



Predicting Antidepressant Citalopram Treatment Response via Changes in Brain Functional Connectivity After Acute Intravenous Challenge

Manfred Klöbl, Gregor Gryglewski, Lucas Rischka, Godber Mathis Godbersen, Jakob Unterholzner, Murray Bruce Reed, Paul Michenthaler, Thomas Vanicek, Edda Winkler-Pjrek, Andreas Hahn, Siegfried Kasper and Rupert Lanzenberger*

Department of Psychiatry and Psychotherapy, Medical University of Vienna, Vienna, Austria

OPEN ACCESS

Edited by:

Yu-Guo Yu,
Fudan University, China

Reviewed by:

Fei He,
Coventry University, United Kingdom
Peter Herman,
Yale University, United States

*Correspondence:

Rupert Lanzenberger
rupert.lanzenberger@meduniwien.ac.at

Received: 21 April 2020

Accepted: 31 August 2020

Published: 06 October 2020

Citation:

Klöbl M, Gryglewski G, Rischka L, Godbersen GM, Unterholzner J, Reed MB, Michenthaler P, Vanicek T, Winkler-Pjrek E, Hahn A, Kasper S and Lanzenberger R (2020) Predicting Antidepressant Citalopram Treatment Response via Changes in Brain Functional Connectivity After Acute Intravenous Challenge. *Front. Comput. Neurosci.* 14:554186. doi: 10.3389/fncom.2020.554186

Introduction: The early and therapy-specific prediction of treatment success in major depressive disorder is of paramount importance due to high lifetime prevalence, and heterogeneity of response to standard medication and symptom expression. Hence, this study assessed the predictability of long-term antidepressant effects of escitalopram based on the short-term influence of citalopram on functional connectivity.

Methods: Twenty nine subjects suffering from major depression were scanned twice with resting-state functional magnetic resonance imaging under the influence of intravenous citalopram and placebo in a randomized, double-blinded cross-over fashion. Symptom factors were identified for the Hamilton depression rating scale (HAM-D) and Beck's depression inventory (BDI) taken before and after a median of seven weeks of escitalopram therapy. Predictors were calculated from whole-brain functional connectivity, fed into robust regression models, and cross-validated.

Results: Significant predictive power could be demonstrated for one HAM-D factor describing insomnia and the total score ($r = 0.45-0.55$). Remission and response could furthermore be predicted with an area under the receiver operating characteristic curve of 0.73 and 0.68, respectively. Functional regions with high influence on the predictor were located especially in the ventral attention, fronto-parietal, and default mode networks.

Conclusion: It was shown that medication-specific antidepressant symptom improvements can be predicted using functional connectivity measured during acute pharmacological challenge as an easily assessable imaging marker. The regions with high influence have previously been related to major depression as well as the response to selective serotonin reuptake inhibitors, corroborating the advantages of the current approach of focusing on treatment-specific symptom improvements.

Keywords: treatment response prediction, major depressive disorder, selective serotonin reuptake inhibitors, functional connectivity, resting-state, functional magnetic resonance imaging

INTRODUCTION

Major depressive disorder (MDD) constitutes to one of the most prevalent psychiatric diseases and is seen as the second leading cause of disability worldwide (Spijker et al., 2004; Ferrari et al., 2013). Lifetime prevalence in industrialized countries are reaching almost 20% (Kessler and Bromet, 2013) and are even increasing (Weinberger et al., 2017). Moreover, ~30% of patients do not respond to standard medication and hence, suffer from depression longer than necessary before other therapeutic approaches are attempted (Souery et al., 1999). High prevalence and non-responder rates make MDD an obvious target for models of treatment success.

The ultimate goal is to find general predictive biomarkers that can be used to help make therapeutic decisions before determining the first intervention (Dunlop and Mayberg, 2014). During the last years, neuroimaging has steadily gained importance in this quest, with a distinct focus on changes in brain networks (Dichter et al., 2015) and processing of emotional stimuli caused by MDD (Langenecker et al., 2018a). Further promising results were also reported using, e.g., reward (Nguyen et al., 2019), emotional conflict (Fonzo et al., 2019) and response inhibition tasks (Tozzi et al., 2020). However, identification of neuroimaging markers or other treatment predictors is hampered by several factors: MDD is a highly heterogeneous disease, which is not only mirrored in the variety of symptom manifestations/trajectories (Drysdale et al., 2017; Hartmann et al., 2018) and neuroscientific findings but also the tools for diagnosis and assessment of severity vary considerably. Standard clinical questionnaires like the Hamilton depression rating scale [HAM-D; Hamilton (1960)] or Beck's depression inventory [BDI; Beck et al. (1961)] cover a broad range of symptoms but the overall depression severity is mostly seen as a mere sum of these.

Moreover, studies trying to identify treatment predictors often have to deal with small sample sizes that hardly allow for checking the generalizability on independent datasets. Hence, potential markers are frequently identified via mere correlation against quantified treatment responses. Due to considerable variability in neuroimaging data processing, correlations found can easily be specific for a given methodological setting [e.g., Borchardt et al. (2016)]. Moreover, findings also display specificity for the respective treatment under investigation (Brakowski et al., 2017). A further obstacle arises from the necessity that a predictive marker should be reliably measurable in the majority of patients. Within the fields of neuroimaging, resting-state (RS) functional magnetic resonance imaging (fMRI) constitutes one of the least demanding examinations, as it only requires subjects to lie still for a few minutes. Compared to task fMRI, it does not depend on performance, requires much less attention from the patients and is easier to standardize.

These characteristics have made RS fMRI a popular modality for investigating the effects of MDD on the brain's function and networks, and especially looking for markers aimed at a certain treatment response. Several studies claimed to have found predictors for various therapeutic approaches such as electroconvulsive therapy (Argyelan et al., 2016), transcranial magnetic stimulation (Avissar et al., 2017; Du et al., 2018;

Weigand et al., 2018), psychotherapy (Crowther et al., 2015), mixed (He et al., 2016; Gong et al., 2018; Hou et al., 2018; Zhu et al., 2018), or single-product drug therapy (Alexopoulos et al., 2012; Fu et al., 2015; Cheng et al., 2017). Even though significant correlations with different RS functional connectivity (FC) metrics were found, out-of-sample validations or assessments of generalizability within the same sample [e.g., cross-validation (CV) or bootstrapping] were not reported in the publications above.

In order to address these challenges, the current work follows a multistage approach: Depressive symptoms within the sample were summarized according to latent sources using factor analysis (FA), which allows for investigation of effects of overall and symptom-specific severity. Using network based statistics [NBS; (Zalesky et al., 2010)], potentially predictable factors were identified by searching for related functionally connected brain networks. Finally, true predictive power was estimated using two validation schemes based on pseudo-independent data sampled from the study population and compared to analytically circular models built upon the previously identified networks, as they are sometimes encountered in the literature.

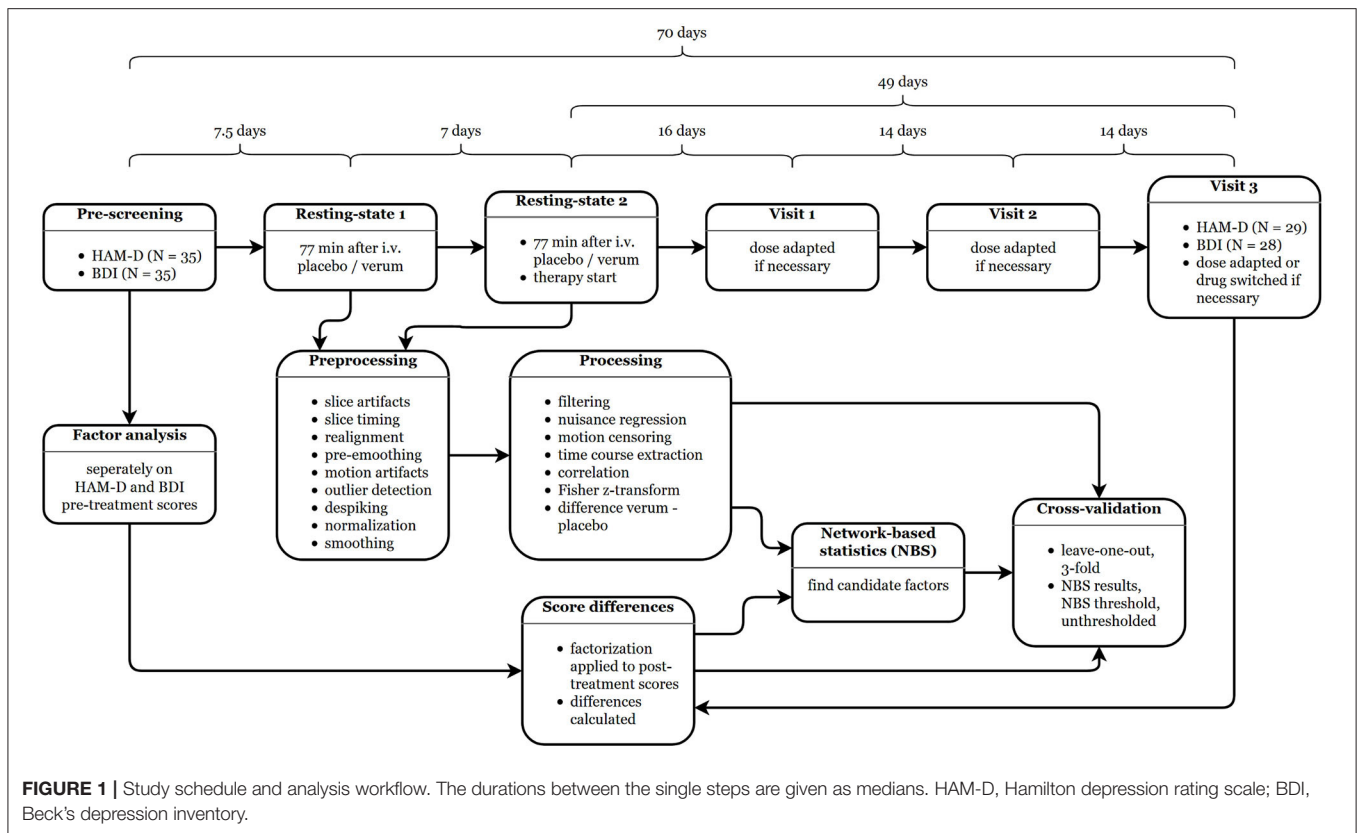
MATERIALS AND METHODS

This study was conducted in accordance with the Declaration of Helsinki, approved by the ethics committee and following the Good Scientific Practice guidelines of the Medical University of Vienna as part of a larger project registered at ClinicalTrials.gov (NCT02711215). Due to highly skew distributions, descriptive statistics are reported as “minimum/median/maximum” throughout this work.

Subjects and Study Design

Thirty five patients suffering from MDD (19.65/26.96/55.28 years old, 16 females) were recruited in the outpatient clinic as well as at the hospital ward of the Department of Psychiatry and Psychotherapy of the Medical University of Vienna. Written informed consent was obtained from all participants. Two MRI examinations were conducted, 3/7/56 days apart, depending on the patients' availability. Inclusion criteria comprised of unipolar/first-episode MDD (according to DSM-IV, HAM-D ≥ 18) but otherwise general health (assessed via physical examination and the Structured Clinical Interview for DSM-IV) and no psychopharmacological treatment within the last three months, excluding the occasional intake of antihistamines and benzodiazepines. Subjects were excluded in case of former or current substance abuse, pregnancy and any MRI contraindications.

At each date, participants received either 8 mg of citalopram (a commonly used selective serotonin reuptake inhibitor—SSRI) diluted in 8 ml saline or a comparable amount of saline as placebo, injected intravenously (Kasper and Muller-Spahn, 2002) over 8 min in a randomized, double-blind, cross-over fashion (i.e., subjects receiving citalopram at the first scan received pure saline in the second one and vice versa). Successful drug application was ascertained by plasma levels drawn from arterial cannula placed in the left or right radial artery while the



subjects were in the scanner (Alshelh et al., 2018). Starting the day after the second MRI examinations, patients received pharmacological treatment for MDD with an initial dose of 5 mg escitalopram (Cipralext—Lundbeck, Copenhagen, Denmark). Racemic citalopram was used for intravenous application since no syringeable escitalopram solution is currently available. In case of missing response (HAM-D decrease <50%), the dose was increased to an adequate level following therapeutic drug monitoring. This was done during visits carried out with ~2-week intervals, again depending on the patients' availability. All subjects received only escitalopram until the third visit and were switched to alternative medication (Duloxetine, Venlafaxin, or Mirtazapin) thereafter if necessary. For a detailed overview of the study schedule and analyses conducted see **Figure 1**.

Psychometric Scoring

At each visit, depression severity was rated using HAM-D and BDI questionnaires, providing expert- and self-ratings due to known differences in the perception of symptom improvements (Lambert et al., 1986). Both questionnaires were previously reported to have high internal consistency and reliability (Bagby et al., 2004; Barkham et al., 2007). For the current analysis, only the assessment at inclusion (pre-scores) and at the last visit, where all subjects received escitalopram (post-scores; conducted 48/70/153 days after enrollment and 41/49/83 days after initiation of treatment),

were used. In the end, the according psychometric data was available in 29 patients for HAM-D and 28 for BDI. For two subjects, post-scores were missing and linearly interpolated from the visit before and after, rounded up (no change in medication).

Factor Analysis of Depression Scores

Given the number of existing but inconclusive factorization approaches for HAM-D and BDI, a FA appropriate for the current population was conducted and afterwards compared to previous results. Factorization was performed using SPSS 22 (IBM, Armonk, New York) on pre-scores only. Post-scores were automatically calculated from loading matrices to avoid influences of treatment response. For optimizing factors, scores with the lowest measure of sampling adequacy (MSA) were removed and FA repeated until the Kaiser-Meyer-Olkin (KMO) criterion was >0.5 (Backhaus, 2003) and the estimation converged. The following settings were used in SPSS: Factors were extracted using principal axis factorization, where the number of factors was determined via the Kaiser criterion (Guttman, 1954; Kaiser and Dickman, 1959), the analysis was based on correlation matrices, factors were orthogonally rotated using Varimax and the final scores were calculated employing the Anderson-Rubin approach. Varimax and the Anderson-Rubin method were used to enforce orthogonality of the factors and to ensure that the scores are uncorrelated (for details, see the SPSS manual).

Resting-State Data Acquisition

Neuroimaging data was recorded using a 3 T Siemens Biograph mMR system (Siemens, Erlangen, Germany). The resting-state scans started 71/77/100 min after infusion of study medication (the time differences resulted from the acquisition of other sequences not presented here). In detail, 243 frames per run were acquired using the following parameters: repetition time 2.44 s, echo time 30 ms, $2.1 \times 2.1 \times 3$ mm voxel size with 0.75 mm gap, 100×100 voxels in-plane, 36 slices, GRAPPA 2. Subjects were instructed to let their mind wander, look at a black crosshair on a gray background and stay awake.

Resting-State Preprocessing

The RS data of the 29 subjects without missing HAM-D scores was preprocessed using MATLAB R2014a (The MathWorks, Natick, Massachusetts), Statistical Parametric Mapping, version 12 (SPM12¹), and ArtRepair, version 5b (Mazaika et al., 2009)². The following steps were performed: (i) correction of transient slice artifacts (ArtRepair), (ii) slice-timing correction (SPM), (iii) realignment (SPM), (iv), reslicing of realigned images (SPM), (v) pre-smoothing with 4 mm full width at half maximum (FWHM) to improve subsequent motion artifact correction (SPM), (vi) motion regression using the model by Grooten et al. (2000) (ArtRepair), (vii) detection of motion outliers (ArtRepair), (viii) despiking (ArtRepair), (ix) normalization to the standard space defined by the Montreal Neurological Institute (MNI) with an isotropic resolution of 2 mm (SPM).

Resting-State Processing

Further RS-specific processing was conducted using in-house MATLAB code. Nuisance regression and frequency filtering were performed within one model as recommended by Hallquist et al. (2013). As nuisance parameters the first five principal components of white matter and cerebrospinal fluid were used, defined as the respective compartments of the Harvard-Oxford atlas³ thresholded at 95%. The passband was limited to 0.01–0.10 Hz. Frames identified as motion-related outliers using a modified version of the `art_global` MATLAB function (ArtRepair) with linear detrending to avoid misclassification of scanner drifts as movement were excluded from regression and further analysis. Mean time series were extracted from non-overlapping spheres of 10 mm diameter around the 264 region of interest centers defined by Power et al. (2011) (56/65/80 voxels), Pearson-correlated and Fisher-z-transformed given the parametric nature of the subsequent analyses. To reduce bias, only voxels present in all datasets were used for time course extraction.

Network-Based Statistics Analysis

NBS (Zalesky et al., 2010) provide a way to tests connected networks (i.e., continuously linked nodes) with edges showing similar effects. The approach is based on connectivity matrices (functional or structural) and can be seen as an equivalent to the established cluster-level as compared to the voxelwise analysis

on volumetric neuroimaging data. In the same manner, within the NBS framework, a general linear model is first estimated for each edge, using the change in the psychometric scores as regressor for the influence of intravenous citalopram on functional connectivity. Since this method allows for detection of widely distributed effects, it was utilized for a meaningful pre-selection of potentially predictable factors taking all connections into account simultaneously.

In detail, differences in connectivity z-matrices between the verum and placebo condition were used as input for NBS. The design matrices specified general linear models comprising of the respective score changes post-pre (HAM-D or BDI sum or factors), sex and age of patients, the mean connectivity averaged over all connections and conditions (Saad et al., 2013) and an intercept term. Connectivity thresholds were calculated as the *t*-value corresponding to $p \leq 0.001$ at the available degrees of freedom. The significance threshold was set to $p \leq 0.10$ to also cover trend-level effects (i.e., overall smaller networks). Extent (network size) and intensity (network strength) summary statistics were calculated for comparison as suggested by the NBS manual. For statistical inference, 10,000 permutations were performed. Results were not further corrected for the number of tests as this step was only conducted for selecting promising factors.

Assessment of Generalizability

The degree of generalizability was estimated using CVs of complete and simplified models (without covariates) in order to mitigate overfitting. Weights for single connections were calculated as absolute value of the partial Pearson correlations (following the unidirectional NBS contrasts) with the respective coefficients being corrected for covariates and subsequently scaled by the standard deviation of all absolute correlations within the training set (models with signed weights were also calculated for comparison). The FC predictor for each subject was defined as the weighted average of difference of the z-matrices. Training set predictors, covariates and offset term were used afterwards to estimate linear regression models (with robust “bisquare” weighting in MATLAB) for the changes in the HAM-D and BDI factors, which were then applied to the test sets. Three selection procedures (NBS results, NBS threshold and direction of contrast, no threshold) and two resampling techniques (leave-one-out CV, LOOCV; 3-fold CV without role reversal of test and training sets, 1,000 redraws—3CV) were used to avoid assessment-specific results. Agreement between actual and predicted data was again estimated via Pearson correlation (offset and scaling are already accounted for in the models). The results were adjusted for the number of factors per score using the Sidak correction.

RESULTS

Demographics

Six out of the 35 patients with depression enrolled in the study had to be excluded from model estimation due to missing post-scores and one for additionally missing BDI data.

¹www.fil.ion.ucl.ac.uk/spm/software/spm12/

²cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html

³fsl.fmrib.ox.ac.uk/fsl/fslwiki

TABLE 1 | Factor analysis of Hamilton depression rating scale (HAM-D) and Beck's depression inventory (BDI) scores.

HAM-D items	Factor number						BDI items	Factor number					
	1	2	3	4	5	6		1	2	3	4	5	6
1	0.021	0.132	0.030	0.576	0.162	0.020	A	0.609	-0.026	0.438	0.458	0.036	0.181
2	-0.051	0.047	0.129	0.608	-0.270	0.100	B	0.175	0.206	0.358	0.110	-0.056	0.111
3	-0.174	0.172	0.265	-0.020	0.684	-0.049	C	0.374	0.674	0.249	0.103	-0.104	-0.050
4	0.209	0.323	-0.384	0.051	0.208	0.424	D	0.053	0.116	0.773	0.212	0.179	0.073
5	0.820	0.203	-0.118	-0.075	-0.147	0.034	E	0.375	0.391	-0.511	-0.054	0.251	0.031
6	0.713	0.018	0.136	0.101	0.074	0.130	F	0.621	0.249	-0.070	0.104	-0.008	-0.021
7	0.141	-0.007	-0.287	0.129	0.442	-0.003	G	0.098	0.806	0.183	0.367	-0.039	0.022
8	0.142	-0.171	0.253	0.031	-0.070	0.653	H	0.161	0.738	-0.037	-0.258	0.239	0.132
9	0.123	0.120	0.462	0.206	0.062	0.004	I	0.578	0.009	0.221	0.087	-0.028	-0.038
10	-0.127	0.019	0.258	0.359	-0.023	-0.420	J	0.658	0.229	-0.067	0.121	0.215	0.261
11	0.165	-0.015	-0.041	0.404	0.141	-0.197	K	0.446	0.272	0.203	-0.051	0.301	0.192
12	0.117	0.821	0.016	-0.028	0.255	-0.184	L	-0.02	0.298	-0.006	0.065	0.385	0.454
14	0.393	0.137	0.279	0.068	-0.019	0.099	M	0.249	0.168	0.180	0.563	-0.161	-0.215
15	0.173	0.321	0.295	-0.016	-0.341	-0.181	N	0.284	0.267	0.435	-0.126	0.116	0.248
16	0.154	0.692	0.091	0.292	-0.113	0.070	O	0.131	-0.01	0.142	0.582	0.330	0.072
17	0.065	-0.016	0.796	0.000	-0.036	0.071	P	-0.001	0.016	-0.042	0.649	-0.313	0.586
							Q	0.260	-0.145	0.379	0.286	0.801	0.068
							S	0.009	-0.056	0.054	0.079	-0.435	0.027
							U	0.135	-0.016	0.217	-0.032	-0.008	0.670

The two highest-loading items are highlighted for each factor. Meaning of single items: 1: depressed mood, 2: feeling of guilt, 3: suicide, 4: insomnia: early in the night, 5: insomnia: middle of the night, 6: insomnia: early hours of the morning, 7: work and activities, 8: retardation, 9: agitation, 10: anxiety psychic, 11: anxiety somatic, 12: somatic symptoms gastro-intestinal, 13: general somatic symptoms (removed), 14: genital symptoms, 15: hypochondriasis, 16: loss of weight, 17: insight; A: sadness, B: pessimism, C: past failure, D: loss of pleasure, E: guilty feelings, F: punishment feelings, G: self-dislike, H: self-criticalness, I: suicidal thoughts or wishes, J: crying, K: agitation, L: loss of interest, M: indecisiveness, N: worthlessness, O: loss of energy, P: changes in sleeping pattern, Q: irritability, R: changes in appetite (removed), S: concentration difficulty, T: tiredness or fatigue (removed), U: loss of interest in sex.

At inclusion, the remaining 29 subjects' HAM-D distribution was 18/21/38 and at the third visit 0/7/26. At this time, 16 subjects fulfilled the criteria for remission ($HAMD \leq 7$) and 19 responded (HAM-D reduction of 50% or more). For the 28 subjects with both BDI scores, the respective values are 12/28/46 before and 0/16/38 after the investigated treatment period. The final datasets contained 14 female participants (13 for BDI models). No significant influences of the time between psychometric assessments or therapy duration on score reductions were found (Pearson and Spearman correlations all $p \geq 0.10$).

Factor Analysis

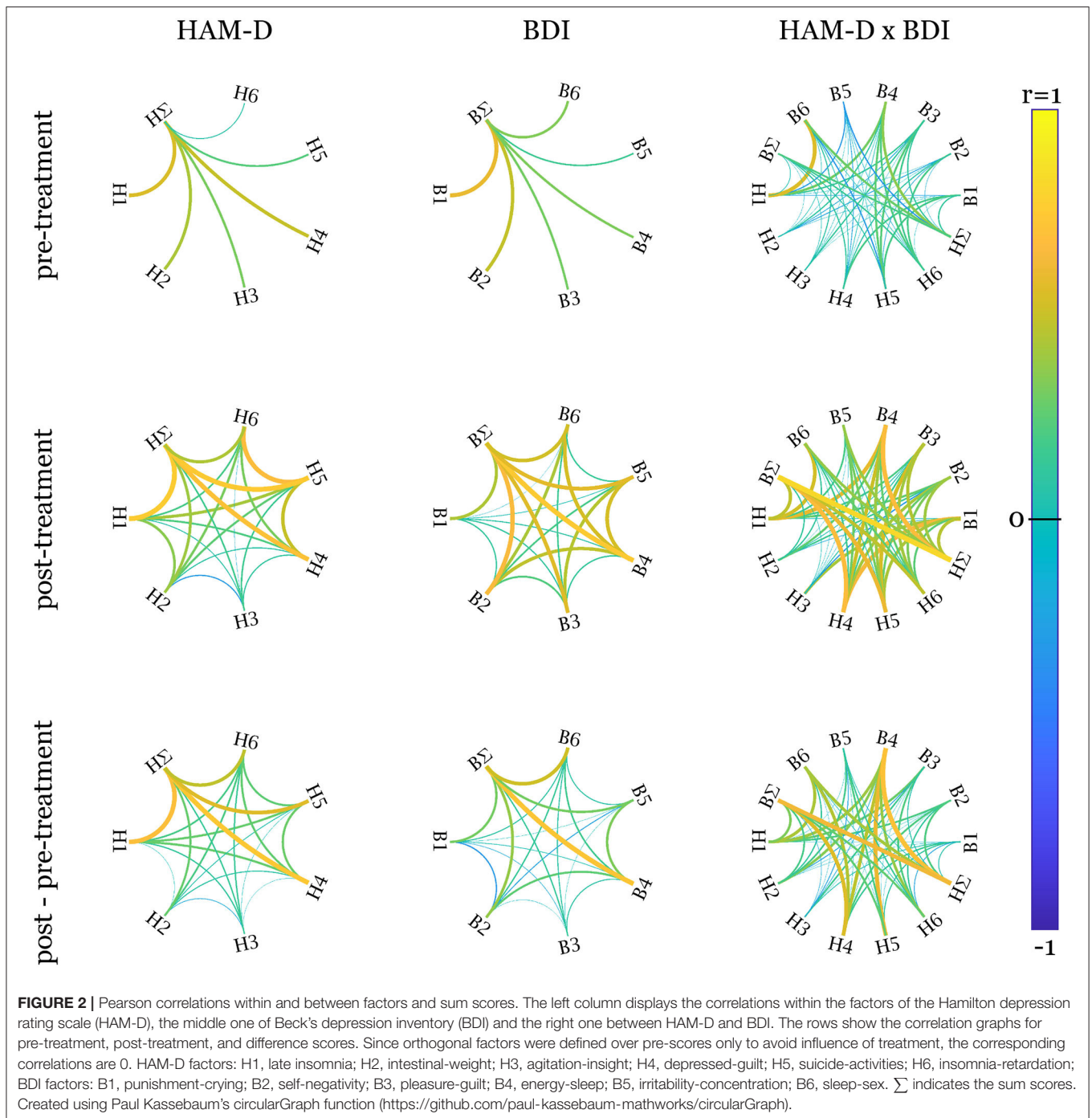
For HAM-D, the pre-scores of all 35 subjects were entered into the FA, 34 could be used for BDI (item enumerations see **Table 1**). After dropping HAM-D item 13, a KMO of 0.53 was reached and six factors explaining 48.50% of variance were identified using the Kaiser criterion. For BDI, a KMO of 0.52 was achieved after removing item T. However, since the procedure did not converge, item R with the next-lowest MSA was also removed, yielding a KMO of 0.56 and also six factors, which explain 57.48% of the overall variance. For easier understanding, the single factors are henceforth addressed according to the two items with the respective highest loadings (factor number in parentheses); HAM-D: late insomnia (1), intestinal-weight (2), agitation-insight (3), depressed-guilt (4), suicide-activities (5), insomnia-retardation (6); BDI: punishment-crying (1), self-negativity (2),

pleasure-guilt (3), energy-sleep (4), irritability-concentration (5), sleep-sex (6).

The correlation structure of factor scores within and between the questionnaires is given in **Figure 2**. Within questionnaires, factor pre-scores are perfectly uncorrelated due to orthogonal Varimax rotation, only show occasionally higher correlations of post- but considerable relationships with sum scores. Between HAM-D and BDI, strong correlations of differences are found for "late insomnia" and "energy-sleep," "depressed-guilt" and "energy-sleep"/"sleep-sex"/BDI sum, "suicide-activities" and "energy-sleep," "insomnia-retardation" and "energy-sleep," HAM-D sum and "energy-sleep"/"sleep-sex"/BDI sum. Sum post-scores also show a high correlation contrary to their pre-scores. Other relationships are especially pronounced between "insomnia" and "sleep" factors of HAM-D and BDI, respectively.

Network-Based Statistics

Overall, 56 tests were conducted [2 scores * (6 factors + 1 sum) * 2 contrast directions * 2 statistics]. No multiplicity correction was applied at this stage since the purpose of these tests was solely the preselection of potentially predictable candidate factors. Extent and intensity statistics were calculated for comparison and yielded the same results. Networks showing uncorrected or trend-level significance were found in 3 scenarios for HAM-D and BDI, respectively. Their structures are depicted in **Figure 3**. Score differences correlated negatively with the difference in connectivity between verum

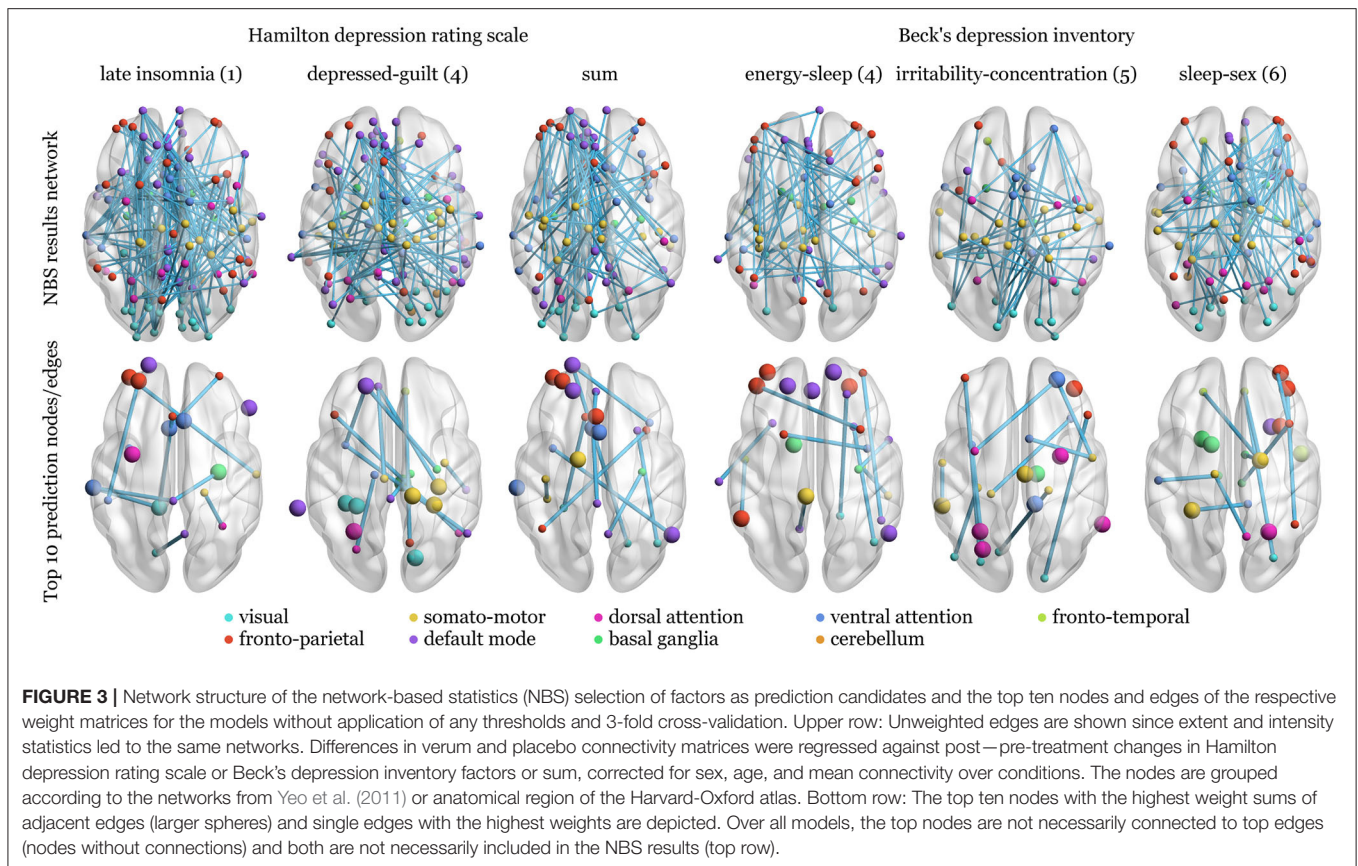


and placebo conditions for all reported cases except for the “irritability-concentration” and “sleep-sex” factors of BDI.

Predictive Power of Citalopram-Induced FC Changes

Predictive power in terms of median Pearson correlation between predicted and actual data is presented in **Table 2** for the full and reduced models. All models showed lower or at most equal predictive power with signed weights (results not shown).

The highest correlation between actual and predicted scores is naturally achieved for connection differences, which were preselected based on the overall data. Being an obvious case of double-dipping, this estimation was only included for reasons of comparison. Whether a threshold was used to suppress connections with low predictive potential has no generally positive influence on the estimates. On the contrary, some predictions show higher agreement using all the connections. Correlations calculated for 3CV models are mostly larger than



those for LOOCV indicating no influence of sample size but potentially influential subjects. **Figures 4, 5** show the distribution of median weights for all tests on full models with regions of interest grouped according to Yeo et al. (2011) and basal ganglia and cerebellum regions from the Harvard-Oxford atlas. Due to the higher number of redraws and variance, median weights are smaller for 3CV with most of them being 0 if the NBS threshold is used. For the threshold-free models with covariates and evaluated using 3CV, the ten nodes and edges with highest sum and individual weights, respectively, are depicted in **Figure 3**. Predictors for HAM-D models feature especially high weights for the ventral attention (VA; e.g., anterior mid-cingulate cortex, left superior temporal and supramarginal gyrus, insula, eye fields), default mode (DM; e.g., frontal cortex, anterior and posterior cingulate cortex, precuneus) and fronto-parietal (FP; e.g., frontal and prefrontal cortex, anterior cingulate cortex) networks. For the BDI models, the FP (mainly mid-frontal cortex, including orbital parts), somato-motor (SM, for the “energy-sleep” factor; e.g., frontal regions, angular gyrus), and DM (e.g., pre- and postcentral gyrus, paracentral lobule, Heschl's gyrus) networks display high weights. Nearly all correlations between real and predicted data are higher for the full models when compared to the reduced ones, excluding the “irritability-concentration” model. The BDI factor “sleep-sex” is not predictable when no covariates and no threshold are

used (the correlation between the predicted and the real data approaches zero).

Since the model for the HAM-D sum score yielded significant results, it was further analyzed in terms of the area under the receiver operating characteristics curve (AUC) and balanced accuracy (BAC). For this, the usual clinical dichotomizations into remission and response were applied to the threshold-free predictions tested with 3CV and 1,000 redraws, as these represent the least restrained model: $AUC_{rem} = 0.73$, $AUC_{resp} = 0.68$, $BAC_{rem} = 0.68$ and $BAC_{resp} = 0.60$. The receiver operating characteristic curves of the *post-hoc* classification performance are presented in **Figure 6**.

DISCUSSION

The results indicate that short-term influences of citalopram on FC indeed contain predictive potential for certain clinician, but not self-ratings of antidepressant SSRI treatment response.

Identification of Factors

An important advantage of FA is meaningful interpretability of results, which is much more difficult for other methods like principal, independent component, or canonical correlation analysis. The latter was also used in Drysdale et al. (2017) in the interest of identifying depression subtypes and FC

TABLE 2 | Median correlation coefficients as estimates of generalizability for models with and without covariates (sex, age, mean global connectivity).

Factor		leave-one-out cross-validation			3-fold cross-validation		
		NBS results	NBS threshold	no threshold	NBS results	NBS threshold	no threshold
Models with covariates	HAM-D: late insomnia (1)	$r = 0.94; p < 1E-13$	$r = 0.45; p < 0.05$	$r = 0.49; p < 0.05$	$r = 0.95; p < 1E-14$	$r = 0.51; p < 0.05$	$r = 0.55; p < 0.01$
	HAM-D: depressed-guilt (4)	$r = 0.83; p < 1E-7$	$r = 0.33; n.s.$	$r = 0.33; n.s.$	$r = 0.88; p < 1E-9$	$r = 0.40; n.s.$	$r = 0.40; n.s.$
	HAM-D: sum	$r = 0.90; p < 1E-10$	$r = 0.41; n.s.$	$r = 0.42; n.s.$	$r = 0.92; p < 1E-11$	$r = 0.51; p < 0.05$	$r = 0.51; p < 0.05$
	BDI: energy-sleep (4)	$r = 0.92; p < 1E-11$	$r = 0.07; n.s.$	$r = 0.19; n.s.$	$r = 0.94; p < 1E-12$	$r = 0.19; n.s.$	$r = 0.19; n.s.$
	BDI: irritability-concentration (5)	$r = 0.85; p < 1E-8$	$r = 0.24; n.s.$	$r = 0.21; n.s.$	$r = 0.88; p < 1E-9$	$r = 0.27; n.s.$	$r = 0.26; n.s.$
	BDI: sleep-sex (6)	$r = 0.88; p < 1E-9$	$r = 0.48; p < 0.05$	$r = 0.40; n.s.$	$r = 0.90; p < 1E-10$	$r = 0.41; n.s.$	$r = 0.40; n.s.$
Models without covariates	HAM-D: late insomnia (1)	$r = 0.83; p < 1E-7$	$r = 0.38; n.s.$	$r = 0.40; n.s.$	$r = 0.88; p < 1E-9$	$r = 0.47; p < 0.05$	$r = 0.48; p < 0.05$
	HAM-D: depressed-guilt (4)	$r = 0.67; p < 0.001$	$r = 0.00; n.s.$	$r = 0.23; n.s.$	$r = 0.79; p < 1E-6$	$r = 0.20; n.s.$	$r = 0.35; n.s.$
	HAM-D: sum	$r = 0.74; p < 0.001$	$r = 0.21; n.s.$	$r = 0.28; n.s.$	$r = 0.82; p < 1E-7$	$r = 0.35; n.s.$	$r = 0.40; n.s.$
	BDI: energy-sleep (4)	$r = 0.83; p < 1E-7$	$r = -0.07; n.s.$	$r = 0.08; n.s.$	$r = 0.88; p < 1E-8$	$r = 0.13; n.s.$	$r = 0.12; n.s.$
	BDI: irritability-concentration (5)	$r = 0.85; p < 1E-7$	$r = 0.26; n.s.$	$r = 0.24; n.s.$	$r = 0.89; p < 1E-9$	$r = 0.36; n.s.$	$r = 0.36; n.s.$
	BDI: sleep-sex (6)	$r = 0.71; p < 0.001$	$r = 0.29; n.s.$	$r = 0.03; n.s.$	$r = 0.78; p < 1E-5$	$r = 0.14; n.s.$	$r = -0.01; n.s.$

Absolute partial Pearson correlation coefficients were used for weighting. Naturally, estimations based on results from network-based statistics (NBS) show the highest agreement since the predictive matrix elements were selected using all of the data. Whether an additional threshold was used to exclude connections with low predictive potential does not make a substantial difference. P-values for the 1,000 runs of 3-fold cross-validation were calculated for $n-2$ degrees of freedom, where n is the number of subjects (assuming all subjects contributed to the result). HAM-D, Hamilton depression rating scale; BDI, Beck's depression inventory; n.s., not significant.

correlates. Even though this approach yields linear combinations of symptoms that already correlate well with those of FC, additional data for independent validation is required. Moreover, in an attempt to replicate the findings in an independent sample, Dinga et al. (2019) also were not able to find similar structures, calling the generalizability of the study into question. Hence, a multistep approach was chosen for the current work keeping psychometric and RS data separated before creating prediction models.

Table 1 shows that negative influences on highest-loaded items are only found for two BDI factors (“guilty feelings” for the “pleasure-guilt” and “concentration difficulty” for the “irritability-concentration” factor). This opposing effect is indeed also observable in the original data: $\rho = -0.32$, $p = 0.07$ for “loss of pleasure” and “guilty feelings” and $\rho = -0.35$, $p < 0.05$ for “irritability” and “concentration difficulty” (Spearman correlation, uncorrected). Several authors have performed exploratory [e.g., Brown et al. (2012), Lee et al. (2018)] and confirmatory [e.g., Morley et al. (2002), Tobias et al. (2017)] FA on BDI data of different patient populations with varying methods and results. In their work, Brown et al. (2012) report only positive correlations between the single BDI items in chronic fatigue patients using a two-factor model accounting for 35.7% of the overall variance. Lee et al. (2018) created a combined model of Beck's anxiety and depression inventories using a five-factor

model (56.2% of combined variance explained) for a mixed group of psychiatric patients. Since those studies used either oblique rotation (allowing for correlation between factors, which was unwanted here) or did not further specify it, direct comparisons of the factors are difficult due to differently oriented coordinate systems.

As for BDI, HAM-D has also been subjected to FA for various scenarios. Broen et al. (2015) conducted a factorization in Parkinson's patients using principal component analysis, oblique rotation and the Kaiser criterion identifying six factors accounting for 59.2% of the total variance. Even though these numbers are very close to those reported here, comparisons beyond the magnitude of results are again impeded by different rotations. Also using principal component analysis and the Kaiser criterion, Olden et al. (2009) identified four latent HAM-D factors for patients suffering from terminal cancer. Since that study also used Varimax rotation, it can be easily derived via standardized regression coefficients that the “late insomnia” and “intestinal-weight” factors have likely equivalents in the “insomnia” and “somatic” factors, respectively (Table 3). A third “depression” factor might be partially represented in “suicide-activities” and “insomnia-retardation.”

The correlation structure of factors in Figure 2 indicates that not only the pre-treatment scores but also improvements in the single factors are, to a high degree, independent. Correlations in

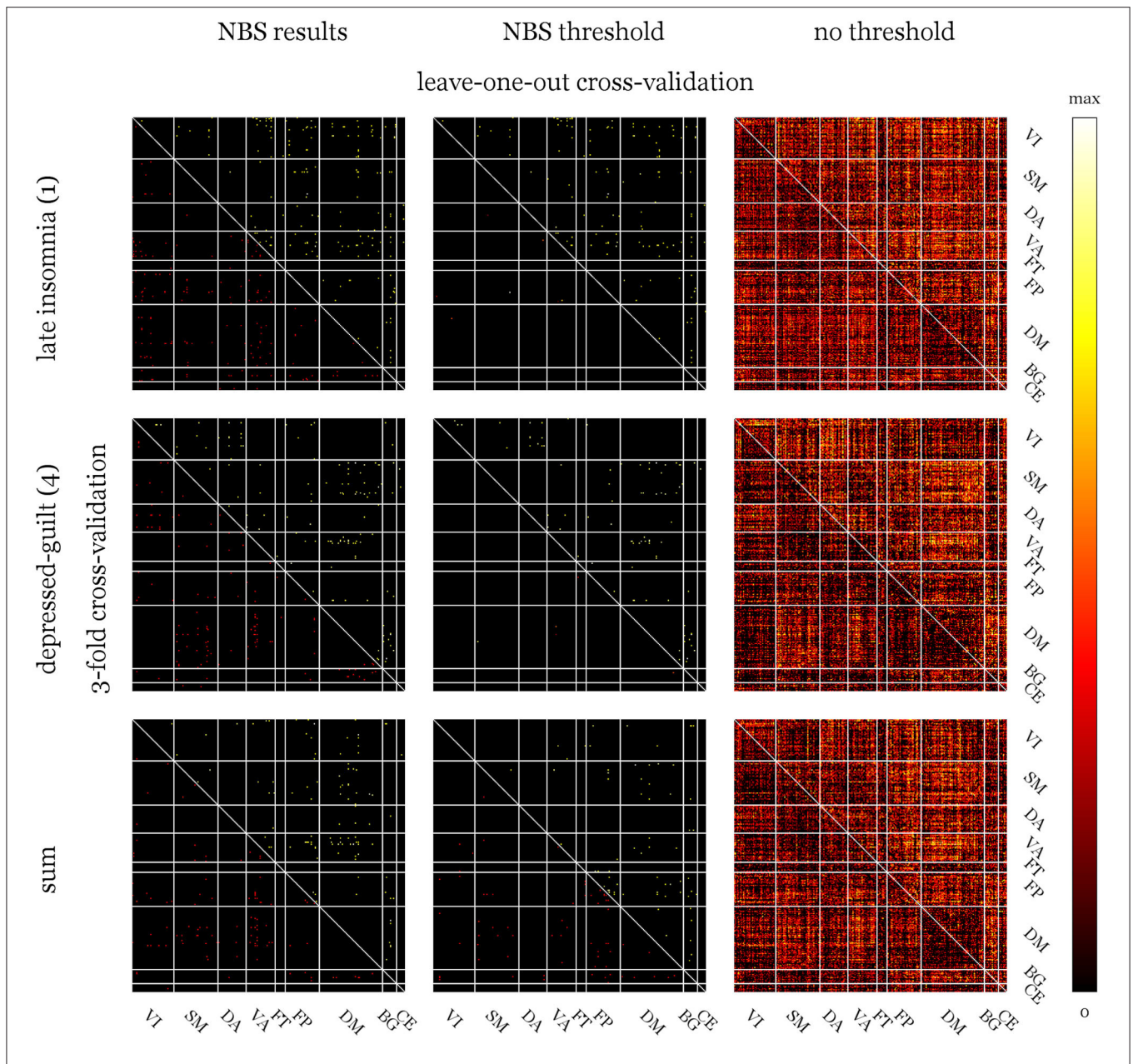


FIGURE 4 | Median weight matrices for full models of derived Hamilton depression rating scale (HAM-D) factors and sum score using three selection strategies. Upper triangles show weights calculated via leave-one-out cross-validation (LOOCV) and lower ones for 3-fold cross validation (3CV) with 1,000 redraws. The left column shows rather equally distributed weights if validation is performed on the pre-selected edges of the network-based statistics (NBS) results. Using the NBS threshold (middle column) to eliminate connections with potentially low predictive capabilities, more of them could be considered in the 1,000 runs of 3CV compared to the number-of-subjects runs of LOOCV. This leads to 0 medians for most 3CV weights. If no threshold was used (right column), the patterns for both methods look quite similar. Weight colors represent the relative influence of the connections on the models and were scaled between 0 and the maximum of each matrix. Yeo atlas: VI, visual; SM, somato-motor; DA, dorsal attention; VA, ventral attention; FT, fronto-temporal; FP, fronto-parietal; DM, default mode; Harvard-Oxford: BG atlas, basal ganglia; CE, cerebellum.

post-scores can be explained by general influences of medication on depressive symptoms. Factors with high loadings for similar items, especially the “insomnia” and “sleep” factors of HAM-D and BDI also display higher correlations emphasizing the appropriateness of their descriptions. Contrary to what could be expected and is reported in Kobak et al. (1990), Reynolds and

Kobak (1995), this finding is not generalizable to sum pre-scores. Correlations of HAM-D and BDI pre-scores are only $r = 0.21$ but $r = 0.82$ for post- and $r = 0.63$ for difference scores. However, this discrepancy regarding the pre-treatment assessment is in line with Lambert et al. (1986) and the reason for investigating both scores. Assuming the expert-ratings are more accurate, this might

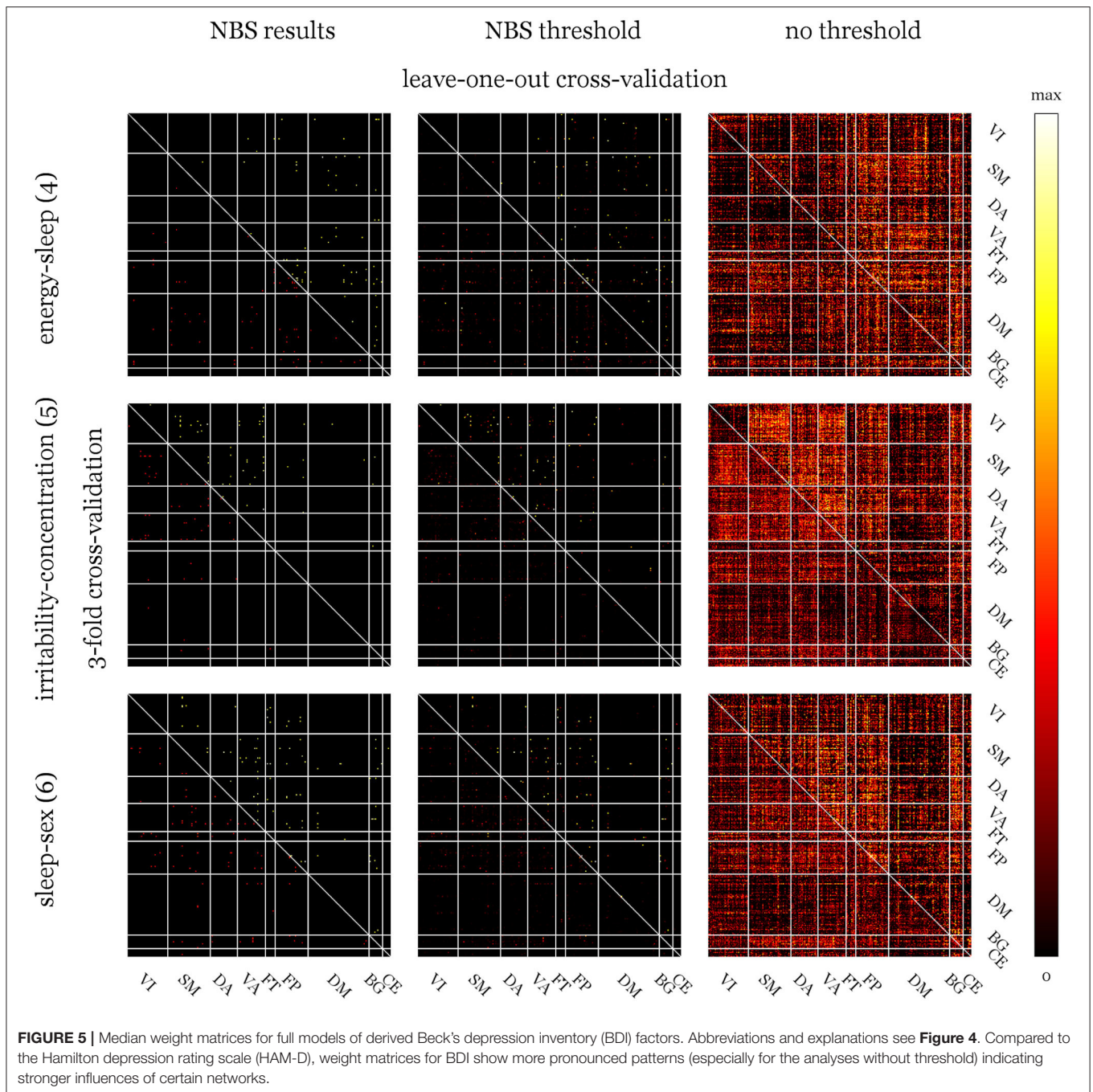


FIGURE 5 | Median weight matrices for full models of derived Beck's depression inventory (BDI) factors. Abbreviations and explanations see **Figure 4**. Compared to the Hamilton depression rating scale (HAM-D), weight matrices for BDI show more pronounced patterns (especially for the analyses without threshold) indicating stronger influences of certain networks.

indicate a biased self-perception with decreasing magnitude over the course of antidepressant treatment that lead to BDI scores being far less predictable from FC. From another perspective, the scores could also be seen as being based on different references: The clinician-based scoring might be influenced by having seen numerous patients suffering from differently severe depressions, whereas the self-rating is based on each patient's perception, which might be affected by a wide range of factors. Furthermore, these findings might also explain NBS results for BDI factors being trend-level significant at most. The partial similarity of

weight distributions of the HAM-D models (see **Figure 4**) is also reflected in the respective correlations of the factors with the sum score.

Predictive Power of Escitalopram Treatment Response Models

The overall aim of this work was to assess whether the response to antidepressant treatment (after a median of 10 weeks) with escitalopram can be predicted using short-term effects of citalopram on FC. To evaluate which aspects of

TABLE 3 | Comparison of the factors extracted from the Hamilton depression rating scale (HAM-D) by Olden et al. (2009) and the current analysis.

Standardized β	late insomnia	intestinal-weight	agitation-insight	depressed-guilt	suicide-activities	insomnia-retardation
Anxiety	-0.30	-0.51	-0.27	-0.07	-0.42	-0.69
Depression	-0.54	-0.16	-0.24	-0.06	0.28	0.27
Insomnia	0.74	-0.09	-0.27	0.06	0.16	0.05
Somatic	0.05	0.68	-0.29	0.12	0.13	-0.05

The table shows regression coefficients β for z-scored columns modeling z-scored rows. "Insomnia" and "Somatic" seem to have a certain agreement with the "late insomnia" and "intestinal-weight" factors. "Depression" might be split up into "suicide-activities" and "insomnia-retardation," where the negative influence of "late insomnia" could be connected to the strong representation of sleep items in two different factors. The exclusively negative coefficients for "Anxiety" are probably related to the comparably low loadings on the respective items (HAM-D items 10 and 11). This interpretation is also corroborated by the fact that the "depressed-guilt" factor shows the highest anxiety-item loadings and least negative β , whereas the "insomnia-retardation" factor behaves exactly the opposite way.

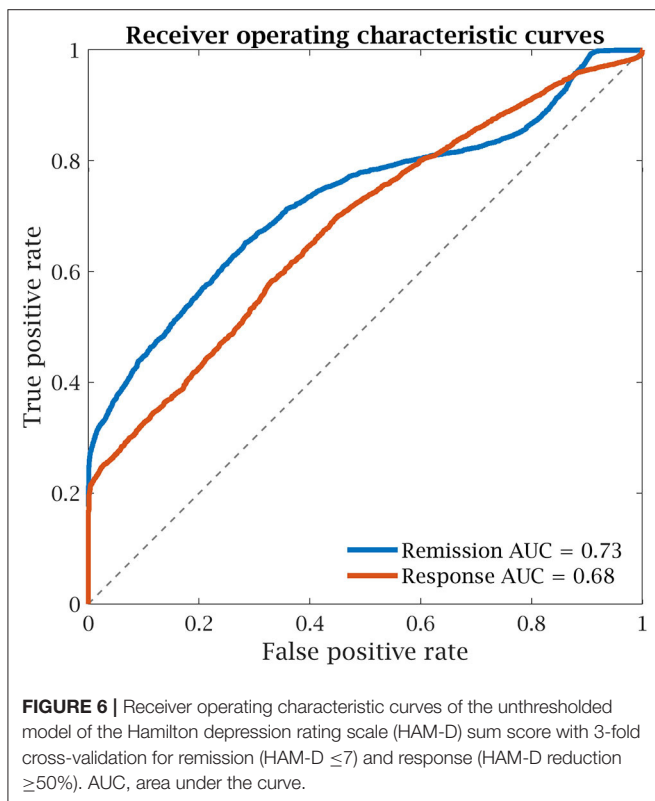


FIGURE 6 | Receiver operating characteristic curves of the unthresholded model of the Hamilton depression rating scale (HAM-D) sum score with 3-fold cross-validation for remission (HAM-D ≤ 7) and response (HAM-D reduction $\geq 50\%$). AUC, area under the curve.

depressive symptomatology are even represented in RS data, NBS analysis of drug influences was conducted. Three networks were identified for HAM-D ("late insomnia," "depressed-guilt" factors and sum score) and BDI changes ("energy-sleep," "irritability-concentration," "sleep-sex" factors) with $p \leq 0.05$ and $p \leq 0.10$ uncorrected for the number of contrasts, respectively. Extent and intensity summary statistics yielded the same results (Figure 3). The networks themselves were of minor importance but afterwards used to demonstrate the effect of biased selection of predictors. Furthermore, the left and right columns of Figures 4, 5, as well as comparing the upper and lower row of Figure 3 indicate that many NBS connections were also assigned higher weights in the unrestricted case using all the connections. Predictive power expressed as Pearson correlation between

estimated and real score reductions was at least twice as high (a factor of four in terms of explained variance) and approaching 1 when only edges of the NBS results were considered (Table 2). Since these were identified over all subjects, the scores that should be predicted already influenced the selection ("double dipping").

Sikora et al. (2016) present a very similar ratio, where FC within the salience network (SN) explains 65% of the placebo response ($r \approx 0.81$) but only shows a predictive power of $r = 0.41$. Concluding predictive capabilities from mere correlations is commonly encountered in literature even though CV or similar approaches could easily be employed to avoid overestimations and add a layer of generalizability.

For the subsequent models, reduced variants without covariates were evaluated to address overfitting (Table 2, lower half). These models performed worse for all but one (BDI: "irritability-concentration") factor, suggesting that the additional information contained in these variables outweighs the increased model complexity. Hence, further investigations were limited to full models. However, the model for the BDI factor "sleep-sex" does not show any predictive capabilities when estimated solely on full correlation matrices, indicating that the result was driven by at least one covariate. Further analyses showed that correlations between predicted and real scores only approach zero when age is not included as a covariate. Indeed, subject age and change in "sleep-sex" factor significantly correlate ($\rho = -0.48, p < 0.01$, Spearman correlation), confirming this assumption. Still, in the preceding NBS analysis a trend-level significant network was found for the factor scores. This might either represent a false-positive result or the effect is only detectable in the corrective presence of an age covariate. However, there could be clinical relevance to this finding since it implies that improvements of depressive symptoms for this factor might be predictable to a certain degree from subject age alone. However, verification in independent samples is needed.

One HAM-D factor describing "late insomnia" and the sum score show significant correlations between predicted and real score differences after correction for multiplicity. Slightly higher agreements for 3CV compared to LOOCV could point toward influential samples. Since for LOOCV the model only changes by one subject between turns, influential ones will be present in the majority of estimates. 3CV, on the other hand, allows for more model variance as one third of the data is left out for each redraw.

Robust “bisquare” weighting was used to mitigate the potential problem of leverage points, but residual effects might be present.

The finding that, besides the sum score, especially changes in a factor describing insomnia are predictable, could be of particular clinical interest. Some drugs with effects on serotonin reuptake/affinity to the serotonin transported are effective in the treatment of sleep disorders (Brietzke et al., 2019). Furthermore, insomnia is one of the possible side effects of escitalopram treatment (Burke, 2002; Waugh and Goa, 2003). Sleep disturbances at baseline themselves also carry predictive information for the antidepressant response (Manglick et al., 2013; Sung et al., 2015), which is partly reflected in the high correlations with the HAM-D sum in **Figure 2** and the fact that the factors were defined based on the pre-treatment score alone. A stronger relationship between the post- and difference scores of the “sleep” factors (energy-sleep and sleep-sex) and the sum score can also be observed for the BDI. Furthermore, bidirectional relationships involving functional and structural connectivity between sleep disorders and MDD are known (Rosenberg et al., 2014; Khazaie et al., 2017). This all suggests a more complex relationship between SSRIs, sleep, and treatment response that demands further investigation. Furthermore, it seems that the predictability of changes in insomnia related to depression is either specific to FC or the influence of citalopram, since this factor could not be predicted in a well-powered study using structural and diffusion MRI (Yang et al., 2018).

Whether a priori thresholding of connections has any positive influence is not clear, but since each matrix element was weighted by the correlation with the regression target, which was also used to define the cutoff, the effect is probably small. When using multiple thresholded runs to build a final model, it needs to be noted that most single potential predictors are below the threshold most of the time (see the lower triangles of the median weight matrices in **Figures 4, 5**, middle columns). For sufficiently large datasets, algorithms that combine regularization, variable selection and prediction such as LASSO regression might provide a more powerful approach. However, in light of the sample size and the failed replication attempt of a promising model combining these steps (Drysdale et al., 2017; Dinga et al., 2019), they were kept separated.

Influence of Functional Nodes and Networks

For the two full HAM-D models showing significant predictive power with or without thresholding, the estimated correlations range from $r = 0.45$ to 0.55 —explaining ~20 to 30% of the variance. Sikora et al. (2016) found a relationship of similar strength ($r = 0.41$) between predicted and real responses to placebo but not antidepressant medication based on connectivity within the SN [included in the VA network of the Yeo atlas (Yeo et al., 2011)], especially the rostral anterior cingulate cortex. **Figures 3, 4** show that connections within and including subregions of the anterior cingulate, which is known for being affected by MDD (Chen et al., 2018; Helm et al., 2018) as well as SSRIs (Arnone et al., 2018; Schranter et al., 2018) contribute highly to the HAM-D predictors. A much stronger correlation

between predicted and true absolute HAM-D scores of healthy and MDD subjects of $r = 0.91$ was reported by Qin et al. (2015). Amongst other regions, the posterior cingulate cortex, precuneus, insula, and basal ganglia were found to be mainly included in treatment-related connections. Those are also present in the nodes and edges with highest weights (**Figure 3**). Applied to the subset of clinically recovered patients, the model of Qin et al. (2015) failed, which is most likely caused by comparably greater variations between healthy subjects and patients. This assumption is supported by the fact that regions with therapy-related connectivity were identified as largely contributive across several of the current models.

A different approach was taken by Karim et al. (2018) for pharmacotherapy in late-life depression and Leaver et al. (2018) for electroconvulsive therapy applied to MDD. Patients were dichotomized into responders and non-responders allowing for classification via support vector machines. The first study used two emotional paradigms and RS fMRI at baseline, the same scans 1 day after a single dose of study medication and Montgomery-Asberg depression rating scale scores taken at baseline and 1 week into treatment. This setup achieved an AUC of 0.77 where important nodes for the support vector machine were largely located at the frontal gyri, including the orbital parts, at task and rest. These regions also display high individual prediction weights for all factors and might represent the influence of MDD on emotion processing (Loeffler et al., 2018). Furthermore, activation of these regions was shown to be indicative for future recurrence of depression (Langenecker et al., 2018b) and can be modulated using SSRIs (Wolf et al., 2018). Leaver et al. (2018) predicted electroconvulsive therapy outcomes with BACs of 58–68%: Across CV splits, FP, motor and superior temporal regions were most often selected as important nodes. The FP network indeed seems to play a universally pivotal role for predicting antidepressant response, whereas the SM network might be of especial importance for certain factors (**Figure 3**).

Komulainen et al. (2018) recently reported influences of escitalopram on FP regions in depressed patients after 1 week and before the onset of symptom improvements. Considering the current results, it might be possible that the FP network shows responses to SSRIs even within hours and speed or magnitude of this effect are, at least partially predictive for the antidepressant response. The motor network is usually not of primary interest in patients with MDD. Still, there are studies showing alterations in MDD (Sarkheil et al., 2020) and also predictive power of SM activation or connectivity for the response to serotonergic medication in Parkinson’s (Ye et al., 2016) and MDD (Klimes-Dougan et al., 2018). Dunlop et al. (2017) investigated the predictive power of subcallosal cingulate cortex connectivity for antidepressant treatment with escitalopram, duloxetine (a serotonin norepinephrine reuptake inhibitor) or cognitive behavioral therapy. Remission after 12 weeks of medication could be identified with an AUC of 0.72 (no validation) based on the sum of subcallosal cingulate cortex FC to the ventromedial and ventrolateral prefrontal cortex/insula and midbrain. The AUC was higher for the sum of the FC scores compared to each one alone, corroborating the strategy of using a combination of nodes as predictor. Despite a certain consistency

in influential connections, it must be noted that multivariate approaches do not directly allow for drawing conclusions on the importance of single features, i.e., one feature with high weight can easily be overruled by multiple others with lower weight. Even though dichotomization of outcomes was performed *post-hoc* in the current study, the results are of similar magnitudes ($AUC = 0.73/0.68$, $BAC = 0.68/0.60$) as those reported by Dunlop et al. (2017).

In a comparison of different machine learning algorithms, Patel et al. (2015) found that age, the mini-mental state examination and structural imaging features were predictive for the diagnosis of late-life depression, whereas for drug response, functional, and structural connectivity yielded the most generalizable results. Patel et al. (2015) also reports that support vector machines are outperformed by alternating decision trees but at least for structural data remain the prevailing machine learning approach for predicting antidepressant treatment response (Patel et al., 2016). This points toward a general underestimation of the predictive capabilities of neuroimaging for depression.

Other than the above-mentioned machine learning algorithms, the current study is not based on non-linear prediction but employs a simpler regression approach due to the moderate sample size. However, using a weighted mean to calculate a single predictor still allows for considering the whole brain at once. Another advantage of the current model is that a continuous outcome can be used avoiding arbitrary cutoffs and loss of statistical power (Kuss, 2013). Compared to studies using only one condition per patient, the influence of subjective traits, which might predict the probability of recovery rather than the response to a certain approach, can be excluded. This is especially important when RS fMRI is utilized where volunteers are usually instructed to let their mind wander. For instance, MDD patients with less pronounced lethargy might more think of physical activity leading to higher SM FC. Considering future studies, non-linearity can be taken into account in two ways: Firstly, the connectivity estimation itself can be based on methods other than the conventional linear Pearson correlation that, e.g., provide higher stability (Geerligs et al., 2016; Meszlényi et al., 2017). Secondly, more complex prediction algorithms can be used such as LASSO regression or the machine learning approaches discussed above. The linear methods used here can be seen as a special case of these approaches, which implies that they could only yield more accurate results given their requirements are met.

Limitations

The main limitation that needs to be addressed is the available number of subjects. Starting at FA, there seems to be no agreement on sample sizes necessary to achieve stable results (Fabrigar et al., 1999; Henson and Roberts, 2006). However, it was shown that respecting quality criteria such as KMO and MSA in combination with a small number of factors already leads to reasonable outcomes (Preacher and MacCallum, 2002; Costello and Osborne, 2005) and FA can be reliably applied to

small sample sizes (De Winter et al., 2009). Within the context of factor optimization, it was also necessary to remove items. Even though this is a common procedure, the respective items vary considerably between studies (Cole et al., 2004). Since no syringeable escitalopram solution exists, the racemic mixture was applied, possibly leading to a less pronounced effect on FC. The relationship between short- and long-term effects of citalopram might also be biased by additional influences on changes in depressive symptoms. These cannot be ruled out since placebo control during the treatment period was not possible. Another source of variation is the large range of time between scans and between psychometric assessments. This was not corrected for due to the already high number of predictors compared to samples but showed no significant relationship and any influence is also reduced by the resampling schemes employed. A higher sample size would also allow for selection of covariates within the model estimation process without considerably compromising stability (Heinze et al., 2018). Future studies should also explicitly estimate the influence of already collected psychometric data, which was out of scope of the current analysis but could constitute an individual baseline with additional measures steering the direction and magnitude of change (Karim et al., 2018).

Conclusion

Using the short-term influence of citalopram on FC compared to placebo (saline), it was possible to predict improvements on one factor and the sum of the HAM-D at $r = 0.45$ to 0.55, validated using pseudo-independent data. The sum score model also performed reasonably well after *post-hoc* dichotomization into remitters and non-remitters ($AUC = 0.73$). Thus, based on the baseline depression severity assessed via the HAM-D and the individual improvements estimated by the model, the score after 7 weeks of escitalopram therapy can be predicted. This constitutes a clear advantage of the regression approach compared to predicting remission or response alone. The same approach, however, failed for BDI, likely due to a mismatch between self-perception of patients and expert perception of clinical symptoms. Including age, sex and a correction for the individual average connectivity was shown to increase predictive power of the models. In spite of in-sample testing, validation of results on independent datasets is needed. Furthermore, prior to an application of the model, a clinical validation is recommended, ideally involving a control group. Since also the HAM-D sum score was predictable, future studies could also take simpler approaches and concentrate on other methodological aspects to further increase predictive power. Additional studies are also needed in light of the different classes of antidepressant medication especially with promising alternatives such as ketamine and psilocybin (Witkin et al., 2018).

DATA AVAILABILITY STATEMENT

The processed data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

This study involving human participants was reviewed and approved by the ethics committee of the Medical University of Vienna. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

The data analysis was conducted by MK with assistance of MR and planned by AH and MK. The overall study was designed by GG, AH, EW-P, RL, and SK and also coordinated by GG. Neuroimaging data was recorded by LR and MK and quality control performed by MR and MK. Medical assistance and patient monitoring were provided by GMG, JU, PM, TV, and GG. Psychometric scoring was performed and validated by the same people. Medical supervision was provided by SK and EW-P. RL is the principal investigator and supervisor of the study. All authors have read and revised the manuscript and agreed to the final version.

REFERENCES

- Alexopoulos, G. S., Hoptman, M. J., Kanelopoulos, D., Murphy, C. F., Lim, K. O., and Gunning, F. M. (2012). Functional connectivity in the cognitive control network and the default mode network in late-life depression. *J. Affect. Disord.* 139, 56–65. doi: 10.1016/j.jad.2011.12.002
- Alshelh, Z., Marciszewski, K. K., Akhter, R., Di Pietro, F., Mills, E. P., Vickers, E. R., et al. (2018). Disruption of default mode network dynamics in acute and chronic pain states. *Neuroimage Clin.* 17, 222–231. doi: 10.1016/j.nicl.2017.10.019
- Argyelan, M., Lencz, T., Kaliora, S., Sarpal, D. K., Weissman, N., Kingsley, P. B., et al. (2016). Subgenual cingulate cortical activity predicts the efficacy of electroconvulsive therapy. *Transl. Psychiatry* 6:e789. doi: 10.1038/tp.2016.54
- Arnone, D., Wise, T., Walker, C., Cowen, P. J., Howes, O., and Selvaraj, S. (2018). The effects of serotonin modulation on medial prefrontal connectivity strength and stability: A pharmacological fMRI study with citalopram. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 84, 152–159. doi: 10.1016/j.pnpb.2018.01.021
- Avissar, M., Powell, F., Ilieva, I., Respingo, M., Gunning, F. M., Liston, C., et al. (2017). Functional connectivity of the left DLPFC to striatum predicts treatment response of depression to TMS. *Brain Stimul.* 10, 919–925. doi: 10.1016/j.brs.2017.07.002
- Backhaus, K. (2003). *Multivariate Analysemethoden: Eine Anwendungsorientierte Einführung*. Berlin; Heidelberg: Springer.
- Bagby, R. M., Ryder, A. G., Schuller, D. R., and Marshall, M. B. (2004). The Hamilton depression rating scale: has the gold standard become a lead weight? *Am. J. Psychiatry* 161, 2163–2177. doi: 10.1176/appi.ajp.161.12.2163
- Barkham, M., Mullin, T., Leach, C., Stiles, W. B., and Lucock, M. (2007). Stability of the CORE-OM and the BDI-I prior to therapy: evidence from routine practice. *Psychol. Psychother.* 80, 269–278. doi: 10.1348/147608306X148048
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., and Erbaugh, J. (1961). An inventory for measuring depression. *Arch. Gen. Psychiatry* 4, 561–571. doi: 10.1001/archpsyc.1961.01710120031004
- Borchardt, V., Lord, A. R., Li, M., Van Der Meer, J., Heinze, H. J., Bogerts, B., et al. (2016). Preprocessing strategy influences graph-based exploration of altered functional networks in major depression. *Hum. Brain Mapp.* 37, 1422–1442. doi: 10.1002/hbm.23111

FUNDING

This work was supported by grants of the Austrian Science Fund (FWF, KLI516), the Vienna Science and Technology Fund (WWTF, CS18-039), and the Else Kröner-Fresenius-Stiftung (EKFS, 2014_A192) awarded to RL. MK and LR are recipients of a DOC fellowship of the Austrian Academy of Sciences at the Department of Psychiatry and Psychotherapy at the Medical University of Vienna.

ACKNOWLEDGMENTS

We thank the study staff of the Neuroimaging Lab at the Department of Psychiatry and Psychotherapy for clinical, scientific, technical, and administrative support. We thank the staff at the Department of Biomedical Imaging and Image-guided Therapy for technical support. Furthermore, we want to thank our students for their assistance throughout the course of this study.

- Brakowski, J., Spinelli, S., Dorig, N., Bosch, O. G., Manoliu, A., Holtforth, M. G., et al. (2017). Resting state brain network function in major depression - depression symptomatology, antidepressant treatment effects, future research. *J. Psychiatr. Res* 92, 147–159. doi: 10.1016/j.jpsychires.2017.04.007
- Brietzke, E., Vazquez, G. H., Kang, M. J. Y., and Soares, C. N. (2019). Pharmacological treatment for insomnia in patients with major depressive disorder. *Expert Opin. Pharmacother.* 20, 1341–1349. doi: 10.1080/14656566.2019.1614562
- Broen, M. P., Moonen, A. J., Kuijff, M. L., Dujardin, K., Marsh, L., Richard, I. H., et al. (2015). Factor analysis of the hamilton depression rating scale in Parkinson's disease. *Parkinsonism Relat. Disord.* 21, 142–146. doi: 10.1016/j.parkreldis.2014.11.016
- Brown, M., Kaplan, C., and Jason, L. (2012). Factor analysis of the beck depression inventory-II with patients with chronic fatigue syndrome. *J. Health Psychol.* 17, 799–808. doi: 10.1177/1359105311424470
- Burke, W. J. (2002). Escitalopram. *Expert Opin. Investig. Drugs* 11, 1477–1486. doi: 10.1517/13543784.11.10.1477
- Chen, L., Wang, Y., Niu, C., Zhong, S., Hu, H., Chen, P., et al. (2018). Common and distinct abnormal frontal-limbic system structural and functional patterns in patients with major depression and bipolar disorder. *Neuroimage Clin.* 20, 42–50. doi: 10.1016/j.nicl.2018.07.002
- Cheng, Y., Xu, J., Arnone, D., Nie, B., Yu, H., Jiang, H., et al. (2017). Resting-state brain alteration after a single dose of SSRI administration predicts 8-week remission of patients with major depressive disorder. *Psychol. Med.* 47, 438–450. doi: 10.1017/S0033291716002440
- Cole, J. C., Motivala, S. J., Dang, J., Lucko, A., Lang, N., Levin, M. J., et al. (2004). Structural validation of the hamilton depression rating scale. *J. Psychopathol. Behav. Assess.* 26, 241–254. doi: 10.1023/B:JJOBA.0000045340.38371.04
- Costello, A. B., and Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract. Assess. Res. Eval.* 10:1–9. doi: 10.7275/yjy1-4868
- Crowther, A., Smoski, M. J., Minkel, J., Moore, T., Gibbs, D., Petty, C., et al. (2015). Resting-state connectivity predictors of response to psychotherapy in major depressive disorder. *Neuropsychopharmacology* 40, 1659–1673. doi: 10.1038/npp.2015.12
- De Winter, J. C., Dodou, D., and Wieringa, P. A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behav. Res.* 44, 147–181. doi: 10.1080/00273170902794206

- Dichter, G. S., Gibbs, D., and Smoski, M. J. (2015). A systematic review of relations between resting-state functional-MRI and treatment response in major depressive disorder. *J. Affect. Disord.* 172, 8–17. doi: 10.1016/j.jad.2014.09.028
- Dinga, R., Schmaal, L., Penninx, B. W. J. H., Van Tol, M. J., Veltman, D. J., Van Velzen, L., et al. (2019). Evaluating the evidence for biotypes of depression: methodological replication and extension of Drysdale et al. *NeuroImage Clin.* 22:101796. doi: 10.1016/j.nicl.2019.101796
- Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., et al. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat. Med.* 23, 28–38. doi: 10.1038/nm.4246
- Du, L., Liu, H., Du, W., Chao, F., Zhang, L., Wang, K., et al. (2018). Stimulated left DLPFC-nucleus accumbens functional connectivity predicts the anti-depression and anti-anxiety effects of rTMS for depression. *Transl. Psychiatry* 7:3. doi: 10.1038/s41398-017-0005-6
- Dunlop, B. W., and Mayberg, H. S. (2014). Neuroimaging-based biomarkers for treatment selection in major depressive disorder. *Dialogues Clin. Neurosci.* 16, 479–490.
- Dunlop, B. W., Rajendra, J. K., Craighead, W. E., Kelley, M. E., Mcgrath, C. L., Choi, K. S., et al. (2017). Functional connectivity of the subcallosal cingulate cortex and differential outcomes to treatment with cognitive-behavioral therapy or antidepressant medication for major depressive disorder. *Am. J. Psychiatry* 174, 533–545. doi: 10.1176/appi.ajp.2016.16050518
- Fabrigar, L. R., Wegener, D. T., Maccallum, R. C., and Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychol. Methods* 4, 272–299. doi: 10.1037/1082-989X.4.3.272
- Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J., et al. (2013). Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med.* 10:e1001547. doi: 10.1371/journal.pmed.1001547
- Fonzo, G. A., Etkin, A., Zhang, Y., Wu, W., Cooper, C., Chin-Fatt, C., et al. (2019). Brain regulation of emotional conflict predicts antidepressant treatment response for depression. *Nat. Hum. Behav.* 3, 1319–1331. doi: 10.1038/s41562-019-0732-1
- Fu, C. H., Costafreda, S. G., Sankar, A., Adams, T. M., Rasenick, M. M., Liu, P., et al. (2015). Multimodal functional and structural neuroimaging investigation of major depressive disorder following treatment with duloxetine. *BMC Psychiatry* 15:82. doi: 10.1186/s12888-015-0457-2
- Geerligs, L., Cam, C., and Henson, R. N. (2016). Functional connectivity and structural covariance between regions of interest can be measured more accurately using multivariate distance correlation. *Neuroimage* 135, 16–31. doi: 10.1016/j.neuroimage.2016.04.047
- Gong, L., Hou, Z., Wang, Z., He, C., Yin, Y., Yuan, Y., et al. (2018). Disrupted topology of hippocampal connectivity is associated with short-term antidepressant response in major depressive disorder. *J. Affect. Disord.* 225, 539–544. doi: 10.1016/j.jad.2017.08.086
- Grootoink, S., Hutton, C., Ashburner, J., Howseman, A. M., Josephs, O., Rees, G., et al. (2000). Characterization and correction of interpolation effects in the realignment of fMRI time series. *Neuroimage* 11, 49–57. doi: 10.1006/nimg.1999.0515
- Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika* 19, 149–161. doi: 10.1007/BF02289162
- Hallquist, M. N., Hwang, K., and Luna, B. (2013). The nuisance of nuisance regression: spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. *Neuroimage* 82, 208–225. doi: 10.1016/j.neuroimage.2013.05.116
- Hamilton, M. (1960). A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* 23, 56–62. doi: 10.1136/jnnp.23.1.56
- Hartmann, A., Von Wietersheim, J., Weiss, H., and Zeeck, A. (2018). Patterns of symptom change in major depression: classification and clustering of long term courses. *Psychiatry Res.* 267, 480–489. doi: 10.1016/j.psychres.2018.03.086
- He, Z., Cui, Q., Zheng, J., Duan, X., Pang, Y., Gao, Q., et al. (2016). Frequency-specific alterations in functional connectivity in treatment-resistant and -sensitive major depressive disorder. *J. Psychiatry Res.* 82, 30–39. doi: 10.1016/j.jpsychires.2016.07.011
- Heinze, G., Wallisch, C., and Dunkler, D. (2018). Variable selection - a review and recommendations for the practicing statistician. *Biom. J.* 60, 431–449. doi: 10.1002/bimj.201700067
- Helm, K., Viol, K., Weiger, T. M., Tass, P. A., Grefkes, C., Del Monte, D., et al. (2018). Neuronal connectivity in major depressive disorder: a systematic review. *Neuropsychiatry Dis. Treat* 14, 2715–2737. doi: 10.2147/NDT.S170989
- Henson, R. K., and Roberts, J. K. (2006). Use of exploratory factor analysis in published research: common errors and some comment on improved practice. *Educ. Psychol. Measure.* 66, 393–416. doi: 10.1177/0013164405282485
- Hou, Z., Gong, L., Zhi, M., Yin, Y., Zhang, Y., Xie, C., et al. (2018). Distinctive pretreatment features of bilateral nucleus accumbens networks predict early response to antidepressants in major depressive disorder. *Brain Imaging Behav.* 12, 1042–1052. doi: 10.1007/s11682-017-9773-0
- Kaiser, H. F., and Dickman, K. (1959). Analytic determination of common factors. *Am. Psychol. Rep.* 14, 425–438.
- Karim, H. T., Wang, M., Andreescu, C., Tudorascu, D., Butters, M. A., Karp, J. F., et al. (2018). Acute trajectories of neural activation predict remission to pharmacotherapy in late-life depression. *Neuroimage Clin.* 19, 831–839. doi: 10.1016/j.nicl.2018.06.006
- Kasper, S., and Muller-Spahn, F. (2002). Intravenous antidepressant treatment: focus on citalopram. *Eur. Arch. Psychiatry Clin. Neurosci.* 252, 105–109. doi: 10.1007/s00406-002-0363-8
- Kessler, R. C., and Bromet, E. J. (2013). The epidemiology of depression across cultures. *Annu. Rev. Public Health* 34, 119–138. doi: 10.1146/annurev-publhealth-031912-114409
- Khazaie, H., Veronese, M., Noori, K., Emamian, F., Zarei, M., Ashkan, K., et al. (2017). Functional reorganization in obstructive sleep apnoea and insomnia: a systematic review of the resting-state fMRI. *Neurosci. Biobehav. Rev.* 77, 219–231. doi: 10.1016/j.neubiorev.2017.03.013
- Klimes-Dougan, B., Westlund Schreiner, M., Thai, M., Gunlicks-Stoessel, M., Reigstad, K., and Cullen, K. R. (2018). Neural and neuroendocrine predictors of pharmacological treatment response in adolescents with depression: a preliminary study. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 81, 194–202. doi: 10.1016/j.pnpb.2017.10.015
- Kobak, K. A., Reynolds, W. M., Rosenfeld, R., and Greist, J. H. (1990). Development and validation of a computer-administered version of the hamilton depression rating scale. *Psychol. Assess. J. Consult. Clin. Psychol.* 2, 56–63. doi: 10.1037/1040-3590.2.1.56
- Komulainen, E., Heikkilä, R., Nummenmaa, L., Raji, T. T., Harmer, C. J., Isometsä, E., et al. (2018). Short-term escitalopram treatment normalizes aberrant self-referential processing in major depressive disorder. *J. Affect. Disord.* 236, 222–229. doi: 10.1016/j.jad.2018.04.096
- Kuss, O. (2013). The danger of dichotomizing continuous variables: a visualization. *Teach. Stat.* 35, 78–79. doi: 10.1111/test.12006
- Lambert, M. J., Hatch, D. R., Kingston, M. D., and Edwards, B. C. (1986). Zung, beck, and hamilton rating scales as measures of treatment outcome: a meta-analytic comparison. *J. Consult. Clin. Psychol.* 54, 54–59. doi: 10.1037/0022-006X.54.1.54
- Langenecker, S. A., Crane, N. A., Jenkins, L. M., Phan, K. L., and Klumpp, H. (2018a). Pathways to neuroprediction: opportunities and challenges to prediction of treatment response in depression. *Curr. Behav. Neurosci. Rep.* 5, 48–60. doi: 10.1007/s40473-018-0140-2
- Langenecker, S. A., Jenkins, L. M., Stange, J. P., Chang, Y. S., Deldonno, S. R., Bessette, K. L., et al. (2018b). Cognitive control neuroimaging measures differentiate between those with and without future recurrence of depression. *Neuroimage Clin.* 20, 1001–1009. doi: 10.1016/j.nicl.2018.10.004
- Leaver, A. M., Wade, B., Vasavada, M., Hellemann, G., Joshi, S. H., Espinoza, R., et al. (2018). Fronto-temporal connectivity predicts ECT outcome in major depression. *Front. Psychiatry* 9:92. doi: 10.3389/fpsy.2018.00092
- Lee, K., Kim, D., and Cho, Y. (2018). Exploratory factor analysis of the beck anxiety inventory and the beck depression inventory-II in a psychiatric outpatient population. *J. Korean Med. Sci.* 33:e128. doi: 10.3346/jkms.2018.33.e128

- Loeffler, L. A. K., Radke, S., Habel, U., Ciric, R., Satterthwaite, T. D., Schneider, F., et al. (2018). The regulation of positive and negative emotions through instructed causal attributions in lifetime depression - a functional magnetic resonance imaging study. *Neuroimage Clin.* 20, 1233–1245. doi: 10.1016/j.nicl.2018.10.025
- Manglick, M., Rajaratnam, S. M., Taffe, J., Tonge, B., and Melvin, G. (2013). Persistent sleep disturbance is associated with treatment response in adolescents with depression. *Aust. N. Z. J. Psychiatry* 47, 556–563. doi: 10.1177/0004867413481630
- Mazaika, P. K., Hoeft, F., Glover, G. H., and Reiss, A. L. (2009). Methods and software for fMRI analysis of clinical subjects. *Neuroimage* 47:S58. doi: 10.1016/S1053-8119(09)70238-1
- Meszlényi, R. J., Hermann, P., Buza, K., Gál, V., and Vidnyánszky, Z. (2017). Resting state fMRI functional connectivity analysis using dynamic time warping. *Front. Neurosci.* 11:75. doi: 10.3389/fnins.2017.00075
- Morley, S., Williams, A. C., and Black, S. (2002). A confirmatory factor analysis of the beck depression inventory in chronic pain. *Pain* 99, 289–298. doi: 10.1016/S0304-3959(02)00137-9
- Nguyen, K. P., Fatt, C. C., Treacher, A., Mellema, C., Trivedi, M. H., and Montillo, A. (2019). Predicting response to the antidepressant bupropion using pretreatment fMRI. *Predict. Intell. Med.* 11843, 53–62. doi: 10.1007/978-3-030-32281-6_6
- Olden, M., Rosenfeld, B., Pessin, H., and Breitbart, W. (2009). Measuring depression at the end of life: is the hamilton depression rating scale a valid instrument? *Assessment* 16, 43–54. doi: 10.1177/1073191108320415
- Patel, M. J., Andreescu, C., Price, J. C., Edelman, K. L., Reynolds, C. F., and Aizenstein, H. J. (2015). Machine learning approaches for integrating clinical and imaging features in late-life depression classification and response prediction. *Int. J. Geriatr. Psychiatry* 30, 1056–1067. doi: 10.1002/gps.4262
- Patel, M. J., Khalaf, A., and Aizenstein, H. J. (2016). Studying depression using imaging and machine learning methods. *NeuroImage Clin.* 10, 115–123. doi: 10.1016/j.nicl.2015.11.003
- Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., et al. (2011). Functional network organization of the human brain. *Neuron* 72, 665–678. doi: 10.1016/j.neuron.2011.09.006
- Preacher, K. J., and MacCallum, R. C. (2002). Exploratory factor analysis in behavior genetics research: factor recovery with small sample sizes. *Behav. Genet.* 32, 153–161. doi: 10.1023/A:1015210025234
- Qin, J., Shen, H., Zeng, L. L., Jiang, W., Liu, L., and Hu, D. (2015). Predicting clinical responses in major depression using intrinsic functional connectivity. *Neuroreport* 26, 675–680. doi: 10.1097/WNR.0000000000000407
- Reynolds, W. M., and Kobak, K. A. (1995). Reliability and validity of the Hamilton depression inventory: a paper-and-pencil version of the Hamilton depression rating scale clinical interview. *Psychol. Assess.* 7, 472–483. doi: 10.1037/1040-3590.7.4.472
- Rosenberg, J., Maximov, I. I., Reske, M., Grinberg, F., and Shah, N. J. (2014). “Early to bed, early to rise”: diffusion tensor imaging identifies chronotype-specificity. *Neuroimage* 84, 428–434. doi: 10.1016/j.neuroimage.2013.07.086
- Saad, Z. S., Reynolds, R. C., Jo, H. J., Gotts, S. J., Chen, G., Martin, A., et al. (2013). Correcting brain-wide correlation differences in resting-state fMRI. *Brain Connect.* 3, 339–352. doi: 10.1089/brain.2013.0156
- Sarkheil, P., Odysseos, P., Bee, I., Zvyagintsev, M., Neuner, I., and Mathiak, K. (2020). Functional connectivity of supplementary motor area during finger-tapping in major depression. *Compr. Psychiatry* 99:152166. doi: 10.1016/j.comppsy.2020.152166
- Schrantee, A., Lucassen, P. J., Booij, J., and Reneman, L. (2018). Serotonin transporter occupancy by the SSRI citalopram predicts default-mode network connectivity. *Eur. Neuropharmacol.* 28, 1173–1179. doi: 10.1016/j.euroneuro.2018.07.099
- Sikora, M., Heffernan, J., Avery, E. T., Mickey, B. J., Zubieta, J. K., and Pecina, M. (2016). Salience network functional connectivity predicts placebo effects in major depression. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 1, 68–76. doi: 10.1016/j.bpsc.2015.10.002
- Souery, D., Amsterdam, J., De Montigny, C., Lecrubier, Y., Montgomery, S., Lipp, O., et al. (1999). Treatment resistant depression: methodological overview and operational criteria. *Eur. Neuropharmacol.* 9, 83–91. doi: 10.1016/S0924-977X(98)00004-2
- Spijker, J., Graaf, R., Bijl, R. V., Beekman, A. T., Ormel, J., and Nolen, W. A. (2004). Functional disability and depression in the general population. *results from the Netherlands mental health survey and incidence study (NEMESIS)*. *Acta Psychiatr. Scand.* 110, 208–214. doi: 10.1111/j.1600-0447.2004.00335.x
- Sung, S. C., Wisniewski, S. R., Luther, J. F., Trivedi, M. H., and Rush, A. J. (2015). Pre-treatment insomnia as a predictor of single and combination antidepressant outcomes: a CO-MED report. *J. Affect. Disord.* 174, 157–164. doi: 10.1016/j.jad.2014.11.026
- Tobias, K. G., Lehrfeld, J., Rosenfeld, B., Pessin, H., and Breitbart, W. (2017). Confirmatory factor analysis of the beck depression inventory-II in patients with advanced cancer: a theory-driven approach. *Palliat. Support. Care* 15, 704–709. doi: 10.1017/S1478951517000724
- Tozzi, L., Goldstein-Piekarski, A. N., Korgaonkar, M. S., and Williams, L. M. (2020). Connectivity of the cognitive control network during response inhibition as a predictive and response biomarker in major depression: evidence from a randomized clinical trial. *Biol. Psychiatry* 87, 462–472. doi: 10.1016/j.biopsych.2019.08.005
- Waugh, J., and Goa, K. L. (2003). Escitalopram: a review of its use in the management of major depressive and anxiety disorders. *CNS Drugs* 17, 343–362. doi: 10.2165/00023210-200317050-00004
- Weigand, A., Horn, A., Caballero, R., Cooke, D., Stern, A. P., Taylor, S. F., et al. (2018). Prospective validation that subgenual connectivity predicts antidepressant efficacy of transcranial magnetic stimulation sites. *Biol. Psychiatry* 84, 28–37. doi: 10.1016/j.biopsych.2017.10.028
- Weinberger, A. H., Gbedemah, M., Martinez, A. M., Nash, D., Galea, S., and Goodwin, R. D. (2017). Trends in depression prevalence in the USA from 2005 to 2015: widening disparities in vulnerable groups. *Psychol. Med.* 48, 1308–1315. doi: 10.1017/S0033291717002781
- Witkin, J. M., Knutson, D. E., Rodriguez, G. J., and Shi, S. (2018). Rapid-acting antidepressants. *Curr. Pharm. Des.* 24, 2556–2563. doi: 10.2174/1381612824666180730104707
- Wolf, D., Klasen, M., Eisner, P., Zepf, F. D., Zvyagintsev, M., Palomero-Gallagher, N., et al. (2018). Central serotonin modulates neural responses to virtual violent actions in emotion regulation networks. *Brain Struct. Funct.* 223, 3327–3345. doi: 10.1007/s00429-018-1693-2
- Yang, J., Zhang, M., Ahn, H., Zhang, Q., Jin, T. B., Li, L., et al. (2018). Development and evaluation of a multimodal marker of major depressive disorder. *Hum. Brain Mapp.* 39, 4420–4439. doi: 10.1002/hbm.24282
- Ye, Z., Rae, C. L., Nombela, C., Ham, T., Rittman, T., Jones, P. S., et al. (2016). Predicting beneficial effects of atomoxetine and citalopram on response inhibition in Parkinson’s disease with clinical and neuroimaging measures. *Hum. Brain Mapp.* 37, 1026–1037. doi: 10.1002/hbm.23087
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165. doi: 10.1152/jn.00338.2011
- Zalesky, A., Fornito, A., and Bullmore, E. T. (2010). Network-based statistic: identifying differences in brain networks. *Neuroimage* 53, 1197–1207. doi: 10.1016/j.neuroimage.2010.06.041
- Zhu, J., Cai, H., Yuan, Y., Yue, Y., Jiang, D., Chen, C., et al. (2018). Variance of the global signal as a pretreatment predictor of antidepressant treatment response in drug-naive major depressive disorder. *Brain Imaging Behav.* 12, 1768–1774. doi: 10.1007/s11682-018-9845-9

Conflict of Interest: Without any relevance to this work, SK received grants/research support, consulting fees and/or honoraria within the last 3 years from Angelini, AOP Orphan Pharmaceuticals AG, Celegne GmbH, Eli Lilly, Janssen-Cilag Pharma GmbH, KRKA-Pharma, Lundbeck A/S, Mundipharma, Neuraxpharm, Pfizer, Sanofi, Schwabe, Servier, Shire, Sumitomo Dainippon Pharma Co. Ltd., and Takeda. RL received travel grants and/or conference speaker honoraria within the last 3 years from Bruker BioSpin MR, Heel, and support from Siemens Healthcare regarding clinical research using PET/MR. He is shareholder of BM Health GmbH since 2019. TV received travel grants and compensation for workshop participation from Pfizer and Eli Lilly and speaker honoraria from Shire.

Preliminary findings of this study were presented at the workshop of the 32nd ECNP congress in Nice, 2019. Further analyses based on these data were presented at the annual meeting of the Organization for Human Brain Mapping, 2019, in Rome.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Klöbl, Gryglewski, Rischka, Godbersen, Unterholzner, Reed, Michenthaler, Vanicek, Winkler-Pjrek, Hahn, Kasper and Lanzenberger. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.