2019; **269:** 949–62.
19 Wittchen H-U, Wunderlich U, Gruschwitz S, Zaudig M. SKID I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen. Interviewheft und Beurteilungsheft. Eine deutschsprachige, erweiterte Bearb. d. amerikanischen Originalversion des SKID I. Göttingen: Hogrefe, 1997 (in German).
20 Rock PL, Roiser JP, Riedel WJ, Blackwell AD. Cognitive impairment in depression: a systematic review and meta-analysis. *Psychol Med* 2014; **44:** 2029–40.
21 Post RM, Roy-Byrne PP, Uhde TW. Graphic representation of the life course of illness in patients with affective disorder. *Am J Psychiatry* 1988; **145:** 844–48.
22 Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; **23:** 56–62.

23 Vogelbacher C, Möbius TWD, Sommer J, et al. The Marburg-Münster Affective Disorders Cohort Study (MACS): a quality assurance protocol for MR neuroimaging data. *Neuroimage* 2018; **172:** 450–60.

24 Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *Neuroimage* 2012; **62:** 782–90.

25 Woolrich MW, Jbabdi S, Patenaude B, et al. Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 2009; **45** (suppl)**:** S173–86.

26 Biringer E, Mykletun A, Sundet K, Kroken R, Stordal KI, Lund A. A longitudinal analysis of neurocognitive function in unipolar depression. *J Clin Exp Neuropsychol* 2007; **29:** 879–91.

27 Hedges EP, See C, Si S, McGuire P, Dickson H, Kempton MJ. Meta-analysis of longitudinal neurocognitive performance in people at clinical high-risk for psychosis. *Psychol Med* 2022; **52:** 2009–16.

28 Marek S, Tervo-Clemmens B, Calabro FJ, et al. Reproducible brain-wide association studies require thousands of individuals. *Nature* 2022; **603:** 654–60.

29 Nakagawa S, Schielzeth H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol Evol* 2013; **4:** 133–42.

30 Lupien SJ, McEwen BS, Gunnar MR, Heim C. Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nat Rev Neurosci* 2009; **10:** 434–45.

31 Jauregui-Huerta F, Ruvalcaba-Delgadillo Y, Gonzalez-Castañeda R, Garcia-Estrada J, Gonzalez-Perez O, Luquin S. Responses of glial cells to stress and glucocorticoids. *Curr Immunol Rev* 2010; **6:** 195–204.

32 Filley CM, Fields RD. White matter and cognition: making the connection. *J Neurophysiol* 2016; **116:** 2093–104.

33 Schärer LO, Krienke UJ, Graf S-M, Meltzer K, Langosch JM. Validation of life-charts documented with the personal life-chart app—a self-monitoring tool for bipolar disorder. *BMC Psychiatry* 2015; **15:** 49.

34 Repple J, Mauritz M, Meinert S, et al. Severity of current depression and remission status are associated with structural connectome alterations in major depressive disorder. *Mol Psychiatry* 2020; **25:** 1550–58.