Schizophrenia Bulletin vol. 44 no. 5 pp. 1021–1034, 2018
doi:10.1093/schbul/sbx114
Advance Access publication August 31, 2017

Predicting Response to Repetitive Transcranial Magnetic Stimulation in Patients With Schizophrenia Using Structural Magnetic Resonance Imaging: A Multisite Machine Learning Analysis

Nikolaos Koutsouleris*[,1], Thomas Wobrock[2,3], Birgit Guse[2], Berthold Langguth[4], Michael Landgrebe[4,5], Peter Eichhammer[4], Elmar Frank[4], Joachim Cordes[6], Wolfgang Wölwer[6], Francesco Musso[6], Georg Winterer[7], Wolfgang Gaebel[6], Göran Hajak[8], Christian Ohmann[9], Pablo E. Verde[9], Marcella Rietschel[10], Raees Ahmed[11], William G. Honer[12], Dominic Dwyer[1], Farhad Ghaseminejad[12], Peter Dechent[13], Berend Malchow[1], Peter M. Kreuzer[4], Tim B. Poeppl[4], Thomas Schneider-Axmann[1], Peter Falkai[1], and Alkomiet Hasan[1]

[1]Department of Psychiatry and Psychotherapy, Klinikum der Universität München, Ludwig-Maximilians-Universität, Munich; [2]Department of Psychiatry and Psychotherapy, Georg-August-University Goettingen; [3]County Hospitals Darmstadt-Dieburg, Groß-Umstadt; [4]Department of Psychiatry and Psychotherapy, University of Regensburg; [5]Department of Psychiatry, Psychosomatics and Psychotherapy, kbo-Lech-Mangfall-Klinik Agatharied, Germany; [6]Department of Psychiatry and Psychotherapy, Heinrich-Heine University, Düsseldorf; [7] Experimental & Clinical Research Center (ECRC), Charite – University Medicine Berlin; [8]European Clinical Research Infrastructure Network (ECRIN), Düsseldorf, Germany (previously: Coordination Centre for Clinical Trials, Heinrich-Heine-University, Düsseldorf); [9]Coordination Centre for Clinical Trials, Heinrich-Heine University, Düsseldorf; [10]Department of Genetic Epidemiology in Psychiatry, Institute of Central Mental Health, Medical Faculty Mannheim, University of Heidelberg; [11]Referat Klinische Studien Management, Georg-August-University Goettingen; [12]Institute of Mental Health, The University of British Columbia, Vancouver, Canada; [13]Department of Cognitive Neurology, Georg-August-University Goettingen

*To whom correspondence should be addressed; Professor for Neurodiagnostic Applications in Psychiatry, Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University, Nussbaumstr. 7, D-80336 Munich, Germany; tel: 0049-(0)-89-4400-55885, fax: 0049-(0)-89-4400-55776, e-mail: nikolaos.koutsouleris@med.uni-muenchen.de

Background: **The variability of responses to plasticity-inducing repetitive transcranial magnetic stimulation (rTMS) challenges its successful application in psychiatric care. No objective means currently exists to individually predict the patients' response to rTMS.** ***Methods:*** **We used machine learning to develop and validate such tools using the pre-treatment structural Magnetic Resonance Images (sMRI) of 92 patients with schizophrenia enrolled in the multisite RESIS trial (http://clinicaltrials.gov, NCT00783120): patients were randomized to either active ($N = 45$) or sham ($N = 47$) 10-Hz rTMS applied to the left dorsolateral prefrontal cortex 5 days per week for 21 days. The prediction target was nonresponse vs response defined by a ≥20% pre-post Positive and Negative Syndrome Scale (PANSS) negative score reduction.** ***Results:*** **Our models predicted this endpoint with a cross-validated balanced accuracy (BAC) of 85% (nonresponse/response: 79%/90%) in patients receiving active rTMS, but only with 51% (48%/55%) in the sham-treated sample. Leave-site-out cross-validation demonstrated cross-site generalizability of the active rTMS predictor despite smaller training samples (BAC: 71%). The predictive pre-treatment pattern involved gray matter density reductions in prefrontal, insular, medio-temporal, and cerebellar cortices, and increments in parietal and thalamic structures. The low BAC of 58% produced by the active rTMS predictor in sham-treated patients, as well as its poor performance in predicting positive symptom courses supported the therapeutic specificity of this brain pattern.** ***Conclusions:*** **Individual responses to active rTMS in patients with predominant negative schizophrenia may be accurately predicted using structural neuromarkers. Further multisite studies are needed to externally validate the proposed treatment stratifier and develop more personalized and biologically informed rTMS interventions.**

Key words: schizophrenia/repetitive transcranial magnetic stimulation/neuroanatomical pattern classification/machine learning/voxel-based morphometry/treatment outcome prediction/response heterogeneity

Introduction

Repetitive transcranial magnetic stimulation (rTMS) offers unique possibilities for the induction of long-term

excitability and plasticity changes at the neural-systems level.[1] Due to these plasticity-inducing properties, rTMS has been systematically evaluated as a novel treatment option for various neuropsychiatric disorders. Encouraging therapeutic effects were reported in stroke,[2,3] depression,[4] and schizophrenia[5]—however with equivocal results particularly in the latter disorder.[5–13] This heterogeneity may result from genetic,[14] neuroanatomical,[15] neurofunctional,[16] connectivity-based,[17] and sociodemographic factors[18] mediating the variability of rTMS effectiveness, but none can yet explain the large variability observed at the single-patient level. The discovery and validation of such rTMS response predictors operating reliably in the individual patient could increase effectiveness and broad applicability of rTMS. Furthermore, identifying neuromarkers of rTMS-related treatment outcomes may elucidate pathophysiological mechanisms that control impaired neural plasticity in mental disorders,[19] thus propelling biologically informed, and hence more targeted noninvasive brain stimulation strategies.

Using longitudinal structural neuroimaging, we recently showed that outcome heterogeneity may emerge from the affected brain's varying capacity to structurally adapt hippocampal-precuneal brain networks following a plasticity-inducing rTMS intervention.[20] This observation may point to a link between the longitudinal variability of treatment-induced dynamic brain changes and the cross-sectional structural and functional brain heterogeneity of the "group of schizophrenias".[21,22] Recently, this heterogeneity has been decomposed using cognitive and electrophysiological clustering, revealing distinct biotypes associated with different degrees of clinical impairment and structural brain alterations.[23] Similarly, a recent study deconvolving depression into 4 biotypes by clustering resting-state functional MRI data showed that these biotypes confer highly different response likelihoods to rTMS.[24] Hence, these links between cross-sectional heterogeneity and therapeutic outcome variability raise the intriguing possibility that pre-treatment neuroanatomical signatures may be harnessed to predict rTMS-induced treatment outcomes in patients with severe mental illnesses.

Machine learning applied to the baseline structural Magnetic Resonance Images (sMRI) of patients with major depression has already demonstrated potential to predict response to electro-convulsive therapy (ECT) at the single-subject level.[25] Similarly, pattern recognition techniques were recently reported to successfully predict therapeutic outcomes of ECT in patients with schizophrenia using resting-state functional MRI.[26] In contrast to classical univariate statistics, machine learning algorithms can sift the complex brain patterns underlying neuropsychiatric disorders for clinically relevant predictive fingerprints.[27,28] However, major unresolved challenges question these translational claims, particularly concerning the current lack of investigations assessing the proposed biomarkers' *specificity* for chosen treatment and outcome endpoints based on randomized clinical trial designs, or validating their *generalizability* across sites, scanners, and populations.[29]

Based on the currently largest study database of patients who received "Repetitive Transcranial Magnetic Stimulation (rTMS) for the Treatment of Negative Symptoms in Schizophrenia (RESIS),"[6] we tested for the first time whether machine learning enables the single-subject prediction of rTMS-induced therapeutic responses in schizophrenia with predominant negative symptoms based on single-timepoint sMRI data recorded prior to treatment start. Leveraging the multisite, randomized, double-blinded, placebo-controlled, multiple-timepoint design of the RESIS trial,[6] we probed the proposed models' treatment and outcome specificity, their temporal stability in differentiating good vs poor responders, as well as their leave-site-out generalizability.

Methods

Study Subjects and Endpoints for Individualized Prediction

Patients with an ICD-10 diagnosis of schizophrenia enrolled in RESIS met the following criteria: Positive and Negative Syndrome Scale (PANSS) Negative Subscore (PANSS-NS) > 20 points, 1 PANSS-NS item ≥4, no PANSS-NS reduction ≥10% in the 14 days before treatment start, and an illness duration of ≥1 year.[6] From the Intention-To-Treat (ITT) population (N = 157), 96 patients had pre-treatment sMRI (active/sham rTMS: N = 45/47) and primary PANSS-NS outcome endpoints defined as follows[6,30]: ΔPANSS–NS% = (PANSS–NS_{T1} – PANSS–NS_{T0}) * 100/(PANSS–NS_{T0} – 7). PANSS-NS response was defined as ≥20% improvement between baseline and day 21.[6,9,30] Accordingly, patients were assigned to response or nonresponse groups, with assignments serving as targets for the machine learning analyses. The same approach defined a PANSS positive (PANSS-PS) endpoint based on the respective pre-post score change.

Intervention

Patients received either 10 Hz active or sham rTMS applied to the left DLPFC according to the EEG-10–20 system (F3-electrode, 5 sessions/wk during the 3-week period, 1000 stimuli/d, 50 stimuli/train) with 110% of the individual resting motor threshold (RMS).[6] Clinical data were recorded before stimulation (baseline/T0) and after day 21 (T1), day 28 (T2), day 45 (T3) and day 105 (T4) (supplementary figure 1). As reported previously, in the ITT population no significant differences in the primary outcome, other clinical outcomes and cognition could be detected between active and sham rTMS.[6,9]

MRI Imaging Data Acquisition and Processing

Structural MR images were obtained on two 3T systems (Siemens Trio) and one 1.5T system (Siemens Sonata) using T1-weighted sequences.[20] All images were quality-controlled (N.K.) and 4 subjects had to be removed due to poor image quality. The sMRI of the remaining 92 patients underwent automated tissue segmentation and high-dimensional stereotactic registration with Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL),[31] producing gray matter density (GMD) images registered to the MNI-152 template and smoothed with an 8 mm Gaussian kernel. We did not modulate the GMD images with the Jacobian determinants obtained from DARTEL as this step may decrease the sensitivity of voxel-based morphometry to detect mesoscopic brain alterations due to "multiplicative noise," ie, the interindividual macroscopic brain shape differences captured by the Jacobian determinants.[32] We tested the effects of modulation in a supplementary analysis (see below). Further image acquisition and preprocessing details are described in the supplementary methods.

Neuroanatomical Pattern Classification

We generated 3 overall predictive models using our freely available machine learning tool NeuroMiner (https://www.pronia.eu/neurominer/): one to predict the PANSS-NS response criterion in patients receiving active rTMS, and—to validate model specificity—one to predict this endpoint in sham-treated patients, and a third model to predict PANSS-PS outcomes in active rTMS. As in our previous work[33,34] and recommended for predictive modeling,[35] we employed repeated nested cross-validation with 10 permutations × 20 folds at the outer cross-validation cycle, and 1 × 19 folds at the inner cycle to prevent information leaking between patients used for training, testing, and validating our models. This enabled the unbiased estimation of these models' generalizability to new patients. Following preprocessing steps were applied to each inner partition of the nested cross-validation cycle (supplementary methods): First, the GMD training data were standardized based on their voxel-level mean and SDs. Standardized data were site-adjusted using an established multivariate correction method[36]: Principal Component Analysis (PCA) reduced the training patients' standardized GMD images to 20–25 principal components (PCs), accounting for 80% of the variance in the images.[37] Then, analysis of variance (ANOVA) fitted 3 predictors encoding the training subjects' site membership to their loadings on each PC. PCs explaining the site-encoding design matrix with $R^2 > .16$ were removed from the training data. Finally, processed training data were scaled feature-wise from 0 to 1. Then, standardization, PCA-based dimensionality reduction, PC removal, and scaling parameters were applied to the respective test and validation subjects' data.

In each training sample, we used the linear Support Vector Machine (SVM)[38,39] to detect a decision boundary that predicted patients' outcome class in given rTMS treatment arm using their site-adjusted PCA scores. All SVM models originating from an inner-cross validation cycle were then combined into an ensemble predictor,[40] which was applied to the respective outer-cycle validation patients. This process was repeated across all outer-cycle partitions of the repeated nested cross-validation design and for each validation patient the obtained SVM decision scores were integrated into a probabilistic ensemble prediction using majority voting. The GMD baseline signature predicting subsequent response vs nonresponse to active rTMS was described in figure 1.

We performed additional analyses to test model significance, generalizability and therapeutic specificity: First, we determined whether the observed prediction performances of the active and sham predictors significantly differed from a null distribution of the respective outcome labels by training and cross-validating SVM models on $n = 1000$ random label permutations.[41] Model significance was defined at $\alpha = .05$ as $P = \sum_{i=1}^{n=1000} (\mathrm{BAC}_{\mathrm{observed}} < \mathrm{BAC}_{\mathrm{permuted}}) / n$, where $\mathrm{BAC} = \frac{(\mathrm{sensitivity} + \mathrm{specificity})}{2}$. Second, we assessed the models' generalizability by implementing a leave-site-out (LSO) approach: each of the 3 study sites was iteratively held back for validation, while the remaining data entered the inner cross-validation cycles. We observed that the LSO classifiers showed lower prediction performances in the active rTMS condition than the pooled classifier (table 2). Neuroanatomical outcome probabilities and Receiver-Operating Characteristics (ROC) plots generated by the 3 LSO classifiers on the held-back sites are presented in figure 2B along with the respective charts of the pooled cross-validation experiment (figure 2A). To assess whether this performance drop was due to residual site effects, or, alternatively due to the lower training sample sizes available in the LSO experiments, we trained LSO predictors on $n = 1000$ permutations of the patients' site memberships. We rejected the null hypothesis of original LSO predictors having an equal or higher prediction performance than the permuted predictors at $\alpha = .05$, where $P = \sum_{i=1}^{n=1000} \left(\mathrm{BAC}_{\mathrm{LSO}_{\mathrm{observed}}} \geq \mathrm{BAC}_{\mathrm{LSO}_{\mathrm{permuted}}}\right) / n$. Third, we tested the treatment and outcome specificity of the active rTMS predictor, by measuring its prediction performance

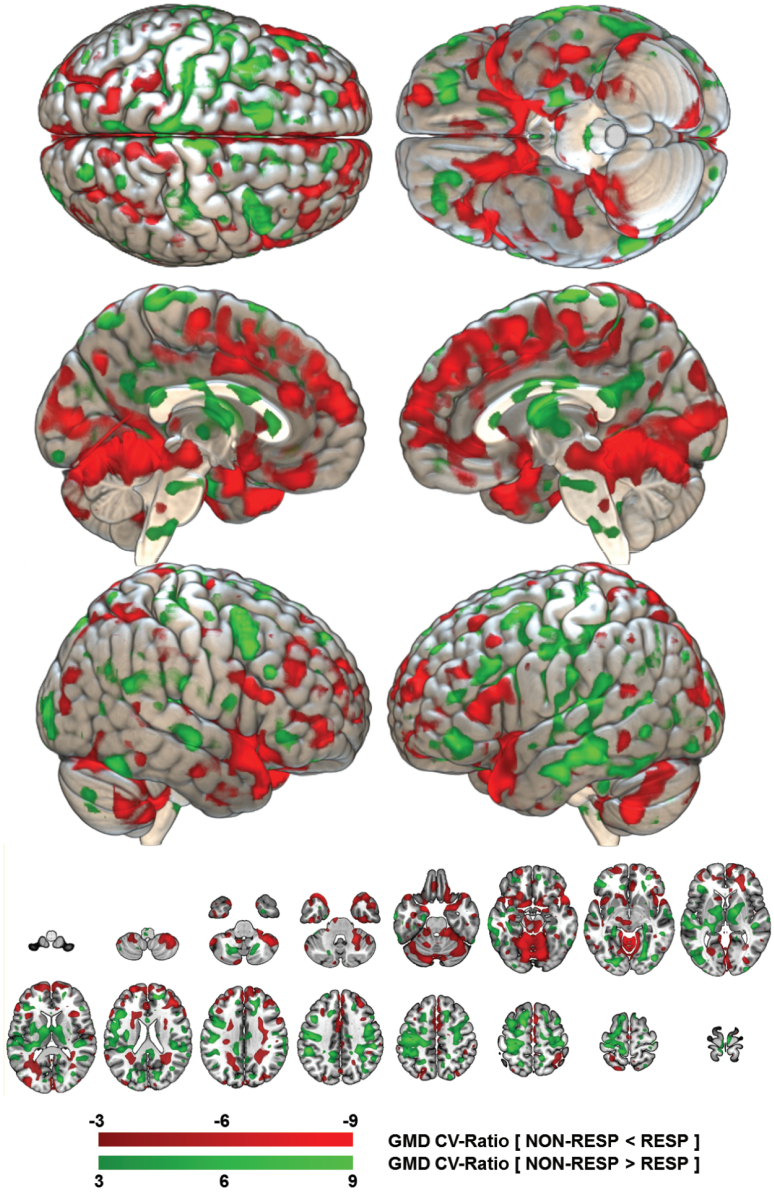

Fig. 1. Reliability of the baseline gray matter density pattern predicting subsequent response vs nonresponse to active repetitive transcranial magnetic stimulation (rTMS). The reliability of the gray matter density (GMD) pattern elements was measured in terms of a Cross-Validation Ratio (CVR) map [CVR = mean(**w**) / standard error(**w**)], where **w** are the weight vectors of the 5111 Support Vector Machine (SVM) models generated in the study's repeated nested cross-validation setup]. The CVR map was thresholded at a CVR of ±3, corresponding to an alpha level of .01 reliable areas of GMD reduction in non-responders (NON-RESP) vs responders (RESP) are shaded in red colours, whereas areas of GMD increments are painted in green. The open-source 3D rendering software MRIcroGL (C. Rohrden) available at https://www.nitrc.org/projects/mricrogl/ was used to overlay the CVR map on the MNI single-subject template.

in *sham*-treated patients, and by comparing it against the PANSS-PS outcome predictor.

Further supplementary analyses were carried out in the active rTMS group to explore whether the analysis of (1) global gray matter, white matter, cerebrospinal fluid, and total intracranial volumes, or (2) modulated voxel-level GMD maps provided similar predictive performance as the analysis of unmodulated, voxel-level GMD maps (supplementary figure 5 and supplementary table 3). In addition, we assessed whether standard mass-univariate analysis methods detected a similar pattern of GMD differences between nonresponders vs responders as the multivariate pattern recognition pipeline described above (supplementary figure 6).

Post hoc Analyses

As only the active rTMS outcome prediction model was significant in the permutation analysis (see Results section), we performed a series of post-hoc analyses using SPSS 23 (IBM; significance level: $\alpha = .05$) to evaluate (1) how well the patients' neuroanatomical nonresponse

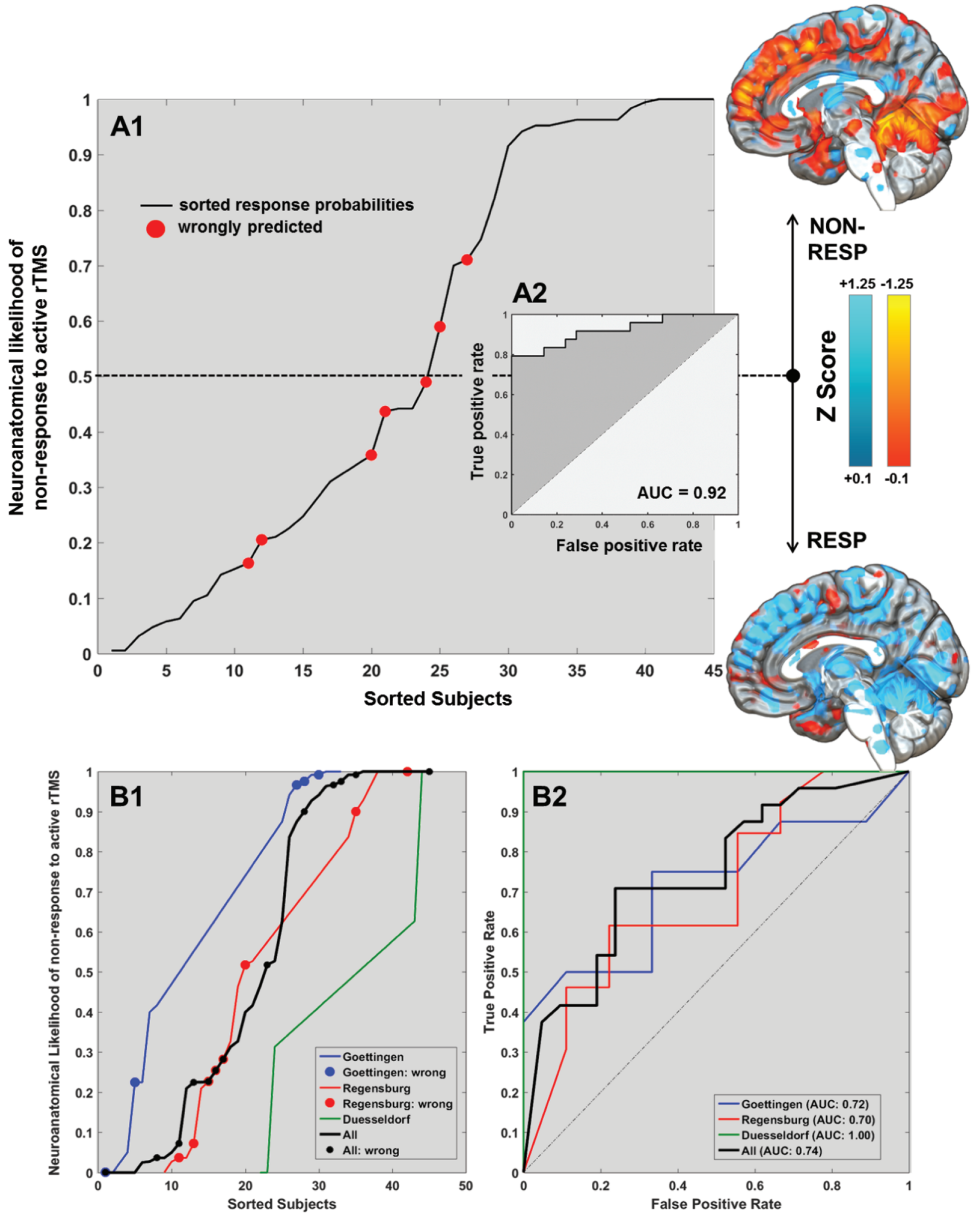

Fig. 2. Sorted neuroanatomical nonresponse likelihoods of patients treated with active repetitive transcranial magnetic stimulation (rTMS) (**A1**). For each patient in study population the active rTMS outcome predictor generated ensemble-based out-of-training probabilities of belonging to the nonresponse group (*y* axis). The figure shows these likelihoods sorted in ascending order, with red-colored subjects indicating misclassifications. Using these outcome predictions, a Receiver-Operating Characteristic (ROC) analysis was conducted (**A2**). Based on the patients identified as support vectors by the Support Vector Machine (SVM) algorithm, voxel-level mean and standard deviation maps were computed and used to standardize patients in the active rTMS group. The standardized maps were separately averaged for the 2 prediction groups to illustrate the quantitative baseline gray matter density (GMD) differences between patients with predicted nonresponses (top) vs responses (bottom) to active rTMS. The bottom panel of the figure shows the sorted outcome probabilities and misclassification (circles) for each of the 3 RESIS sites (**B1**) as well as the respective ROCs and Areas-under-the-Curve (AUC; **B2**).

likelihoods generated by this classifier predicted their continuous PANSS *score changes* measured between T0 and T1, and (2) whether the classifier's stratification effect observed at T1 delineated rTMS after-effects that extended to the T3 and T4 time points. For the former analysis, we calculated the Spearman's ρ correlation between the patients' neuroanatomical nonresponse likelihoods and PANSS-NS change scores (figure 3A). For the latter, we performed a linear mixed model analysis by modeling the effects of the within-subject factor "TIME" (T0, T3 and T4) and the fixed between-subject factor "PREDICTED OUTCOME" (response vs nonresponse prediction) on the dependent PANSS-NS score variable. We did not include the T1 (day 21) and T2 (day 28) examinations into the mixed model to avoid re-fitting the predicted outcome trajectories to the target endpoint of the machine learning

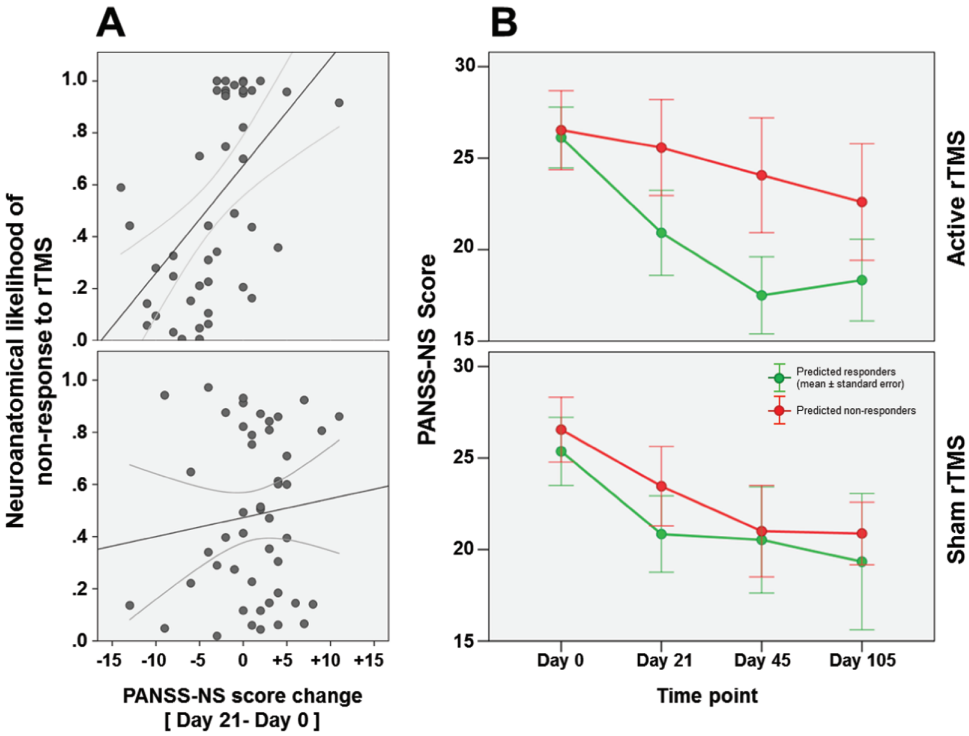

Fig. 3. (A) Correlation analyses between patients' neuroanatomical likelihood of nonresponse to active (top) and sham (bottom) TMS and PANSS-NS absolute score changes measured between T0 and T1. (B) Descriptive absolute PANSS-NS trajectory graphs of patients with predicted nonresponse (red) vs response (green) to active (top) vs sham (bottom) repetitive transcranial magnetic stimulation (rTMS) spanning from baseline to day 105 follow-up examinations. Here, the T1 timepoint (day 21) was included in the figure for visualization purposes.

analysis and its closest follow-up interval (1-wk difference). The factor "'SITE" was entered as random factor in the design. Additionally, we assessed the specificity of the PANSS-NS outcome predictor's longitudinal stratification effect by comparing it against mixed models which used the PANSS positive, GAF, and MADRS scores at T0, T3 and T4, instead of the PANSS-NS as dependent variables (supplementary analyses). Finally, we performed a supplementary regression analysis to explore whether the neuroanatomical response likelihoods of the patients randomized to active rTMS were significantly ($\alpha = .05$) associated with the longitudinal hippocampal-precuneal volume increases reported previously.[20]

Results

Sociodemographic and Clinical Data

PANSS-NS did not differ between treatment groups at baseline and at T1 ($t_{(90)} \leq 0.853$, $P \geq .396$) as previously shown in the ITT population.[6] Similarly, we did find not find other clinical baseline differences except for higher PANSS-PS scores in active rTMS and similar trend effects in the PANSS general and PANSS total scores (table 1). Treatment responses observed in our MRI sample also agree with the ITT analysis, with both treatment groups improving similarly over time (PANSS: all $F \geq 10.51$, all $P \leq .002$; MADRS: $F_{(1,89)} = 17.27$, $P < .001$; GAF: $F_{(1,83)} = 16.24$, $P < .001$; supplementary table 2). Distributions of PANSS-NS responders and nonresponders were equal across rTMS conditions (active vs sham rTMS responders/nonresponders: 21/24 vs 22/25; $\chi^2_{(1)} < 0.001$, $P = .989$). PANSS-PS scores were significantly higher in the active rTMS condition at baseline ($t_{(87)} = 2.565$, $P = .012$), but not at day 21 ($t_{(87)} = 0.876$, $P = .383$).[6]

Neuroanatomical Prediction Performance

The MRI-based outcome classifier trained on the active rTMS group correctly separated PANSS-NS nonresponders from responders with a cross-validated balanced accuracy (BAC; sensitivity, specificity) of 84.8% (79.2%, 90.2%; table 2), resulting in Area-Under-the-Curve of 0.92 (figure 2) and a positive Likelihood Ratio of 8.3. Hence, a positive/negative prediction increased a patient's nonresponse/response likelihoods by +37.1%/+32.5%, summing up to a gain in prognostic accuracy of +69.6%. In contrast, the supplementary prediction analysis using global and total-intracranial volumes provided a cross-validated BAC of 66.7% (Sensitivity: 66.7%, Specificity: 66.7%; supplementary table 3). The lowest prediction performance (BAC: 60.1%, Sensitivity: 58.3%, Specificity: 61.9%) was observed when modulated GM maps were used instead of GMD maps.

The neuroanatomical pattern predicting subsequent nonresponse to active rTMS particularly involved relative GMD reductions in the (1) dorsomedial and ventromedial

Table 1. Sociodemographic and Clinical Differences at Baseline Between Patients Treated With Active vs Sham rTMS

	Active rTMS (n = 45)		Sham rTMS (n = 47)		Active vs Sham		
					χ^2	*df*	*P*
Sociodemographic variables							
Gender (male: female)	39:6		37:10		1.010	1	.315[a]
Site (Goettingen: Regensburg: Duesseldorf)	17:22:6		17:22:8		0.242	2	.886[a]
Hand preference (right: not right)	39:5		39:6		0.080	1	.778[a]
	Mean	SD	Mean	SD	*t*		
Age (y)	34.00	9.93	35.64	9.37	0.814	90	.418[b]
Education (y)	11.43	1.91	11.34	2.13	0.215	89	.830[b]
rTMS functional and anatomical parameters							
Left Resting Motor Threshold (RMT)	46.72	10.31	47.95	11.78	0.519	85	.605[b]
Scalp-to-cortex distance BA 9 (mm)	16.27	2.16	16.68	1.92	−0.951	90	.344[b]
Scalp-to-cortex distance BA 46 (mm)	16.28	2.43	16.90	2.50	−1.211	90	.299[b]
Severity of illness and treatment							
PANSS negative symptoms	26.31	4.45	25.91	4.42	0.428	90	.669[b]
PANSS positive symptoms	14.41	4.33	12.36	3.15	2.565	87	.012*[b]
PANSS general symptoms	42.36	9.46	38.73	9.92	1.766	87	.081[b]
PANSS total	83.23	14.23	77.27	14.76	1.939	87	.056[b]
Global Assessment of Functioning	52.07	12.27	52.37	11.77	0.115	83	.908[b]
Antipsychotic Dose (CPZ mg)	598.60	451.14	596.60	494.49	0.020	86	.984[b]
Depression severity							
Montgomery–Åsberg Depression Rating Scale (MADRS)	14.48	5.48	13.83	5.59	0.557	89	.579[b]

Note: PANSS, Positive and Negative Syndrome Scale; CPZ: Chlorpromazine equivalents, BA: Brodmann Area; rTMS, repetitive transcranial magnetic stimulation. Sociodemographic and Clinical Differences Were Assessed Using Independent *t* tests and Chi-square tests in SPSS 23.
[a]*P* value obtained from Chi-square test on independence.
[b]*P* value obtained from independent *t* test. *$P < 0.05$.

prefrontal, frontopolar and cingulate cortices, (2) the insular, opercular, temporopolar and medial temporal cortices, and (3) the cerebellum (figure 1). Increased baseline GMD predicting nonresponse was observed in the left-hemispheric somatosensory and parietal cortices with extensions to the lateral temporal and premotor structures, as well as in the thalamic nuclei, bilaterally. This pattern of relative GMD reductions and increments was much more extended than the differences detected by the univariate GMD analysis (supplementary figure 6). There, we found only cerebellar and circumscribed medial prefrontal and frontopolar GMD reductions to be associated with subsequent nonresponse to rTMS. Permutation analysis showed that the BAC generated by the machine learning-based pattern was significant at $P < .001$. Furthermore, this pattern specifically separated nonresponders from responders in the active, but *not* in the sham treatment group: the trained active rTMS predictor's BAC measured 57.5% (56.0%, 59.1%; table 2) in the latter patients. Moreover, our machine learning pipeline produced an accurate predictor for the PANSS-NS endpoint, but *not* for PANSS-PS (BAC = 35.9%; table 2). Finally, the active rTMS predictor was *not* significantly moderated by sociodemographic and clinical variables, antipsychotic treatment intensity (chlorpromazine equivalents) at baseline and over the treatment period, type of psychopharmacological treatment (clozapine, antidepressants, mood stabilizers), as well as the scalp-to-cortex distances of Brodmann Areas 9 and 46, or the left resting motor threshold (RMT) (table 3).

When validating the active rTMS model's LSO generalizability, we measured an overall BAC (sensitivity, specificity) of 71.1% (70.8%, 71.4%; $P < .001$; table 2). The lowest/highest BAC of 64.1%/100% (each $P < .05$) was observed when the 22/6 patients from Regensburg/Duesseldorf were used for LSO validation (49%/13.3% of the active rTMS sample). These performances did not differ from the BACs obtained in the 1000 random permutations of the patients' site membership (all $P > .35$, table 2) indicating that the BAC drop between pooled and LSO analysis was due to smaller training samples and not residual site effects. Finally, in contrast to the active rTMS outcome classifier, the predictive model trained on the sham rTMS sample produced a BAC of 51.3% (48.0%, 54.6%), which was nonsignificant and non-generalizable (table 2).

Neuroanatomical Nonresponse Likelihoods Relating to Outcome Trajectories

The neuroanatomical response probabilities obtained from the active rTMS predictor explained 34% of the variance in the PANSS-NS changes measured

Table 2. Validated Predictive Performances of Active and Sham rTMS Response Classifiers

Machine Learning Analyses	N_{tr}	TP	TN	FP	FN	Sens	Spec	BAC	PPV	NPV	PSI	LR+	DOR	*P*
Active rTMS: PANSS-NS outcome predictor														
Pooled repeated nested cross-validation (P-CV)	45	19	19	2	5	79.2	90.2	84.8	90.5	79.2	69.6	8.3	69.1	<.001*[a]
Leave-site-out repeated nested cross-validation (LSO-CV)														
Full sample	45	16	16	5	8	70.8	71.4	71.1	73.9	68.2	42.1	2.50	6.1	<.001*[a]
Goettingen left-out	28	6	6	3	2	75.0	66.7	70.8	66.7	75.0	41.7	2.25	5.1	.029*[a]
Regensburg left-out	23	8	6	3	5	61.5	66.7	64.1	72.7	54.5	27.3	1.85	3.4	.042*[a]
Duesseldorf left-out	39	3	3	0	0	100	100	100	100	100	100	—	—	.007*[a]
Non-inferiority analysis of LSO-CV vs P-CV based on 1000 permutations of the site labels: Mean (SD) performance of permuted LSO predictors														
Full sample (SD)	45	17.0 (2.5)	13.6 (2.6)	7.4 (2.6)	7.3 (2.5)	70.8 (10.3)	66.7 (12.4)	67.9 (7.0)	69.6 (7.5)	66.7 (8.1)	37.5 (14.2)	2.00	4.0	.725[b]
Goettingen (SD)	28	6.0 (1.3)	5.1 (1.5)	2.9 (2.2)	2.6 (1.9)	72.7 (16.9)	66.7 (21.4)	70.1 (10.4)	72.7 (16.8)	66.7 (17.2)	41.7 (20.8)	2.18	4.8	.582[b]
Regensburg (SD)	23	8.0 (2.0)	6.5 (1.8)	3.7 (2.6)	3.8 (2.7)	72.1 (19.1)	66.7 (20.8)	68.2 (9.0)	71.4 (15.3)	66.7 (14.8)	39.3 (18.1)	2.16	4.8	.359[b]
Duesseldorf (SD)	39	2.0 (0.9)	2.0 (1.0)	0.9 (0.9)	0.9 (0.9)	75.0 (23.3)	75.0 (26.9)	80.0 (18.1)	75.0 (26.3)	66.7 (27.4)	50.0 (35.1)	4.00	16.0	.979[b]
Validation of the active rTMS predictor in sham-treated patients (active rTMS predictor trained in P-CV)	45	14	13	9	11	56.0	59.1	57.5	60.9	54.2	15.0	1.37	1.87	—
Active rTMS: PANSS-PS outcome predictor														
Pooled repeated nested cross-validation	44	3	15	14	12	20.0	51.7	35.9	17.6	55.6	-26.8	0.41	0.17	—
Sham rTMS: PANSS-NS outcome predictor														
Pooled repeated nested cross-validation	47	12	12	10	13	48.0	54.6	51.3	54.6	48.0	2.55	1.06	1.12	.441
Leave-site-out repeated cross-validation														
Full sample	47	12	8	14	13	48.0	36.4	42.2	46.2	38.1	-15.8	0.75	0.57	.913
Goettingen left-out	30	7	0	8	2	77.8	0.0	38.9	46.7	0.0	-53.3	0.78	0.60	.895
Regensburg left-out	25	4	4	5	9	44.4	30.8	37.6	44.4	30.8	-24.8	0.56	0.31	.955
Duesseldorf left-out	39	1	4	1	2	33.3	80.0	56.7	50.0	66.7	16.7	1.67	2.78	.299

Note: BAC, balanced accuracy; DOR, Diagnostic Odds Ratio; FN, false negatives; FP, false positives; LR+, Positive Likelihood ratio; NPV, Negative Predictive Value; PPV, Positive Predictive Value; PSI, Prognostic Summary Index (PSI = PPV + NPV − 100), *Sens* Sensitivity, *Spec* Specificity, *TN* true negatives, *TP* true positives; rTMS, repetitive transcranial magnetic stimulation. Positive/negative predictions refer to treatment nonresponse/response after 3 weeks of active or sham rTMS. Hence sensitivity measures the classifiers' capacity to correctly identify patients with nonresponse to the respective treatment as such. For both the active and sham rTMS predictor 2 types of validation analyses were carried out: Pooled Repeated Nested Cross-Validation (P-CV), and Leave-Site-Out Repeated Nested Validation (LSO-CV), which iteratively trained the respective predictor 2 of the 3 sites and then applied the trained predictor to the held-back site. In the active rTMS condition 3 further analyses were conducted: (1) analysis of performance non-inferiority of LSO-CV vs P-CV experiments comparing the effects of reduced training sample size vs site/scanner effects on classification performance, (2) sham rTMS validation analysis: the active rTMS predictor trained in the P-CV experiment was applied to the sham-treated patients to assess the specificity of the predictive models for the active treatment, and (3) a P-CV analysis of the single-patient predictability of a PANSS-PS outcome criterion.

[a]*P* value obtained from permutation analysis of P-CV/ LSO-CV predictors trained on respective original label distributions vs 1000 respective predictors trained on random label permutations.

[b]*P* value obtained from permutation analysis of LSO-CV predictors trained on respective original site label distributions vs 1000 respective predictors trained on random permutations of the patients' site labels. *$P < 0.05$.

Table 3. Sociodemographic and Clinical Differences at Baseline Between Patients Predicted With a Day 21-Response vs Nonresponse to Active rTMS

	Active rTMS (Predicted Nonresponse) (n = 21)		Active rTMS (Predicted Response) (n = 24)		Nonresponse vs Response		
					χ^2	*df*	*P*
Sociodemographic variables							
Gender (male: female)	19:2		20:4		0.495	1	.482
Site (Goettingen: Regensburg: Düsseldorf)	7:11:3		10:11:3		0.331	2	.848
Hand preference (right: non-right)	18:2		21:3		0.068	1	.795
	Mean	SD	Mean	SD	*t*		
Age (y)	35.81	11.72	32.42	7.97	1.148	43	.257
Education (y)	11.45	1.99	11.42	1.87	0.057	42	.955
rTMS functional and anatomical parameters							
Left Resting Motor Threshold (RMT)	48.86	9.79	44.68	10.60	1.340	41	.188
Scalp-to-cortex distance BA 9 (mm)	16.13	1.90	16.40	2.40	−0.416	43	.679
Scalp-to-cortex distance BA 46 (mm)	15.76	2.12	16.74	2.63	−1.359	43	.181
Illness severity and functioning at baseline							
PANSS negative symptoms	26.52	4.94	26.13	4.10	−0.297	43	.768
PANSS positive symptoms	13.95	4.15	14.83	4.53	−0.665	42	.510
PANSS general symptoms	43.24	10.22	41.57	8.86	0.582	42	.564
PANSS total	83.71	15.55	82.78	13.24	0.214	41	.831
Global assessment of functioning	49.40	13.93	54.50	10.30	−1.359	40	.182
Psychopharmacological treatment							
Antipsychotic dose (CPZ mg) at T0	629.83	398.99	574.16	495.60	0.388	39	.700
Cumulative antipsychotic dose (CPZ mg) between T0 and T1	13226.39	8378.73	12112.25	10426.69	0.369	39	.714
					χ^2	*df*	*P*
Treated with clozapine (yes:no)	2:19		6:18		1.835	1	.176
Treated with antidepressants (yes:no)	10:11		7:17		1.622	1	.233
Treated with mood stabilizers (yes:no)	3:18		1:23		1.416	1	.234
Depression severity	Mean	SD	Mean	SD	*t*		
Montgomery–Åsberg Depression Scale (MADRS)	14.95	5.88	14.08	5.22	0.518	42	.607

Note: PANSS, Positive and Negative Syndrome Scale; CPZ, Chlorpromazine Equivalents, BA, Brodmann Area; rTMS, repetitive transcranial magnetic stimulation. Sociodemographic, Clinical and Psychopharmacological Treatment Differences Were Assessed Using Independent *t* and Chi-square tests in SPSS 23.

between T0 and T1 (ρ = 0.583, $P < .0001$; figure 3A). Furthermore, the mixed-model analysis applied to the absolute PANSS-NS scores of this patient group showed a significant main effect of predicted outcome group on the PANSS-NS trajectories (F = 13.6, $P < .001$; figure 3B). The trajectories analyzed in the mixed model did not include the day 21 (T1) target of the neuroanatomical classifier nor its closest day 28 (T2) time point, but spanned T0, T3 and T4, which indicates that the response vs nonresponse stratification was stable up to 2 months after the completion of treatment. Additional mixed-model analyses showed that patients stratified into PANSS-NS responders vs nonresponders also differed significantly in their GAF trajectories (supplementary figure 4). Similar stratification effects were not observed for the respective PANSS-PS or MADRS courses (supplementary material).

Discussion

This is to our knowledge the first study reporting the successful application of MRI-based machine learning to the prediction of individual responses to the rTMS treatment of neuropsychiatric illness. The predictive

model provided accurate outcome estimates in 8.4 of 10 patients and the employed methods facilitated robust generalizability to new study sites despite the significantly lower training sample sizes in the leave-site-out analysis. Importantly, we did not find any sociodemographic, clinical, or psychopharmacological treatment parameters to confound the active rTMS outcome classifier's stratification effects. These effects were not limited to a short-term separation between negative symptom responders vs responders but extended for 3 months into the post-treatment naturalistic follow-up period. We also found that the PANSS-NS stratification effects generalized to the patients' global functioning trajectories, suggesting that relative prefronto-temporo-limbic GMD reductions prior to treatment start moderate both a subsequent improvement of negative symptoms and functioning induced by rTMS to the left DLPFC. This neuroanatomical baseline variance in patients with schizophrenia may represent an important, so far unknown biological factor driving the equivocal efficacy reported by previous rTMS studies in the field.[5]

Our validation analyses showed that the active rTMS outcome predictor was not only accurate but also therapeutically *specific* because the model did not distinguish *globally* between good vs poor symptom courses, as shown by our negative findings in the patients' psychotic and depressive symptoms domains. The low prediction performance of the active rTMS predictor in *sham*-treated patients further corroborated the detected structural brain pattern's high predictive specificity for verum effects on the negative symptom trajectories. Moreover, the contrast between the accurate PANSS-NS and the weak PANSS-PS outcome classifier supports the predictive specificity of the detected neuroanatomical marker for the left-hemispheric DLPFC stimulation and its downstream neural system-level effects. This interventional specificity agrees with 2 meta-analyses suggesting a superiority of (1) the left DLPFC as stimulation target for negative symptoms compared to other locations,[42] and (2) the left temporo-parietal junction over other stimulation targets for the rTMS treatment of psychotic symptoms, ie, auditory verbal hallucinations.[43] The former meta-analysis extracted further potential predictors of DLPFC-rTMS efficacy on negative symptoms: a baseline PANSS-NS ≥ 20, RMS stimulation intensity of 110%, rTMS frequency of 10 Hz and treatment duration of ≥3 weeks, as well as a duration of illness <8 years. Despite the implementation of all these recommendations in RESIS,[6] we did not find any superiority of active vs sham treatment on clinical and cognitive outcomes.[6,9] Our current findings suggest that these previous negative results may have been mediated by a neuroanatomically determined response variability, which obscured active treatment efficacy in the randomized clinical trial design. Hence, prospective clinical trials are needed to test whether matching of suitable patients to rTMS can enhance treatment outcomes based on pre-treatment MRI-based patient stratification.

A recent study investigated the predictive value of the neuroanatomical properties of the left temporal lobe on rTMS efficacy in treating refractory auditory verbal hallucinations. The authors reported that the Scalp-to-Cortex Distance (SCD) and local GMD of this structure were predictive of treatment response.[15] In contrast, we did not find any significant effects of the SCD over Brodmann Areas 9 and 46 on active rTMS treatment outcomes or on the structural adaptions of the left hippocampal-precuneal networks in our present or previous work (table 3). In turn, we explored whether changes in brain structure following active rTMS previously reported for a subset of 33 RESIS patients[20] could be predicted by our neuroanatomical model (supplementary results). In this subsample, we observed a significant association ($F = 4.09$, $P = .015$) between the patients' PANSS-NS outcome probabilities generated by our active rTMS predictor (figure 2) and subsequent left hippocampal volume increases. This finding suggests that the model's neuroanatomical decision rule, which involved reduced prefronto-limbic and cerebellar GMD as well as increased premotor-thalamic GMD at baseline, not only successfully stratified different PANSS-NS courses, but also predicted longitudinal brain changes in the limbic system induced by active rTMS applied to the left DLPFC. Still, the relationship between the predictor's mesoscopic[32] structural brain pattern and the underlying mechanistic surrogate remains unclear. Recent evidence suggested that the T1-weighted images analyzed here (and by the majority of structural neuroimaging studies) are significantly influenced by the microstructural properties of the brain tissue such as myelination, iron and water content.[44] Further imaging studies using parallel quantitative MRI protocols may disentangle the contributions of these histopathological factors and thus shed light on the structural mechanisms determining a patient's capacity to respond to rTMS.

Interestingly, early pre-post studies investigating rTMS-induced short-term effects on cerebral blood flow and oxygenation in healthy controls and patients with major depression revealed activation patterns that qualitatively overlapped with the prefrontal, temporal and thalamic areas of the neuroanatomical prediction pattern reported here.[45–47] Further system-level overlaps can be traced when comparing our pattern with a recent study assessing possible links between baseline resting-state fMRI connectivity, clinical phenotypes of depression and subsequent response to rTMS in a large sample of 1188 patients, of whom 154 were treated with rTMS.[24] This study detected 4 distinct depression biotypes that differentially predicted rTMS efficacy: Biotype 1 characterized by anhedonia, psychomotor retardation, and orbitofrontal as well as limbic hypoconnectivity showed the highest response likelihood to rTMS (82.5%) in contrast to biotype 4 (29.6%) which was characterized by hyperconnectivity in fronto-striatal

networks and higher anxiety scores. Evidence for an impaired anticorrelated coupling between the DLPFC-based Central Executive Network (CEN) and the medial prefrontal, frontopolar and medial parietal regions of the Default Mode Network (DMN) has been provided by studies investigating the neural correlates of rTMS efficacy in depression.[48,49] In fact, high-frequency rTMS may attenuate abnormally elevated within-DMN connectivity[50] and restore the anticorrelated activation patterns of DMN and CEN.[51] Recently, these findings have been extended to schizophrenia showing DMN suppression deficits in cognitively impaired vs non-impaired patients with first-episode psychosis[52] as well as altered connectivity patterns between different DMN subnetworks and the CEN.[53] Latter work also provided evidence that these alterations are associated with the severity of negative symptoms in schizophrenia and can be used to separate patients from healthy controls with >75% accuracy.[53] If we relate the pattern of DMN-related GMD reductions and motor-thalamic increments observed in our study to this previous evidence, we may speculate that rTMS responders have a more functionally hyperactive DMN at baseline and thus increased GMD in these regions compared to nonresponders who show extended structural deficits in DMN-related hubs, reducing their capacity for adaptive brain responses following rTMS.[20]

In the context of these hypotheses, the heterogeneity of neuroanatomical and clinical phenotypes underlying schizophrenia,[22,23] and their overlaps with early-onset depression,[33] point to complex biotypes conferring differential response likelihoods to noninvasive brain stimulation techniques through still unknown interactions between molecular, neuroanatomical, neurofunctional and neurophysiological factors. The present results linking baseline neuroanatomical variation in schizophrenia with adaptive behavioral and structural brain changes following rTMS suggest the value of future in-depth mechanistic investigations analyzing large-scale patient samples using combinations of multi-modal MRI, brain stimulation and multivariate analysis techniques.

The predictive performance of the active rTMS predictor reported here is comparable to those of a recent non-controlled open-label neuroimaging study predicting the outcome of depressive and schizophrenic patients following ECT.[25,26] However, to the best of our knowledge, our study is the first to demonstrate the feasibility of a reliable MRI-based prediction of rTMS-treatment outcomes based on the superior evidence levels provided by a state-of-the-art controlled RCT design. This design enabled us to benchmark the leave-site-out generalizability, the interventional (active vs sham) and outcome domain specificity, as well as the temporal stability of the model's predictions. The results obtained in these validation analyses suggest that the high response variability to noninvasive brain stimulation observed in many trials using rTMS for the treatment of depression[54,55] or schizophrenia[5] may be successfully mapped to pre-treatment structural brain patterns. So far, variability rates up to 50% in rTMS protocols have hampered the overall clinical efficacy of rTMS in research and clinical care, challenging the prediction of rTMS in any individual patient.[56] Our findings suggest that MRI-based machine learning derived from a careful RCT design may foster a better understanding of and the ability to predict how different disease-related brain phenotypes contribute to the individual patient's capacity to respond to brain stimulation. Further validation of the identified predictive pattern in larger, but currently unavailable external patient databases and future biomarker-stratified clinical trials[57] could be important next steps to reduce the response variability, increase the clinical efficacy and propel the availability of rTMS as a valid treatment option for patients with predominant negative schizophrenia.

Supplementary Material

Supplementary data is available at *Schizophrenia Bulletin* online.

Funding

The Repetitive Transcranial Magnetic Stimulation (rTMS) for the Treatment of Negative Symptoms in Schizophrenia (RESIS) trial was supported by the Deutsche Forschungsgemeinschaft Grant No. FA–210/1. The funding source was not involved in the study design, the collection and analysis of the data, or the writing of the manuscript.

Acknowledgments

N.K. has received paid speakership from Otsuka at 3 occasions. T.W. has received paid speakerships from Alpine Biomed, AstraZeneca, Bristol Myers Squibb, Eli Lilly, I3G, Janssen Cilag, Novartis, Lundbeck, Roche, Sanofi-Aventis, Otsuka and Pfizer, and has accepted travel or hospitality not related to a speaking engagement from AstraZeneca, Bristol-Myers-Squibb, Eli Lilly, Janssen Cilag, and Sanofi-Synthelabo; and has received restricted research grants from AstraZeneca, Cerbomed, I3G and AOK (health insurance company). B.L. received honoraria and speakers' fees from ANM, Astra Zeneca, Autifony, Lundbeck, Merz, Magventure, Novartis, Pfizer and Servier, research funding from the Tinnitus Research Initiative, the German Research Foundation, the German Bundesministerium für Bildung und Forschung, the American Tinnitus Association, Astra Zeneca and Cerbomed, funding for equipment from Magventure and Deymed and travel and accommodation payments from Lilly, Lundbeck, Servier and Pfizer. J.C. was a member of an advisory board of Roche, accepted travel or hospitality not related to a speaking engagement from

Servier, support for symposia from Inomed, Localite, Magventure, Roche, Mag & More, NeuroConn, Syneika, FBI Medizintechnik, Spitzer Arzneimittel and Diamedic. W.W. has received paid speakerships from Bristol-Myers Squibb, Essex Pharma, Janssen-Cilag, Lilly Deutschland, and Pfizer Neuroscience. He is a member of the Neuroscience Academy of Roche Pharma. W.G. has received symposia support from Janssen-Cilag GmbH, Neuss, Lilly Deutschland GmbH, Bad Homburg and Servier, Munich. He is a member of the Faculty of the Lundbeck International Neuroscience Foundation (LINF), Denmark. G.H. has received payments as speaker, consultant, author or for research funding during the last 5 years from Actelion, Affectis, Astra-Zeneca, Bayerische Motorenwerke, Bundesministerium für Bildung und Forschung, Bundesministerium für Strahlenschutz, Bristol-Meyers Squibb, Cephalon, Daimler Benz, Deutsche Forschungsgesellschaft, Elsevier, EuMeCom, Essex, Georg Thieme, Gerson Lerman Group Council of Healthcare Advisors, GlaxoSmithKline, Janssen-Cilag, Lilly, Lundbeck, Meda, Merck, Merz, Novartis, Pfizer, Proctor & Gamble, Sanofi-Aventis, Schering-Plough, Sepracor, Servier, Springer, Urban & Fischer, and Volkswagen. W.G.H. is an unpaid member of the Advisory Board of In Silico Biosciences, and a paid consultant to Otsuka/Lundbeck, Roche, Novartis, Eli Lilly, MDH Consulting, and the Canadian Agency on Drugs and Technology in Health. P.F. was honorary speaker for Janssen-Cilag, Astra-Zeneca, Eli Lilly, Bristol Myers-Squibb, Lundbeck, Pfizer, Bayer Vital, SmithKline Beecham, Wyeth, and Essex. During the last 5 years, but not presently, he was a member of the advisory boards of Janssen-Cilag, AstraZeneca, Eli Lilly, and Lundbeck. A.H. has been invited to scientific meetings by Lundbeck, Janssen-Cilag, and Pfizer, and he received a paid speakership from Desitin, Otsuka and Lundbeck. He was member of an advisory board of Roche. B.G., M.L., P.E., E.F., F.M., G.W., C.O., P.E.V., M.R., R.A., D.D., F.G., P.D., B.M., P.M.K., T.B.P., and T.S-A. had no conflict of interest.

References

1. Gersner R, Kravetz E, Feil J, Pell G, Zangen A. Long-term effects of repetitive transcranial magnetic stimulation on markers for neuroplasticity: differential outcomes in anesthetized and awake animals. *J Neurosci*. 2011;31:7521–7526.
2. Nardone R, Höller Y, Langthaler PB, et al. rTMS of the prefrontal cortex has analgesic effects on neuropathic pain in subjects with spinal cord injury. *Spinal Cord*. 2017;55:20–25.
3. Smith MC, Stinear CM. Transcranial magnetic stimulation (TMS) in stroke: Ready for clinical practice? *J Clin Neurosci*. 2016;31:10–14.
4. Blumberger DM, Maller JJ, Thomson L, et al. Unilateral and bilateral MRI-targeted repetitive transcranial magnetic stimulation for treatment-resistant depression: a randomized controlled study. *J Psychiatry Neurosci*. 2016;41:E58–E66.
5. Hasan A, Strube W, Palm U, Wobrock T. Repetitive noninvasive brain stimulation to modulate cognitive functions in schizophrenia: a systematic review of primary and secondary outcomes. *Schizophr Bull*. 2016;42(Suppl 1):S95–S109.
6. Wobrock T, Guse B, Cordes J, et al. Left prefrontal high-frequency repetitive transcranial magnetic stimulation for the treatment of schizophrenia with predominant negative symptoms: a sham-controlled, randomized multicenter trial. *Biol Psychiatry*. 2015;77:979–988.
7. Li Z, Yin M, Lyu XL, Zhang LL, Du XD, Hung GC. Delayed effect of repetitive transcranial magnetic stimulation (rTMS) on negative symptoms of schizophrenia: findings from a randomized controlled trial. *Psychiatry Res*. 2016;240:333–335.
8. Mogg A, Purvis R, Eranti S, et al. Repetitive transcranial magnetic stimulation for negative symptoms of schizophrenia: a randomized controlled pilot study. *Schizophr Res*. 2007;93:221–228.
9. Hasan A, Guse B, Cordes J, et al. Cognitive effects of high-frequency rTMS in schizophrenia patients with predominant negative symptoms: results from a multicenter randomized sham-controlled trial. *Schizophr Bull*. 2016;42:608–618.
10. Dlabac-de Lange JJ, Bais L, van Es FD, et al. Efficacy of bilateral repetitive transcranial magnetic stimulation for negative symptoms of schizophrenia: results of a multicenter double-blind randomized controlled trial. *Psychol Med*. 2015;45:1263–1275.
11. Hoffman RE, Boutros NN, Hu S, Berman RM, Krystal JH, Charney DS. Transcranial magnetic stimulation and auditory hallucinations in schizophrenia. *Lancet*. 2000;355:1073–1075.
12. Vercammen A, Knegtering H, Liemburg EJ, den Boer JA, Aleman A. Functional connectivity of the temporo-parietal region in schizophrenia: effects of rTMS treatment of auditory hallucinations. *J Psychiatr Res*. 2010;44:725–731.
13. Loo CK, Sainsbury K, Mitchell P, Hadzi-Pavlovic D, Sachdev PS. A sham-controlled trial of left and right temporal rTMS for the treatment of auditory hallucinations. *Psychol Med*. 2010;40:541–546.
14. Silverstein WK, Noda Y, Barr MS, et al. Neurobiological predictors of response to dorsolateral prefrontal cortex repetitive transcranial magnetic stimulation in depression: a systematic review. *Depress Anxiety*. 2015;32:871–891.
15. Nathou C, Simon G, Dollfus S, Etard O. Cortical anatomical variations and efficacy of rtms in the treatment of auditory hallucinations. *Brain Stimul*. 2015;8:1162–1167.
16. Hamada M, Murase N, Hasan A, Balaratnam M, Rothwell JC. The role of interneuron networks in driving human motor cortical plasticity. *Cereb Cortex*. 2013;23:1593–1605.
17. Nettekoven C, Volz LJ, Leimbach M, et al. Inter-individual variability in cortical excitability and motor network connectivity following multiple blocks of rTMS. *Neuroimage*. 2015;118:209–218.
18. Aguirre I, Carretero B, Ibarra O, et al. Age predicts low-frequency transcranial magnetic stimulation efficacy in major depression. *J Affect Disord*. 2011;130:466–469.
19. Falkai P, Rossner MJ, Schulze TG, et al. Kraepelin revisited: schizophrenia from degeneration to failed regeneration. *Mol Psychiatry*. 2015;20:671–676.
20. Hasan A, Wobrock T, Guse B, et al. Structural brain changes are associated with response of negative symptoms

to prefrontal repetitive transcranial magnetic stimulation in patients with schizophrenia. *Mol Psychiatry*. 2017;22:857–864.

21. Bleuler E. *Dementia Praecox Oder Gruppe Der Schizophrenien*. Leipzig, Germany: Deuticke; 1911.
22. Zhang T, Koutsouleris N, Meisenzahl E, Davatzikos C. Heterogeneity of structural brain changes in subtypes of schizophrenia revealed using magnetic resonance imaging pattern analysis. *Schizophr Bull*. 2015;41:74–84.
23. Clementz BA, Sweeney JA, Hamm JP, et al. Identification of distinct psychosis biotypes using brain-based biomarkers. *Am J Psychiatry*. 2016;173:373–384.
24. Drysdale AT, Grosenick L, Downar J, et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med*. 2017;23:28–38.
25. Redlich R, Opel N, Grotegerd D, et al. Prediction of individual response to electroconvulsive therapy via machine learning on structural magnetic resonance imaging data. *JAMA Psychiatry*. 2016;73:557–564.
26. Li P, Jing RX, Zhao RJ, et al. Electroconvulsive therapy-induced brain functional connectivity predicts therapeutic efficacy in patients with schizophrenia: a multivariate pattern recognition study. *NPJ Schizophr*. 2017;3:21.
27. Orrù G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A. Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci Biobehav Rev*. 2012;36:1140–1152.
28. Iniesta R, Stahl D, McGuffin P. Machine learning, statistical learning and the future of biological research in psychiatry. *Psychol Med*. 2016;46:2455–2465.
29. Hahn T, Nierenberg AA, Whitfield-Gabrieli S. Predictive analytics in mental health: applications, guidelines, challenges and perspectives. *Mol Psychiatry*. 2017;22:37–43.
30. Leucht S, Leucht C, Huhn M, et al. Sixty years of placebo-controlled antipsychotic drug trials in acute schizophrenia: systematic review, bayesian meta-analysis, and meta-regression of efficacy predictors [published online ahead of print May 25, 2017]. *Am J Psychiatry*. doi:10.1176/appi.ajp.2017.
31. Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage*. 2007;38:95–113.
32. Radua J, Canales-Rodríguez EJ, Pomarol-Clotet E, Salvador R. Validity of modulation and optimal settings for advanced voxel-based morphometry. *Neuroimage*. 2014;86:81–90.
33. Koutsouleris N, Meisenzahl EM, Borgwardt S, et al. Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers. *Brain*. 2015;138:2059–2073.
34. Koutsouleris N, Riecher-Rössler A, Meisenzahl EM, et al. Detecting the psychosis prodrome across high-risk populations using neuroanatomical biomarkers. *Schizophr Bull*. 2015;41:471–482.
35. Varoquaux G, Raamana PR, Engemann DA, Hoyos-Idrobo A, Schwartz Y, Thirion B. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. *Neuroimage*. 2017;145:166–179.
36. Dyrba M, Ewers M, Wegrzyn M, et al.; EDSD study group. Robust automated detection of microstructural white matter degeneration in Alzheimer's disease using machine learning classification of multicenter DTI data. *PLoS One*. 2013;8:e64925.
37. Hansen LK, Larsen J, Nielsen FA, et al. Generalizable patterns in neuroimaging: how many principal components? *Neuroimage*. 1999;9:534–544.
38. Chang C-C, Lin C-J. *LIBSVM: A Library for Support Vector Machines*. ACM Transactions on Intelligent Systems and Technology (TIST); 2011;2:27.
39. Vapnik VN. An overview of statistical learning theory. *IEEE Trans Neural Netw*. 1999;10:988–999.
40. Polikar R. Ensemble based systems in decision making. *IEEE Circuits and Systems*. 2006;6:21–45.
41. Golland P, Fischl B. Permutation tests for classification: towards statistical significance in image-based studies. *Inf Process Med Imaging*. 2003;18:330–341.
42. Shi C, Yu X, Cheung EF, Shum DH, Chan RC. Revisiting the therapeutic effect of rTMS on negative symptoms in schizophrenia: a meta-analysis. *Psychiatry Res*. 2014;215:505–513.
43. Slotema CW, Blom JD, van Lutterveld R, Hoek HW, Sommer IE. Review of the efficacy of transcranial magnetic stimulation for auditory verbal hallucinations. *Biol Psychiatry*. 2014;76:101–110.
44. Lorio S, Kherif F, Ruef A, et al. Neurobiological origin of spurious brain morphological changes: a quantitative MRI study. *Hum Brain Mapp*. 2016;37:1801–1815.
45. Speer AM, Kimbrell TA, Wassermann EM, et al. Opposite effects of high and low frequency rTMS on regional brain activity in depressed patients. *Biol Psychiatry*. 2000;48:1133–1141.
46. Nahas Z, Lomarev M, Roberts DR, et al. Unilateral left prefrontal transcranial magnetic stimulation (TMS) produces intensity-dependent bilateral effects as measured by interleaved BOLD fMRI. *Biol Psychiatry*. 2001;50:712–720.
47. Speer AM, Willis MW, Herscovitch P, et al. Intensity-dependent regional cerebral blood flow during 1-Hz repetitive transcranial magnetic stimulation (rTMS) in healthy volunteers studied with H215O positron emission tomography: I. Effects of primary motor cortex rTMS. *Biol Psychiatry*. 2003;54:818–825.
48. Fox MD, Buckner RL, White MP, Greicius MD, Pascual-Leone A. Efficacy of transcranial magnetic stimulation targets for depression is related to intrinsic functional connectivity with the subgenual cingulate. *Biol Psychiatry*. 2012;72:595–603.
49. Fox MD, Liu H, Pascual-Leone A. Identification of reproducible individualized targets for treatment of depression with TMS based on intrinsic connectivity. *Neuroimage*. 2013;66:151–160.
50. Eldaief MC, Halko MA, Buckner RL, Pascual-Leone A. Transcranial magnetic stimulation modulates the brain's intrinsic activity in a frequency-dependent manner. *Proc Natl Acad Sci USA*. 2011;108:21229–21234.
51. Liston C, Chen AC, Zebley BD, et al. Default mode network mechanisms of transcranial magnetic stimulation in depression. *Biol Psychiatry*. 2014;76:517–526.
52. Zhou L, Pu W, Wang J, et al. Inefficient DMN suppression in schizophrenia patients with impaired cognitive function but not patients with preserved cognitive function. *Scientific Reports*. 2016;6:21657. doi:10.1038/srep21657.
53. Wang H, Zeng LL, Chen Y, Yin H, Tan Q, Hu D. Evidence of a dissociation pattern in default mode subnetwork functional connectivity in schizophrenia. *Sci Rep*. 2015;5:14655.

54. Brunoni AR, Chaimani A, Moffa AH, et al. Repetitive transcranial magnetic stimulation for the acute treatment of major depressive episodes: a systematic review with network meta-analysis. *JAMA Psychiatry* 2017;74: 143–152.
55. Gaynes BN, Lloyd SW, Lux L, et al. Repetitive transcranial magnetic stimulation for treatment-resistant depression: a systematic review and meta-analysis. *J Clin Psychiatry*. 2014;75:477–489; quiz 489.
56. Hamada M, Rothwell JC. Neurophysiology of rTMS: important caveats when interpreting the results of therapeutic interventions. In: Platz T, ed. *TherapeuticrTMS Neurology*. Cham, Heidelberg, New York, Dordrecht, London: Springer International Publishing; 2016:1–10.
57. Dunn G, Emsley R, Liu H, Landau S. Integrating biomarker information within trials to evaluate treatment mechanisms and efficacy for personalised medicine. *Clin Trials*. 2013;10:709–719.