Received: 31 October 2022 | Revised: 29 May 2023 | Accepted: 4 June 2023

DOI: 10.1002/hbm.26405

RESEARCH ARTICLE

WILEY

Widespread intra-axonal signal fraction abnormalities in bipolar disorder from multicompartment diffusion MRI: Sensitivity to diagnosis, association with clinical features and pharmacologic treatment

Erick Jorge Canales-Rodríguez[1,2,3] | Norma Verdolini[1,2,4,5] | Silvia Alonso-Lana[1,2,6] | María Llanos Torres[7] | Francesco Panicalli[8] | Isabel Argila-Plaza[1] | Elena Rodriguez-Cano[1,2] | Irene Montoro[9,10,11] | Beatriz Garcia-Ruiz[9,10,11] | Esther Jimenez[2,4] | Cristina Varo[4] | Anna Lluch[10,11] | Caterina del Mar Bonnin[2,4,12,13] | Silvana Maluf[14] | Marc Pujol[15] | Núria Jaurrieta Guarner[14] | Salvador Sarró[1,2] | Eduard Vieta[2,4] | Elisabet Vilella[2,9,10,11] | Raymond Salvador[1,2] | Edith Pomarol-Clotet[1,2]

[1]FIDMAG Research Foundation, Germanes Hospitalàries, Sant Boi de Llobregat, Barcelona, Spain

[2]Mental Health Research Networking Center (CIBERSAM), Madrid, Spain

[3]Signal Processing Laboratory (LTS5), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

[4]Bipolar and Depressive Disorders Unit, Hospital Clinic, Institute of Neurosciences, University of Barcelona, IDIBAPS, Barcelona, Spain

[5]Local Health Unit Umbria 1, Department of Mental Health, Mental Health Center of Perugia, Perugia, Italy

[6]Research Center and Memory Clinic, Fundació ACE, Institut Català de Neurociències Aplicades, Universitat Internacional de Catalunya, Barcelona, Spain

[7]CSMiA de Gràcia. Hospital Mare de Déu de la Mercè, Barcelona, Spain

[8]Benito Menni CASM, Sant Boi de Llobregat, Barcelona, Spain

[9]Hospital Universitari Institut Pere Mata, Reus, Spain

[10]Institut d'Investigació Sanitària Pere Virgili (IISPV), Reus, Spain

[11]Universitat Rovira i Virgili (URV), Reus, Spain

[12]Psychiatry Department, Institut d'Investigació Biomèdica-Sant Pau (IIB-SANT PAU), Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

[13]Department of Psychiatry and Forensic Medicine, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain

[14]Unidad de Agudos, Hospital de Sagrat Cor de Martorell, Barcelona, Spain

[15]Department of Radiology, Hospital de Sant Rafael, Barcelona, Spain

Correspondence
Erick Jorge Canales-Rodríguez, FIDMAG Signal Processing Laboratory (LTS5), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.
Email: ejcanalesr@gmail.com; erick.canalesrodriguez@epfl.ch

Abstract

Despite diffusion tensor imaging (DTI) evidence for widespread fractional anisotropy (*FA*) reductions in the brain white matter of patients with bipolar disorder, questions remain regarding the specificity and sensitivity of *FA* abnormalities as opposed to other diffusion metrics in the disorder. We conducted a whole-brain voxel-based

Raymond Salvador and Edith Pomarol-Clotet are joint last author (equal contribution, shared senior authorship).

Funding information
"la Caixa" Foundation, Grant/Award Numbers: 100010434, LCF/PR/GN18/50310006; Departament d'Innovació, Universitats i Empresa, Generalitat de Catalunya, Grant/Award Numbers: 2014SGR1573, 2014SGR398, 2017SGR1365; H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: 754550; Instituto de Salud Carlos III, Grant/Award Numbers: CD18/00029, CM17/00258, CPII13/00018, CPII16/00018, MV18/00054, PI14/01148, PI15/00277, PI15/00283, PI15/00852, PI18/00805, PI18/00810, PI18/00877, PI21/00525; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: PZ00P2_185814

multicompartment diffusion MRI study on 316 participants (i.e., 158 patients and 158 matched healthy controls) employing four diffusion metrics: the mean diffusivity (*MD*) and *FA* estimated from DTI, and the intra-axonal signal fraction (*IASF*) and microscopic axonal parallel diffusivity (*Dpar*) derived from the spherical mean technique. Our findings provide novel evidence about widespread abnormalities in other diffusion metrics in BD. An extensive overlap between the *FA* and *IASF* results suggests that the lower *FA* in patients may be caused by a reduced intra-axonal volume fraction or a higher macromolecular content in the intra-axonal water. We also found a diffuse alteration in *MD* involving white and grey matter tissue and more localised changes in *Dpar*. A Machine Learning analysis revealed that *FA*, followed by *IASF*, were the most helpful metric for the automatic diagnosis of BD patients, reaching an accuracy of 72%. Number of mood episodes, age of onset/duration of illness, psychotic symptoms, and current treatment with lithium, antipsychotics, antidepressants, and antiepileptics were all significantly associated with microstructure abnormalities. Lithium treatment was associated with less microstructure abnormality.

KEYWORDS
bipolar disorder, brain, diffusion tensor imaging, spherical mean technique, tissue microstructure

1 | INTRODUCTION

Bipolar disorder (BD) is a chronic and recurrent mental health condition affecting >1% of people worldwide (Vieta et al., 2018). There are various BD subtypes characterised by episodes of mania or hypomania alternated with depressive episodes, causing unpredictable changes in mood and behaviour. Although there are effective medicines for treating the acute phase of the illness and preventing relapse, much can be done to improve the lives of people suffering from this condition. Early and accurate diagnosis is difficult due to heterogenous clinical presentation and complex underlying factors determining onset. As inappropriate interventions may destabilise the course and outcome of the disease (Berk et al., 2006), identifying specific biomarkers for BD is crucial to enhance the proper early treatment of patients (Vieta et al., 2018) and monitoring the effects of both conventional and new personalised treatments.

Among many research lines on the neuroanatomical substrates of BD, magnetic resonance imaging (MRI) has been one of the most productive. The ENIGMA consortium has conducted large-scale collaborative analyses of structural MRI data involving thousands of patients and healthy controls (HC; Thompson et al., 2014). These mega- and meta-analyses have found widespread abnormalities in patients with BD, revealing previously undetected associations (Ching et al., 2020). Thus, cortical grey matter was reported to be bilaterally thinner in BD ($N = 6503$ individuals), especially in the frontal, temporal, and parietal regions (Hibar et al., 2018). This pattern of reduced cortical thickness was associated with a longer duration of illness, even after controlling for age. Commonly prescribed medications like lithium, antipsychotics, and antiepileptic drugs were significantly associated with cortical thickness and surface area (Hibar et al., 2018). Another study ($N = 4698$ individuals) using hippocampal subfield volumetry found smaller volumes in 9 out of 12 anatomically and functionally distinct hippocampal subfields potentially involved in the pathophysiology of BD (Haukvik et al., 2020). Interestingly, while lithium users did not show volume differences compared with HC, non-users did. Additionally, antiepileptic and antipsychotic treatment was associated with reduced volumes (Haukvik et al., 2020).

Given the established abnormalities of cortical (Hibar et al., 2018) and subcortical brain volumes in BD (Hibar et al., 2016), the question of whether these abnormalities are present in the white matter (WM) microstructure has been addressed by various research groups (e.g., see Canales-Rodríguez et al., 2013; Vederine et al., 2011). Recently, in the largest diffusion tensor imaging (DTI) study of BD ($N = 3033$ individuals) conducted by the ENIGMA working group (Favre et al., 2019), widespread fractional anisotropy (*FA*) reductions were found—within and beyond fronto-limbic regions—across 29 out of 44 WM regions of interest. The strongest effects were located in the corpus callosum and cingulum. A shorter illness duration, late onset, and lithium treatment were associated with higher *FA*. In contrast, no significant *FA* differences were associated with illness severity, antidepressant treatment, history of psychotic symptoms, or between BD-I and BD-II subtypes (Favre et al., 2019). Even though this multicenter study offers several advantages compared with previous smaller-scale DTI studies, some open questions remain regarding the sensitivity and specificity of *FA*: Is *FA* the most sensitive diffusion metric in BD? How can we explain the observed changes in terms of biological parameters?

DTI models each voxel as a single homogeneous compartment. However, brain WM contains at least two distinct water compartment types that can be detected—the intra- and extra-axonal spaces—

differentially contributing to the observed diffusion MRI signal. Consequently, DTI-derived metrics like *FA* are sensitive but not specific to changes in tissue microstructure parameters such as the volume fractions of the intra- and extra-axonal spaces and fibre dispersion (Beaulieu, 2002). In recent years, various multicompartment diffusion MRI models have been proposed to isolate the signal contribution from different compartments, providing more specific diffusion-based features. Among these advanced methods, the spherical mean technique (SMT; Kaden et al., 2016) has been of notable interest since, in addition to modelling the diffusion signals from the intra- and extra-axonal spaces, it provides metrics unconfounded by fibre crossings and orientation dispersion, which can be estimated from multishell diffusion MRI data acquired in clinical scanners. Furthermore, the SMT is more suitable than Neurite Orientation Dispersion and Density Imaging method (NODDI; Zhang et al., 2012), another similar multicompartment model, as the diffusivity parallel to the axons is estimated from the data and does not need to be fixed by hand during the fitting (Kaden et al., 2016).

For the first time, here we apply a whole-brain voxel-based approach combining DTI and SMT diffusion-derived metrics to reveal specific brain tissue microstructure abnormalities in BD. We also replicate the *FA*-based analyses previously reported by the ENIGMA multicenter study (Favre et al., 2019) on a homogeneous and considerably large sample of participants ($n = 316$) scanned in the same MRI machine. After comparing the SMT-based features between HCs and patients with BD, we also evaluate the sensitivity of various diffusion metrics for the automatic diagnosis of BD through machine learning algorithms and cross-validation. Finally, we look for differences between BD subtypes and the possible effect of duration of illness, age at onset, history of psychosis, and current pharmacological treatment with lithium, antidepressant, antipsychotic, and antiepileptic drugs.

2 | METHODS

2.1 | Sample

A total of 316 Caucasian participants consisting of 158 euthymic patients with BD and 158 HC matching on age and estimated Intelligence Quotient (IQ) were recruited. Exclusion criteria for all participants included left-handedness, personal history of neurological disease, brain injury, presence of cardiovascular risk factors for cognitive impairment (i.e., ictus), suffering from a severe medical condition (e.g., cancer), alcohol/substance abuse/dependence in the 12 months before participation, claustrophobia, presence of contraindications for MRI scanning (i.e., cardiac pacemaker, metal fragment/prostheses, pregnancy, surgical intervention within 1 month before scanning) and electroconvulsive therapy treatment in the previous 12 months before recruitment. Patients with a first episode of the disease were excluded, as well as those with an onset later than 55 years old. Additional exclusion criteria for the HC included any personal or first-degree relative history of a psychiatric disorder and current or previous psychopharmacological treatment prescription. All participants were required to be aged between 18 and 60 and to have an estimated intelligence quotient (IQ) >85, which was assessed using the Word Accentuation Test (Test de Acentuación de Palabras, TAP; Del Ser et al., 1997), a Spanish word reading test conceptually similar to the British National Adult Reading Test (Nelson & Willison, 1991) and the American Wide Range of Achievement Test (Wilkinson, 1993). The TAP is based on the pronunciation of low-frequency words whose accents have been removed, it has been standardised to the WAIS-III, and its scores can be converted into full-scale IQ estimates. The normal range is defined for TAP > 14 (Gomar et al., 2011).

Exhaustive information on clinical features and current pharmacologic treatments was available for 138 patients. Medication intake was registered at the time of scanning, as data on past use was not available. Half of the patients ($n = 69$) were on treatment with lithium, $n = 77$ were taking antipsychotics, $n = 57$ were taking antidepressants, and $n = 81$ were on treatment with antiepileptic drugs. $n = 101$ had a diagnosis of BD-I, and the remaining $n = 37$ of BD-II. $n = 84$ had a history of psychosis. This subsample of 138 patients was used to find associations between diffusion metrics and the clinical and medication variables. Other variables of interest were the number of episodes, age, sex, duration of illness, and age of onset, defined as the age of the first mood episode. The number of mood episodes was based on hospitalisation reports and self-reports when admission was not required. Since it was not always possible to determine the exact number of mood episodes, we employed a three-level scale (i.e., 1, 2, 3) with patients that had suffered <5, between 5 and 10, and with >10 episodes, respectively.

All participants were recruited by three institutions ($n = 100$ from FIDMAG Research Foundation, $n = 129$ from Hospital Universitari Institut Pere Mata, and $n = 87$ from Hospital Clínic of Barcelona) and were scanned using the same 3T MRI machine and acquisition sequences. The study was carried out following the latest version of the Declaration of Helsinki and was reviewed by the ethical committee of the three institutions. Written informed consent was obtained from all participants.

2.2 | MRI data acquisition and preprocessing

Multishell diffusion MRI data were acquired for each participant using a 3T Ingenia CX scanner (Philips Medical Systems, software version: 5.4) located at the Barcelonaβeta Brain Research Center (Pasqual Maragall Foundation, Barcelona, Spain) with a standard 32-channel head coil and the following sequence parameters: Field-of-view = 230 × 230 mm; voxel-size = 2.05 × 2.05 mm^2; repetition time (TR) = 10.1 s; echo time (TE) = 103 ms; flip angle = 90°; number-of-slices = 64; slice-thickness = 2.1 mm; number of averages = 1; acceleration factor = 2; number of shells = 3; *b*-values = 625, 1250, and 2500 s/mm^2; number of diffusion gradient directions = 100; number of b0 (i.e., *b*-value = 0) images = 7, plus 1 b0 with reverse phase to correct for spatial

distortions. For each subject, the diffusion MRI data were corrected to remove noise (Veraart et al., 2016) and Gibbs ringing artefacts (Kellner et al., 2016) by using the MRtrix3 software (https://www.mrtrix.org/; Tournier et al., 2019), as well as susceptibility, eddy-current, and head motion distortions using the topup (Andersson et al., 2003) and eddy (Andersson & Sotiropoulos, 2016) toolboxes in FSL (https://fsl.fmrib.ox.ac.uk/fsl/; Smith et al., 2004).

2.3 | Models and descriptors

Two data modelling techniques were applied to assess different properties of the self-diffusion process of water in brain tissues: (1) DTI (Basser & Pierpaoli, 1996) and (2) SMT (Kaden et al., 2016). The *FA* and mean diffusivity (*MD*) metrics were obtained from the diffusion tensors, estimated using the *dtifit* program included in FSL, and employing the diffusion volumes with *b*-values = 1250 s/mm^2. On the other hand, the intra-axonal signal fraction (*IASF*) and the microscopic diffusivity parallel to the axons within the intra- and extracellular spaces (denoted by *Dpar*) were computed from all the diffusion volumes using the SMT software (https://github.com/ekaden/smt; Kaden et al., 2016). Each of these four rotationally-invariant scalar contrasts reveals complementary information about the local diffusion process. While *FA* is a metric designed to quantify the anisotropy of Gaussian diffusion processes in regions of parallel fibres, *MD* quantifies the mean magnitude of diffusion within the voxel and is sensitive to the overall presence of obstacles to diffusion, regardless of orientation. *IASF* is a proxy for the (%) diffusion signal arising from neurites and axons, and hence, it quantifies the *apparent* volume fraction occupied by axons within the voxel. Finally, *Dpar* is the mean parallel diffusivity inside the axons and their extra-axonal surroundings (Kaden et al., 2016).

2.4 | Multimodal spatial normalisation

After skull removal, all brain images were normalised using the high-resolution nonlinear registration Symmetric Normalization (SyN) algorithm (Avants et al., 2008). All images were simultaneously used for coregistration, producing multimodal study-specific templates that were later applied in the final normalisation. A single deformation field was estimated for each participant by considering all the estimated metrics. Then, this deformation field was used to register each image to the template space, as described in Canales-Rodríguez et al. (2013). Consequently, our results are not affected by metric-specific normalisation bias (i.e., miss-registration among different metrics) because all maps from each participant were normalised using the same deformation field. The analyses included GM and WM tissue for all metrics except *FA*, which was restricted to a brain mask with *FA* > 0.2. Before running the statistical analyses, all images were spatially filtered with a Gaussian kernel of $\sigma = 2$ mm.

2.5 | Statistical analyses

Differences in demographic characteristics among the groups were examined using χ^2 tests (categorical variables) and ANOVAs (continuous variables).

Our primary focus was to examine group differences between patients with BD and HC for each diffusion metric. To this end, we created a general linear model (GLM) where a binary indicator of diagnosis (i.e., 0 = controls, 1 = patients) was the predictor of interest, and age and sex were included as covariates.

We further examined patient-specific clinical characteristics, including pharmacological treatment (i.e., lithium, antidepressants, antipsychotics, and antiepileptic drugs), diagnosis subtype (i.e., BD-I and BD-II), history of psychosis, age of onset, number of episodes, and duration of illness. For each pharmacological variable, a binary indicator was created where a given subject was either 1 (i.e., receiving the medication) or 0 (i.e., not receiving the medication). Likewise, diagnosis subtype and history of psychosis were modelled as dichotomous variables (i.e., BD-I/BD-II = 1/0 and yes/no = 1/0, respectively).

On the other hand, the age of onset, duration of illness and the number of episodes were treated as continuous variables, and their GLMs were built to identify linear associations (i.e., positive or negative correlations) with the diffusion metrics. The number of episodes was recorded as an ordinal variable with three scales (i.e., 1, 2, 3), and linearity was assumed in the models as if it were a continuous variable. Age and sex were included as covariates in all the analyses (Favre et al., 2019) to obtain results not confounded by inter-group differences in these variables and to separate the global effect of age on the duration of illness, age of onset and the number of episodes.

All statistical analyses were carried out by independently fitting the described GLMs for each brain voxel of each diffusion metric and applying non-parametric permutation tests as implemented in the 'randomise' FSL tool (Jenkinson et al., 2012; Smith et al., 2004). The Threshold-Free Cluster Enhancement (TFCE) method (Smith & Nichols, 2009) was employed using 5000 permutations. Only significant results at $p = .05$, corrected for multiple comparisons, are reported. Anatomical locations of the significant WM and GM regions were reported considering the following atlases included in FSL (Mori et al., 2008; Oishi et al., 2008): 'JHU ICBM-DTI-81 White-Matter Labels'; 'JHU White-Matter Tractography Atlas'; 'MNI Structural Atlas'; 'Harvard-Oxford Cortical Structural Atlas'; 'Harvard-Oxford Subcortical Structural Atlas'; 'Cerebellar Atlas in MNI152 space after normalisation with FNIRT', as well as the Anatomical Automatic Labelling atlas (Tzourio-Mazoyer et al., 2002) included in the MRIcron software.

Finally, a post hoc analysis was performed to assess the extent to which the findings from the previous analyses overlap and to explore whether the effects of pharmacologic treatment and clinical features are associated with differences between HC and BD patients. First, for each diffusion metric, we identified the clusters showing significant results in the inter-group comparison (HC vs BD) and the analyses testing for the effects of medication and clinical variables in the BD group. Then, we computed the overlap between the significant

clusters to create regions of interest (ROIs). Only relevant ROIs with more than 50 voxels were considered. These steps were repeated for each of the medication and clinical variables. Subsequently, we calculated the mean value per ROI for each healthy control included in the HC versus BD analysis, and each patient with BD included in the analyses testing for medication and clinical effects, respectively. Finally, we tested for statistically significant differences between each pair of groups at the ROI level: HC versus BD-yes, HC versus BD-no, and BD-yes versus BD-no, where BD-yes and BD-no indicate the two BD subgroups taking and not taking medication. For the analyses employing the variables age of onset and duration of illness, we report their linear associations with the diffusion metrics, in line with their previous statistical designs, respectively.

2.6 | Machine learning algorithms

To assess the efficacy of each diffusion metric for diagnostic prediction in BD, we evaluated the discriminative power of four classifiers: Ridge and Lasso regression, the Gaussian Process Classifier (GPC), and the Random Forest classifier, which were selected for their computational efficiency and excellent performance in a previous study with similar data (Salvador et al., 2017). All possible combinations of algorithms and diffusion metrics (i.e., $n = 4 \times 4$) were considered in pairwise classifications of the matched samples of healthy control participants and patients with BD. Specifically, the libraries *kernlab*, *glmnet*, and *randomForest* available for the R statistical software (https://www.r-project.org/) were used for the machine learning analyses. To have a non-biased performance assessment, we applied the classifiers by implementing a stratified 10-fold cross-validation procedure, as described by Nunes et al. (2020) and Salvador et al. (2017). The optimal regularization values for the Ridge and Lasso classifiers were selected through internal cross-validation within the training samples. For the GPC and the Random Forest classifiers, default hyperparameter values were used (apart from the number of trees in the Random Forest, which was set to 1000 as in two previous studies [Salvador et al., 2017, 2019]). A simple bootstrap procedure was applied to the individual test predictions to obtain the 95% confidence intervals for the accuracies.

Finally, we also built algorithms that combined information from all four measures (FA, IASF, Dpar, and MD) in a single classifier. Specifically, we applied the three approaches that had provided the highest accuracies in a multimodal study previously carried out in a sample of patients with schizophrenia (Salvador et al., 2019). On the one hand, we made simple classification updates by considering the average of the four prediction probabilities provided by the four measures. Alternatively, the highest of the four probabilities (in absolute value) was evaluated for prediction. Finally, a two-step Ridge classifier was also applied, as described in Salvador et al. (2019), based on selecting the 20% most relevant voxels from each measure map and then carrying out a second fitting merging the voxels chosen from the four modalities.

3 | RESULTS

3.1 | Demographic and clinical information

Demographic data for patients ($n = 158$) and HC ($n = 158$) are shown in Table 1. The groups were matched for age, sex, and estimated premorbid IQ (TAP score).

Table 2 shows the patient-specific clinical characteristics for the subgroup of patients ($n = 138$) without missing data for any of the variables reported. The subsequent statistical analyses, undertaken to examine associations among the diffusion metrics and the clinical characteristics, are based on this subsample of patients. There were no inter-group differences in age and sex between patients taking lithium or antipsychotics and those not taking them, and between patients with a different diagnosis subtype. Conversely, there were sex differences (only at the uncorrected level) between patients taking antidepressants, antiepileptic drugs, and those with a history of psychosis. The age of patients taking antidepressants was significantly higher, corrected for multiple comparisons, than those not taking them.

Table S1 reports the linear correlation among the demographic and clinical characteristics of patients with BD listed in Tables 1 and 2. From the resulting 28 pairwise correlations, four were significant, including the associations between psychotic symptoms and diagnosis subtype (positive correlation), lithium and antiepileptic users (negative correlation), antidepressant users and duration of illness (positive correlation), and duration of illness and age of onset (negative correlation).

3.2 | Inter-group comparison

Statistical analyses revealed a bilateral and widespread pattern of lower *FA* and *IASF* and higher *MD* values in patients compared with HCs. Moreover, a more restricted pattern of higher *Dpar* was also found (see Figure 1). Most significant *FA* differences were located in the body of the corpus callosum (CC), cerebral peduncle, WM regions adjacent to temporal and parietal lobes, corticospinal tract, and inferior longitudinal fasciculus. Further information on these results is provided in Table S2 of the SM. The main peaks of the *MD* statistical map were located in the genu and body of the CC, superior and posterior corona radiata, as well as the insula, Heschl's gyrus, cerebellum, superior frontal orbital cortex, supramarginal gyrus, and middle temporal gyrus (see Table S3, Figure 1). The most significant WM reductions in *IASF* were observed in the body and splenium of the CC, cingulum bundle, and medial lemniscus. In GM, the strongest effects were found in the cerebellum, inferior temporal gyrus, lateral occipital cortex, frontal pars triangularis, and medial frontal gyrus (see Table S4, Figure 1). Finally, increased *Dpar* was found in nine clusters located in the GM, with main peaks in the frontal lobe, cingulum, olfactory, calcarine, insula, Heschl's gyrus, rolandic operculum, temporal and occipital lobes, fusiform, and lingual gyrus (see Table S5).

TABLE 1 Demographic characteristics of bipolar and control participants.

	Patients ($n = 158$)	Controls ($n = 158$)	p-value
Age (years)	44.8 ± 9.2	44.1 ± 9.1	.45
Sex (male/female)	62/96	63/95	.91
Premorbid IQ (TAP)	24.4 ± 3.7	24.7 ± 3.1	.49
Age of onset (years)	25.9 ± 8.8		
Duration of illness (years)	18.8 ± 9.5		

Note: Values are reported as means and standard deviations. The whole sample was composed of individuals of Caucasian ancestry.

TABLE 2 Clinical characteristics of patients with bipolar disorder without missing data ($n = 138$).

	Yes	No	p-value
Lithium (*n*)	69	69	
Age (years)	44.6 ± 9.1	45.3 ± 8.6	.61
Sex: n (% female)	40 (58%)	46 (67%)	.29
Antidepressants (*n*)	57	81	
Age (years)	48.1 ± 7.3	42.7 ± 9.1	**<.001***
Sex: *n* (% female)	42 (74%)	44 (54%)	**.021**
Antipsychotics (*n*)	77	61	
Age (years)	44.7 ± 8.8	45.2 ± 8.9	.71
Sex: *n* (% female)	49 (64%)	37 (61%)	.72
Antiepileptics (*n*)	81	57	
Age (years)	45.3 ± 8.6	44.4 ± 9.2	.58
Sex: *n* (% female)	56 (69%)	30 (53%)	**.049**
History of psychosis (*n*)	84	54	
Age (years)	44.2 ± 9.3	46.1 ± 8.0	.2
Sex: *n* (% female)	46 (55%)	40 (74%)	**.022**
Diagnosis subtype (*n*)	107 (BD-I)	31 (BD-II)	
Age (years)	44.4 ± 9.1	46.8 ± 7.7	.19
Sex: *n* (% female)	64 (60%)	22 (71%)	.26

Note: Values are given as means and standard deviations. Uncorrected p-values ≤.05 are in bold, and Bonferroni corrected p-values ≤0.05 are indicated by asterisks.

3.3 | Machine learning: Diagnosis of BD

Levels of accuracy extracted from the validation subsets, provided by the four algorithms applied to the four individual metrics, are shown in Figure 2 and Table 3. In all algorithm-metric combinations, similar levels of accuracy, well above the 0.5 value expected by random assignment, were achieved. In all cases, estimated accuracies were above 0.6, and in all four algorithms, *FA* provided the best classifications, reaching an absolute maximum of 0.72 (95% CI: 0.67–0.77) with the Lasso. This value was significantly higher than the accuracies provided by the other three metrics with the Lasso (*FA* vs *MD*: McNemar's $\chi^2 = 7.93$, $df = 1$, p-value = .005; *FA* vs *IASF*: McNemar's $\chi^2 = 4.69$, $df = 1$, p-value = .030; *FA* vs *Dpar*: McNemar's $\chi^2 = 7.38$, $df = 1$, p-value = .007). No statistically significant differences between metrics were observed for the other three algorithms.

We report two supplementary analyses in the SM to understand how the four diffusion metrics correlate and the degree to which they are distinct. In Figure S1, we created scatter plots for all combinations of the four diffusion metrics. The results indicate that the metrics are not linearly related overall, and they carry common and complementary information about the brain tissue microstructure. On the other hand, in Figure S2, we report levels of redundancy in terms of shared predictive information in the machine learning analysis between pairs of diffusion metrics. Although there was some redundancy in the information provided by the different measures, they all contained relevant predictive information. Still, the three multimodal algorithms that combined information from the four diffusion measures did not significantly improve the accuracy levels given by the unimodal classifiers. Specifically, while both average and maximal probability classifiers did not reach the 72% accuracy obtained by the Lasso with the FA, the two-step Ridge only brought a marginal improvement up to 73% (0.731; 95% bootstrap confidence interval: 0.684–0.778).

3.4 | Associations with pharmacologic treatment

Within the BD subgroup of 138 subjects (see Section 3.1), we found a significantly higher *FA* and *IASF* and a reduced *MD* in patients receiving versus not receiving lithium (Figure 3a). The most significant *FA* results were observed in three clusters located in the thalamus, brainstem, and superior longitudinal fasciculus. The *IASF* analysis revealed increased values in four clusters reaching the calcarine, cingulum gyrus, and cingulum bundle. A more extended pattern of significant statistical differences was obtained for the *MD* metric. Nine clusters with main peaks at the calcarine, postcentral gyrus, occipital cortex, insula, frontal pars triangularis, parietal, fusiform, and precuneus were detected (Table S6).

In contrast, we found a significantly lower *IASF* and higher *MD* in patients taking antidepressants (Figure 3b). While the *IASF* results were located in one cluster in the right hemisphere, reaching the temporal fusiform cortex and the inferior longitudinal fasciculus, a widespread bilateral pattern of *MD* differences was obtained. Eight significant clusters involving the following main regions were detected: the body of CC, cingulum bundle, WM adjacent to precentral gyrus, occipital cortex, lingual, postcentral, gyrus rectus, temporal, inferior frontal orbital, and supramarginal cortices (Table S7).

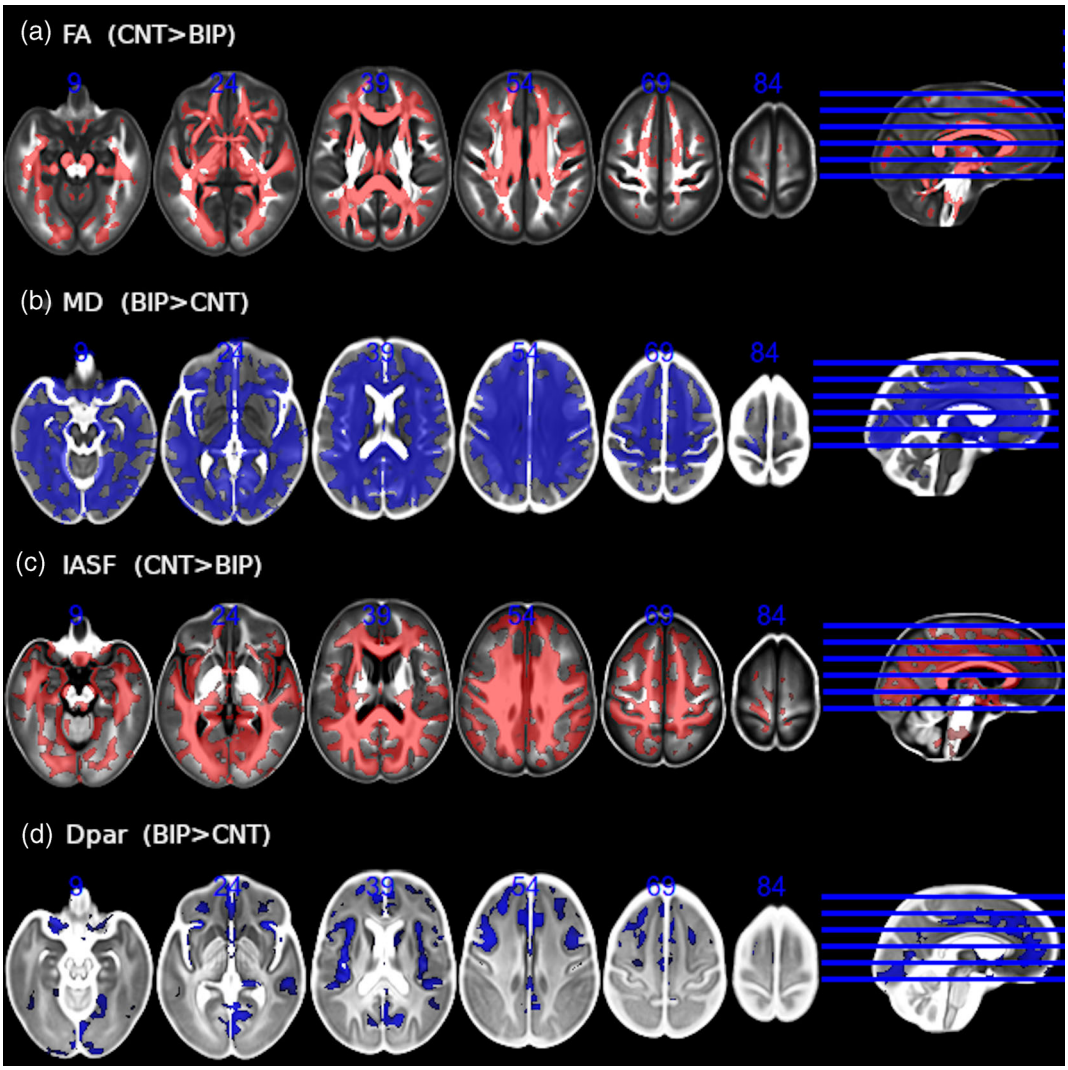

FIGURE 1 Inter-group comparison between patients with bipolar disorder (BD) and healthy controls (HC), corrected for multiple comparisons at $p = .05$. (a) fractional anisotropy (*FA*) findings. Regions showing significant reduction in the BD group are shown in red. (b) mean diffusivity (*MD*) findings. Areas showing a significantly higher *MD* in BD patients are displayed in blue. (c) intra-axonal signal fraction (*IASF*). Regions showing significant reduction in the BD group are depicted in red. (d) microscopic diffusivity parallel to axons within the intra- and extra-cellular spaces (*Dpar*) findings. Regions with increased *Dpar* in patients with BD are shown in blue. The right side of the images represents the right side of the brain. The anatomical images in the background are the multimodal study-specific templates built for each metric, respectively.

We also found significantly reduced *FA* and increased *Dpar* in patients taking antipsychotics (Figure 3c). The *FA* was significantly lower in one cluster in the left parahippocampal gyrus. The *Dpar* was significantly higher in five clusters, with main peaks in the calcarine, lingual, left insula, and the triangular part of the inferior frontal gyrus (Table S8).

Significantly higher *Dpar* and *MD* and a lower *FA* and *IASF* were observed in patients taking antiepileptics (Figure 3d). *Dpar* was reduced in three main clusters. The first one reached the calcarine, cerebellum, lingual, and precuneus in the right hemisphere. The second cluster was located in the left insula and amygdala, and the third cluster was located in the right middle frontal gyrus. Another three main clusters of widespread differences in *MD* were detected. In the left brain hemisphere, the first cluster reached the fusiform, cerebellum, vermis, and cingulum, while the cuneus, lingual, precentral gyrus, and calcarine were significantly different in the right hemisphere. The second and third clusters appeared in the left hemisphere, with main peaks in the angular and parietal cortices and the insula and fusiform gyrus, respectively (Table S9). On the other hand, the *FA* analysis revealed a significant reduction in three main WM regions. The first one involved the superior cerebellar peduncle and the body of the CC bilaterally. The other two regions were located in the left superior longitudinal fasciculus and the frontal lobe. Finally, *IASF* was lower in one area involving the calcarine and cuneus, bilaterally, the right lingual and superior occipital, and the left cingulum. Additional differences were observed in one cluster in the right hemisphere, with main peaks in the supplementary motor area, cingulum, precentral and superior frontal gyri, and WM adjacent to the frontal lobe. The third

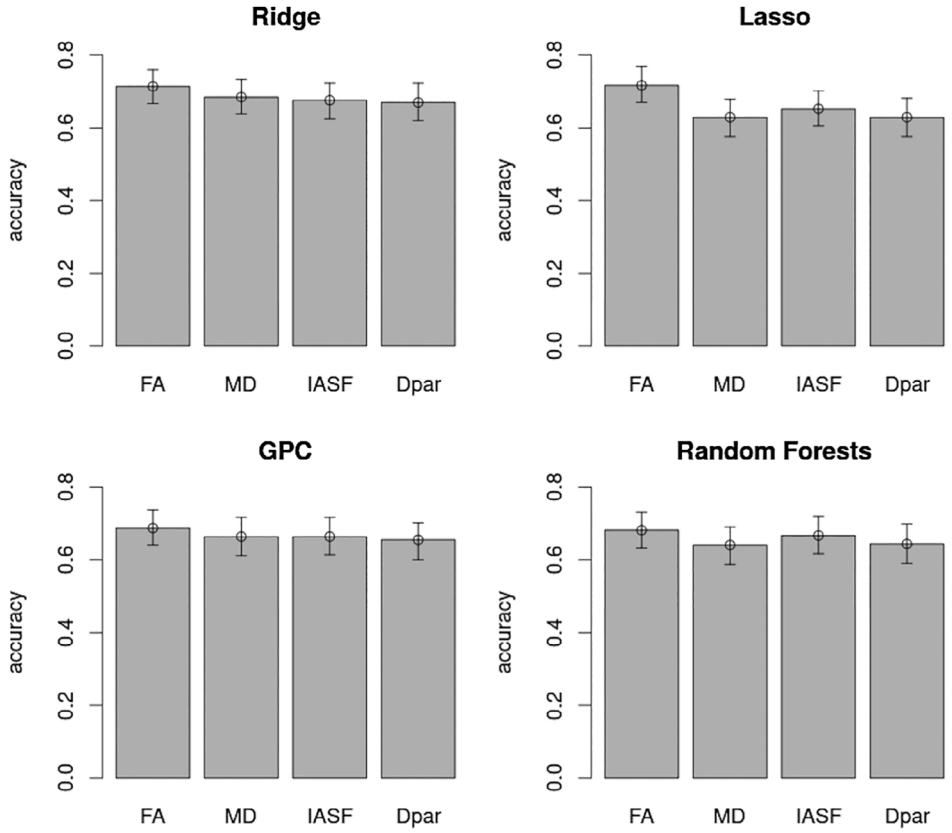

FIGURE 2 Accuracy levels (i.e., the fraction of correctly classified subjects in the validation subsets) obtained by the four different algorithms applied to the four metrics. 95% confidence intervals based on bootstrapping are also shown for each accuracy estimate. GPC, Gaussian Process Classifier.

TABLE 3 Mean classification accuracy levels for the four algorithms and the four diffusion metrics.

	FA	*MD*	*IASF*	*Dpar*
Ridge	0.71	0.68	0.68	0.67
Lasso	0.72	0.63	0.65	0.63
GPC	0.69	0.66	0.66	0.65
Random Forests	0.68	0.64	0.67	0.64

region was located in the left hemisphere, including the supplementary motor area, the superior medial frontal cortex, and adjacent WM (Table S9).

Figure 4 presents the results of post-hoc analyses conducted to investigate whether the pharmacologic treatments are associated with the differences between HC and BD patients reported in Figure 1. Patients receiving lithium treatment demonstrated diffusion metrics (FA: 969 voxels; MD: 26003 voxels; and IASF: 2535 voxels) that were closer to those of HC, in comparison to patients not receiving lithium. On the other hand, patients treated with antiepileptics (Dpar: 8746 voxels; FA: 12656 voxels; MD: 72103 voxels; and IASF: 9179 voxels) exhibited more affected diffusion metrics than those not receiving such treatment. Similarly, the subgroup of patients taking antidepressants had higher MD values (35,838 voxels), while patients taking antipsychotics had a slightly higher Dpar value (881 voxels) than those not taking these medications. The differences between patients receiving and not receiving the medications were statistically significant for all diffusion metrics.

3.5 | Associations with clinical features

Results from clinical features revealed a higher *FA* in patients without psychotic symptoms located in the forceps minor of the CC. In patients with psychotic symptoms, higher *IASF* and *Dpar* were observed in the cerebellum, reaching the cerebellar peduncle and medial lemniscus (Figure 5a, Table S10).

The correlation analyses between diffusion metrics and age of onset and duration of illness led to statistical maps involving the same brain regions. A positive correlation between *FA* and *IASF* with the age of onset and a negative correlation with the duration of illness was found (Figure 5b,c). While *IASF* findings were located in two clusters in the supramarginal and superior temporal cortex (Tables S11 and S12), results from *FA* were detected in four and five similar clusters in both analyses, respectively. The anatomical regions showing more significant associations were the genu of CC bilaterally, the right anterior corona radiata and external capsule, and the left fornix, cingulum, body of CC, and WM adjacent to the temporal lobe (Tables S11 and S12).

As the age of onset and duration of illness are significantly correlated ($Corr = -0.432$, $p < 1.2e{-}7$, see Table S1), we repeated the analyses, including both variables in the same general linear model, and age and gender as covariates. These analyses, aimed at identifying which variable was driving the observed associations, did not yield statistically significant results.

The number of episodes was positively correlated with *Dpar* in two clusters in the WM on the left hemisphere (Figure 5c), mainly

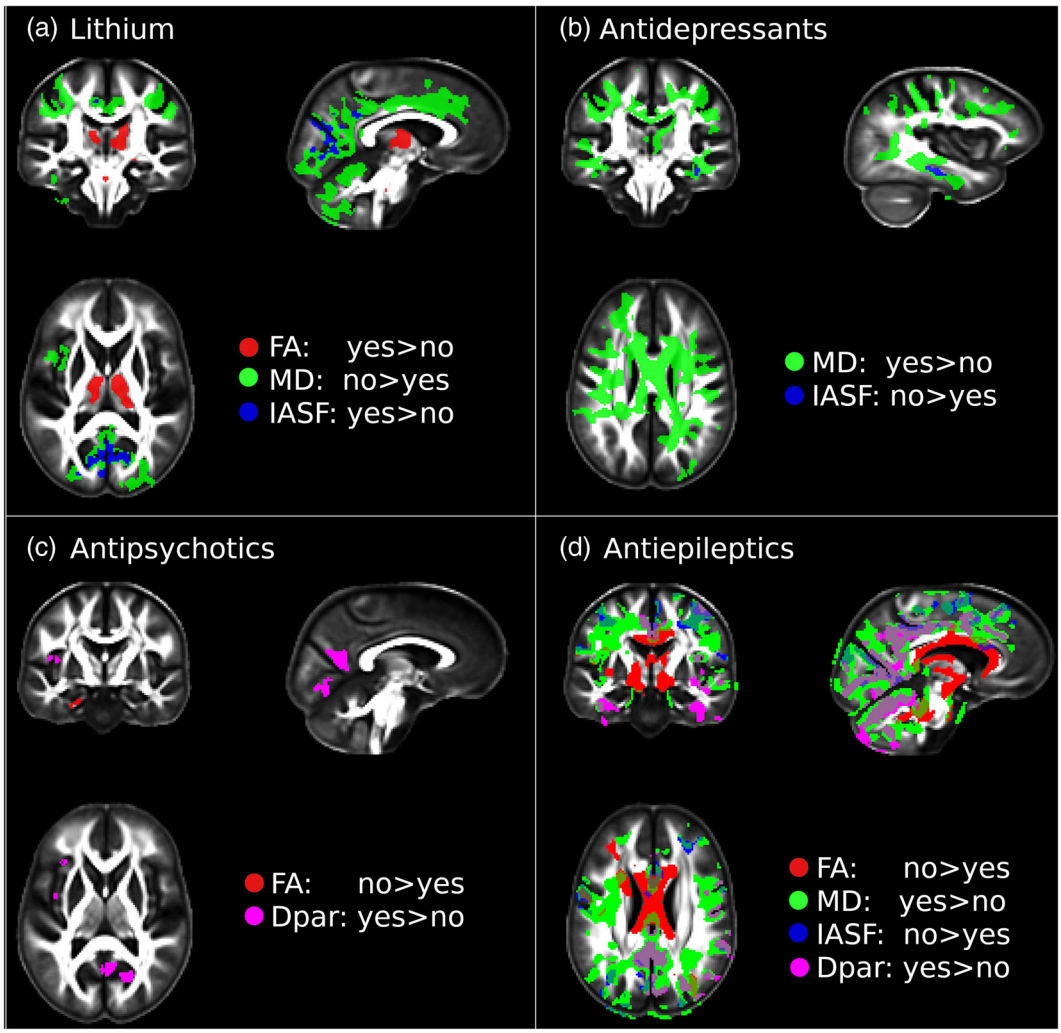

FIGURE 3 Associations of pharmacologic treatments with diffusion MRI microstructure in patients with bipolar disorder. (a) lithium. (b) antidepressants. (c) antipsychotics, and (d) antiepileptics. The different diffusion MRI metrics are colour coded. Fractional anisotropy (*FA*) findings are displayed in red; results from the mean diffusivity (*MD*) in green; the intra-axonal signal fraction (*IASF*) is displayed in blue, and findings from microscopic diffusivity parallel to the axons within the intra- and extra-cellular spaces (*Dpar*) are shown in magenta. The anatomical image in the background is the study-specific *FA* template. The right side of the images represents the right side of the brain. For each panel, the terms 'yes' and 'no' indicate whether patients received pharmacologic treatment.

involving the corticospinal tract, internal capsule, cingulum bundle, optic radiation, and the posterior segment of the arcuate fasciculus (Table S13).

We did not find significant differences between BD subtypes (i.e., BD-I vs. BD-II diagnosis) with any of the four diffusion metrics analysed.

In Figure 6, post hoc analyses were conducted to explore the associations between clinical variables and the observed differences between healthy controls (HC) and bipolar disorder (BD) patients in Figure 1. Among patients without psychotic symptoms, the IASF values (640 voxels) were comparable to those of HC, while patients with psychotic symptoms had statistically significant reductions in IASF values. Additionally, patients with an early age of onset exhibited lower FA (10,660 voxels) and IASF (413 voxels) compared with those with a late onset, whose diffusion metrics were more similar to HC. Furthermore, patients with longer illness duration demonstrated reduced FA (7603 voxels) and IASF (283 voxels) compared with those with a shorter duration.

4 | DISCUSSION

We conducted a whole-brain voxelwise diffusion MRI study using DTI and SMT (i.e., a multicompartment microscopic biophysical model) to reveal the anatomical regions affected in BD compared with healthy controls. Using a machine learning analysis, we also evaluated the individual classification accuracies of the four employed diffusion metrics (i.e., *FA* and *MD* estimated from DTI, and *IASF* and *Dpar* estimated from SMT) for the automatic diagnosis of BD. Finally, we also tested the effects of subtype diagnosis, duration of illness, age of onset,

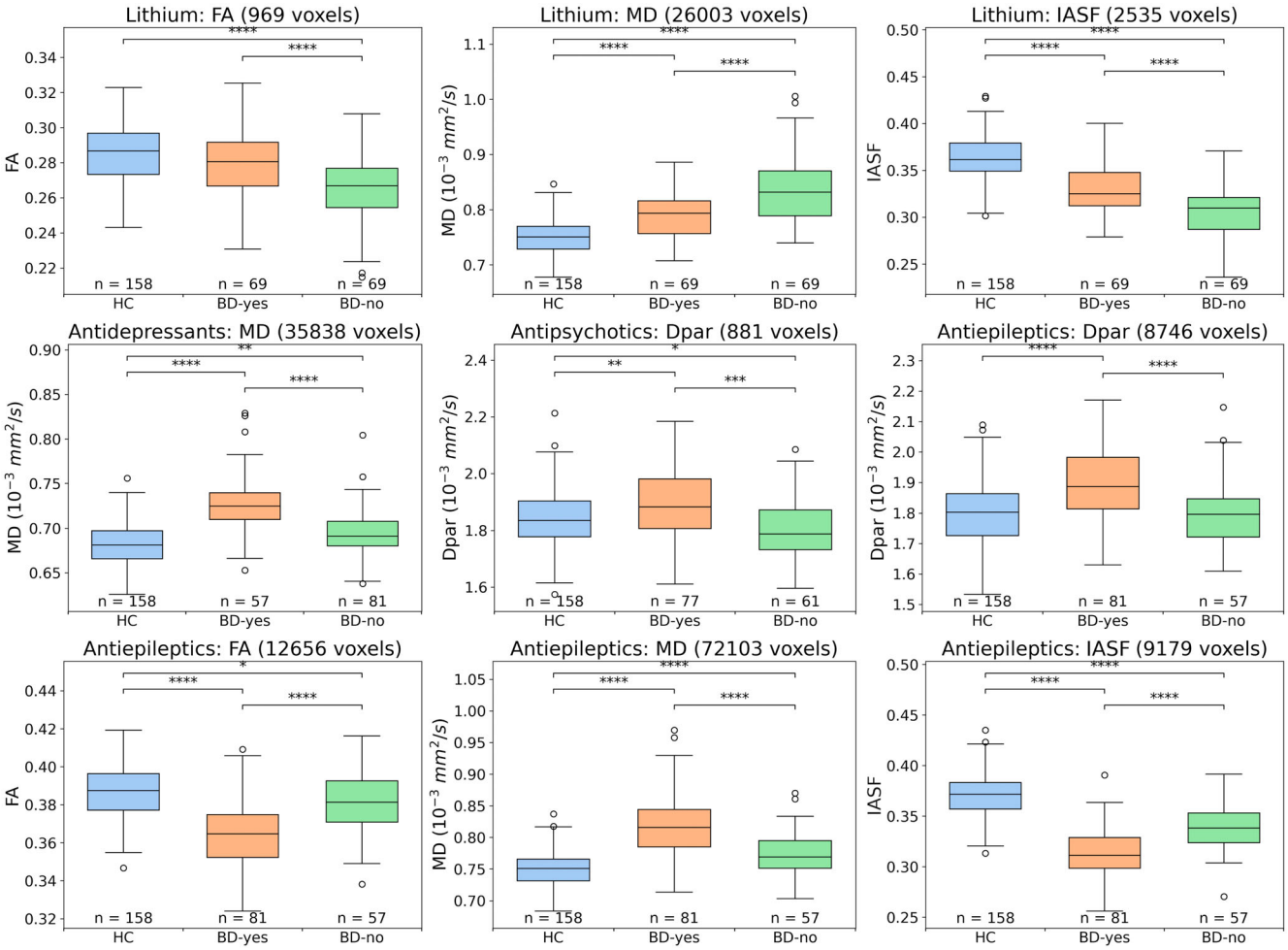

FIGURE 4 Post-hoc analysis of overlapping regions of interest (ROI) with statistically significant results in both the healthy controls (HC) versus bipolar disorder (BD) comparisons (Figure 1) and the analyses testing for medication effects (Figure 3). For each diffusion metric and medication, we tested for statistical differences between each pair of groups at the ROI level: HC versus BD-yes, HC versus BD-no, and BD-yes versus BD-no, where BD-yes and BD-no are the two BD subgroups on and off medication. The number of voxels in the resulting ROI per panel and the uncorrected *p*-values from the three *t*-tests performed in each case are reported. Only results corrected for multiple comparisons using Bonferroni (i.e., $p \leq .05/3 = .0167$) are shown. The significance levels are presented as follows *.0167 > p > = .005; **.005 > $p \geq .0005$; ***.0005 > $p \geq .00005$; ****$p < .00005$.

history of psychosis, and pharmacological treatment with lithium, antidepressant, antipsychotic, and antiepileptic drugs on diffusion MRI microstructure.

We found widespread reduced *FA* and *IASF* and increased *MD* in patients with BD. Conversely, the statistically significant regions identified by *Dpar* were more anatomically restricted (see Figure 1). On the one hand, the most significant WM statistical differences were located in the genu, body, and splenium of the CC, cerebral peduncle, cingulum bundle, WM regions adjacent to temporal and parietal lobes, corticospinal tract, inferior longitudinal fasciculus, and medial lemniscus. On the other hand, the strongest effects in GM were located in the Heschl's gyrus, cerebellum, superior frontal orbital, frontal pars triangularis, medial frontal gyrus, supramarginal gyrus, middle and inferior temporal gyrus, lateral occipital cortex, calcarine, insula, rolandic operculum fusiform, and lingual (see Tables S2–S5).

Notably, our *FA* results agree with those recently reported by the large-scale ENIGMA-DTI-BD study (Favre et al., 2019), where a widespread reduction in the WM of patients with BD was found, as well as with other studies that found both limbic and non-limbic tract disruptions (Canales-Rodríguez et al., 2013; Emsell et al., 2013; Phillips & Swartz, 2014). Besides, our GM results align well with mega- and meta-analyses from the ENIGMA consortium reporting an extensive bilateral reduction of the cortical mantle in BD, especially in frontal, temporal, and parietal regions (Hibar et al., 2018), which has been found in other small-scale studies as well, for example, see Madre et al. (2020).

The widespread alterations involving both WM and GM regions found here and in previous ENIGMA studies challenge the existing models of BD (Ching et al., 2020). While the alterations of WM fascicles and GM structures belonging to the limbic system and fronto-limbic connections agree with the standard models of BD, as pointed

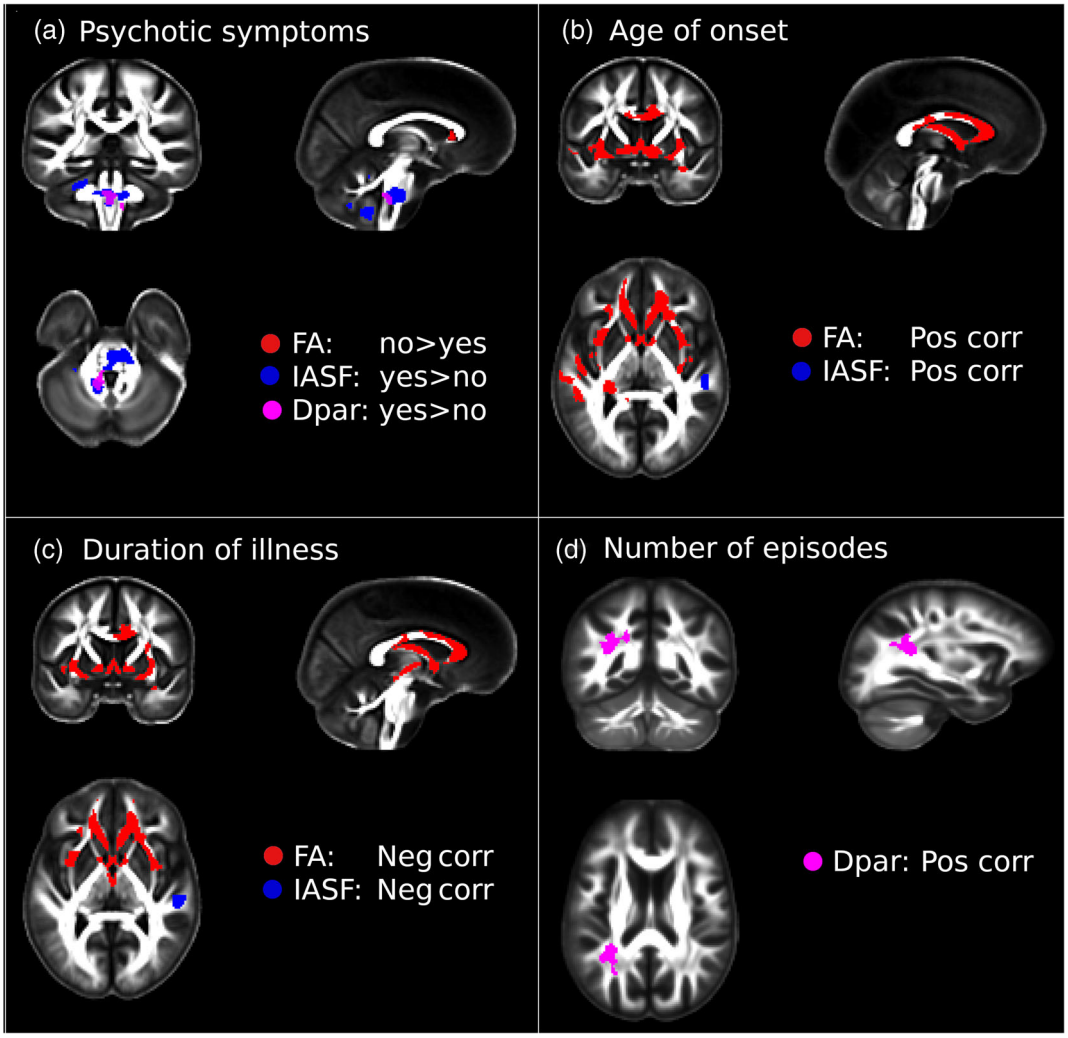

FIGURE 5 Associations between clinical features and diffusion MRI microstructure in patients with bipolar disorder. (a) Psychotic symptoms. The terms 'yes' and 'no' indicate patients with and without psychotic symptoms, respectively. (b) Age of onset. (c) Duration of illness, and (d) number of episodes. The different diffusion MRI metrics are colour coded. Fractional anisotropy (*FA*) findings are displayed in red, results from the intra-axonal signal fraction (IASF) are displayed in blue, and results from the microscopic diffusivity parallel to the axons within the intra- and extra-cellular spaces (*Dpar*) are shown in magenta. The anatomical image in the background is the study-specific *FA* template. The right side of the images represents the right side of the brain.

out by (Ching et al., 2020), the role of the CC in BD must be elucidated yet. The current BD models should be extended to include the inter-hemispheric connections as features of interest. However, this may be intricate as inter-hemispheric abnormalities are not specific to BD, and similar widespread microstructure abnormalities have been previously reported by ENIGMA groups in patients with schizophrenia (Kelly et al., 2018) and major depressive disorder (MDD; van Velzen et al., 2020).

Apart from the *FA*, which is the most frequently used WM diffusion metric and which was the one analysed in the large-scale ENIGMA-DTI-BD study, in our work, we also considered the *MD*, which is sensitive to changes in both WM and GM tissue. Also, importantly, and in contrast to DTI measures, the two SMT metrics used here are not sensitive to fibre crossings and fibre dispersion, which are present in a large proportion of brain voxels (Jeurissen et al., 2013). In addition, SMT metrics can also discriminate between diffusion MRI signals arising from the intra- and extra-axonal compartments.

Although the number of subjects in our study is smaller than in previous multicenter ENIGMA studies, it is large enough (i.e., $n = 316$) to detect widespread microstructure differences. Also, in contrast to these ENIGMA studies, our sample was well-matched for age, sex, and estimated premorbid IQ, minimising potential confounding effects from these variables and avoiding the multicenter heterogeneity caused by using data from different scanning devices and MRI sequences (Radua et al., 2020). Another advantage of our study is the whole-brain voxelwise approach which, compared with the ROI-based methods used by ENIGMA, can reveal focal microstructure abnormalities undetected at the ROI level (Abé et al., 2022).

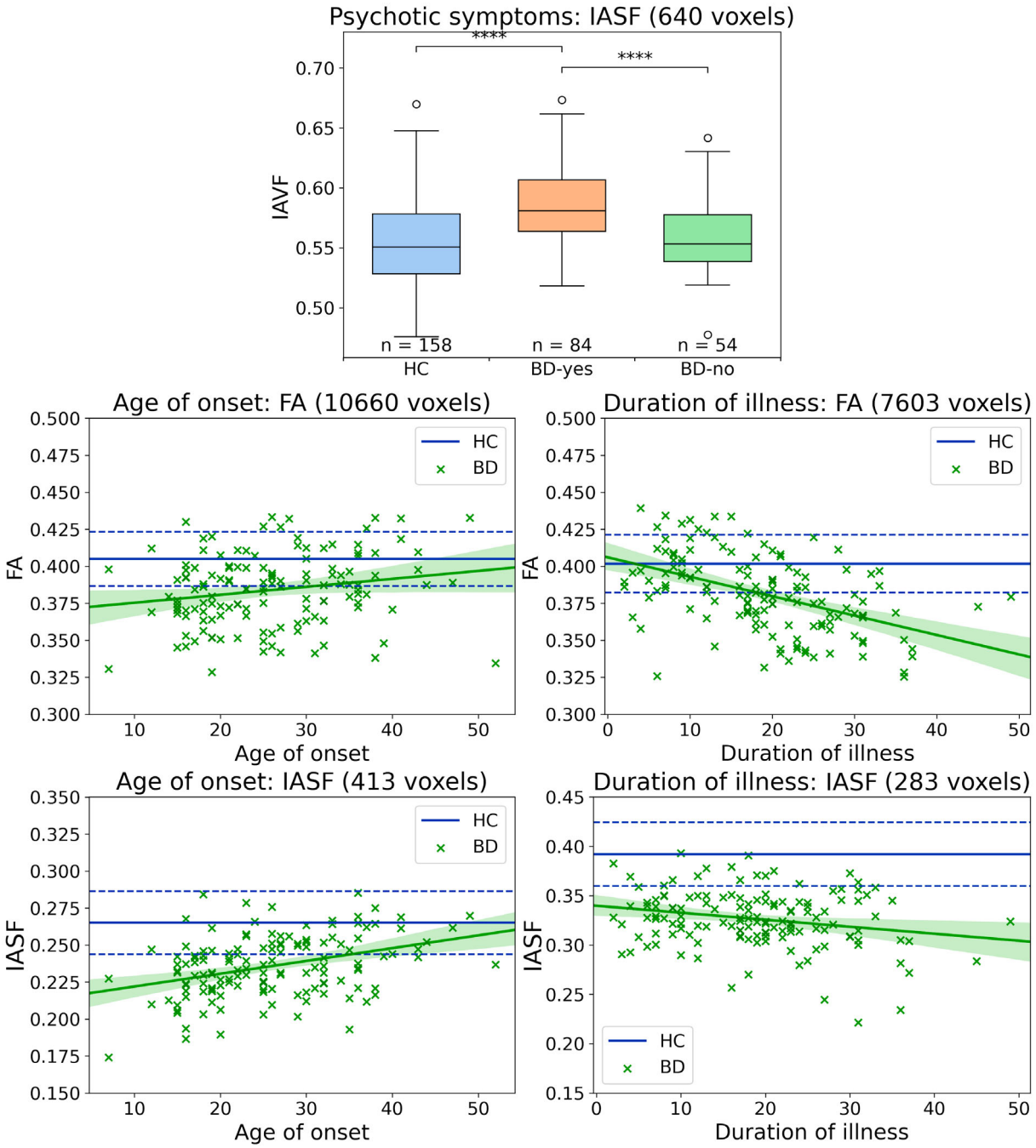

FIGURE 6 Post-hoc analysis of overlapping regions of interest (ROI) with statistically significant results both in the healthy controls (HC) versus bipolar disorder (BD) comparisons (Figure 1) and in the analyses testing the effects of clinical variables (Figure 5). The top panel shows results corresponding to the intra-axonal signal fraction (IASF) and psychotic symptoms. Statistical differences at the ROI level were tested between the following pairs of groups: HC versus BD with psychotic symptoms (BD-yes), HC versus BD without psychotic symptoms (BD-no), and BD-yes versus BD-no. Only results corrected for multiple comparisons using Bonferroni (i.e., $p \leq .05/3 = .0167$) are shown. The significance levels for the uncorrected p-values from the three t-tests are given as follows: *$.0167 > p \geq .005$; **$.005 > p \geq .0005$; ***$.0005 > p \geq .00005$; ****$p < .00005$. The middle and bottom panels show the fractional anisotropy (FA) and IASF results, respectively. The scatter plots show the values of FA and IASF as a function of the age of onset and duration of illness in the BD patients, as well as the regression lines fitting the data and their 95% confidence intervals. The FA and IASF values of the HC subjects are represented by three horizontal blue lines: the middle line indicates the mean value, and the upper and lower lines are the values above and below one standard deviation from the mean, respectively. The number of voxels in the resulting ROI is reported for each panel.

4.1 | How can we explain the observed changes in terms of biological parameters?

The widespread overlapping between *FA* and *IASF* abnormalities supports the hypothesis that a lower intra-axonal signal fraction is underlying the smaller WM *FA* values observed in patients. Lower intra-axonal volumes (and smaller *IASF* values) can either be due to a reduced number of axons or reduced axonal diameters. In both scenarios, lower *IASF* will be linked to a higher extra-axonal signal fraction owing to a relative increase in the number of water molecules in the extra-axonal space. In turn, these microstructure changes may be explained by progressive brain damage (Shioya et al., 2015), genetic factors (Patel et al., 2021), or neuroinflammation processes (Tuozzo et al., 2018).

Although a reduced intra-axonal volume may be the most plausible hypothesis for the observed results, an excessive accumulation of macromolecules in the intra-axonal water may reduce the intra-axonal T2 relaxation time and, consequently, the *IASF*. As diffusion MRI models do not commonly model the T2 relaxation process (e.g., see Veraart et al., 2018), changes in T2 cannot be disentangled from intra-axonal volume changes. Indeed, differential expression of cytoskeletal components of the axon-myelin unit has been reported in postmortem samples of patients with BD, which could modify the T2 time; for a detailed review, see Valdés-Tovar et al. (2022) and references therein. Furthermore, there is evidence supporting the neurotoxic effects of mood episodes in BD, which are associated with altered blood levels of biomarkers related to inflammation, oxidative stress, and neurotrophins, inducing a decrease in neurite density and cell viability (Wollenhaupt-Aguiar et al., 2016).

Even though we cannot currently provide a final answer about whether the observed changes are dominated by a reduction in the intra-axonal volume or an increase of macromolecules in the intra-axonal water, our results can be used as a guide for future studies aimed at resolving this essential question. Thus, while the intra-axonal T2 can be estimated using diffusion-relaxation MRI techniques (Barakovic, Tax, et al., 2021; McKinnon & Jensen, 2019), the inner axon diameter may be calculated by employing advanced diffusion MRI microstructure models and strong diffusion gradients (Assaf et al., 2008; Barakovic, Girard, et al., 2021; Daducci et al., 2015; Dyrby et al., 2013; Veraart et al., 2020).

4.2 | Is FA the most sensitive diffusion metric in BD?

A key goal of neuroimaging research in BD is identifying quantitative biomarkers that increase diagnostic precision and pave the way for personalised treatments (Phillips & Swartz, 2014). If, hypothetically, a main biological factor dominated the microstructural abnormalities in BD, a diffusion MRI model quantifying such factor while removing other confounding effects would provide very specific information about these microstructural abnormalities. In this study, we compared four metrics in terms of their performance to classify patients and controls (see Figure 2 and Table 3). Surprisingly, *FA*, a diffusion MRI metric modulated by many different microstructure features (including the axon packing density, fibre orientation dispersion, fibre crossings, and myelin volume), provided better classification scores than more specific microstructure metrics. This result suggests that, although these specific microstructure metrics may provide a more definite description of brain features, these features may not be as related to bipolarity as *FA*. Still, differences in classification accuracy between *FA* and the other metrics were moderate.

Interestingly, classification accuracies reported here are higher than individual site accuracies from previous ENIGMA BD studies based on T1w data and cortical thickness (Nunes et al., 2020). In part, this could be explained by a lower spatial resolution provided ROI averages compared with our voxel-based approach. However, after pooling data from 3020 individuals from 13 sites (i.e., mean sensitivity of 66%; Nunes et al., 2020), accuracies were closer to ours. Similar accuracies with cortical thickness were also observed in a previous study using data at the vertex level (Salvador et al., 2017). In such a study, the highest accuracy (63%) was obtained for the GM volume estimated from voxel-based morphometry. Still, this performed worse than our *FA* findings (68%–72%, Table 3). Indeed, *FA*, *IASF* (65%–68%) and *MD* (63%–68%) may be combined with metrics derived from T1w data to build multimodality classifiers for the improved automatic classification of BD patients (e.g., see Salvador et al., 2019). Another potentially helpful metric to enhance the accuracy is the myelin water fraction, which can be estimated from multicomponent T2 relaxometry techniques and multi-echo T2 data (Piredda et al., 2021, Yu et al., 2021, Canales-Rodríguez et al., 2021a, 2021b, 2021c) . Nevertheless, in our study, we did not observe a significant improvement in the performance when the four diffusion metrics were employed together in the classification.

It is important to recognise that machine learning (ML)-based classification methods in psychiatry have a larger goal than distinguishing between patients with bipolar disorder (BD) and healthy controls. In clinical practice, a more useful outcome would be to distinguish between patients with different psychiatric disorders and their subtypes. Achieving a high classification accuracy between healthy controls and patients with BD is not sufficient for this primary objective, but it is certainly necessary.

4.3 | Association between pharmacologic treatment and brain tissue microstructure

In the patient group, lithium intake was associated with higher *FA* and *IASF*, and reduced *MD*. *FA* and *IASF* results were anatomically restricted (in the thalamus, brainstem, superior longitudinal, calcarine, cingulum gyrus and cingulum bundle), while *MD* showed a more diffuse pattern reaching several GM regions (Figure 3a and Table S6). Notably, in a previous ENIGMA study of subcortical volumes in BD, lithium treatment was associated with larger thalamic volumes compared with non-treated patients (Hibar et al., 2016). Our *FA* findings are also concordant with results from the large-scale ENIGMA-

DTI-BD study (Favre et al., 2019). Previous studies have suggested that lithium confers neuroprotective effects on the WM (Abramovic et al., 2018) and GM (Hafeman et al., 2012; Hajek et al., 2012; Hibar et al., 2018; Lyoo et al., 2010; Lyskowski et al., 1982) and attenuates the microstructure differences associated with BD, along with its known neurogenetic effects (Kato, 2022; Rybakowski, 2022).

We found that antidepressant treatment was associated with an alteration in the tissue microstructure of patients, manifested by a lower *IASF* and a higher *MD* (Figure 3b, Table S7). A late-life depression study in the elderly examined the association of worsening WM grade from initial to follow-up MRI scans and antidepressant use between the scans. In that study, all antidepressants were associated with worsening WM during the studied period (Steffens et al., 2008). Our results may be related to a recent ENIGMA study in MDD patients, where no *FA* differences were observed between antidepressant users and non-users (van Velzen et al., 2020). However, that study also found a lower *FA* in antidepressant non-users versus HCs, while no such differences were detected between antidepressant users versus HCs. The latter result may suggest that WM deficits may be lessened by antidepressant treatment (van Velzen et al., 2020), which is not concordant with our findings. Nevertheless, there are methodological differences that could explain the observed discrepancies. First, results from MDD and BD patients may be different. Second, the reported antidepressant effect was only observed via an indirect comparison with HCs, as the direct comparison between users and non-users did not reveal statistical differences. Third, we found significant results in the *IASF* and *MD* metrics, which were not used in that previous study. Four, as in our study, antidepressant use was examined at the time of scanning, as data on past antidepressant use was unavailable (van Velzen et al., 2020). Finally, the age of patients taking antidepressants was significantly higher in our study than those not having them (Table 2), and there was a significant correlation between antidepressant use and duration of illness (Table S1). Even though age was included as a covariate in our model (to attenuate the effect of these associations), results could still be confounded by other variables like disease severity (van Velzen et al., 2020).

Patients receiving antipsychotics showed a reduced *FA* and increased *Dpar* when compared with patients not receiving antipsychotics (Figure 3c). While *FA* was significantly lower in the left parahippocampal gyrus, *Dpar* was higher in the calcarine, lingual, left insula, and inferior frontal gyrus (Table S8). Our results may indicate that antipsychotic treatment affects the tissue microstructure, in agreement with a previous study that found an association between antipsychotic treatment and reduced GM volumes (Haukvik et al., 2020). This result is consistent with the large-scale ENIGMA-DTI-BD study (Favre et al., 2019), which found lower *FA* in patients on antipsychotic treatment. However, the differences were observed at a different location (i.e., within the genu of the CC). Similarly, a surface-based morphometry ENIGMA study found reduced cortical surface area associated with BD patients taking antipsychotic treatment (Hibar et al., 2018).

When comparing antiepileptic users and non-users, we found significant differences in the four diffusion metrics involving both WM and GM regions (Figure 3d, Table S9), suggesting a widespread atrophic effect of antiepileptic treatment on tissue microstructure. Previous evidence indicates that antiepileptic treatment reduces the cortical thickness (Hibar et al., 2018) and hippocampal volumes (Haukvik et al., 2020). In WM, the large-scale ENIGMA-DTI-BD study (Favre et al., 2019) found a reduced *FA* in several tracts and average global *FA* in patients taking anticonvulsants.

The post-hoc analyses presented in Figure 4 indicated associations between pharmacologic treatments and the differences observed between HC and BD patients reported in Figure 1. Specifically, treatment with Lithium was associated with a normalisation of the differences, while treatment with other medications (e.g., antiepileptics, antidepressants, and antipsychotics) was associated with an accentuation of the differences. Notably, the differences between patients taking and not taking these medications were statistically significant for all diffusion metrics.

4.4 | Association between clinical features and brain tissue microstructure

In patients with psychotic symptoms (compared with those not having them), a lower *FA* was observed in a small cluster in the genu of the CC. (Figure 5a, Table S10). In disagreement with our results, the large-scale ENIGMA-DTI-BD study (Favre et al., 2019) did not report any significant relationship between *FA* and history of psychotic symptoms. Nevertheless, our results agree with a multicenter tractography study that found that patients with such a clinical profile had a lower generalised *FA* in the CC (Sarrazin et al., 2014). On the other hand, we found a higher *IASF* and *Dpar* in patients with versus without psychotic symptoms in the cerebellum, cerebellar peduncle and medial lemniscus. However, these findings cannot be compared with previous studies as this is the first work studying these associations in BD. Our results suggest that a history of psychotic symptoms may be related to an increase in the intra-axonal volume in WM and GM regions in the cerebellum. Yet, we are not able to suggest a plausible biological mechanism to explain this finding, and we cannot discard that some of our significant results are false positives, especially with smaller clusters. This finding should be replicated in future studies.

The analyses quantifying the effect of age of onset and duration of illness produced similar results but with opposite signs. While a positive correlation between *FA* (located in the genu and body of the CC, the right anterior corona radiata and external capsule, and the left fornix, cingulum, and WM adjacent to the temporal lobe) and *IASF* (in the supramarginal and superior temporal cortex) was detected with the age of onset, a negative correlation was observed for the duration of illness (see Figure 5b,c, Tables S11 and S12). This result could be explained by the high negative correlation between the age of onset and duration of illness. To identify which variable was more strongly associated with the diffusion metrics, we repeated the analyses by including both variables in the same general linear model, also covarying for age and sex. However, no significant statistical differences were found. This may be attributable to the high (negative) correlation

levels between the two variables, which may have brought collinearity issues and higher noise (i.e., larger standard errors) in the estimated model coefficients. Hence, with the current data, we cannot determine if both variables exert a similar effect on the diffusion metrics or if one dominates the association. Our results agree with the large-scale ENIGMA-DTI-BD study (Favre et al., 2019), which found a higher *FA* associated with shorter illness duration and later disorder onset.

When assessing the effect of the number of episodes on the diffusion metrics, we only found a significant positive association with *Dpar* in a WM cluster in the left brain hemisphere. (Figure 5c, Table S13). As in the large-scale ENIGMA-DTI-BD study (Favre et al., 2019), no significant *FA* alterations were associated with the number of mood episodes. Finally, we observed no significant difference between patients with BD type-I and BD type-II diagnoses, in agreement with other ENIGMA BD studies (e.g., Favre et al., 2019; Hibar et al., 2016, 2018).

The post hoc analyses depicted in Figure 6 showed associations between psychotic symptoms, duration of illness, and early age of onset, and the differences observed between HC and BD patients in Figure 1. Specifically, these clinical variables were associated with an accentuation of the differences. It is important to note that these findings indicate associations rather than establishing causal relationships.

4.5 | Conclusions and limitations

Our findings confirm the widespread WM *FA* alterations in BD recently reported by the largest multicenter DTI study in BD carried out so far (Favre et al., 2019) and provide novel evidence about widespread abnormalities in other diffusion metrics not evaluated in that study. The large anatomical overlap between *FA* and *IASF* abnormalities suggests that the lower *FA* observed in patients could be caused by a reduced intra-axonal volume fraction or a higher macromolecular content in the intra-axonal water. Future studies shall be conducted to investigate these two hypotheses. Moreover, we found a widespread alteration in *MD* involving WM and GM tissue, agreeing with previous structural multicenter ENIGMA studies that found widespread abnormalities in the cortical mantle in BD. In addition, our Machine Learning analysis revealed that *FA* is the most helpful metric for the automatic diagnosis of BD patients, reaching an accuracy of 72%, higher than those reported in previous multicenter studies using structural MRI data (i.e., 66%). The *IASF* and *MD* metrics provided a similar but slightly reduced performance. In the patients' group, we found that a longer duration of illness, earlier age of onset, having psychotic symptoms, and suffering a higher number of mood episodes (after accounting for age and sex) was associated with increased abnormalities. No significant differences between patients diagnosed in the BD-I and BD-II subtypes were observed. We found that lithium, antidepressants, antipsychotics, and antiepileptics are significantly associated with the diffusion metrics. While lithium appeared to be the only medication associated with attenuating microstructure differences related to BD, users of antiepileptics showed more pronounced microstructure alterations. It is important to note that these findings indicate associations rather than establishing causal relationships.

To some extent, the presence of significant correlations between the studied variables may prevent a direct and easy interpretation of results. Indeed, indirect effects may be produced by the significant linear associations between lithium and antiepileptic users (negative correlation), duration of illness and age of onset (negative correlation), antidepressant users and duration of illness or age (positive correlation), and psychotic symptoms and diagnosis subtype (positive correlation). Other non-studied (latent) variables also could confound our results. It remains to be investigated how the number and duration of episodes and the cumulative doses of each medication are related to longitudinal brain tissue microstructure changes in BD. Obtaining reliable information is not always possible since not all episodes require hospitalisations and part of the available information is based on self-reports. Pharmacologic treatment was examined at the time of scanning because data on past use was not available. Surely, these questions may be better answered by using longitudinal designs (e.g., Abé et al., 2022) and clinical trials (e.g., Hoertel et al., 2013). In the statistical analysis, the number of episodes was treated as a continuous variable to identify linear positive or negative associations with the diffusion metrics. However, as defined in this study, the number of episodes is an ordinal variable with three ordered categories that are not necessarily equidistant. Therefore, our results may be caused by non-modelled non-linear effects. Furthermore, it is important to acknowledge that our study focused exclusively on patients with bipolar disorder. Therefore, the generalisability of our findings to other populations or psychiatric disorders remains uncertain. Finally, it should be noted that our study population consisted exclusively of individuals of Caucasian descent. Consequently, the applicability of our findings to patients of other ethnicities is uncertain, and caution should be exercised when attempting to generalise our findings to populations of diverse ethnic backgrounds.

AUTHOR CONTRIBUTIONS

Erick Jorge Canales-Rodríguez: Conceptualisation, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Visualization, Roles/Writing – original draft, Writing – review & editing. **Norma Verdolini**: Data curation, Investigation, Writing – review & editing. **Silvia Alonso-Lana**: Data curation, Investigation, Writing – review & editing. María **Llanos Torres**: Data curation, Investigation, Writing – review & editing. **Francesco Panicalli**: Data curation, Investigation, Writing – review & editing. **Isabel Argila-Plaza**: Data curation, Investigation, Writing – review & editing. **Elena Rodriguez-Cano**: Data curation, Investigation, Writing – review & editing. **Irene Montoro**: Data curation, Investigation, Writing – review & editing. **Beatriz Garcia-Ruiz**: Data curation, Investigation, Writing – review & editing. **Esther Jimenez**: Data curation, Investigation, Writing – review & editing. **Cristina Varo**: Data curation, Investigation, Writing – review & editing. **Anna Lluch**: Data curation, Investigation, Writing – review & editing. **Caterina del Mar Bonnin**: Data curation, Investigation, Writing – review & editing. **Silvana Maluf**: Data curation, Investigation, Writing – review & editing.

Marc Pujol: Data curation, Investigation, Writing – review & editing. **Núria Jaurrieta Guarner**: Data curation, Investigation, Writing – review & editing. **Salvador Sarró**: Data curation, Investigation, Writing – review & editing. **Eduard Vieta**: Data curation, Funding acquisition, Investigation, Project administration, Resources, Writing – review & editing. **Elisabet Vilella**: Data curation, Funding acquisition, Investigation, Project administration, Resources, Writing – review & editing. **Raymond Salvador**: Conceptualisation, Funding acquisition, Investigation, Methodology, Validation, Writing – review & editing. **Edith Pomarol-Clotet**: Conceptualisation, Data curation, Funding acquisition, Investigation, Project administration, Resources, Writing – review & editing.

ACKNOWLEDGMENTS

This work was supported by the Generalitat de Catalunya: 2014SGR1573, 2014SGR398 and 2017SGR1365 and by the Centro de Investigación Biomédica en Red de Salud (CIBERSAM). Also by several grants funded by the Instituto de Salud Carlos III and the Spanish Ministry of Science, Innovation, and Universities (co-funded by the European Regional Development Fund/European Social Fund 'Investing in your future'): Río Hortega Contract (CM17/00258 to NV), Sara Borrell Research Contract (CD18/00029 to EC-R), Miguel Servet Research Contract (CPII16/00018 to EP-C and CPII13/00018 to RS), Research Mobility programme (MV18/00054 to EP-C), Research Projects (PI15/00277 to EC-R, PI15/00852, PI15/00283, PI18/00805 to EV, PI18/00877 and PI21/00525 to RS, PI14/01148, PI18/00810 to EP-C). The study has been supported by a BITRECS project conceded to NV. BITRECS project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement No. 754550 and from 'La Caixa' Foundation (ID 100010434), under the agreement LCF/PR/GN18/50310006. EC-R was supported by the Swiss National Science Foundation, Ambizione grant PZ00P2_185814. The funding organisations played no role in the study design, data collection, analysis, or manuscript approval. We finally thank all the study participants.

CONFLICT OF INTEREST STATEMENT

Dr Vieta has received grants and served as a consultant, advisor, or CME speaker for the following entities (unrelated to the present work): AB-Biotics, Abbott, Allergan, Angelini, Biogen, Boehringer-Ingelheim, Dainippon Sumitomo Pharma, Ferrer, Gedeon Richter, Janssen, Lundbeck, Otsuka, Sage, Sanofi-Aventis, Takeda, and Viatris. Dr. Verdolini has received financial support for CME activities and travel funds from the following entities (unrelated to the present work): Angelini, Janssen-Cilag, Lundbeck, Otsuka. The other authors have nothing to disclose.

DATA AVAILABILITY STATEMENT

The MRI data supporting this study's findings are available from the corresponding author upon reasonable request. The raw MRI data are not publicly available due to privacy or ethical restrictions. Non-MRI data supporting and summarizing this study's findings are available in this article's supplementary material.

ORCID

Erick Jorge Canales-Rodríguez https://orcid.org/0000-0001-6421-2633

Raymond Salvador https://orcid.org/0000-0001-5557-1562

REFERENCES

Abé, C., Ching, C. R. K., Liberg, B., Lebedev, A. V., Agartz, I., Akudjedu, T. N., Alda, M., Alnæs, D., Alonso-Lana, S., Benedetti, F., Berk, M., Bøen, E., Bonnin del, C. M., Breuer, F., Brosch, K., Brouwer, R. M., Canales-Rodríguez, E. J., Cannon, D. M., Chye, Y., … Landén, M. (2022). Longitudinal structural brain changes in bipolar disorder: A multicenter neuroimaging study of 1232 individuals by the ENIGMA bipolar disorder working group. *Biological Psychiatry*, *91*, 582–592.

Abramovic, L., Boks, M. P. M., Vreeker, A., Verkooijen, S., van Bergen, A. H., Ophoff, R. A., Kahn, R. S., & van Haren, N. E. M. (2018). White matter disruptions in patients with bipolar disorder. *European Neuropsychopharmacology*, *28*, 743–751.

Andersson, J. L. R., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*, *20*, 870–888.

Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, *125*, 1063–1078.

Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., & Basser, P. J. (2008). AxCaliber: A method for measuring axon diameter distribution from diffusion MRI. *Magnetic Resonance in Medicine*, *59*, 1347–1354. https://doi.org/10.1002/mrm.21577/full

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*, 26–41.

Barakovic, M., Girard, G., Schiavi, S., Romascano, D., Descoteaux, M., Granziera, C., Jones, D. K., Innocenti, G. M., Thiran, J.-P., & Daducci, A. (2021). Bundle-specific axon diameter index as a new contrast to differentiate white matter tracts. *Frontiers in Neuroscience*, *15*, 646034.

Barakovic, M., Tax, C. M. W., Rudrapatna, U., Chamberland, M., Rafael-Patino, J., Granziera, C., Thiran, J. P., Daducci, A., Canales-Rodríguez, E. J., & Jones, D. K. (2021). Resolving bundle-specific intra-axonal T2 values within a voxel using diffusion-relaxation tract-based estimation. *NeuroImage*, *227*, 117617.

Basser, P. J., & Pierpaoli, C. (1996). Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of Magnetic Resonance. Series B*, *111*, 209–219.

Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system – A technical review. *NMR in Biomedicine*, *15*, 435–455.

Berk, M., Berk, L., Moss, K., Dodd, S., & Malhi, G. S. (2006). Diagnosing bipolar disorder: How can we do it better? *The Medical Journal of Australia*, *184*, 459–462.

Canales-Rodríguez, E. J., Alonso-Lana, S., Verdolini, N., Sarró, S., Feria, I., Montoro, I., Garcia-Ruiz, B., Jimenez, E., Varo, C., Albacete, A., Argila-Plaza, I., Lluch, A., Bonnin, C. M., Vilella, E., Vieta, E., Pomarol-Clotet, E., & Salvador, R. (2021c). Age– and gender–related differences in brain tissue microstructure revealed by multi–component T2 relaxometry. *Neurobiology of Aging*, *106*, 68–79.

Canales-Rodríguez, E. J., Pizzolato, M., Piredda, G. F., Hilbert, T., Kunz, N., Pot, C., Yu, T., Salvador, R., Pomarol-Clotet, E., Kober, T., Thiran, J. P., & Daducci, A. (2021a). Comparison of non–parametric T2 relaxometry methods for myelin water quantification. *Medical Image Analysis*, *69*, 101959.

Canales-Rodríguez, E. J., Pizzolato, M., Yu, T., Piredda, G. F., Hilbert, T., Radua, J., Kober, T., & Thiran, J. P. (2021b). Revisiting the T2 spectrum imaging inverse problem: Bayesian regularized non–negative least squares. *Neuroimage*, *244*, 118582.

Canales-Rodríguez, E. J., Pomarol-Clotet, E., Radua, J., Sarró, S., Alonso-Lana, S., del Mar, B. C., Goikolea, J. M., Maristany, T., García-

Álvarez, R., Vieta, E., McKenna, P., & Salvador, R. (2013). Structural abnormalities in bipolar euthymia: A multicontrast molecular diffusion imaging study. *Biological Psychiatry*, *76*, 239–248.

Ching, C. R. K., Hibar, D. P., Gurholt, T. P., Nunes, A., Thomopoulos, S. I., Abé, C., Agartz, I., Brouwer, R. M., Cannon, D. M., de Zwarte, S. M. C., Eyler, L. T., Favre, P., Hajek, T., Haukvik, U. K., Houenou, J., Landén, M., Lett, T. A., McDonald, C., Nabulsi, L., … Andreassen, O. A. (2022). What we learn about bipolar disorder from large-scale neuroimaging: Findings and future directions from the ENIGMA bipolar disorder working group. *Human Brain Mapping*, *43*, 56–82.

Daducci, A., Canales-Rodríguez, E. J., Zhang, H., Dyrby, T. B., Alexander, D. C., & Thiran, J.-P. P. (2015). Accelerated microstructure imaging via convex optimization (AMICO) from diffusion MRI data. *NeuroImage*, *105*, 32–44.

del Ser, T., Gonzalez-Montalvo, J. I., Martinez-Espinosa, S., Delgado-Villapalos, C., & Bermejo, F. (1997). Estimation of premorbid intelligence in Spanish people with the word accentuation test and its application to the diagnosis of dementia. *Brain and Cognition*, *33*, 343–356.

Dyrby, T. B., Søgaard, L. V., Hall, M. G., Ptito, M., & Alexander, D. C. (2013). Contrast and stability of the axon diameter index from microstructure imaging with diffusion MRI. *Magnetic Resonance in Medicine*, *70*, 711–721.

Emsell, L., Leemans, A., Langan, C., van Hecke, W., Barker, G. J., McCarthy, P., Jeurissen, B., Sijbers, J., Sunaert, S., Cannon, D. M., & McDonald, C. (2013). Limbic and callosal white matter changes in euthymic bipolar I disorder: An advanced diffusion magnetic resonance imaging tractography study. *Biological Psychiatry*, *73*, 194–201.

Favre, P., Pauling, M., Stout, J., Hozer, F., Sarrazin, S., Abé, C., Alda, M., Alloza, C., Alonso-Lana, S., Andreassen, O. A., Baune, B. T., Benedetti, F., Busatto, G. F., Canales-Rodríguez, E. J., Caseras, X., Chaim-Avancini, T. M., Ching, C. R. K., Dannlowski, U., Deppe, M., … Houenou, J. (2019). Widespread white matter microstructural abnormalities in bipolar disorder: Evidence from mega- and meta-analyses across 3033 individuals. *Neuropsychopharmacology*, *44*, 2285–2293.

Gomar, J. J., Ortiz-Gil, J., McKenna, P. J., Salvador, R., Sans-Sansa, B., Sarro, S., Guerrero, A., & Pomarol-Clotet, E. (2011). Validation of the word accentuation test (TAP) as a means of estimating premorbid IQ in Spanish speakers. *Schizophrenia Research*, *128*, 175–176.

Hafeman, D. M., Chang, K. D., Garrett, A. S., Sanders, E. M., & Phillips, M. L. (2012). Effects of medication on neuroimaging findings in bipolar disorder: An updated review. *Bipolar Disorders*, *14*, 375–410.

Hajek, T., Bauer, M., Pfennig, A., Cullis, J., Ploch, J., O'Donovan, C., Bohner, G., Klingebiel, R., Young, L. T., Macqueen, G. M., & Alda, M. (2012). Large positive effect of lithium on prefrontal cortex N-acetylaspartate in patients with bipolar disorder: 2-centre study. *Journal of Psychiatry & Neuroscience*, *37*, 185–192.

Haukvik, U. K., Gurholt, T. P., Nerland, S., Elvsåshagen, T., Akudjedu, T. N., Alda, M., Alnæs, D., Alonso-Lana, S., Bauer, J., Baune, B. T., Benedetti, F., Berk, M., Bettella, F., Bøen, E., Bonnín, C. M., Brambilla, P., Canales-Rodríguez, E. J., Cannon, D. M., Caseras, X., … Agartz, I. (2020). In vivo hippocampal subfield volumes in bipolar disorder—A mega-analysis from the enhancing neuro imaging genetics through meta-analysis bipolar disorder working group. *Human Brain Mapping*, *43*, 385–398.

Hibar, D. P., Westlye, L. T., Doan, N. T., Jahanshad, N., Cheung, J. W., Ching, C. R. K. K., Versace, A., Bilderbeck, A. C., Uhlmann, A., Mwangi, B., Krämer, B., Overs, B., Hartberg, C. B., Abe, C., Dima, D., Grotegerd, D., Sprooten, E., Ben, E., Jimenez, E., … Andreassen, O. A. (2018). Cortical abnormalities in bipolar disorder: An MRI analysis of 6503 individuals from the ENIGMA bipolar disorder working group. *Molecular Psychiatry*, *23*, 932–942.

Hibar, D. P., Westlye, L. T., van Erp, T. G. M., Rasmussen, J., Leonardo, C. D., Faskowitz, J., Haukvik, U. K., Hartberg, C. B., Doan, N. T., Agartz, I., Dale, A. M., Gruber, O., Krämer, B., Trost, S., Liberg, B., Abé, C., Ekman, C. J., Ingvar, M., Landén, M., … Andreassen, O. A. (2016). Subcortical volumetric abnormalities in bipolar disorder. *Molecular Psychiatry*, *21*, 1710–1716.

Hoertel, N., Le, S. Y., Lavaud, P., Dubertret, C., & Limosin, F. (2013). Generalizability of clinical trial results for bipolar disorder to community samples: Findings from the National Epidemiologic Survey on alcohol and related conditions. *The Journal of Clinical Psychiatry*, *74*, 265–270.

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). Fsl. *NeuroImage*. *62*, 782–790.

Jeurissen, B., Leemans, A., Tournier, J. D., Jones, D. K., & Sijbers, J. (2013). Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human Brain Mapping*, *34*, 2747–2766.

Kaden, E., Kelm, N. D., Carson, R. P., Does, M. D., & Alexander, D. C. (2016). Multi-compartment microscopic diffusion imaging. *NeuroImage*, *139*, 346–359.

Kato, T. (2022). Mechanisms of action of anti-bipolar drugs. *European Neuropsychopharmacology*, *59*, 23–25.

Kellner, E., Dhital, B., Kiselev, V. G., & Reisert, M. (2016). Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic Resonance in Medicine*, *76*, 1574–1581. https://doi.org/10.1002/mrm.26054

Kelly, S., Jahanshad, N., Zalesky, A., Kochunov, P., Agartz, I., Alloza, C., Andreassen, O. A., Arango, C., Banaj, N., Bouix, S., Bousman, C. A., Brouwer, R. M., Bruggemann, J., Bustillo, J., Cahn, W., Calhoun, V., Cannon, D., Carr, V., Catts, S., … Donohoe, G. (2018). Widespread white matter microstructural differences in schizophrenia across 4322 individuals: Results from the ENIGMA schizophrenia DTI working group. *Molecular Psychiatry*, *23*, 1261–1269.

Lyoo, I. K., Dager, S. R., Kim, J. E., Yoon, S. J., Friedman, S. D., Dunner, D. L., & Renshaw, P. F. (2010). Lithium-induced gray matter volume increase as a neural correlate of treatment response in bipolar disorder: A longitudinal brain imaging study. *Neuropsychopharmacology*, *35*, 1743–1750.

Lyskowski, J., Nasrallah, H. A., Dunner, F. J., & Bucher, K. (1982). A longitudinal survey of side effects in a lithium clinic. *The Journal of Clinical Psychiatry*, *43*, 284–286.

Madre, M., Canales-Rodríguez, E. J., Fuentes-Claramonte, P., Alonso-Lana, S., Salgado-Pineda, P., Guerrero-Pedraza, A., Moro, N., Bosque, C., Gomar, J. J., Ortíz-Gil, J., Goikolea, J. M., Bonnin, C. M., Vieta, E., Sarró, S., Maristany, T., McKenna, P. J., Salvador, R., & Pomarol-Clotet, E. (2020). Structural abnormality in schizophrenia versus bipolar disorder: A whole brain cortical thickness, surface area, volume and gyrification analyses. *NeuroImage: Clinical*, *25*, 102131.

McKinnon, E. T., & Jensen, J. H. (2019). Measuring intra-axonal T2 in white matter with direction-averaged diffusion MRI. *Magnetic Resonance in Medicine*, *81*, 2985–2994.

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., Faria, A. V., Mahmood, A., Woods, R., Toga, A. W., Pike, G. B., Neto, P. R., Evans, A., Zhang, J., Huang, H., Miller, M. I., van Zijl, P., & Mazziotta, J. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *NeuroImage*, *40*, 570–582.

Nelson, H. E., & Willison, J. R. (1991). *The revised national adult reading test*. NFER-Nelson.

Nunes, A., Schnack, H. G., Ching, C. R. K., Agartz, I., Akudjedu, T. N., Alda, M., Alnæs, D., Alonso-Lana, S., Bauer, J., Baune, B. T., Bøen, E., del Bonnin, C. M., Busatto, G. F., Canales-Rodríguez, E. J., Cannon, D. M., Caseras, X., Chaim-Avancini, T. M., Dannlowski, U., Díaz-Zuluaga, A. M., … Hajek, T. (2020). Using structural MRI to identify bipolar disorders—13 site machine learning study in 3020 individuals from the ENIGMA bipolar disorders working group. *Molecular Psychiatry*, *25*, 2130–2143.

Oishi, K., Zilles, K., Amunts, K., Faria, A., Jiang, H., Li, X., Akhter, K., Hua, K., Woods, R., Toga, A. W., Pike, G. B., Rosa-Neto, P., Evans, A., Zhang, J., Huang, H., Miller, M. I., van Zijl, P. C. M., Mazziotta, J., & Mori, S. (2008). Human brain white matter atlas: Identification and assignment of common anatomical structures in superficial white matter. *NeuroImage*, *43*, 447–457.

Patel, Y., Parker, N., Shin, J., Howard, D., French, L., Thomopoulos, S. I., Pozzi, E., Abe, Y., Abé, C., Anticevic, A., Alda, M., Aleman, A., Alloza, C., Alonso-Lana, S., Ameis, S. H., Anagnostou, E., McIntosh, A. A., Arango, C.,

Arnold, P. D., … Paus, T. (2021). Virtual histology of cortical thickness and shared neurobiology in 6 psychiatric disorders. *JAMA Psychiatry*, *78*, 47–63.

Phillips, M. L., & Swartz, H. A. (2014). A critical appraisal of neuroimaging studies of bipolar disorder: Toward a new conceptualization of underlying neural circuitry and a road map for future research. *American Journal of Psychiatry*, *171*, 829–843.

Piredda, G. F., Hilbert, T., Canales-Rodríguez, E. J., Pizzolato, M., von Deuster, C., Meuli. R., Pfeuffer, J., Daducci, A., Thiran, J. P., & Kober, T. (2021). Fast and high–resolution myelin water imaging: Accelerating multi–echo GRASE with CAIPIRINHA. *Magnetic Resonance in Medicine*, *85*, 209–222.

Radua, J., Vieta, E., Shinohara, R., Kochunov, P., Quidé, Y., Green, M. J., Weickert, C. S., Weickert, T., Bruggemann, J., Kircher, T., Nenadić, I., Cairns, M. J., Seal, M., Schall, U., Henskens, F., Fullerton, J. M., Mowry, B., Pantelis, C., Lenroot, R., … Pineda-Zapata, J. (2020). Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *Neuroimage*, *218*, 116956.

Rybakowski, J. K. (2022). Lithium. *European Neuropsychopharmacology*, *57*, 86–87.

Salvador, R., Canales-Rodríguez, E., Guerrero-Pedraza, A., Sarró, S., Tordesillas-Gutiérrez, D., Maristany, T., Crespo-Facorro, B., McKenna, P., & Pomarol-Clotet, E. (2019). Multimodal integration of brain images for MRI-based diagnosis in schizophrenia. *Frontiers in Neuroscience*, *13*, 1203.

Salvador, R., Radua, J., Canales-Rodríguez, E. J., Solanes, A., Sarroa, S., Goikolea, J. M. J. M., Valiente, A., Montea, G. C., Del Natividad, C. M. D. C., Guerrero-Pedraza, A., Moro, N. N., Fernaândez-Corcuera, P., Amann, B. L., Maristany, T., Vieta, E., Mckenna, P. J., & Pomarol-Clote, E. (2017). Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. *PLoS One*, *12*, e0175683.

Sarrazin, S., Poupon, C., Linke, J., Wessa, M., Phillips, M., Delavest, M., Versace, A., Almeida, J., Guevara, P., Duclap, D., Duchesnay, E., Mangin, J. F., Le Dudal, K., Daban, C., Hamdani, N., D'Albis, M. A., Leboyer, M., & Houenou, J. (2014). A multicenter tractography study of deep white matter tracts in bipolar I disorder: Sychotic features and interhemispheric disconnectivity. *JAMA Psychiatry*, *71*, 388–396.

Shioya, A., Saito, Y., Arima, K., Kakuta, Y., Yuzuriha, T., Tanaka, N., Murayama, S., & Tamaoka, A. (2015). Neurodegenerative changes in patients with clinical history of bipolar disorders. *Neuropathology*, *35*, 245–253.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., de Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., de Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*, S208–S219.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*, 83–98.

Steffens, D. C., Chung, H., Krishnan, K. R. R., Longstreth, W. T., Carlson, M., & Burke, G. L. (2008). Antidepressant treatment and worsening white matter on serial cranial magnetic resonance imaging in the elderly: The cardiovascular health study. *Stroke*, *39*, 857–862.

Thompson, P. M., Stein, J. L., Medland, S. E., Hibar, D. P., Vasquez, A. A., Renteria, M. E., Toro, R., Jahanshad, N., Schumann, G., Franke, B., Wright, M. J., Martin, N. G., Agartz, I., Alda, M., Alhusaini, S., Almasy, L., Almeida, J., Alpert, K., Andreasen, N. C., … Drevets, W. (2014). The ENIGMA consortium: Large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging and Behavior*, *8*, 153–182.

Tournier, J. D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C. H., & Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, *202*, 116137.

Tuozzo, C., Lyall, A. E., Pasternak, O., James, A. C. D., Crow, T. J., & Kubicki, M. (2018). Patients with chronic bipolar disorder exhibit widespread increases in extracellular free water. *Bipolar Disorders*, *20*, 523–530.

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

Valdés-Tovar, M., Rodríguez-Ramírez, A. M., Rodríguez-Cárdenas, L., Sotelo-Ramírez, C. E., Camarena, B., Sanabrais-Jiménez, M. A., Solís-Chagoyán, H., Argueta, J., & López-Riquelme, G. O. (2022). Insights into myelin dysfunction in schizophrenia and bipolar disorder. *World Journal of Psychiatry*, *12*, 264–285.

van Velzen, L. S., Kelly, S., Isaev, D., Aleman, A., Aftanas, L. I., Bauer, J., Baune, B. T., Brak, I. V., Carballedo, A., Connolly, C. G., Couvy-Duchesne, B., Cullen, K. R., Danilenko, K. V., Dannlowski, U., Enneking, V., Filimonova, E., Förster, K., Frodl, T., Gotlib, I. H., … Schmaal, L. (2020). White matter disturbances in major depressive disorder: A coordinated analysis across 20 international cohorts in the ENIGMA MDD working group. *Molecular Psychiatry*, *25*, 1511–1525.

Vederine, F. E., Wessa, M., Leboyer, M., & Houenou, J. (2011). A meta-analysis of whole-brain diffusion tensor imaging studies in bipolar disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *35*, 1820–1826.

Veraart, J., Novikov, D. S., Christiaens, D., Ades-aron, B., Sijbers, J., & Fieremans, E. (2016). Denoising of diffusion MRI using random matrix theory. *NeuroImage*, *142*, 394–406.

Veraart, J., Novikov, D. S., & Fieremans, E. (2018). TE dependent diffusion imaging (TEdDI) distinguishes between compartmental T2 relaxation times. *NeuroImage*, *182*, 360–369.

Veraart, J., Nunes, D., Rudrapatna, U., Fieremans, E., Jones, D. K., Novikov, D. S., & Shemesh, N. (2020). Noninvasive quantification of axon radii using diffusion MRI. *eLife*, *9*, e49855. https://doi.org/10.7554/eLife.49855

Vieta, E., Berk, M., Schulze, T. G., Carvalho, A. F., Suppes, T., Calabrese, J. R., Gao, K., Miskowiak, K. W., & Grande, I. (2018). Bipolar disorders. *Nature Reviews Disease Primers*, *4*, 1–16.

Wilkinson, G. S. (1993). *Wide range achievement test* (3rd ed.). Wide Range.

Wollenhaupt-Aguiar, B., Pfaffenseller, B., de Chagas, V. S., Castro, M. A. A., Passos, I. C., Kauer-Sant'Anna, M., Kapczinski, F., & Klamt, F. (2016). Reduced neurite density in neuronal cell cultures exposed to serum of patients with bipolar disorder. *The International Journal of Neuropsychopharmacology*, *19*, pyw051.

Yu, T., Canales-Rodríguez, E. J., Pizzolato, M., Piredda, G. F., Hilbert, T., Fischi-Gomez, E., Weigel, M., Barakovic, M., Cuadra, M. B., Granziera, C., Kober, T., & Thiran, J. P. (2021). Model–informed machine learning for multi–component t2 relaxometry. *Medical Image Analysis*, *69*, 101940.

Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., & Alexander, D. C. (2012). NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, *61*, 1000–1016.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Canales-Rodríguez, E. J., Verdolini, N., Alonso-Lana, S., Torres, M. L., Panicalli, F., Argila-Plaza, I., Rodriguez-Cano, E., Montoro, I., Garcia-Ruiz, B., Jimenez, E., Varo, C., Lluch, A., del Mar Bonnin, C., Maluf, S., Pujol, M., Guarner, N. J., Sarró, S., Vieta, E., Vilella, E., … Pomarol-Clotet, E. (2023). Widespread intra-axonal signal fraction abnormalities in bipolar disorder from multicompartment diffusion MRI: Sensitivity to diagnosis, association with clinical features and pharmacologic treatment. *Human Brain Mapping*, *44*(12), 4605–4622. https://doi.org/10.1002/hbm.26405