Psychological Medicine

cambridge.org/psm

Effects of cumulative illness severity on hippocampal gray matter volume in major depression: a voxel-based morphometry study

Dario Zaremba[1], Verena Enneking[1], Susanne Meinert[1], Katharina Förster[1], Christian Bürger[1], Katharina Dohm[1], Dominik Grotegerd[1], Ronny Redlich[1], Bruno Dietsche[2], Axel Krug[2], Tilo Kircher[2], Harald Kugel[3], Walter Heindel[3], Bernhard T Baune[4], Volker Arolt[1] and Udo Dannlowski[1]

[1]Department of Psychiatry and Psychotherapy, University of Muenster, Muenster, Germany; [2]Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany; [3]Department of Clinical Radiology, University of Muenster, Muenster, Germany and [4]Discipline of Psychiatry, University of Adelaide, Adelaide, Australia

Original Article

Cite this article: Zaremba D *et al* (2018). Effects of cumulative illness severity on hippocampal gray matter volume in major depression: a voxel-based morphometry study. *Psychological Medicine* **48**, 2391–2398. https://doi.org/10.1017/S0033291718000016

Received: 22 June 2017
Revised: 23 November 2017
Accepted: 21 December 2017
First published online: 8 February 2018

Key words:
Illness severity; hippocampus; depression; voxel-based morphometry; hospitalization; illness duration; magnetic resonance imaging

Author for correspondence:
Udo Dannlowski, E-mail: dannlow@uni-muenster.de

Abstract

Background. Patients with major depression show reduced hippocampal volume compared to healthy controls. However, the contribution of patients' cumulative illness severity to hippocampal volume has rarely been investigated. It was the aim of our study to find a composite score of cumulative illness severity that is associated with hippocampal volume in depression.

Methods. We estimated hippocampal gray matter volume using 3-tesla brain magnetic resonance imaging in 213 inpatients with acute major depression according to DSM-IV criteria (employing the SCID interview) and 213 healthy controls. Patients' cumulative illness severity was ascertained by six clinical variables via structured clinical interviews. A principal component analysis was conducted to identify components reflecting cumulative illness severity. Regression analyses and a voxel-based morphometry approach were used to investigate the influence of patients' individual component scores on hippocampal volume.

Results. Principal component analysis yielded two main components of cumulative illness severity: *Hospitalization* and *Duration of Illness*. While the component *Hospitalization* incorporated information from the intensity of inpatient treatment, the component *Duration of Illness* was based on the duration and frequency of illness episodes. We could demonstrate a significant inverse association of patients' *Hospitalization* component scores with bilateral hippocampal gray matter volume. This relationship was not found for *Duration of Illness* component scores.

Conclusions. Variables associated with patients' history of psychiatric hospitalization seem to be accurate predictors of hippocampal volume in major depression and reliable estimators of patients' cumulative illness severity. Future studies should pay attention to these measures when investigating hippocampal volume changes in major depression.

Introduction

Morphological changes in the brain of patients with major depression have been reported by numerous neuroimaging studies and were confirmed in recent meta-analyses (Campbell *et al.* 2004; Videbech & Ravnkilde, 2004; Koolschijn *et al.* 2009; Schmaal *et al.* 2016). Decreased gray matter volumes of the hippocampus, the anterior cingulate cortex and parts of the dorsolateral and dorsomedial prefrontal cortex are among the most reliable alterations (Arnone *et al.* 2012; Bora *et al.* 2012; Du *et al.* 2012). Because of its susceptibility to stressful life experiences such as childhood maltreatment, which are known to increase the risk for the development of major depression (Gilbert *et al.* 2009), the hippocampus has been the focus of structural magnetic resonance imaging (MRI) studies in affective disorders (Dannlowski *et al.* 2012; Stratmann *et al.* 2014; Opel *et al.* 2016).

Although decreased hippocampal volume in major depression is a robust and reliable finding, the effect size of −0.14 standard deviations compared to healthy controls is rather small (Schmaal *et al.* 2016). One further problem repeatedly noted in meta-analyses is the high heterogeneity of sample and clinical characteristics. Whereas some studies included primarily young patients early in the disease process or patients with few illness episodes (Posener *et al.* 2003; Lange & Irle, 2004), others focused on late-life or recurrent depression (Sheline *et al.* 1999; Taylor *et al.* 2005). Variance in hippocampal volume due to clinical heterogeneity has long been treated as an undesirable source of noise. The relevance of these clinical parameters for structural alterations, however, became obvious in a meta-analysis (McKinnon *et al.*

CAMBRIDGE UNIVERSITY PRESS

2009) that demonstrated that decreased volume was predominantly observed in subsamples of patients with more than two years of illness duration or in those with more than one episode. In conjunction with findings on lack of effects of acute symptom severity, these results implicate that gray matter decrease in patients with major depression seems to be the result of protracted illness exposure rather than acute depression. On the cellular level, progressive depression is thought to involve maladaptive neuroplastic changes such as dendritic regression, loss of hippocampal neurons and inhibition of neurogenesis (Pittenger & Duman, 2008). All of these processes have also been observed in response to chronic stress and are likely to be mediated by a hypersecretion of glucocorticoids (Sapolsky, 2000; Conrad, 2008).

The first study to find evidence in favor of this hypothesis has been published 20 years ago and found a negative association of the untreated duration of depression and the volume of the hippocampus in $n = 10$ women (Sheline *et al.* 1996), a result that was replicated by the same group a few years later in $n = 24$ patients (Sheline *et al.* 1999). Only one other study could demonstrate a relationship between the cumulative lifetime duration of depression and hippocampal volume (Frodl *et al.* 2008), whereas other replication attempts failed (Frodl *et al.* 2002; Lloyd *et al.* 2004). A comparable yet slightly different measure than the cumulative lifetime duration of depression is illness duration, which is defined as the elapsed time since onset of the first depressive episode. While some authors reported a negative linear association between illness duration and hippocampal volume (Bell-McGinty *et al.* 2002), others found the relationship to be better approximated by a logarithmic function (MacQueen *et al.* 2003). Yet, others did not find any associations at all (Hickie *et al.* 2005; Frodl *et al.* 2006). Furthermore, the clinical measure of lifetime depressive episodes in association with hippocampal volume showed inconsistent findings between studies. One study and one meta-analysis support the evidence for a negative association (Videbech & Ravnkilde, 2004; Stratmann *et al.* 2014), whereas other studies failed to replicate such an association (Bremner *et al.* 2000; Bell-McGinty *et al.* 2002; MacQueen *et al.* 2003).

One major limitation of these studies is the focus on single, highly selective clinical variables. To obtain a better approximation of patients' cumulative illness severity, it would be preferable to include multiple characteristics of illness severity on a composite score; however, previous studies have not derived a composite measure of cumulative illness severity. A composite measure has the advantage of integrating information from multiple variables and comprises a data-driven weighting of these variables.

To address this, we incorporated information from six clinical variables to investigate the influence of patients' cumulative illness severity on hippocampal volume. Instead of analyzing these variables separately, we conducted a principal component analysis to identify latent components that best characterize cumulative illness severity. These components were tested as predictors of hippocampal volume using voxel-based morphometry in a sample of 213 inpatients with major depression. Additionally, we included a sample of 213 healthy controls to explore differences in hippocampal volume compared with patients. Our study design enabled us to investigate the following objectives: first, to replicate findings of hippocampal volume reductions in patients with major depression compared to healthy controls; second, to investigate whether patients' cumulative illness severity correlates with hippocampal volume.

Materials and methods

Subjects

All inpatients from the Department of Psychiatry and Psychotherapy of the University Hospital in Muenster with a diagnosis of major depression were screened against study inclusion and exclusion criteria by attending therapists. Two hundred sixteen acutely depressed inpatients met criteria and were included. Three patients had to be excluded during preprocessing of MRI data (see the subsection 'Voxel-based morphometry), leaving 213 subjects in the patient sample for statistical analyses. As a control group, we selected 213 healthy controls from an ongoing study investigating the neurogenetics of affective disorders, who were matched to patients according to sex, age, and years of education (for details, see Table 1). Subjects for the healthy control sample were recruited in response to local newspaper ads and public notices. The present sample was independent of our previous investigations on morphometric correlates in depression (Stratmann *et al.* 2014).

Common exclusion criteria were any history of neurological illness, medical condition (e.g. cancer, chronic inflammatory, or autoimmune diseases, and infections), head trauma or unconsciousness, alcohol or substance dependence, psychotic disorders, prior electroconvulsive therapy, and usual MRI contraindications. Regular blood tests for inpatients ensured the absence of substance or alcohol use. All subjects underwent Structured Clinical Interview (SCID-I) for DSM-IV to obtain clinical diagnoses in patients and to ensure no history of psychiatric illness in controls (Wittchen *et al.* 1997). Comorbid anxiety, eating, and somatoform disorders in the patient sample were no exclusion criteria (for details, see Table 1). Acute symptom severity was verified by the 21-item Hamilton Depression Rating Scale (HDRS; Hamilton, 1960) and the Beck Depression Inventory (BDI; Beck *et al.* 1987). All patients had a minimum score of 17 points on the HDRS, whereas healthy controls were allowed to have a maximum score of five points on the HDRS. To assess patients' cumulative illness severity, the following six variables were evaluated in structured clinical interviews by trained raters: number of lifetime depressive episodes, time since first lifetime depressive symptoms, time since the first lifetime psychiatric symptoms, cumulative lifetime duration of depression, cumulative lifetime duration of psychiatric hospitalization, and number of lifetime psychiatric hospitalizations. Most patients were medicated at the time of scanning (for details, see Table 1), while controls were free from any psychotropic medication. To evaluate the potential impact of psychotropic medication in patients, each substance was coded as absent = 0, low = 1 (equal or lower average dose), or high = 2 (>average dose), relative to the midpoint of the daily dose range recommended by Physician's-Desk-Reference. We calculated a medication load index for each patient by summing all individual medication, as used in previous studies (Redlich *et al.* 2015a). The study was approved by the Ethics Committee at the University of Muenster and all participants gave written informed consent prior to commencing any study procedures.

Image acquisition

T1-weighted high-resolution anatomical images of the head were acquired (Gyroscan Intera 3T, Philips Medical Systems, the Netherlands) using a three-dimensional fast gradient echo sequence (turbo field echo), repetition time = 7.4 ms, echo time = 3.4 ms, flip angle = 9°, two signal averages, inversion prepulse every 814.5 ms, acquired over a field of view of 256 mm (feet–

https://doi.org/10.1017/S0033291718000016 Published online by Cambridge University Press

Table 1. Sociodemographic and clinical characteristics of our study sample consisting of 213 acutely depressed patients and 213 healthy controls

	MD mean ± S.D. (range)	HC mean ± S.D. (range)	*p* Value[a]
Sociodemographic characteristics			
Sex (f/m)	114/99	114/99	
Age	38.28 ± 12.01 (18–63)	38.31 ± 11.95 (20–58)	0.985
Years of education	14.70 ± 2.46 (9–23)	14.54 ± 1.96 (10–21)	0.453
Questionnaires			
HDRS	23.34 ± 4.71 (17–42)	0.76 ± 1.19 (0–5)	<0.001
BDI	27.57 ± 8.82 (9–53)	1.08 ± 1.41 (0–5)	<0.001
Clinical characteristics			
Number of lifetime depressive episodes	4.45 ± 5.34 (1–40)	n.a.	n.a.
Time since first lifetime depressive symptoms (months)	101.46 ± 102.52 (1–492)	n.a.	n.a.
Time since first lifetime psychiatric symptoms (months)	127.25 ± 119.62 (3–540)	n.a.	n.a.
Cumulative lifetime duration of depression (months)	30.58 ± 35.54 (1–192)	n.a.	n.a.
Cumulative lifetime duration of psychiatric hospitalization (weeks)	10.68 ± 14.19 (1–81)	n.a.	n.a.
Number of lifetime psychiatric hospitalizations	1.95 ± 1.61 (1–9)	n.a.	n.a.
Medical characteristics[b]			
Medication load index	2.49 ± 1.39 (0–8)	n.a.	n.a.
SNRI	98	n.a.	n.a.
Antipsychotics	80	n.a.	n.a.
SSRI	58	n.a.	n.a.
NaSSA	44	n.a.	n.a.
Tricyclic antidepressants	6	n.a.	n.a.
Mood-stabilizers	13	n.a.	n.a.
Others	26	n.a.	n.a.
None	11	n.a.	n.a.
Depression subtype[c]			
Melancholic	168	n.a.	n.a.
Atypical	13	n.a.	n.a.
Not specified	32	n.a.	n.a.
Lifetime comorbidities[b]			
None	99	n.a.	n.a.
Social phobia	28	n.a.	n.a.
Panic disorder with agoraphobia	25	n.a.	n.a.
Specific phobia	12	n.a.	n.a.
Dysthymia	11	n.a.	n.a.
Generalized anxiety disorder	10	n.a.	n.a.
Obsessive-compulsive disorder	9	n.a.	n.a.
Eating disorder	8	n.a.	n.a.
Posttraumatic stress disorder	8	n.a.	n.a.
Panic disorder without agoraphobia	7	n.a.	n.a.
Agoraphobia without history of panic disorder	7	n.a.	n.a.
Somatoform disorder	3	n.a.	n.a.

MD, major depression ($n = 213$); HC = healthy controls ($n = 213$); HDRS, Hamiton Depression Rating Scale; BDI, Beck Depression Inventory.
[a]*p* Values according to *t* tests.
[b]Multiple entries per patient possible.
[c]Based on DSM-IV-TR criteria.

https://doi.org/10.1017/S0033291718000016 Published online by Cambridge University Press

head) × 204 mm (anterior–posterior) × 160 mm (right–left), frequency encoding in feet to head direction, phase encoding in AP and RL directions, reconstructed to voxels of 0.5 mm × 0.5 mm × 0.5 mm.

Voxel-based morphometry

Hippocampal gray matter information was analyzed using the VBM8-toolbox (http://dbm.neuro.uni-jena.de/vbm). Pre-processing of T1-weighted images was performed using default parameters, as described in previous studies (Dannlowski *et al.* 2015; Redlich *et al.* 2015b). Processing steps included bias-correction, tissue classification and normalization to MNI-space using linear (12-parameter affine) and non-linear transformations including high-dimensional DARTEL-normalization. As suggested in the VBM8 manual, normalized gray matter segments were modulated by non-linear components to compensate for the loss of information in absolute volume through spatial normalization. This step involved multiplying the spatially normalized gray matter segments by its relative volume before and after spatial normalization and thus removed confounding effects of different brain sizes. Non-linear modulation applied the correction for differences in total brain size directly to the data instead of including total intracranial volume in statistical models. Data quality of gray matter images was verified by implemented VBM8-functions using the covariance structure of each image with all other images. Modulated gray matter images were smoothed with a Gaussian kernel of 8 mm FWHM.

Statistical analyses

We conducted a principal component analysis to identify latent components underlying the six clinical variables (number of lifetime depressive episodes, time since first lifetime depressive symptoms, time since the first lifetime psychiatric symptoms, cumulative lifetime duration of depression, cumulative lifetime duration of psychiatric hospitalization, and number of lifetime psychiatric hospitalizations). Component analysis and extraction were computed with IBM SPSS Statistics 23. Based on the Kaiser–Guttman criterion, only components with an eigenvalue (EV) >1.0 were extracted.

Finally, we computed component scores for each extracted factor, which represented patients' individual placement on this component, using a regression approach. To see if extracted factor scores were associated with acute symptom severity at the time of scanning, we computed non-parametric Spearman Rho correlations with HDRS scores.

Analyses of gray matter volume were calculated using statistical parametric mapping software (SPM12, Wellcome Department of Cognitive Neurology, London, UK; http://www.fil.ion.ucl.ac.uk/spm), using an absolute threshold masking of 0.1. To investigate differences in hippocampal volumes between patients and controls, we performed an ANCOVA with diagnosis as a between-subject factor (patients *v.* healthy controls), controlling for age, sex and years of education (objective 1). Although patients and controls were well matched for sociodemographic variables, we decided to include age, sex, and years of education as covariates in our model to account for small residual covariate imbalance between groups (Stuart, 2010). Second, multiple regression analyses restricted to the patient sample were applied to investigate potential associations between hippocampal volume and obtained component scores (objective 2). Each regression model comprised the particular component score, controlling for age, sex, BDI, and medication load index. To exclude confounding effects of antipsychotic medication and comorbid psychiatric illness, we repeated analyses with both variables as additional covariates (yes/no) to the regression models. Equivalent regression models were also conducted for each of our six clinical variables separately to demonstrate effects independent of component selection.

Using a region-of-interest (ROI) approach, we restricted analyses to the bilateral hippocampal and parahippocampal gyrus as defined by the AAL-atlas (Tzourio-Mazoyer *et al.* 2002), implemented in the WFU pickatlas (Maldjian *et al.* 2003). Significance thresholds for multiple testing were obtained at the cluster-level by threshold-free cluster enhancement as a non-parametric approach, which is implemented in the TFCE-toolbox (http://dbm.neuro.uni-jena.de/tfce, Version 110). We established a conservative FWE-corrected threshold of $p < 0.05$ for the ROI obtained by 5000 permutations per test. Additionally, we performed exploratory whole-brain analyses at $p < 0.001$, uncorrected, with a cluster threshold of $k = 50$ voxels.

Results

Principal components of cumulative illness severity

Three patients had to be excluded for anatomical abnormalities, identified as extreme outliers in the check data quality function. The final sample for statistical analyses comprised $n = 213$ patients and $n = 213$ healthy controls.

Following principal component analysis, two components were extracted and retained ($EV_{component1} = 3.47$, $EV_{component2} = 1.01$). Both components together accounted for 75% of the variance (58% by the first component, 17% by the second component) and were highly correlated ($r = 0.52$). To enhance component interpretability, we performed an oblique promax rotation. Rotated factor loadings yielded a clear allocation of each variable to one of the two components (all factor loadings >0.60; see online supplementary Table S1). Variables associated with the first component were number of lifetime depressive episodes, time since first lifetime depressive symptoms, time since first lifetime psychiatric symptoms, and cumulative lifetime duration of depression. This component was termed *Duration of Illness* (Cronbach's $\alpha = 0.68$). Variables associated with the second component were cumulative lifetime duration of psychiatric hospitalization and number of lifetime psychiatric hospitalizations. This component was termed *Hospitalization* ($r_S = 0.80$). Whereas *Hospitalization* factor scores were significantly correlated with HDRS scores ($r_S = 0.159$, $p = 0.020$), *Duration of Illness* factor scores showed a tendency in the same direction but failed to reach significance ($r_S = 0.132$, $p = 0.054$).

Hippocampal volume differences between patients and healthy controls

The ANCOVA showed a tendency of reduced hippocampal gray matter volumes in patients compared to healthy controls, which did not survive rigorous FWE-correction (Cohen's $d = -0.35$, $p_{FWE} = 0.075$). The opposite contrast (patients>healthy controls) revealed no significant clusters in the hippocampus.

Results from exploratory whole-brain analyses are listed in online supplementary Table S2. Notably, gray matter reductions

https://doi.org/10.1017/S0033291718000016 Published online by Cambridge University Press

in the right hippocampal formation, fusiform gyrus, and left cerebellum emerged. The only region increased in patients compared with healthy controls was the left middle and inferior occipital gyrus.

Principal component Duration of Illness *and hippocampal volume*

There was no significant association of hippocampal gray matter volume with *Duration of Illness* component scores ($r = -0.17$, N.S.). Out of four variables represented by *Duration of Illness*, only the number of lifetime depressive episodes showed a significant negative association with right hippocampal gray matter volume ($x = 40$, $y = -24$, $z = -15$, $k = 245$, $p_{FWE} = 0.014$, $r = -0.25$). There was no correlation of hippocampal gray matter volume with time since first lifetime depressive symptoms ($r = -0.13$, N.S.), time since first lifetime psychiatric symptoms ($r = -0.15$, N.S.) nor with cumulative lifetime duration of depression ($r = -0.18$, N.S.). Exploratory whole-brain analyses showed a negative association between *Duration of Illness* and gray matter volume in a cluster in the right insula (see online supplementary Table S2).

Principal component Hospitalization *and hippocampal volume*

Multiple regression analyses in the patient sample revealed a significant negative association of bilateral hippocampal gray matter volume with *Hospitalization* component scores (left: $x = -22$, $y = -24$, $z = -23$, $k = 525$, $p_{FWE} = 0.021$, $r = -0.25$; right: $x = 22$, $y = -12$, $z = -12$, $k = 186$, $p_{FWE} = 0.029$, $r = -0.23$, see Fig. 1). Effects of *Hospitalization* remained highly significant after entering antipsychotic medication and comorbid psychiatric illness as additional covariates (left: $x = -22$, $y = -24$, $z = -23$, $k = 466$, $p_{FWE} = 0.021$, $r = -0.25$; right: $x = 22$, $y = -10$, $z = -12$, $k = 160$, $p_{FWE} = 0.037$, $r = -0.21$). Regression analyses based on the two raw variables revealed a significant negative association of left hippocampal gray matter volume with cumulative lifetime duration of psychiatric hospitalization ($x = -33$, $y = -33$, $z = -12$, $k = 422$, $p_{FWE} = 0.021$, $r = -0.24$) and a tendency for a negative correlation with number of lifetime psychiatric hospitalizations (left: $x = -22$, $y = -24$, $z = -23$, $k = 87$, $p_{FWE} = 0.070$, $r = -0.22$; right: $x = 24$, $y = -10$, $z = -12$, $k = 84$, $p_{FWE} = 0.065$, $r = -0.23$).

Exploratory whole-brain analyses showed a negative association of *Hospitalization* and gray matter volume in clusters, which comprised parts of the temporal, frontal and occipital lobe as well as the hippocampus (see online supplementary Table S2).

As the main effect of group (patients<healthy controls) did not reach significance, we wanted to investigate if there was a hippocampal reduction if we only compared patients with high *Hospitalization* factor scores (upper half determined by a median split, $n = 107$) to healthy controls. Therefore, we performed an additional ANCOVA with diagnosis as a between-subject factor (patients with high *Hospitalization* factor scores *v.* healthy controls), controlling for age, sex and years of education. This yielded a highly significant effect, which indicated reduced bilateral hippocampal volumes in these patients compared to healthy controls (left: $x = -30$, $y = -24$, $z = -26$, $k = 781$, Cohen's $d = -0.45$, $p_{FWE} = 0.009$; right: $x = 22$, $y = -27$, $z = -15$, $k = 935$, Cohen's $d = -0.60$, $p_{FWE} = 0.001$).

Discussion

The present data highlight the influence of cumulative illness severity on hippocampal gray matter volume in patients with major depression. To the best of our knowledge, this was the first study to identify principal components based on multiple clinical variables, which characterize patients' cumulative illness severity. The two components *Duration of Illness* and *Hospitalization* explained a significant amount of heterogeneity in cumulative illness severity. Furthermore, we could demonstrate robust inverse effects of the component *Hospitalization* on bilateral hippocampal gray matter volume.

In contrast to previous studies, we found only a trend of gray matter reductions in patients with major depression compared to healthy controls. Yet, those patients with a severe course of illness did show significant hippocampal volume reductions compared to healthy controls, which indicates that hippocampal alterations might be limited to severe depression. This observation is in line with previous findings on most pronounced reductions in

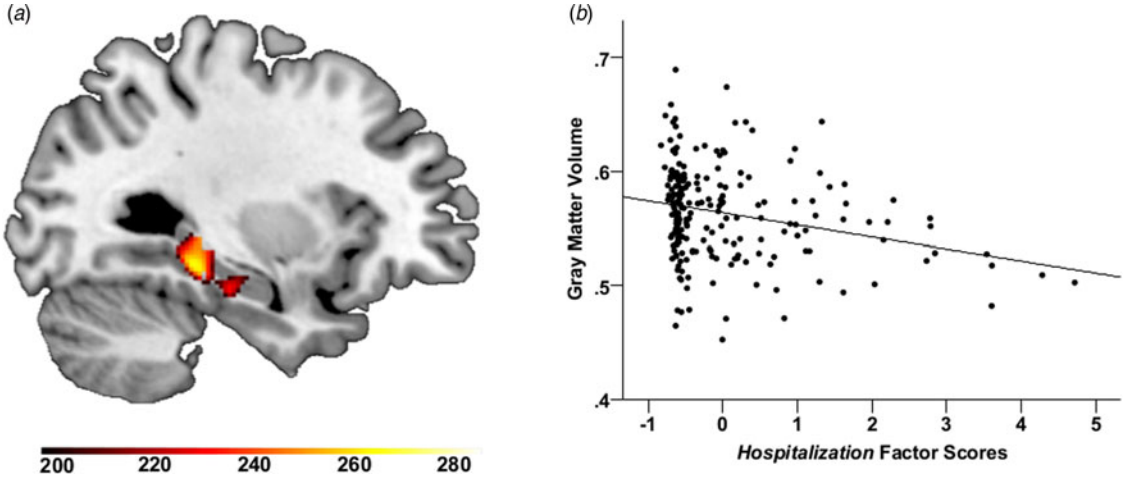

Fig. 1. Effect of Hospitalization component scores on hippocampal gray matter volume in patients with major depression. (*a*) Sagittal view ($x = -31$) depicting gray matter volume negatively associated with Hospitalization component scores. Color bar: TFCE values. (*b*) Scatter plot depicting gray matter volume at $x = -22$, $y = -24$, $z = -23$ correlated with Hospitalization component scores within the patient sample. Continuous line: regression slope.

https://doi.org/10.1017/S0033291718000016 Published online by Cambridge University Press

recurrent depression (Stratmann *et al.* 2014; Schmaal *et al.* 2016), early-onset depression (Schmaal *et al.* 2016), and patients with at least two years of illness (McKinnon *et al.* 2009). If only these extreme groups are considered, effect sizes of hippocampal volume reduction are almost always higher.

The component *Hospitalization* showed a strong association with gray matter values in the hippocampus, a result that could also be replicated given its underlying raw variable cumulative lifetime duration of psychiatric hospitalization. Negative associations of gray matter volume and hospitalization indices have previously been demonstrated in currently depressive inpatients (Axelson *et al.* 1993). In outpatients, however, these associations were absent (Bremner *et al.* 2000, 2002). Unfortunately, we did not include outpatients in our study, which would have been interesting to enhance the generalizability of results. Compared with clinical variables such as illness duration or the number of lifetime depressive episodes, variables related to the intensity (i.e. number and/or duration) of inpatient treatment are still underrepresented in previous research. Given the pathophysiological relevance of these variables for hippocampal volume and their high reliability, future studies should focus more on variables based on inpatient treatment to investigate the neurobiological effects of cumulative illness severity. It should be noted that the effect of *Hospitalization* could be demonstrated, although acute symptom severity as indicated by the BDI was controlled for in the analysis. This is in line with previous results indicating that the effect of cumulative illness severity on hippocampal gray matter volume is independent of symptom severity at the time of scanning (McKinnon *et al.* 2009). Various studies assume that decreased hippocampal volume in patients with major depression is a product of cumulative illness severity (Bell-McGinty *et al.* 2002; MacQueen *et al.* 2003; Frodl *et al.* 2006, 2008; McKinnon *et al.* 2009; Zou *et al.* 2010; Stratmann *et al.* 2014). Given the cellular processes involved in hippocampal atrophy (Sapolsky, 2000), it seems likely that neurotoxic damage in the brain is delayed and not immediately associated with acute symptom severity.

The lack of an association between hippocampal volume and the component *Duration of Illness* was surprising in the light of the neurotoxicity hypothesis. Yet, this result is in line with previous studies, which did not find any association between self-reported illness duration and hippocampal volume (Bremner *et al.* 2000; Frodl *et al.* 2002, 2006; Lloyd *et al.* 2004). On the other hand, methodological problems due to low accuracy of clinical variables might have driven this finding. This assumption is supported by the rather low reliability of *Duration of Illness*. Recalling the onset and duration of illness is very demanding for patients due to the often gradual development of depressive symptoms, the long course of illness over decades and the episodic nature of major depression (Wittchen *et al.* 1989; Patten *et al.* 2012). Furthermore, acutely depressed patients are known for their susceptibility to autobiographical memory biases and show a tendency to summarize categories of events rather than retrieving single episodes (Williams *et al.* 2007). Information about past hospitalizations, however, can easily be recollected because hospital admissions reflect selective incidents in patients' lives (Andrews *et al.* 1999). These differences in reliability might have driven a two-component structure, where accurate and reliable items loaded on the component *Hospitalization* and less accurate or biased items on the component *Duration of Illness*.

The significant negative correlation of the number of lifetime depressive episodes and hippocampal volume has already been shown by previous studies (Videbech & Ravnkilde, 2004; Stratmann *et al.* 2014). This might be attributed to its intermediate position between *Duration of Illness* and *Hospitalization*, which was demonstrated by moderate factor loadings on both components (see online supplementary Table S1). Due to the aforementioned problems with recollection biases, hospitalization indices might nevertheless be more reliable than the number of lifetime depressive episodes to estimate cumulative illness severity.

Given the beneficial effects of antidepressant treatment on hippocampal growth in animal models and human patients (Malberg *et al.* 2000; Surget *et al.* 2008; Arnone *et al.* 2013), it is noteworthy that despite a long history of inpatient treatment and probably longer history of pharmacotherapy, treated patients still tend to show decreased hippocampal volume. This might result from the effect that patients with more and longer hospitalizations are among those with a severe and chronic disease course of illness, who are more likely to be unresponsive to standard antidepressant treatment. Unfortunately, although we collected medication data, we have no information on the level of individual treatment resistance in this sample following a standardized assessment such as the Thase–Rush Treatment-Resistant Depression Staging Method (Thase & Rush, 1997). One might assume that the negative effects of cumulative illness severity on hippocampal gray matter volume were stronger than the compensatory effects of antidepressant medication, which could have led to an underestimation of the true effect sizes of *Hospitalization*. The effects of antidepressant medication might have been either too small to be observed in our study or were covered by stronger opposing effects of *Hospitalization*. Further, we acknowledge that the medication load index we computed was based on information from the current inpatient treatment only. Thus, our medication load indices reflected medication at the time of scanning but not the cumulative lifetime medication load. To exclude confounds due to psychopharmacological effects, more information about patients' past medication would have been necessary. Although our key findings were not substantially influenced by medication load index or antipsychotic medication, these results need replication in unmedicated patients and longitudinal samples where medication is recorded.

Another potential confounder might have resulted from the inclusion of patients with non-affective psychiatric comorbidities. As decreased hippocampal volume has been demonstrated in various disorders, including anxiety disorders (Smith, 2005), our results might be biased by psychiatric comorbidity. As it was important to us to have a large and representative sample of patients with MDD in our study, we decided not to exclude patients with comorbid psychiatric disorders from analyses. However, we accounted for comorbid psychiatric illness by entering an additional covariate to our model, which did not alter the effects of *Hospitalization* on hippocampal volume.

It should be stated that the component *Hospitalization* just met the Kaiser–Guttman criterion with an EV of 1.01. Furthermore, this component might be criticized because of psychometric problems of two-item-scales. It is generally suggested that extracted components should be based on at least three different variables (Velicer & Fava, 1998). However, due to the limited number of variables available, this recommendation could not be met in our study. As both items that underlied *Hospitalization* were highly correlated, the component might still be considered reliable based on information from only two items.

A limitation of our study is the difficulty to objectively assess measures of illness severity and chronicity and to rely on self-reported measures to characterize patients' cumulative illness

https://doi.org/10.1017/S0033291718000016 Published online by Cambridge University Press

severity. Since self-reported variables are often biased by memory inaccuracies or tendencies towards social desirability, it would be advisable to use additional information sources such as reports from relatives, recent and former therapists or medical records. The inquiry about clinical course variables might be further improved by sophisticated assessment techniques such as the life-charting methodology (Post *et al.* 1988), which offers autobiographical anchor points to enhance recollection accuracy. However, most previous studies focused solely on patients' self-report and only a few explicitly reported information from life-charting (Sheline *et al.* 1999, 2003; Lloyd *et al.* 2004) or psychiatric records (Lloyd *et al.* 2004). Furthermore, we were the first to perform principal component analyses on clinical course data, so that our observed two-component-solution needs further replication from independent samples and validation using confirmatory approaches.

The major strength of our study is the integration of multiple clinical variables on composite scores of cumulative illness severity using data-driven principal component analyses. We show that hospitalization indices are well suited to characterize patients' cumulative illness severity because they might be less prone to memory biases. In sum, the present study clearly shows the importance of patients' cumulative illness severity when it comes to decreased hippocampal volume in major depression. We conclude that structural characteristics of the brains of depressed patients seem to be susceptible to cumulative illness severity. Future studies should consider hospitalization indices to get an accurate and reliable approximation of patients' cumulative illness severity. To further evaluate the prognostic power of hospitalizations, longitudinal designs are necessary, which aim at a characterization of patients during different phases of their illness. For a clinical application of our results, it would be interesting to study the plasticity of gray matter reductions trough psychopharmacological and psychotherapeutic interventions. First studies demonstrated hippocampal normalization and growth following electroconvulsive therapy (Redlich *et al.* 2016), treatment with citalopram (Arnone *et al.* 2013) and remission of a current episode (Hou *et al.* 2012; Phillips *et al.* 2015). Thus, the implications of hippocampal changes in major depression for treatment outcome and recovery are yet to be fully understood.

Acknowledgements. This work was supported by the German Research Foundation (U.D., grant numbers DA1151/5-1, DA1151/5-2, and SFB-TRR58 C09), (T.K., grant number KI588/14-1), (A.K., grant number KO4291/3-1); Innovative Medizinische Forschung (U.D., grant numbers DA120903, DA111107, and DA211012); the Rolf Dierichs-Stiftung (U.D., grant number ZUW80037); and the Interdisciplinary Center for Clinical Research (IZKF) of the medical faculty of Muenster (U.D., grant number Dan3/012/17).

Declaration of Interest. V. A. is a member of the advisory board of or has given presentations on behalf of, the following companies: Astra-Zeneca, Janssen-Organon, Lilly, Lundbeck, Servier, Pfizer, Otsuka, and Trommsdorff. These affiliations are of no relevance to the work described in the manuscript. H. Kugel has received consultation fees from MR:comp GmbH, Testing Services for MR Safety. This cooperation has no relevance to the work that is covered in the manuscript. All other authors report no conflicts of interest, financial or otherwise.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291718000016

References

Andrews G, Aantsey K, Brodaty H, Issakidis C and Luscombe G (1999) Recall of depressive episode 25 years previously. *Psychological Medicine* **29**, 787–791.

Arnone D, McIntosh AM, Ebmeier KP, Munafò MR and Anderson IM (2012) Magnetic resonance imaging studies in unipolar depression: systematic review and meta-regression analyses. *European Neuropsychopharmacology* **22**, 1–16.

Arnone D, McKie S, Elliott R, Juhasz G, Thomas EJ, Downey D *et al.* (2013) State-dependent changes in hippocampal grey matter in depression. *Molecular Psychiatry* **18**, 1265–1272.

Axelson DA, Doraiswamy PM, McDonald WM, Boyko OB, Tupler LA, Patterson LJ *et al.* (1993) Hypercortisolemia and hippocampal changes in depression. *Psychiatry Research* **47**, 163–173.

Beck A, Steer R and Brown G (1987) Beck depression inventory manual. *Archives of General Psychiatry* **4**, 561–571.

Bell-McGinty S, Butters MA, Meltzer CC, Greer PJ, Reynolds CF and Becker JT (2002) Brain morphometric abnormalities in geriatric depression: long-term neurobiological effects of illness duration. *The American Journal of Psychiatry* **159**, 1424–1427.

Bora E, Fornito A, Pantelis C and Yücel M (2012) Gray matter abnormalities in major depressive disorder: a meta-analysis of voxel based morphometry studies. *Journal of Affective Disorders* **138**, 9–18.

Bremner JD, Narayan M, Anderson ER, Staib LH, Miller HL and Charney DS (2000) Hippocampal volume reduction in major depression. *American Journal of Psychiatry* **157**, 115–118.

Bremner JD, Vythilingam M, Vermetten E, Nazeer A, Adil J, Khan S *et al.* (2002) Reduced volume of orbitofrontal cortex in major depression. *Biological Psychiatry* **51**, 273–279.

Campbell S, Marriott M, Nahmias C and MacQueen GM (2004) Lower hippocampal volume in patients suffering from depression: a meta-analysis. *The American Journal of Psychiatry* **161**, 598–607.

Conrad CD (2008) Chronic stress-induced hippocampal vulnerability: the glucocorticoid vulnerability hypothesis. NIH Public Access *Reviews in the Neurosciences* **19**, 395–411.

Dannlowski U, Kugel H, Grotegerd D, Redlich R, Suchy J, Opel N *et al.* (2015) NCAN cross-disorder risk variant Is associated With limbic gray matter deficits in healthy subjects and major depression. *Neuropsychopharmacology* **40**, 2510–2516.

Dannlowski U, Stuhrmann A, Beutelmann V, Zwanzger P, Lenzen T, Grotegerd D *et al.* (2012) Limbic scars: long-term consequences of childhood maltreatment revealed by functional and structural magnetic resonance imaging. *Biological Psychiatry* **71**, 286–293.

Du M-Y, Wu Q-Z, Yue Q, Li J, Liao Y, Kuang W-H *et al.* (2012) Voxelwise meta-analysis of gray matter reduction in major depressive disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **36**, 11–16.

Frodl T, Jäger M, Smajstrlova I, Born C, Bottlender R, Palladino T *et al.* (2008) Effect of hippocampal and amygdala volumes on clinical outcomes in major depression: a 3-year prospective magnetic resonance imaging study. *Journal of Psychiatry & Neuroscience* **33**, 423–430.

Frodl T, Meisenzahl EM, Zetzsche T, Born C, Groll C, Jäger M *et al.* (2002) Hippocampal changes in patients with a first episode of major depression. *The American Journal of Psychiatry* **159**, 1112–1118.

Frodl T, Schaub A, Banac S, Charypar M, Jäger M, Kümmler P *et al.* (2006) Reduced hippocampal volume correlates with executive dysfunctioning in major depression. *Journal of Psychiatry & Neuroscience* **31**, 316–323.

Gilbert R, Widom CS, Browne K, Fergusson D, Webb E and Janson S (2009) Burden and consequences of child maltreatment in high-income countries. *Lancet* **373**, 68–81.

Hamilton M (1960) A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry* **23**, 56–62.

Hickie I, Naismith S, Ward PB, Turner K, Scott E, Mitchell P *et al.* (2005) Reduced hippocampal volumes and memory loss in patients with early- and late-onset depression. *The British Journal of Psychiatry* **186**, 197–202.

Hou Z, Yuan Y, Zhang Z, Bai F, Hou G and You J (2012) Longitudinal changes in hippocampal volumes and cognition in remitted geriatric depressive disorder. *Behavioural Brain Research* **227**, 30–35.

Koolschijn PC, van Haren NE, Lensvelt-Mulders GJ, Hulshoff Pol HE and Kahn RS (2009) Brain volume abnormalities in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. *Human Brain Mapping* **30**, 3719–3735.

Lange C and Irle E (2004) Enlarged amygdala volume and reduced hippocampal volume in young women with major depression. *Psychological Medicine* **34**, 1059–1064.

Lloyd AJ, Ferrier IN, Barber R, Gholkar A, Young AH and O'Brien JT (2004) Hippocampal volume change in depression: late- and early-onset illness compared. *The British Journal of Psychiatry* **184**, 488–495.

MacQueen GM, Campbell S, McEwen BS, Macdonald K, Amano S, Joffe RT *et al.* (2003) Course of illness, hippocampal function, and hippocampal volume in major depression. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 1387–1392.

Malberg JE, Eisch AJ, Nestler EJ and Duman RS (2000) Chronic antidepressant treatment increases neurogenesis in adult rat hippocampus. *Journal of Neuroscience* **20**, 9104–9110.

Maldjian JA, Laurienti PJ, Kraft RA and Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage* **19**, 1233–1239.

McKinnon MC, Yucel K, Nazarov A and MacQueen GM (2009) A meta-analysis examining clinical predictors of hippocampal volume in patients with major depressive disorder. *Journal of Psychiatry & Neuroscience* **34**, 41–54.

Opel N, Zwanzger P, Redlich R, Grotegerd D, Dohm K, Arolt V *et al.* (2016) Differing brain structural correlates of familial and environmental risk for major depressive disorder revealed by a combined VBM/pattern recognition approach. *Psychological Medicine* **46**, 277–290.

Patten SB, Williams J V, Lavorato DH, Bulloch AG, D'Arcy C and Streiner DL (2012) Recall of recent and more remote depressive episodes in a prospective cohort study. *Social Psychiatry and Psychiatric Epidemiology* **47**, 691–696.

Phillips JL, Batten LA, Tremblay P, Aldosary F and Blier P (2015) A prospective, longitudinal study of the effect of remission on cortical thickness and hippocampal volume in patients with treatment-resistant depression. *The International Journal of Neuropsychopharmacology* **18**, 1–9.

Pittenger C and Duman RS (2008) Stress, depression, and neuroplasticity: a convergence of mechanisms. *Neuropsychopharmacology* **33**, 88–109.

Posener JA, Wang L, Price JL, Gado MH, Province MA, Miller MI *et al.* (2003) High-dimensional mapping of the hippocampus in depression. *The American Journal of Psychiatry* **160**, 83–89.

Post RM, Roy-Byrne P and Uhde TW (1988) Graphic representation of the life course of illness in patients with affective disorder. *The American Journal of Psychiatry* **145**, 844–848.

Redlich R, Dohm K, Grotegerd D, Opel N, Zwitserlood P, Heindel W *et al.* (2015a) Reward processing in unipolar and bipolar depression: a functional MRI study. *Neuropsychopharmacology* **40**, 2623–2631.

Redlich R, Grotegerd D, Opel N, Kaufmann C, Zwitserlood P, Kugel H *et al.* (2015b) Are you gonna leave me? Separation anxiety is associated with increased amygdala responsiveness and volume. *Social, Cognitive and Affective Neuroscience* **10**, 278–284.

Redlich R, Opel N, Grotegerd D, Dohm K, Zaremba D, Bürger C *et al.* (2016) Prediction of individual response to electroconvulsive therapy via machine learning on structural magnetic resonance imaging data. *JAMA Psychiatry* **73**, 557–564.

Sapolsky RM (2000) The possibility of neurotoxicity in the hippocampus in major depression: a primer on neuron death. *Biological Psychiatry* **48**, 755–765.

Schmaal L, Veltman DJ, van Erp TGM, Sämann PG, Frodl T, Jahanshad N *et al.* (2016) Subcortical brain alterations in major depressive disorder: findings from the ENIGMA major depressive disorder working group. Nature Publishing Group *Molecular Psychiatry* **21**, 806–812.

Sheline YI, Gado MH and Kraemer HC (2003) Untreated depression and hippocampal volume loss. *American Journal of Psychiatry* **160**, 1516–1518.

Sheline YI, Sanghavi M, Mintun MA and Gado MH (1999) Depression duration but not age predicts hippocampal volume loss in medically healthy women with recurrent major depression. *The Journal of Neuroscience* **19**, 5034–5043.

Sheline YI, Wang PW, Gado MH, Csernansky JG and Vannier MW (1996) Hippocampal atrophy in recurrent major depression. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 3908–3913.

Smith ME (2005) Bilateral hippocampal volume reduction in adults with post-traumatic stress disorder: a meta-analysis of structural MRI studies. *Hippocampus* **15**, 798–807.

Stratmann M, Konrad C, Kugel H, Krug A, Schöning S, Ohrmann P *et al.* (2014) Insular and hippocampal gray matter volume reductions in patients with major depressive disorder. *PLoS ONE* **9**, e102692.

Stuart EA (2010) Matching methods for causal inference: a review and a look forward. NIH Public Access *Statistical Science* **25**, 1–21.

Surget A, Saxe M, Leman S, Ibarguen-Vargas Y, Chalon S, Griebel G *et al.* (2008) Drug-dependent requirement of hippocampal neurogenesis in a model of depression and of antidepressant reversal. *Biological Psychiatry* **64**, 293–301.

Taylor WD, Steffens DC, Payne ME, MacFall JR, Marchuk DA, Svenson IK *et al.* (2005) Influence of serotonin transporter promoter region polymorphisms on hippocampal volumes in late-life depression. *Archives of General Psychiatry* **62**, 537–544.

Thase ME and Rush AJ (1997) When at first you don't succeed: sequential strategies for antidepressant nonresponders. *The Journal of Clinical Psychiatry* **58**, 23–29.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N *et al.* (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* **15**, 273–289.

Velicer WF and Fava JL (1998) Affects of variable and subject sampling on factor pattern recovery. *Psychological Methods* **3**, 231–251.

Videbech P and Ravnkilde B (2004) Hippocampal volume and depression: a meta-analysis of MRI studies. *American Journal of Psychiatry* **161**, 1957–1966.

Williams JMG, Barnhofer T, Crane C, Herman D, Raes F, Watkins E *et al.* (2007) Autobiographical memory specificity and emotional disorder. *Psychological Bulletin* **133**, 122–148.

Wittchen H-U, Burke JD, Semler G, Pfister H, Cranach M Von and Zaudig M (1989) Recall and dating of psychiatric symptoms. American Medical Association *Archives of General Psychiatry* **46**, 437.

Wittchen H-U, Wunderlich U, Gruschwitz S and Zaudig M (1997) *SKID-I. Strukturiertes Klinisches Interview für DSM-IV*. Goettingen: Hogrefe.

Zou K, Deng W, Li T, Zhang B, Jiang L, Huang C *et al.* (2010) Changes of brain morphometry in first-episode, drug-naïve, non-late-life adult patients with major depression: an optimized voxel-based morphometry study. *Biological Psychiatry* **67**, 186–188.