1

2

3

4

5

# Neural excitation/inhibition imbalance and the treatment of severe depression

8

9

10

## Authors

Freek ten Doesschate[1][2]*, Willem Bruin[2], Peter Zeidman[3], Christopher C. Abbott[4], Miklos Argyelan[5], Annemieke Dols[6][7][8], Louise Emsell [9], Philip F.P. van Eijndhoven[10], Eric van Exel[6][7][8], Peter C.R. Mulders[11], Katherine Narr[12], Indira Tendolkar[10], Didi Rhebergen[6][7][8], Pascal Sienaert[13], Mathieu Vandenbulcke[14], Joey Verdijk [1], Mike van Verseveld[1], Hauke Bartsch[15], Leif Oltedal[15][16][17][18], Jeroen A. van Waarde[1] & Guido A. van Wingen[2][19]

## Affiliations

18 *1 Department of Psychiatry, Rijnstate hospital, Arnhem, the Netherlands*

19 *2 Amsterdam UMC, University of Amsterdam, Department of Psychiatry, Amsterdam Neuroscience, Amsterdam, the*

20 *Netherlands*

21 *3 Wellcome Centre for Human Neuroimaging, 12 Queen Square, London, WC1N 3AR, UK*

22 *4 Department of Psychiatry, University of New Mexico School of Medicine, Albuquerque, New Mexico*

23 *5 Center for Psychiatric Neuroscience at the Feinstein Institute for Medical Research, New York, New York*

24 *6 GGZ inGeest Specialized Mental Health Care, Department of Old Age Psychiatry, Oldenaller 1, 1081 HJ Amsterdam,*

25 *the Netherlands*

26 *7 Amsterdam UMC, Vrije Universiteit, Psychiatry, Amsterdam Public Health Research Institute, the Netherlands*

27 *8 Amsterdam UMC, Vrije Universiteit, Psychiatry, Amsterdam Neuroscience, the Netherlands*

28 *9 Katholieke Universiteit Leuven, University Psychiatric Center Katholieke Universiteit Leuven, Leuven, Belgium*

29 *10 Donders Institute for Brain, Cognition and Behavior, Department of Psychiatry, Nijmegen, the Netherlands*

30 *11 Department of Psychiatry, Radboud University Medical Centre, Huispost 961, P.O. Box 9101, 6500 HB Nijmegen, the*

31 *Netherlands*

32 *12 Departments of Neurology, Psychiatry, and Biobehavioral Sciences, University of California, Los Angeles, California*

33 *13 Academic Center for ECT and Neuromodulation (AcCENT), University Psychiatric Center KU Leuven (Catholic*

34 *University of Leuven), Leuven, Belgium*

35 *14 Katholieke Universiteit Leuven, University Psychiatric Center Katholieke Universiteit Leuven, Leuven, Belgium*

36 *15 Mohn Medical Imaging and Visualization Center, Department of Radiology, Haukeland University Hospital, Bergen,*

37 *Norway*

38 *16 Department of Clinical Medicine, University of Bergen, Bergen, Norway*

39 *17 Center for Multimodal Imaging and Genetics, University of California, San Diego, La Jolla, California*

40 *18 Department of Radiology, University of California, San Diego, La Jolla, California*

41 *19 Amsterdam Brain and Cognition, University of Amsterdam, the Netherlands*

## Corresponding author

44 F. ten Doesschate (Ftendoesschate@Rijnstate.nl)

# Abstract

An influential hypothesis holds that depression is related to a neural excitation/inhibition imbalance, but its role in the treatment of depression remains unclear. Here, we show that unmedicated patients with severe depression demonstrated reduced inhibition of brain-wide resting-state networks relative to healthy controls. Patients using antidepressants showed inhibition that was higher than unmedicated patients and comparable to controls, but they still suffered from severe depression. Subsequent treatment with electroconvulsive therapy (ECT) reduced depressive symptoms, but its effectiveness did not depend on changes in network inhibition. Concomitant pharmacotherapy increased the effectiveness of ECT, but only when the strength of neural inhibition before ECT was within the normal range and not when inhibition was excessive. These findings suggest that reversing the excitation/inhibition imbalance may not be sufficient nor necessary for the effective treatment of severe depression, and that brain-state informed pharmacotherapy management may enhance the effectiveness of ECT.

# Main

With more than 250 million people affected, depression is one of the leading causes of disability worldwide[1]. Patients suffer weeks to sometimes years of low mood, anhedonia, sleep problems, weight loss, and — in more severe cases — suicidality, motor retardation and psychotic features. Although 70% of patients show a positive response to (extensive) treatments with pharmacotherapy and psychotherapy[2], approximately half of them eventually relapse and experience one or more recurrent episodes in their lifetime[2,3]. Patients suffering from severe depressive episodes that are not responsive to initial treatments may benefit from electroconvulsive therapy (ECT). With a remission rate of 48-65%, ECT is currently the most effective treatment for pharmacotherapy-resistant depression[4]. Nevertheless, it is clear that for a substantial group of depressed patients an effective treatment is still lacking. Enhancing our understanding on the neural mechanisms of depression and its treatments is indispensable to improve treatment effectiveness and to develop novel treatments.

Healthy functioning of cortical brain circuits depends on processes that promote homeostasis throughout the network, which is regulated by excitatory and inhibitory neurotransmission[5]. One of the proposed neural mechanisms of depression states that symptoms are caused by deficits in inhibitory neurotransmission, resulting in an excitation/inhibition imbalance[6–9]. In line with this hypothesis, deficits in the principal inhibitory neurotransmitter γ-aminobutyric acid (GABA) seem to cause depression-like behavior in animal models[10,11]. Furthermore, GABA concentrations are reduced in depressed patients but not in remitted patients[12]. Accordingly, this hypothesis also predicts that reversing the excitation/inhibition imbalance in depressed patients will lead to a reduction of depressive symptoms[7,9,10]. Indeed, even though classical antidepressants

84  primarily target other neurotransmitters, these agents seem to stimulate GABA-ergic

85  neurotransmission as well[6,7,9,13,14]. Furthermore, benzodiazepines (i.e., a $GABA_A$ receptor

86  agonist) provide early antidepressant effects when combined with antidepressants[15]. Thus,

87  this hypothesis poses that the neural excitation/inhibition imbalance takes a central

88  position in the pathophysiology of depression and its treatment.

89

90  Similar to pharmacotherapy, the effectiveness of ECT may also depend on mechanisms

91  that are closely related to the neural excitation/inhibition balance[16]. ECT uses a brief

92  electrical stimulus to the cranium in order to elicit seizure activity which propagates

93  throughout the brain[17]. For a beneficial treatment response, the administered stimulus

94  should significantly exceed the individuals' seizure threshold[18], thereby providing strong

95  excitation of the brain. Furthermore, ECT effectiveness depends on the generalization of

96  the induced seizure[19], which is accomplished through the opposing forces of excitatory

97  seizure activity and inhibitory restraints[20]. Also, ECT effectiveness is associated with the

98  strength of the post-ictal suppression of induced seizure activity[21]. This phenomenon is

99  marked by strong inhibitory processes terminating the excitatory seizure dynamics. Thus,

100 the working mechanisms of ECT seem to be tightly linked to the excitation/inhibition

101 balance, indicating that disturbances in this balance may influence ECT effectiveness.

102

103 In order to investigate the excitation/inhibition balance in depression and its relation to

104 treatment effects, we used data from a large multicenter international consortium that

105 included patients with severe depression who were treated with pharmacotherapy and

106 subsequent ECT. We used spectral dynamic causal modeling (spDCM) of resting-state

107 functional magnetic resonance imaging (RS-fMRI) data[22]. spDCM is able to accurately

108 estimate extrinsic (i.e., between-region) and intrinsic (i.e., regional self-inhibition) effective

109 connectivity (Figure 1)[22–24]. Well-tuned regional self-inhibition allows for sustained activity

110   in the cortex while maintaining a balance between excitation and inhibition throughout the

111   network. The strength of intrinsic self-inhibition can thus be interpreted as a proxy for the

112   regional excitation/inhibition balance[25]. Recently, a study using magnetoencephalography

113   showed that DCM can quantify the increase in tonic inhibition by a GABA reuptake

114   inhibitor[26]. Here, we use spDCM to study the effects of pharmacotherapy and ECT on

115   extrinsic and intrinsic effective connectivity in severely depressed patients. We provide

116   evidence indicating that reversing the neural excitation/inhibition imbalance is, in and of

117   itself, not sufficient nor necessary for an antidepressant response in patients suffering from

118   severe depression. Our findings also suggest that concomitant pharmacotherapy

119   increases the effectiveness of ECT if the strength of inhibition remains within a healthy

120   range, but not when the degree of inhibition is excessive.

121

122

123

124

125

126

127
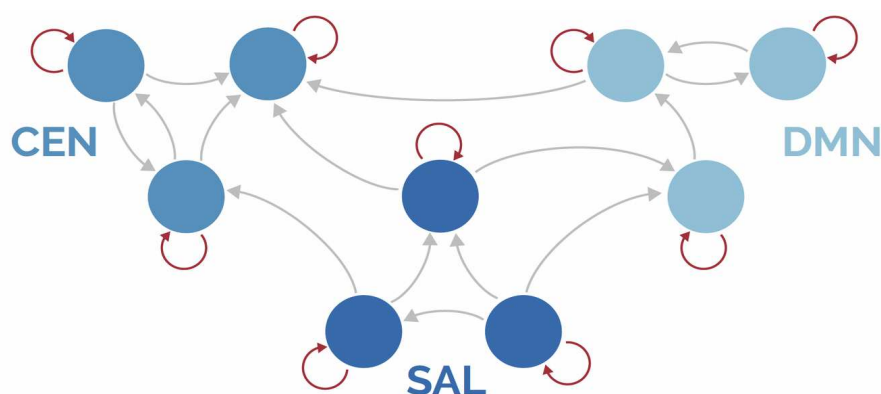


128   **Figure 1 Schematic representation of a directed cortical network.** The nodes (blue circles) represent

129   cortical regions of brain networks that interact through effective connections (gray arrows). The self-

130   connections (red arrows) regulating the strength of regional inhibition. Optimal strength of self-connections

131   allows for sustained activity in the cortex while maintaining a balance between excitation and inhibition

132   throughout the cortical network. The selected networks in this study are the central executive network (CEN),

133   salience network (SAL), and default mode network (DMN).

# Results

**Reversal of the excitation/inhibition imbalance is not sufficient for pharmacotherapy effectiveness in severe depression**

Data were obtained from the Global ECT-MRI Research Collaboration (GEMRIC) database[27]. Participants from seven sites across Europe and North America were divided into four groups: healthy controls (n = 59), unipolar depressed patients without current use of antidepressants or benzodiazepines (patients might use other types of medication, but for convenience were labeled as 'unmedicated group'; n = 60), unipolar depressed patients using antidepressants (n = 56), and unipolar depressed patients using benzodiazepines (n = 63). The latter two groups included patients using both types of medication (n = 24), this considerable overlap was accounted for in the analysis (see SI Methods). Demographic and clinical characteristics of each group are reported in Table 1.

|  | Statistic | HC (n = 59) | NM (n = 60) | Overlap = 24 AD (n = 56) | BZ (n = 63) |
|---|---|---|---|---|---|
| **Age** [a,c,d,e,f] | Mean (+- 1SD) | 48.1 (15.1) | 48.7 (16.1) | 63.1 (11.5) | 54.9 (14.3) |
| **Sex** | Count (%) | 26 (44.1) | 25 (41.7) | 19 (33.9) | 29 (46) |
| **HAM-D** [e] | Mean (+- 1SD) | NA | 23.8 (6.1) | 25.2 (6.6) | 27.6 (7.1) |
| **Antipsychotics use** [d,e] | Count (%) | NA | 4 (6.7) | 29 (51.8) | 25 (39.7) |
| **Lithium carbonate use** | Count (%) | NA | 1 (1.7) | 4 (7.1) | 2 (3.2) |
| **Recurrent depression** | Count (%) | NA | 56 (93.3) | 53 (94.6) | 59 (93.7) |
| **Psychotic features** [d,e] | Count (%) | NA | 8 (13.3) | 17 (30.4) | 10 (15.9) |

**Table 1 Overview of demographic and clinical information.** Mean (±SD) and count (%) statistics of variables for each specific group. Analyses on group differences were performed using student's t-test or $X^2$-test, where appropriate. HC = healthy controls; AD = patients using antidepressants; BZ = patients using benzodiazepines; NM = patients without current AD or BZ medication use; HAM-D = Hamilton Rating Scale for depression; [a] $P_{HC \, vs \, AD} < 0.1$; [b] $P_{HC \, vs \, NM} < 0.1$; [c] $P_{HC \, vs \, BZ} < 0.1$; [d] $P_{NM \, vs \, AD} < 0.1$; [e] $P_{NM \, vs \, BZ} < 0.1$; [f] $P_{AD \, vs \, BZ} < 0.1$.

154 We investigated whether these groups differed in effective connectivity using RS-fMRI and

155 spDCM (see SI Methods)[22]. On the individual level, spDCM modeled causal interactions

156 between neuronal populations located in the regions of interest (ROIs). A forward model

157 generated simulated RS-fMRI timeseries based on these neural models. The cross-

158 spectra of these simulated timeseries were compared to cross-spectra of observed RS-

159 fMRI timeseries. Specifically, Bayesian statistical methods identified neural connectivity

160 parameters that maximized model accuracy (the similarity of the simulated and observed

161 cross-spectra), while minimizing complexity (the effective number of parameters). Using

162 parametric empirical Bayes, the connectivity parameters were then taken to the group

163 level and modeled using a general linear model (GLM; for model designs, see SI Methods

164 and SI Figures). Thereby, spDCM enabled inference on effective (i.e., directed)

165 connectivity between brain regions (Figure 1). Additionally, it allowed inference on the

166 strength of regional inhibitory self-connections, which could be interpreted as local

167 regulation of the excitation/inhibition ratio in order to maintain balance throughout the

168 network (Figure 1)[25]. Fourteen ROIs were selected, located at the core of the default mode

169 network (DMN), salience network (SN) and central executive network (CEN; see SI Figure

170 1). These networks are implicated in the pathophysiology of depression[28]. The analyses

171 yielded no evidence for effects in between-region connectivity, so we focused on the

172 effects in self-connections below. By comparing the evidence for the full connectivity GLM

173 against specific reduced GLMs (SI Methods), we examined effects in inhibitory self-

174 connections of individual ROIs. Also, we examined effects at the network-level, which

175 tested if effects on self-inhibition were present in none of the networks (i.e., evidence for

176 absence of an effect), in a specific network, or widespread across networks (see SI Figure

177 3 and SI Methods). The degree of evidence was expressed as the probability of an effect

178 in a (combination of) connection(s) given the data (i.e., the posterior probability, written as

179   p(m|Y)). Evidence was considered strong if the posterior probability was 95% or higher

180   (i.e., p(m|Y) > 0.95).

181

182   Because multiple GLM designs could be used to examine cross-sectional group

183   differences, we tested the two most convenient designs and selected the one which

184   maximized model evidence (SI Figure 2). The winning model design consisted of separate

185   regressors for each patient group, while healthy controls were modeled as the reference

186   group (i.e., intercept). The results of the selected model showed that self-inhibition was

187   reduced across networks in unmedicated patients compared to healthy controls ($p(m_4|Y)$ =

188   1.0). Patients using benzodiazepines showed an opposite pattern of increased self-

189   inhibition across networks compared to healthy controls ($p(m_4|Y)$ = 1.0). Using Bayesian

190   statistics allowed us to establish the evidence for a null finding. In this case, the results

191   suggest that self-inhibition in patients using antidepressants was comparable to healthy

192   controls, but the level of evidence was not conclusive ($p(m_0|Y)$ = 0.53). The analyses for

193   effects on individual nodes were in line with the above (Figure 2). These findings were

194   mostly robust against adjustments for potential confounding variables (i.e., age, sex,

195   depression severity, presence of recurrent depression, presence of psychotic depression,

196   use of lithium carbonate, use of antipsychotics, and site variables; SI Figure 4). The only

197   substantial difference in the adjusted analyses appeared in patients using

198   benzodiazepines, showing increased self-inhibition exclusively in specific network nodes

199   rather than a widespread increase across networks.
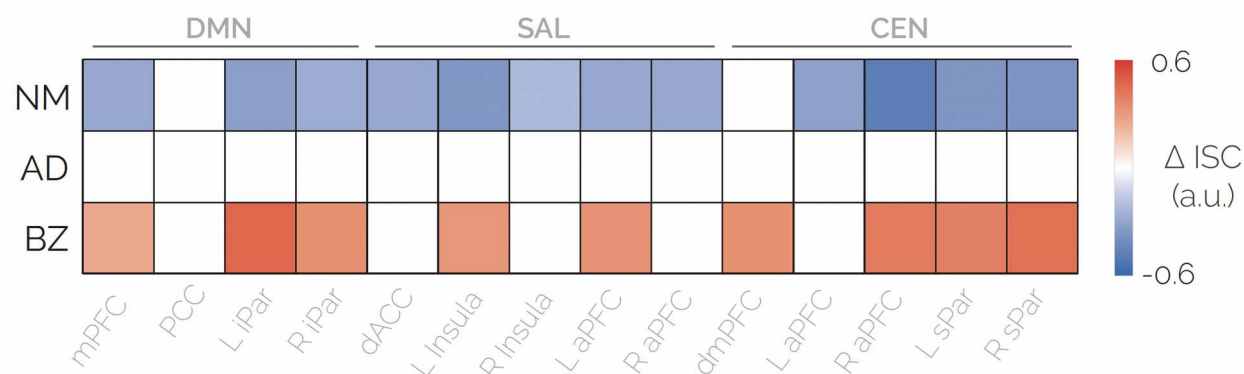
**Figure 2 Group differences in region-specific inhibitory self-connections.** Results of the analysis using automatic search over reduced models to obtain probabilities for specific connectivity parameters. Colored squares indicate very strong evidence ($p$(Y|m) > 0.99) for increased (red) or reduced (blue) self-inhibition between the specified patient group and healthy controls. ΔISC = difference in inhibitory self-connection strength; AD = patients using antidepressants; BZ = patients using benzodiazepines NM= patients not using AD or BZ. DMN = default mode network; SAL = salience network; CEN = central executive network. L= left; R = right; mPFC = medial prefrontal cortex; PCC = posterior cingulate cortex; iPar = inferior parietal cortex; dACC = dorsal anterior cingulate cortex; aPFC = anterior prefrontal cortex; dmPFC = dorsomedial prefrontal cortex; sPar = superior parietal cortex.

We used post-hoc pairwise comparisons to test for differences between patient groups (SI Methods and SI Figure 5). Compared to unmedicated patients, patients using antidepressants and/or benzodiazepines showed increased inhibitory self-connection strengths in all nodes across networks (p(m$_4$|Y) = 1.0 for both groups). Compared to patients using antidepressants, benzodiazepines users showed increased inhibitory self-connections in specific nodes (particularly in the CEN; SI Figure 5). The analyses were adjusted for the same covariates as in the group comparison model described above and adjusted for the number of post-hoc comparisons.

220   In sum, these results yielded that, compared to unmedicated patients, antidepressants

221   users showed a higher degree of self-inhibition which was more similar to healthy controls.

222   Importantly, even though inhibitory deviations (from healthy controls) seemed to be

223   normalized in patients using antidepressants, all patient groups still suffered from severe

224   depressive symptoms (Table 1).

225

226   Next, an analysis was conducted on associations between inhibitory self-connections and

227   depression severity in the distinct patient groups, as measured by the Hamilton

228   Depression Rating Scale (HAM-D[29]). The results showed that more severe depression was

229   associated with reduced self-inhibition across networks in unmedicated patients ($p(m_4|Y) =$

230   $1.0$; SI Figure 6). Except from a few individual nodes, no clear associations between

231   depression severity and inhibitory self-connections were found in patients using

232   antidepressants and/or benzodiazepines ($p(m_0|Y) = 0.51$ and $p(m_0|Y) = 0.43$, respectively).

233   The results were robust against adjustments for potential confounders (i.e., age, sex,

234   presence of recurrent depression, presence of psychotic depression, use of lithium

235   carbonate, use of antipsychotics, and site variables; SI Figure 6).

236

237   **Pharmacotherapy-related effects of excitation/inhibition imbalance on ECT**

238   **effectiveness**

239   Next, we examined whether excitation/inhibition imbalance at baseline was associated

240   with ECT effectiveness. We first extracted the inhibitory self-connection parameters of all

241   ROIs and identified the first principal component as a summary measure for regional self-

242   inhibition of cortical networks throughout the brain. Using multiple regression models, we

243   associated this summary measure to ECT effectiveness (i.e., the change in HAM-D scores

244   from before to after ECT). The analysis was adjusted for baseline HAM-D scores and other

245   known predictors of ECT effectiveness (i.e., presence of psychotic features, age, number

246   of ECT-sessions with bilateral electrode positioning, and site variables; see SI Methods).

247   The results showed a quadratic association between inhibitory self-connection and HAM-D

248   change in the total sample ($F(1,143) = 16.2$, $p < 0.001$, Cohen's $f^2 = 0.11$; Figure 3a).

249   Further analyses in the subgroups revealed that the quadratic relation was explained by

250   medication related differences. Whereas no linear association was found in currently

251   unmedicated patients ($F(1,52) = 0.72$, $p = 0.40$), patients using concomitant

252   benzodiazepines and/or antidepressants showed a negative linear association between

253   inhibitory strength and HAM-D change ($F(1,47) = 11.58$, $p = 0.001$ for antidepressants

254   users; $F(1,56) = 11.50$, $p = 0.001$ for benzodiazepines users; see Figures 3b and 3c). The

255   associations between neural inhibitory strength and ECT effectiveness in patients using

256   pharmacotherapy showed a medium effect size (Cohen's $f^2 = 0.21$ for benzodiazepines

257   users; Cohen's $f^2 = 0.25$ for antidepressants users), indicating that excessive inhibitory

258   strength might be an important predictor of ECT non-response. Since the regression

259   models were adjusted for potential confounders, the degree of variation in ECT

260   effectiveness explained by inhibition seemed to be independent from other known

261   predictors.

262

263   For the analysis of clinical data, we used a multiple regression model to examine the

264   association between relevant predictors — including antidepressants and benzodiazepines

265   use — and ECT effectiveness (see SI Methods). In line with previous research[30], the

266   analyses revealed that the use of antidepressants enhanced ECT effectiveness ($F(1,158)$

267   $= 7.32$, $p = 0.008$, Cohen's $f^{2} = 0.05$). Therefore, our findings regarding the negative

268   effects of inhibitory strength on ECT effectiveness mediated by antidepressants use

269   seemed counterintuitive. Further analyses showed that the use of medication enhanced
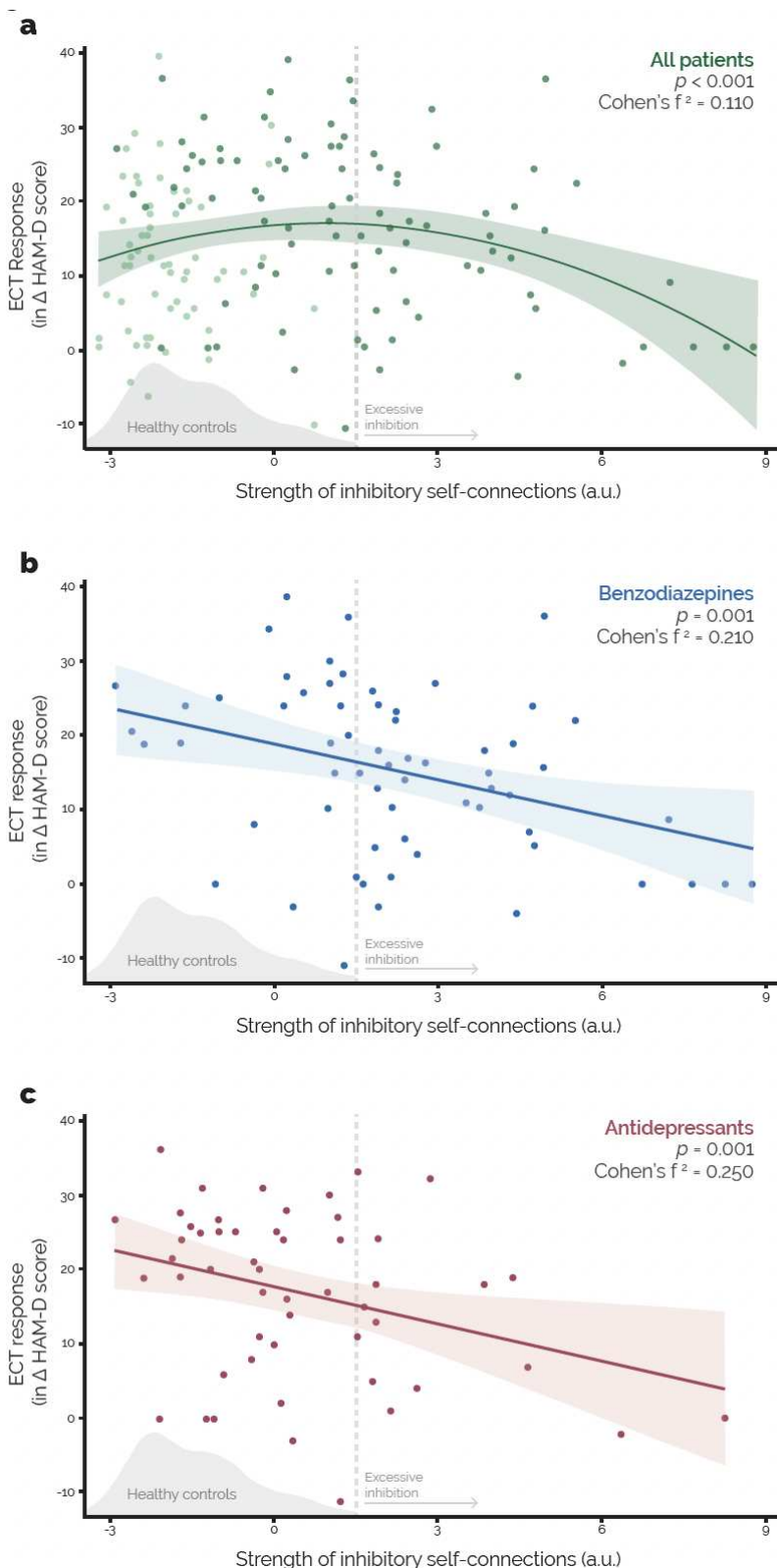
**Figure 3 Associations between baseline inhibitory self-connection strength and ECT effectiveness**

Regression plots on the association between the first principal component of inhibitory self-connections (arbitrary units) and the antidepressant effectiveness of ECT (change in HAM-D score before and after the ECT-course). The figure shows the regression plot for all patients (a), including both unmedicated patients (lighter green) and medicated patients (darker green). Additionally, separate regression plots are shown for patients using benzodiazepines (b) and patients using antidepressant (c). For each regression analysis, p-values and effect sizes for the displayed association are included adjusted for covariates. A medium effect size is indicated by $0.15 \leq f^2 \geq 0.35$. In each plot, a density plot is included to show the distribution of self-inhibition in healthy controls. The strength of patient's self-inhibition that appeared beyond the range of healthy controls is identified as excessive.

300    ECT effectiveness as long as the inhibitory strength was within the range of healthy

301    controls (Figure 3b and 3c). This appeared true for both patients using antidepressants

302    (mean $\Delta$HAMD = 18.2±11.1 SD; t(75) = 3.02, $p$ = 0.003) and benzodiazepines (mean

303    $\Delta$HAMD = 19.6±12.5 SD; t(37) =  2.83, $p$ = 0.007), compared to unmedicated patients

304    (mean $\Delta$HAMD = 11.6±10.0 SD). However, excessive inhibition (defined as an inhibitory

305    strength beyond the range of healthy controls) was present in 27.8% of antidepressants

306    users and 60.3% of benzodiazepines users. None of the unmedicated patients showed

307    excessive inhibition. Regarding the change in HAMD-scores, patients that showed

308    excessive inhibition while using antidepressants (mean $\Delta$HAMD = 12.5±9.7 SD; t(22) =

309    0.30, $p$ = 0.765) and/or benzodiazepines (mean $\Delta$HAMD = 12.3±9.6 SD; t(81) = 0.33, $p$ =

310    0.738) did not differ from patients without concomitant medication use. This indicated that

311    the positive effects of concomitant pharmacotherapy on ECT effectiveness were nullified if

312    the degree of self-inhibition was excessive.

313

314    **Influence of ECT on excitation/inhibition balance and associations with clinical**

315    **effectiveness**

316    We used a longitudinal design to examine changes in excitation/inhibition ratios after the

317    ECT-course, compared to baseline, and its associations with antidepressive effectiveness.

318    We used a multisession resting-state spDCM analysis in a subsample of participants for

319    whom longitudinal data were available ($n_{total}$ = 113; $n_{patient}$ = 89; $n_{HC}$ = 24). Patients were

320    scanned before and after the ECT-course. Healthy controls were also scanned twice,

321    approximately four weeks apart. First, a model was run to test for effects of ECT on

322    inhibitory self-connections in patients compared to longitudinal effects in controls. This

323    model also tested for associations between change in inhibitory self-connections and

324    treatment effectiveness (i.e., change in HAM-D; SI Figure 7). The analysis yielded weak

325    evidence suggesting that longitudinal effects of ECT in patients did not differ from

326    longitudinal effects in healthy controls ($p(m_0|Y) = 0.67$). The model did yield conclusive

327    evidence showing that ECT effectiveness was not associated with changes in inhibitory

328    self-connections in any of the networks ($p(m_0|Y) = 0.99$). These results were robust against

329    adjustments for potential confounding variables (i.e., age, sex, baseline depression

330    severity, presence of recurrent depression, presence of psychotic depression, lithium

331    carbonate use, antipsychotic use, number of ECT-sessions, and site variables). Also,

332    analysis on ECT response (i.e., a binary variable indicating >50% reduction of HAM-D

333    scores after ECT) yielded very similar results (SI Figure 8). Because the patient groups

334    (defined by medication use) showed different baseline inhibitory self-connection strengths,

335    we tested whether longitudinal effects differed between groups. The results showed that

336    ECT effectiveness was not associated with changes in inhibitory self-connections in any of

337    the patient groups ($p(m_0|Y) = 1.0$ for each group; SI Figure 9).

338

## Discussion

340    In this study, we investigated the role of the excitation/inhibition imbalance in the treatment

341    of severe depression. The cross-sectional results revealed that unmedicated patients

342    showed reduced inhibition compared to healthy controls. Also, more severe depression

343    was associated with reduced self-inhibition in this group. Patients using antidepressants

344    and/or benzodiazepines showed increased inhibition compared to unmedicated patients.

345    The degree of inhibition in patients using antidepressants seemed comparable to healthy

346    controls, whereas patients using benzodiazepines showed increased inhibition compared

347    to controls. In these medicated patients, the degree of inhibition was not associated with

348    depression severity. In line with previous findings[12–14], these findings suggest that

349    pharmacotherapy reverses the reduced inhibition that is associated with depression.

350 However, patients in this study all suffered from perpetuating severe depressive

351 symptoms, indicating that the reversal of the excitation/inhibition imbalance in patients

352 using pharmacotherapy, in and of itself, did not result in a persistent antidepressive effect.

353 Additionally, the subsequent ECT effectiveness was not associated with changes in the

354 strength of self-inhibition. Finally, we report that many patients using pharmacotherapy

355 showed a degree of inhibition beyond the range of healthy controls. If the degree of

356 inhibition was within the healthy range, concomitant antidepressants and/or

357 benzodiazepines use was associated with increased ECT effectiveness. Remarkably, the

358 presence of pre-treatment excessive neural inhibition nullified these positive effects of

359 concomitant pharmacotherapy on the outcome of ECT.

360

361 An influential hypothesis poses that deficits in neural excitation/inhibition balance

362 represent a fundamental aspect of depression and that reversing these deficits may be an

363 important property to the antidepressive effect of existing treatments[6–9]. In line with this

364 hypothesis, unmedicated severely depressed patients showed reduced inhibitory self-

365 connections compared to healthy controls. But contrary to this hypothesis, our results

366 suggest that reversal of inhibitory deficits by using antidepressants and/or

367 benzodiazepines is, in and of itself, not a sufficient condition for an antidepressive

368 treatment effect in patients with severe depression. This indicates that either the

369 effectiveness of pharmacotherapy in severe depression did not depend on reversing the

370 inhibitory deficits, or that additional conditions were needed for an antidepressant

371 response (i.e., the absence/presence of other biological, psychological, or social

372 phenotypes). One of such additional conditions could be that an increase in inhibition

373 should be accompanied by an increase in excitation for an antidepressant effect[31]. Also,

374 changes in self-inhibition were not associated with the effectiveness of subsequent ECT,

indicating that reversal of the excitation/inhibition balance neither was a necessary condition for the effective treatment of severe depression. Based on these findings, the previously stated hypothesis might be partially altered. Although the excitation/inhibition imbalance still seemed a key aspect of the neural etiology of severe depression, its reversal by treatment might not be a sufficient nor a necessary condition for an antidepressive effect.

Recent studies showed antidepressant effects of a positive allosteric modulator of GABA$_A$ receptors in post-partum depression[32] and major depressive disorder[33]. In contrast to our findings, these findings suggest that reversing the inhibitory deficit would be a sufficient condition for an antidepressant effect in severe depression. However, the results of these trials may reflect antidepressant effects that are not maintained at the longer-term. The antidepressive effectiveness in the trial on major depressive disorder vanished within six weeks after initiating the pharmacotherapy[34], and the trial on post-partum depression used a follow-up period of only four weeks. Such short-term antidepressant effects are also observed in the treatment of depressed patients with benzodiazepines, a well-known GABA$_A$ modulator[15]. In our study, data on the duration of current pharmacotherapy use was not available. However, severely depressed patients indicated for ECT are typically characterized by long-term pharmacotherapy use (i.e., at least multiple months, but most likely years). Therefore, our findings probably reflect the clinical and neural state of patients in a later phase of pharmacotherapy, as opposed to the early phase studied in the mentioned studies[32,33]. We therefore hypothesize that, even though reversing the excitation/inhibition imbalance might induce short-term antidepressant effects, it might not be sufficient for a long-term clinical response.

400  Our findings on the effects of pharmacotherapy on the excitation/inhibition balance might

401  be specific to more severe and treatment-resistant subtypes of depression and its

402  treatments. Preliminary evidence suggests that severe depression seems to differ from

403  milder subtypes of depression with respect to GABA-ergic dynamics[35,36]. For instance, a

404  transcranial magnetic stimulation (TMS) study indicated that $GABA_B$ receptors seem to be

405  affected by all subtypes of depression, whereas changes in $GABA_A$ receptors might be

406  unique to treatment-resistant depression[37]. Therefore, the pharmacotherapy-related effects

407  on the excitation/inhibition imbalance in our study might be established by changes in

408  specific GABA receptor types or subunits that were insufficient for achieving effectiveness

409  in these treatment-resistant patients, while changes in other receptor types or subunits

410  might. However, the lack of an association between ECT effectiveness and changes in

411  self-inhibition suggests that changes in GABA-ergic dynamics are not necessary for an

412  antidepressant response in severe depression. Also, in other subtypes of depression,

413  reversing the excitation/inhibition imbalance may be sufficient for an antidepressive

414  response. Therefore, further research is needed to test whether our findings generalize

415  across other subtypes of depression.

416

417  We also provide evidence suggesting that an optimized excitation/inhibition balance at

418  baseline is associated with increased ECT effectiveness, whereas too little or too much

419  inhibition can result in a lower antidepressive response. One possible explanation of this

420  effect is that excessive inhibition due to pharmacotherapy may reduce the propagation of

421  the induced seizure. For decades, it has been suggested that seizure propagation

422  between distant regions of the brain is pivotal for ECT effectiveness, including cortical–

423  thalamocortical interactions and direct cortical–cortical connections[19]. Consequently,

424  excessive inhibition might obstruct spreading of seizure activity to essential parts of the

425      patients' brain. In support of this, a study using TMS showed that baclofen, a GABA$_B$

426      receptor agonist, reduced interhemispheric signal propagation[38]. Also, previous studies

427      showed that the strength of the post-ictal suppression is associated with ECT

428      effectiveness[21,39]. Post-ictal suppression is an electroencephalographic phenomenon

429      characterized by an ictal slow-wave phase, associated with inhibitory restraints, after

430      which seizure activity is terminated. The interhemispheric coherence in ictal slow-wave

431      activity has been associated with greater ECT effectiveness[21,40], indicating that individual

432      differences in the seizure termination processes may explain the variation in treatment

433      effectiveness between patients. These termination processes may be suboptimal in

434      patients if the degree of self-inhibition is reduced. In light of the above, our findings denote

435      that an optimal excitation/inhibition balance at baseline increases the effectiveness of ECT

436      by allowing for a proper propagation and termination of the induced seizure.

437

438      ECT did not seem to affect the strength of neural inhibition in our patients, when compared

439      to changes in inhibition over time in healthy controls (although evidence was weak). Some

440      magnetic resonance spectroscopy (MRS) studies showed increase of inhibitory GABA

441      concentrations over the ECT-course, but adding healthy controls diminished these

442      significant findings[41]. Additionally, we provide strong evidence that ECT effectiveness did

443      not depend on changes in regional self-inhibition. This is compatible with previous null

444      findings in studies using frequentist statistics[41.] However, the classical statistics used in

445      these studies were not able to provide conclusive evidence on the absence of an

446      association between GABA concentrations and ECT effectiveness. Using a Bayesian

447      approach, we provide very strong evidence indicating that the effectiveness of ECT did not

448      depend on changes in the excitation/inhibition balance[42].

449

450    In our study, use of pharmacotherapy influenced the relation between neural inhibition and

451    ECT effectiveness. In daily practice, combining ECT with antidepressants is regularly

452    attempted to further enhance effectiveness[43]. Also, a recent meta-analysis of randomized

453    controlled trials showed that the use of antidepressants (i.e., as a dichotomous variable)

454    increased ECT effectiveness when considering the entire group of patients[30]. However,

455    based on our findings, this statement may be nuanced by adding that the positive effects

456    of antidepressants can be nullified if the degree of inhibition (associated with

457    pharmacotherapy use) is excessive. The strength of neural inhibition in medicated patients

458    is probably due to a combination of pharmacotherapy dosage effects and variability in drug

459    metabolism between patients[44]. Consequently, the administered dosage of concomitant

460    antidepressants and, in particular, benzodiazepines may affect the outcome of ECT. This

461    may also explain the contradictory results in the current literature, in which studies

462    reported positive, negative and absence of effects of concomitant benzodiazepines use on

463    ECT response[45]. Additionally, higher dosages of concomitant used benzodiazepines (i.e.,

464    lorazepam) were associated with shorter seizure durations during ECT[46]. To mitigate these

465    effects, the competitive benzodiazepine-antagonist flumazenil can be used. Although, it

466    was not known whether or which patients received flumazenil in our study, this could not

467    prevent that patients with excessive inhibition beforehand showed lower ECT

468    effectiveness. Thus, based on our findings we hypothesize that, while the use of

469    concomitant pharmacotherapy (as a dichotomous variable) may enhance ECT

470    effectiveness, higher dosages may result in a poorer clinical response. But, to provide

471    conclusive evidence on this hypothesis, randomized clinical trials have to be conducted, to

472    test dose-dependent (and preferably plasma-level controlled) effects of benzodiazepines

473    and antidepressants on ECT effectiveness. To the best of our knowledge, no such trials

474    have been reported in the literature yet.

475

476    Our findings may guide future research on treatments for severe depression. The findings

477    suggested that reversing the excitation/inhibition imbalance by pharmacotherapy might not

478    be sufficient in severe depression. This may inform the direction of drug development for

479    depression, especially regarding agents targeting GABA. Additionally, we identified

480    excessive inhibition as a potential novel biomarker of ECT ineffectiveness in patients using

481    benzodiazepines or antidepressants. In our study, the effect size of this biomarker was

482    comparable to the presence of psychotic features (data not shown), which is currently one

483    of the most powerful predictors of ECT outcome[47]. Another advantage of this biomarker is

484    that it may be modifiable by adjusting the pharmacotherapy dosages in the individual ECT-

485    patient. Thereby, our findings suggest that pharmacotherapy management informed by the

486    excitation/inhibition ratio may prevent excessive inhibition and potentially enhance the

487    effectiveness of ECT. Alternatively, patients treated with ECT may simply benefit from

488    lower dosages of pharmacotherapy during the ECT-course. This may especially hold for

489    patients using benzodiazepines, which showed higher risks of excessive inhibition in this

490    study. However, replication neuroimaging and clinical studies are needed before

491    implementation of these findings in clinical guidelines.

492

493    Several limitations have to be taken into account when interpreting our findings. First, the

494    effects of pharmacotherapy on the excitation/inhibition balance were studied using cross-

495    sectional data and an observational design. These findings should be replicated in

496    longitudinal prospective studies to rule out potential selection bias and effects of

497    confounding variables not tested in the current study (e.g., comorbid psychiatric disorders,

498    level of anxiety, concomitant use of other medication). Also, data on the timing of

499    pharmacotherapy (dis)continuation was unavailable, which should be taken into account

500 by future studies. Additionally, effects of pretreatment with flumazenil could not be

501 evaluated. Furthermore, our study was observational and included severely depressed

502 patients that failed to respond to pharmacotherapy and were indicated for ECT. This is a

503 highly selected and severely ill population to test the general neural mechanisms of

504 pharmacotherapy. Therefore, our findings may not necessarily generalize to other

505 subtypes of depression. Moreover, we established neural inhibition/excitation on the level

506 of cortical networks, each node containing millions of neurons. This restrained our level of

507 inference and made conclusions on specific cellular processes impossible (e.g.,

508 neurotransmitter concentrations, effects on specific receptor types or subunits). Finally, no

509 (sufficient) data was available to test other effects of interest, such as dose-dependency of

510 pharmacotherapy effects. Future studies have to be conducted to overcome these

511 limitations.

512

513 In conclusion, our study provides novel evidence on the role of excitation/inhibition

514 imbalance in treatments for severe depression. We show that reversing the

515 excitation/inhibition imbalance may not be sufficient nor necessary for treatment

516 effectiveness using pharmacotherapy or ECT in severe depression. Furthermore, the

517 presence of excessive neural inhibition nullified the positive effects of concomitant

518 pharmacotherapy on the antidepressant effectiveness of ECT. These findings should be

519 replicated by future studies to assess generalizability to other samples and milder

520 subtypes of depression. Nonetheless, this work provides a modest but important step for

521 understanding the working mechanisms of treatments for severe depression and

522 enhancing their effectiveness.

523

524

## Methods

**Participants and treatment**

Data were obtained from the Global ECT-MRI Research Collaboration (GEMRIC)[27]. All patients in this study were indicated for ECT and were treated according to internationally accepted guidelines. The data used for this study were acquired at seven sites across Europe and North America. Participants were divided into four groups: healthy controls (n = 59), treatment-resistant patients without current use of antidepressants and benzodiazepines (noAD/BZ; n = 60), treatment-resistant patients using antidepressants (AD; n = 56), and treatment-resistant patients using benzodiazepines (BZ; n = 63). The latter two groups showed considerable overlap (n = 24), which was accounted for in the analyses. All patients had a clinical diagnosis of unipolar major depressive disorder, classified according to the International Statistical Classification of Diseases and Related Health Problems (ICD-10). Patients without current use of benzodiazepines and antidepressants did receive at least one trial of pharmacotherapy before. Descriptive data on demographics (age, sex) and clinical variables (depression severity, presence of psychotic features, presence of recurrence) were available. Depression severity was assessed using the Hamilton Depression Rating Scale (HAM-D) or Montgomery-Åsberg Depression Rating Scale (MADRS), depending on the sites' preferences. Ratings of the MADRS were converted to HAM-D scores, using the following formula[48]:

$HAMD = -1.58 + 0.86 * MADRS.$

Patients were excluded if (1) bipolar disorder was present[49], (2) BZ or AD use data were incomplete, and (3) if neuroimaging data were unavailable or of insufficient quality.

The ADs that were examined in this study consisted of selective-serotonin reuptake inhibitors (SSRIs), selective-noradrenaline reuptake inhibitors (SNRIs), tricyclic

550 antidepressants (TCAs) and monoamine oxidase inhibitors (MAOIs). In the group of

551 patients currently using AD, 34% used SSRIs, 36% used SNRIs, 27% used TCAs and 3%

552 used MAOIs. The specific type of BZ, and dosages for AD and BZ, were not available in

553 the database. Regarding used electrode placements, 126 patients received ECT with right

554 unilateral (RUL) and 62 patients with a bilateral (BL) electrode placement. The average

555 number of ECT-sessions per course was 12.4 ± 5.5 SD.

556

557 **fMRI acquisition, preprocessing, and quality control**

558 Structural and functional resting-state magnetic resonance imaging (MRI) data were

559 acquired at seven sites (for scanning parameters per site, see SI Table 4). The fMRI

560 preprocessing steps were identical for each participant. First, structural and

561 functional images were reoriented and brain extraction was performed using

562 Advanced Normalization Tools (ANTs v2.2.0). Images were coregistered using ANTs

563 and FMRI Software Library (FSL v5.0.10) using boundary-based registration. After

564 discarding the first two volumes for each fMRI series, FSL's MCFLIRT was used to

565 apply head movement correction by realigning the volumes to the middle volume

566 using six parameters for rigid body transformations. Functional images were spatially

567 smoothed using a Gaussian kernel with 5 mm full-width at half-maximum. For motion

568 correction, an ICA-based strategy for Automatic Removal of Motion Artifacts (ICA-

569 AROMA) was used. The noise components estimated by ICA-AROMA were used to

570 compute denoised cosines for high pass filtering (f=0.009). Furthermore, mean white

571 matter (WM) and cerebrospinal fluid (CSF) timeseries were computed as additional

572 nuisance variables. Temporal high-pass filtering was used to remove low-frequency

573 drifts (< 0.01 Hz) and images were registered to 4 mm isotropic voxel size. Both the

574 denoised cosines and WM/CSF nuisance variables were used to denoise the fMRI

575    data and perform high pass filtering in one single step[50] ANTs were used for

576    normalization of transformation matrices to MNI space using 2 mm standard

577    templates.

578    After the preprocessing, a quality control procedure was performed. First, participants that

579    showed a high degree of motion were identified using three criteria: (1) any

580    rotation/translation exceeding 4 mm/degrees, (2) the average framewise displacement

581    (FD) exceeding 0.3 mm, and (3) less than 4 minutes of RS-fMRI data unaffected by motion

582    (FD<0.25mm). If any of the criteria applied to a participant's RS-fMRI data, the participant

583    was excluded from the study. Then, the data were visually inspected to confirm high-

584    quality coregistration and normalization of the T1w and RS-fMRI scans. Finally, the EPI

585    signal-to-noise and field of view were assessed to identify potential signal dropout or other

586    artifacts. Again, participants with low quality preprocessed data were excluded from the

587    analysis. Finally, RS-fMRI of sufficient quality was available for a total of 214 participants

588    with respect to cross-sectional baseline data. Regarding longitudinal data, RS-fMRI data of

589    sufficient quality was available for 113 participants.

590

591    **First-level spDCM analysis**

592    Spectral dynamic causal modeling (spDCM) for RS-fMRI was used to estimate effective

593    connectivity parameters[22]. Analyses were conducted in Statistical Parametric Mapping

594    (SPM12, revision 7771) with DCM12.5 (revision 7497). spDCM used a generative forward

595    model with two components. The first one described how neuronal populations causally

596    interact. The second component mapped neuronal activity from the first model to observed

597    (crossspectra of) hemodynamic responses. Thereby, spDCM was able to estimate

598    directed connectivity between brain regions as well as regional self-connections. The

599 construct validity of spDCM has been shown by its ability to accurately recover effective

600 connectivity in simulated data[22,23], and specific effective connections corresponding to well-

601 documented neural pathways in rodents[24]. The self-connections were assumed to be

602 inhibitory, which was in line with in vivo animal studies[25]. By regulating the strength of

603 inhibition, the self-connections allowed for sustained activity in the cortex while maintaining

604 a balance between excitation and inhibition throughout the network. A recent

605 magnetoencephalography study using DCM showed that a $GABA_A$ reuptake inhibitor

606 altered the strength of inhibitory (self-)connections in humans[26], indicating that the DCM

607 framework could be used to study the effect of therapeutic interventions on (self-)inhibition.

608

609 We used the following procedure. First, the RS-fMRI timeseries were extracted from

610 fourteen regions of interest (ROIs). These regions formed the core of three networks (i.e.,

611 default mode network (DMN), salience network (SAL) and central executive network

612 (CEN)) that are of central importance to depression[28] and psychopathology in general[51].

613 Time series were extracted from a sphere with 5 mm radius centered around MNI

614 coordinates extracted from previous research (see SI Figure 1)[52]single representative time

615 series was established by retaining the first component of a principal components analysis

616 across voxels in each sphere. For baseline analysis, a full spDCM was estimated for each

617 participant individually (i.e., all 196 connections in matrix A of the neural model, henceforth

618 the 'A-matrix', were turned on). Default settings were used for the individual-level spDCM

619 analyses and no experimental inputs were specified (i.e., all 196 connections in the 'B-

620 matrix' turned off). The expected values and covariances of the estimated parameters of

621 the 'A-matrix' were then taken to the group-level analysis.

622

623 For the longitudinal analysis, a multisession resting-state spDCM approach was used.

624 Only the longitudinal effect of ECT on inhibitory self-connections were estimated.

625 Therefore, a full 'A-matrix' was specified and an identity matrix was used as the 'B-matrix'.

626 In this case, the 'A-matrix' provided the averaged connectivity over both sessions and the

627 'B-matrix' modelled the change between sessions. Because we were interested in

628 changes associated with ECT, the estimated parameters of the 'B-matrix' were taken to

629 the group-level analysis.

630

631 **Parametric Empirical Bayes**

632 Group-level effects could be estimated using the Parametric Empirical Bayesian (PEB)

633 framework[53]. A PEB general linear model (GLM) was able to infer effects at multiple levels

634 (e.g., group effects or between-session effects) by combining the expected values and

635 covariances from all individual-level spDCM parameters. Additionally, the PEB model was

636 able to finesse spDCM parameters at the individual level by using group-level posteriors

637 as empirical priors[54]. PEB models were used for studying associations between baseline

638 self-inhibition and ECT effectiveness (see Analysis on baseline associations with ECT

639 effectiveness, below).

640

641 For each PEB model (see below for PEB model designs), effects on inhibitory self-

642 connections in individual nodes and effects on combinations of self-connections of multiple

643 nodes were examined. For effects on individual self-connections, Bayesian Model

644 Reduction (BMR) was used to automatically search over reduced models with certain

645 connectivity parameters 'turned off' (i.e., fixed at their prior mean of zero). In this method,

646 all possible reduced models were compared, and models with connectivity parameters that

647 did not contribute to model evidence were iteratively discarded. This process was

648 continued until removing parameters reduced the model evidence. Subsequently, a

649 Bayesian model average was calculated over the 256 models with the largest model

650 evidence in the final iteration (or over the remaining models, if less than 256 models

651 survived). Using this 'pruning' and averaging approach allowed inference on the posterior

652 probability of specific connectivity parameters. A threshold for posterior probabilities

653 indicating strong evidence was used for each second-level parameter (p(m|Y) > 0.95).

654

655 For the analysis on specific groups of connections, BMR was used to estimate the model

656 evidence for pre-specified reduced connectivity models. Subsequently, the model

657 evidence between the reduced models was compared to that of an empty model (i.e., all

658 parameters 'turned off') to determine which reduced model best fitted the observed data[54].

659 Five pre-specified reduced models were used, in which models included self-connections

660 in none of the networks, one of the networks (DMN, CEN or SAL), or all of the networks

661 (DMN, CEN and SAL; see SI Figure 3). Thereby, this procedure tested whether

662 associations between variables of interest (e.g., depression severity) and self-connections

663 parameters were absent, network-specific or widespread throughout brain networks.

664 Again, a threshold of a posterior probability indicating (at least) strong evidence in favor of

665 a model was used ($p$(m|Y) > 0.95).

666

667 **PEB model designs**

668 Multiple PEB GLMs were run to test for group differences in inhibitory self-connections at

669 baseline, associations with depression severity, and changes in inhibitory self-connections

670 before and after ECT.

671

First, an analysis on group differences between healthy controls and patient groups was conducted. Group differences could be modeled using multiple GLM designs. We used a model selection procedure by comparing the two most convenient designs with respect to model evidence. A GLM in which all patient groups (unmedicated, AD, BZ) were modeled with a separate binary regressor (and healthy controls as the intercept, i.e., reference group) showed greatest model evidence, and was selected for the analysis on group differences (see SI Figure 2 for the models and difference in model evidence). Additionally, an identical model with extra regressors was run to test if the results of the initial model were robust against effects of potential confounding variables. The added regressors consisted of age, sex, baseline HAM-D score, presence of recurrent depression, presence of psychotic depression, lithium carbonate use, antipsychotic use and site variables. The age and sex variables were mean centered, such that the reference comparison group (i.e., the intercept) represented healthy control subjects of a group average age, averaged over males and females. Due to the collinearity constraint in a GLM design, not all sites variables could be included as covariates (i.e., the sum of all site variables equals the intercept). Therefore, site variables that added substantially to the model evidence were selected as potential confounders. These variables were selected by comparing the model evidence between the model with demographic and clinical covariates with an identical model plus one of the site variables. If a site variable explained considerable variation in the data, the model fit increased. However, adding a covariate also added extra complexity to the model, which reduced model evidence. This procedure tested whether the increase in the goodness-of-fit by adding the site variable outweighed the increase in complexity. Site variables that increased model evidence were included as a potential confounder in the GLM (SI Figure 3).

Because the model above did not allow direct comparisons between patient groups, post-hoc pairwise comparisons between the AD, BZ and unmedicated patient groups were conducted. For each comparison, a GLM was created that included patients of the specified patient groups (see SI Figure 5 for model design). The same potential confounding variables as in the model above were also included in the model. Additionally, AD use was included in the comparison between patients using BZ and unmedicated patients, and BZ use was included in the comparison between patients using AD and unmedicated patients. The latter was done to account for the overlap between the AD and BZ groups. For post-hoc analyses, a threshold that indicated strong evidence according to Bayesian statistics was used, adjusted for multiple comparisons using Bonferroni ($p(m|Y) > 0.983$).

Associations with depression severity were analyzed using a GLM that included patients only (see SI Figure 6 for the model design). This model tested if there were effects of depression severity on inhibitory self-connections in the complete group (i.e., main effect of baseline HAM-D) or specific patient groups (i.e., patient group by baseline HAM-D interactions). Aside from the unadjusted model, a model was run to test if the results were robust against potential confounders. The same demographic and clinical covariates were included as in the adjusted model for group comparison. Also, an identical procedure was used to select relevant site variables (i.e., selecting site variables that increased model evidence).

Finally, PEB models were used to test for longitudinal effects of ECT. Parameters of the first-level longitudinal analyses were entered into group-level PEB models. The first GLM tested for longitudinal effects in patients compared to healthy controls, and for associations

722 with ECT effectiveness (SI Figure 7). Again, we subsequently tested if the results were

723 robust against potential confounding variables. Based on previous studies[47,55] and clinical

724 practice, the following demographic and clinical variables were included in the GLM as

725 potential confounders: age, sex, baseline HAM-D score, presence of recurrent depression,

726 presence of psychotic depression, lithium carbonate use, antipsychotic use, number of

727 ECT-sessions, and site variables. Site variables were selected using an identical

728 procedure as described in the analysis on group differences. Furthermore, an additional

729 model was run for effects of treatment response (i.e., a binary regressor indicating HAM-D

730 reduction greater than 50% compared to baseline), instead of the continuous measure for

731 ECT effectiveness used in the model above. Finally, because patient groups showed

732 different baseline strengths in inhibitory self-connections, we also tested for longitudinal

733 effects of ECT in the distinct patient groups (SI Figure 9 for the model design). The latter

734 model also tested if ECT effectiveness in distinct patient groups was associated with

735 changes in inhibitory self-connections.

736

737 **Analysis on baseline associations with ECT effectiveness**

738 For the clinical analysis, a multiple regression model was run to test the effect of clinical

739 variables on ECT effectiveness, which was defined as the change in HAM-D scores before

740 and after the ECT-course (HAM-D$_{PRE}$ – HAM-D$_{POST}$; i.e., the dependent variable). The

741 included clinical variables were the use of BZ, the use of AD, baseline HAM-D scores,

742 presence of psychotic features, age, number of ECT-sessions with BL electrode

743 placement, and seven binary variables indicating site membership. These variables were

744 selected because those were identified as predictors of ECT effectiveness in previous

745 research[47] and in clinical practice.

746

747   In order to relate the excitation/inhibition ratio at baseline to ECT effectiveness, multiple

748   regression models were used in R[56]. A different approach was used for this analysis

749   compared to the analyses described above (see PEB design models), because we aimed

750   to provide effect size measures that were comparable to the existing literature on ECT

751   outcome predictors (i.e., provided by classic frequentist statistics, rather than Bayesian

752   statistics). First, a PEB GLM (with only the intercept) was used to establish the posterior

753   probability distribution of each of the 14 self-connection parameters for each individual.

754   The mean of the probability distribution for each self-connection was extracted for each

755   participant, and the first principal component (as obtained by principal component analysis

756   using the *prcomp* package) was used as summary measure of the excitation/inhibition

757   balance. Subsequently, a classical (non-Bayesian) multiple regression model was used to

758   examine the relation between excitation/inhibition balance (i.e., the independent variable)

759   and ECT effectiveness. Both linear and quadratic effects of the excitation/inhibition

760   balance on ECT effectiveness were tested. All variables that showed a significant effect in

761   the above-described clinical analysis were included as a covariate in the model (i.e.,

762   baseline HAM-D score, age, presence of psychotic features, number of ECT-sessions with

763   BL electrode placement, and two binary variables indicating site membership). For post-

764   hoc analyses, identical models were built that only included patients of the separate

765   groups (i.e., noAD/BZ, AD and BZ). The same covariates were entered into the model.

766   Because the AD and BZ groups showed considerable overlap, BZ use was added as a

767   covariate in model of the AD group, and AD use was added in the model of the BZ group.

768   This was done to ensure that the effects in de AD and BZ group were independent from

769   BZ and AD use, respectively. Cohen's $f^2$ was used as a measure of effect size, in

770   which $0.15 \leq f^2 \geq 0.35$ indicates a medium effect size[57].

771

# Acknowledgements

The authors are very grateful to all patients who participated in this study. Additionally, we would like to thank all professionals that contributed to treatment delivery and acquisition of the data.

# Author contributions

F.t.D., W.B., C.A., M.A., A.D., L.E., P.F.P.v.E., E.v.E., P.C.R.M., K.N., I.T., D.R., P.S., M.V., J.V., M.v.V., H.B., L.O., J.A.v.W., & G.A.v.W contributed to data acquisition.  F.t.D. primarily performed data analysis. W.B., P.Z., J.A.v.W., & G.A.v.W also made substantial contributions to data analysis. F.t.D., J.A.v.W. & G.A.v.W drafted the manuscript. W.B., P.Z, C.A., M.A., A.D., L.E., P.F.P.v.E., E.v.E., P.C.R.M., K.N., I.T., D.R., P.S., M.V., J.V., M.v.V., H.B., & L.O. substantially revised the manuscript.

# Competing interests

No potential competing interest was reported by the authors

# References

1.  James, S. L. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1789–1858 (2018).

2.  Rush, A. J. *et al.* Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR* D report. *Am. J. Psychiatry* **163**, 1905–1917 (2006).

3.   Eaton, W. W. *et al.* Population-based study of first onset and chronicity in major depressive disorder. *Arch. Gen. Psychiatry* **65**, 513–520 (2008).

4.   Heijnen, W. T., Birkenhäger, T. K., Wierdsma, A. I. & van den Broek, W. W. Antidepressant pharmacotherapy failure and response to subsequent electroconvulsive therapy: a meta-analysis. *J. Clin. Psychopharmacol.* **30**, 616–619 (2010).

5.   Turrigiano, G. Too many cooks? Intrinsic and synaptic homeostatic mechanisms in cortical circuit refinement. *Annu. Rev. Neurosci.* **34**, 89–103 (2011).

6.   Krystal, J. H. *et al.* Glutamate and GABA systems as targets for novel antidepressant and mood-stabilizing treatments. *Mol. Psychiatry* **7**, S71–S80 (2002).

7.   Luscher, B., Shen, Q. & Sahir, N. The GABAergic deficit hypothesis of major depressive disorder. *Mol. Psychiatry* **16**, 383–406 (2011).

8.   Fee, C., Banasr, M. & Sibille, E. Somatostatin-positive gamma-aminobutyric acid interneuron deficits in depression: cortical microcircuit and therapeutic perspectives. *Biol. Psychiatry* **82**, 549–559 (2017).

9.   Duman, R. S., Sanacora, G. & Krystal, J. H. Altered connectivity in depression: GABA and glutamate neurotransmitter deficits and reversal by novel treatments. *Neuron* **102**, 75–90 (2019).

10.  Luscher, B. & Fuchs, T. GABAergic control of depression-related brain states. in *Advances in pharmacology* vol. 73 97–144 (Elsevier, 2015).

11.  Möhler, H. The GABA system in anxiety and depression and its therapeutic potential. *Neuropharmacology* **62**, 42–53 (2012).

12.  Godfrey, K. E. M., Gardner, A. C., Kwon, S., Chea, W. & Muthukumaraswamy, S. D. Differences in excitatory and inhibitory neurotransmitter levels between depressed patients and healthy controls: A systematic review and meta-analysis. *J. Psychiatr. Res.* **105**, 33–44 (2018).

13.  Sanacora, G., Mason, G. F., Rothman, D. L. & Krystal, J. H. Increased occipital cortex GABA concentrations in depressed patients after therapy with selective serotonin reuptake inhibitors. *Am. J. Psychiatry* **159**, 663–665 (2002).

14.  Bhagwagar, Z. *et al.* Increased brain GABA concentrations following acute administration of a selective serotonin reuptake inhibitor. *Am. J. Psychiatry* **161**, 368–370 (2004).

15.  Ogawa, Y. *et al.* Antidepressants plus benzodiazepines for adults with major depression. *Cochrane Database Syst. Rev.* (2019).

16. Voineskos, D. *et al.* The relationship between cortical inhibition and electroconvulsive therapy in the treatment of major depressive disorder. *Sci. Rep.* **6**, 1–8 (2016).

17. Enev, M. *et al.* Imaging onset and propagation of ECT-induced seizures. *Epilepsia* **48**, 238–244 (2007).

18. Sackeim, H. A. *et al.* Effects of stimulus intensity and electrode placement on the efficacy and cognitive effects of electroconvulsive therapy. *N. Engl. J. Med.* **328**, 839–846 (1993).

19. McNally, K. A. & Blumenfeld, H. Focal network involvement in generalized seizures: new insights from electroconvulsive therapy. *Epilepsy Behav.* **5**, 3–12 (2004).

20. Schevon, C. A. *et al.* Evidence of an inhibitory restraint of seizure activity in humans. *Nat. Commun.* **3**, 1–11 (2012).

21. Perera, T. D. *et al.* Seizure expression during electroconvulsive therapy: relationships with clinical outcome and cognitive side effects. *Neuropsychopharmacology* **29**, 813–825 (2004).

22. Friston, K. J., Kahan, J., Biswal, B. & Razi, A. A DCM for resting state fMRI. *Neuroimage* **94**, 396–407 (2014).

23. Razi, A., Kahan, J., Rees, G. & Friston, K. J. Construct validation of a DCM for resting state fMRI. *Neuroimage* **106**, 1–14 (2015).

24. Bernal-Casas, D., Lee, H. J., Weitz, A. J. & Lee, J. H. Studying brain circuit function with dynamic causal modeling for optogenetic fMRI. *Neuron* **93**, 522–532 (2017).

25. Bastos, A. M. *et al.* Canonical microcircuits for predictive coding. *Neuron* **76**, 695–711 (2012).

26. Adams, N. E. *et al.* GABA-ergic dynamics in human frontotemporal networks confirmed by pharmaco-magnetoencephalography. *J. Neurosci.* **40**, 1640–1649 (2020).

27. Oltedal, L. *et al.* The Global ECT-MRI Research Collaboration (GEMRIC): Establishing a multi-site investigation of the neural mechanisms underlying response to electroconvulsive therapy. *NeuroImage Clin.* **14**, 422–432 (2017).

28. Mulders, P. C., van Eijndhoven, P. F., Schene, A. H., Beckmann, C. F. & Tendolkar, I. Resting-state functional connectivity in major depressive disorder: a review. *Neurosci. Biobehav. Rev.* **56**, 330–344 (2015).

29. Hamilton, M. A. X. Development of a rating scale for primary depressive illness. *Br. J. Soc. Clin. Psychol.* **6**, 278–296 (1967).

30. Pluijms, E. M., Kamperman, A. M., Hoogendijk, W. J. G., Birkenhäger, T. K. & van den Broek, W. W. Influence of an adjuvant antidepressant on the efficacy of

electroconvulsive therapy: A systematic review and meta-analysis. *Aust. New Zeal. J. Psychiatry* 0004867420952543 (2020).

31. Fagerholm, E. D. *et al.* Fine-tuning neural excitation/inhibition for tailored ketamine use in treatment-resistant depression. *Transl. Psychiatry* **11**, 1–10 (2021).

32. Meltzer-Brody, S. *et al.* Brexanolone injection in post-partum depression: two multicentre, double-blind, randomised, placebo-controlled, phase 3 trials. *Lancet* **392**, 1058–1070 (2018).

33. Gunduz-Bruce, H. *et al.* Trial of SAGE-217 in patients with major depressive disorder. *N. Engl. J. Med.* **381**, 903–911 (2019).

34. ten Doesschate, F., van Waarde, J. A. & van Wingen, G. A. Non-superiority of zuranolone (SAGE-217) at the longer-term. *J. Affect. Disord.* (2021).

35. Price, R. B. *et al.* Amino acid neurotransmitters assessed by proton magnetic resonance spectroscopy: relationship to treatment resistance in major depressive disorder. *Biol. Psychiatry* **65**, 792–800 (2009).

36. Sanacora, G. *et al.* Subtype-specific alterations of γ-aminobutyric acid and glutamatein patients with major depression. *Arch. Gen. Psychiatry* **61**, 705–713 (2004).

37. Levinson, A. J. *et al.* Evidence of cortical inhibitory deficits in major depressive disorder. *Biol. Psychiatry* **67**, 458–464 (2010).

38. Hui, J. *et al.* Pharmacological mechanisms of interhemispheric signal propagation: a TMS-EEG study. *Neuropsychopharmacology* **45**, 932–939 (2020).

39. Suppes, T. *et al.* Is postictal electrical silence a predictor of response to electroconvulsive therapy? *J. Affect. Disord.* **41**, 55–58 (1996).

40. Ten Doesschate, F., van Wingen, G. A., de Pont, B. J. H. B., Arns, M. & van Waarde, J. A. The longitudinal effects of electroconvulsive therapy on ictal interhemispheric coherence and its associations with treatment outcome: a naturalistic cohort study. *Clin. EEG Neurosci.* **50**, 44–50 (2019).

41. Erchinger, V. J., Ersland, L., Aukland, S. M., Abbott, C. C. & Oltedal, L. Magnetic Resonance Spectroscopy in Depressed Subjects Treated With Electroconvulsive Therapy—A Systematic Review of Literature. *Front. Psychiatry* **12**, 326 (2021).

42. Keysers, C., Gazzola, V. & Wagenmakers, E.-J. Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nat. Neurosci.* **23**, 788–799 (2020).

43. Sackeim, H. A. *et al.* Effect of concomitant pharmacotherapy on electroconvulsive therapy outcomes: short-term efficacy and adverse effects. *Arch. Gen. Psychiatry* **66**, 729–737 (2009).

44.  Oz, M. D., Ozdemir, F. & Suzen, H. S. The Influence of Genetic Variations and Drug Interactions Based on Metabolism of Antidepressants and Anticonvulsants. *Curr. Drug Metab.* (2020).

45.  Janjua, A. U., Dhingra, A. L., Greenberg, R. & McDonald, W. M. The efficacy and safety of concomitant psychotropic medication and electroconvulsive therapy (ECT). *CNS Drugs* **34**, 509–520 (2020).

46.  Boylan, L. S. *et al.* Determinants of seizure threshold in ECT: benzodiazepine use, anesthetic dosage, and other factors. *J. ECT* **16**, 3–18 (2000).

47.  Van Diermen, L. *et al.* Prediction of electroconvulsive therapy response and remission in major depression: meta-analysis. *Br. J. psychiatry* **212**, 71–80 (2018).

48.  Heo, M., Murphy, C. F. & Meyers, B. S. Relationship between the Hamilton depression rating scale and the montgomery-Åsberg depression rating scale in depressed elderly: a meta-analysis. *Am. J. Geriatr. Psychiatry* **15**, 899–905 (2007).

49.  Mason, G. F. *et al.* 304. Cortical GABA reduced in unipolar but not bipolar depression. *Biol. Psychiatry* **47**, S92 (2000).

50.  Pruim, R. H. R. *et al.* ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage* **112**, 267–277 (2015).

51.  Menon, V. Large-scale brain networks and psychopathology: a unifying triple network model. *Trends Cogn. Sci.* **15**, 483–506 (2011).

52.  Raichle, M. E. The restless brain. *Brain Connect.* **1**, 3–12 (2011).

53.  Friston, K. J. *et al.* Bayesian model reduction and empirical Bayes for group (DCM) studies. *Neuroimage* **128**, 413–431 (2016).

54.  Zeidman, P. *et al.* A guide to group effective connectivity analysis, part 2: Second level analysis with PEB. *Neuroimage* **200**, 12–25 (2019).

55.  Oltedal, L. *et al.* Volume of the human hippocampus and clinical response following electroconvulsive therapy. *Biol. Psychiatry* **84**, 574–581 (2018).

56.  Team, R. C. R: A language and environment for statistical computing. (2013).

57.  Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D. & Mermelstein, R. J. A practical guide to calculating Cohen's f2, a measure of local effect size, from PROC MIXED. *Front. Psychol.* **3**, 111 (2012).