UNIVERSITA' VITA-SALUTE SAN RAFFAELE

CORSO DI DOTTORATO DI RICERCA IN NEUROSCIENZE COGNITIVE

COGNITIVE IMPAIRMENT IN MOOD DISORDERS: NEUROPSYCHOLOGY, MULTIMODAL BRAIN IMAGING, AND THE EFFECT OF NEUROINFLAMMATION

DoS: Prof. Francesco Benedetti

Tesi di DOTTORATO di RICERCA di Federico Calesella

matr. 017372

Ciclo di dottorato XXXVI

SSD MED/25

Anno Accademico 2022/2023

UNIVERSITA' VITA-SALUTE SAN RAFFAELE

CORSO DI DOTTORATO DI RICERCA IN NEUROSCIENZE COGNITIVE

COGNITIVE IMPAIRMENT IN MOOD DISORDERS: NEUROPSYCHOLOGY, MULTIMODAL BRAIN IMAGING, AND THE EFFECT OF NEUROINFLAMMATION

DoS: Prof. Francesco Benedetti

Tesi di DOTTORATO di RICERCA di Federico Calesella

matr. 017372

Ciclo di dottorato XXXVI

SSD MED/25

Anno Accademico 2022/2023

RELEASE OF PHD THESIS

I, the undersigned	Federico Calesella
Registration number	017372
Born in	Milano
On	15/07/1994

Author of the PhD Thesis titled

COGNITIVE IMPAIRMENT IN MOOD DISORDERS: NEUROPSYCHOLOGY, MULTIMODAL BRAIN IMAGING, AND THE EFFECT OF NEUROINFLAMMATION

☒ AUTHORIZE the public release of the thesis

☐ DO NOT AUTHORIZE the public release of the thesis for 24 months

from the PhD thesis date, specifically

from ……/……/…… to ……/……/……

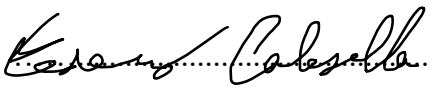
Because:

☐ The whole project or parts of it may be the subject of a patent application;

☐ Parts of the thesis have been or are being submitted to a publisher or are in the press;

☐ The thesis project is financed by external bodies that have rights over it and its publication.

Reproduction of the thesis in whole or in part is forbidden

Date 30/10/2023 Signature 

DECLARATION

This thesis has been:

- composed by myself and has not been used in any previous application for a degree. Throughout the text I use both ‘I’ and ‘We’ interchangeably.

- has been written according to the editing guidelines approved by the University.

Permission to use images and other material covered by copyright has been sought and obtained.

All the results presented here were obtained by myself.

All sources of information are acknowledged by means of reference.

ACKNOWLEDGEMENTS

This thesis has been produced with the financial support of the University Vita-Salute San Raffaele (Milan, Italy), which offered me a doctoral scholarship, and with the support of Unit of Psychiatry and Clinical Psychobiology, Division of Neuroscience, IRCCS San Raffaele Scientific Institute (Milano, Italy), which provided all the necessary means required for this project, in collaboration with CERMAC (Centro Eccellenza Risonanza Magnetica ad Alto Campo, Scientific Institute Ospedale San Raffaele, Milan).

ABSTRACT

The burden of mood disorders continues to rise, with around a third of patients not responding to antidepressants. An accurate and timely diagnosis is essential for proper treatment administration and clinical course. However, due to overlapping symptomatology, approximately 60% of bipolar disorder (BD) patients are initially misdiagnosed as being affected by major depressive disorder (MDD). Furthermore, both disorders also affect cognitive functions independently of mood state, with a significant impact on patients' psychosocial functioning and quality of life. Identifying reliable biomarkers to differentiate between MDD and BD and to detect cognitive deficits at the individual level is crucial to develop personalized treatment strategies.

In this work a wide set of data was collected for 358 depressed patients (139 MDD and 219 BD), including white matter (WM) diffusivity measures, voxel-based morphometry measures, and resting-state functional connectivity (rsFC) features, along with peripheral inflammatory markers (PIMs) and polygenic risk scores. Each feature was entered separately into a machine learning predictive pipeline to differentiate between MDD and BD patients, as well as between cognitively intact and impaired patients.

WM diffusivity, rsFC measures, and PIMs were able to differentiate between MDD and BD patients with accuracy ranging from 62% to 69%. Cognitive deficits were significantly detected only by functional grey matter measures with 67% of accuracy. Results from diagnosis differentiation suggest that BD patients exhibit a subtly greater compromission in WM integrity over a widespread network compared to MDD. Additionally, MDD and BD patients showed distinct seed-based rsFC patterns within the reward system, likely mirroring the behavioural differences between depression and mania. The two diagnostic groups also showed a differential inflammatory profile, with MDD patients also recruiting a regulatory mechanism, contrarily to BD. The identification of patients with cognitive impairment, instead, may rely on a disruption in the cross-talk between the default mode, the executive control, and the salience networks. Overall, our results demonstrate the feasibility of distinguishing BD from MDD patients and of identifying cognitively impaired patients at the individual level using neuroimaging features.

TABLE OF CONTENTS

ACRONYMS AND ABBREVIATIONS

ABC: atlas-based connectivity

ACC: anterior cingulate cortex

aCompCor: anatomical component correction

ACR: anterior corona radiata

AD: axial diffusivity

ADHD: attention deficit hyperactivity disorder

AG: angular gyrus

AIC: anterior limb of internal capsule

ASD: autism-spectrum disorders

ATR: anterior thalamic radiation

AUC: area under the curve

BA: balanced accuracy

BACS: Brief Assessment of Cognition in Schizophrenia

basic FGF: basic fibroblast growth factor

BCC: body of corpus callosum

BD: bipolar disorder

BD-I: type I bipolar disorder

BD-II: type II bipolar disorder

BET: brain extraction tool

BMI: body mass index

CAD: coronary artery disease

CAT12: Computational Anatomy Toolbox

CB: cingulum bundle

CC: corpus callosum

CCL: C-C motif ligand

CG: cingulate gyrus

CH: thalamic portion of cingulum

COX: cyclooxygenase

CP: cerebellar peduncle

CR: corona radiata

CRP: C-reactive protein

CSF: cerebrospinal fluid

CST: corticospinal tract

CV: cross-validation

CXCL = C-X-C motif chemokine ligand

dACC: dorsal anterior cingulate cortex

DAN: dorsal attention network

dlPFC: dorsolateral prefrontal cortex

DMN: default mode network

dmPFC: dorsomedial prefrontal cortex

DRCs: dual regression components

DSM-5: Diagnostic and Statistical Manual of Mental Disorders (5th edition)

DTI: diffusion tensor imaging

EC: external capsule

EN: elastic net

ENIGMA: Enhancing Neuro Imaging Genetics Through Meta Analysis Consortium

ES: equivalent score

FA: fractional anisotropy

fALFF: fractional amplitude of low-frequency fluctuations

FC: functional connectivity

fMRI: functional magnetic resonance imaging (fMRI)

FO: frontal operculum

FOF: fronto-occipital fasciculus

FPCN: fronto-parietal control network

FSL: FMRIB Software Library

FWHM: full width at half maximum

FX: fornix

GCC: genu of corpus callosum

G-CSF: granulocyte colony stimulating factor

GM: grey matter

GM-CSF: granulocyte macrophage colony stimulating factor

GWAS: genome-wide association studies

HALFFpipe: Harmonized AnaLysis of Functional MRI pipeline

HC: healthy controls

HDRS-21: 21-Hamilton Depression Rating Scale

HGF: hepatocyte growth factor

HRC: Haplotype Reference Consortium

IBD: identity by descendent

IC: internal capsule

ICP: inferior cerebellar peduncle

IFGPo: opercular part of the inferior frontal gyrus

IFGPt: triangular part of the inferior frontal gyrus

IFN: interferon

IFOF: inferior fronto-occipital fasciculus

IL: interleukin

IL-1Ra: interleukin-1 receptor antagonist

ILF: inferior longitudinal fasciculus

IPG: inferior parietal gyrus

IPL: inferior parietal lobule

ITG: inferior temporal gyrus

LM: medial lemniscus

LOO: leave-one-out

MAF: minor allele frequency

MCP: middle cerebellar peduncle

M-CSF: macrophage colony stimulating factor

MCG: middle cingulate gyrus

MD: mean diffusivity

MDD: major depressive disorder

mFC: medial frontal cortex

MFG: middle frontal gyrus

MIF: macrophage migration inhibitory factor

MKL: multiple kernel learning

ML: machine learning

MNI: Montreal Neurological Institute

MRI: magnetic resonance imaging

MTG: middle temporal gyrus

NPV: negative predictive value

OCP: occipital pole

OFG: orbitofrontal gyrus

OFuG: occipital fusiform gyrus

PAG: periaqueductal grey

PCG: posterior cingulate gyrus

PCT: pontine crossing tract

PDGF-BB: platelet-derived growth factor-BB

PFC: prefrontal cortex

PGC: Psychiatric Genomics Consortium

PI: posterior insula

PIC: posterior limb of internal capsule

PIM: peripheral inflammatory marker

PoCG: postcentral gyrus

PPV: positive predictive value

PRoNTo: Pattern Recognition for Neuroimaging Toolbox

PRS: polygenic risk score

PTR: posterior thalamic radiation

QC: quality control

rACC: rostral anterior cingulate cortex

RD: radial diffusivity

RDoC: Research Domain Criteria

ReHo: regional homogeneity

RIC: retrolenticular part of internal capsule

ROI: region of interest

rsFC: resting-state functional connectivity

rsfMRI: resting-state functional magnetic resonance imaging

SBC: seed-based connectivity

SCF: stem cell factor

SCGF-β: stem cell growth factor β

SCP: superior cerebellar peduncle

SCR: superior corona radiata

Sens: sensitivity

SFG: superior frontal gyrus

SFOF: superior fronto-occipital fasciculus

SLF: superior longitudinal fasciculus

SMG: supramarginal gyrus

SNP: single nucleotide polymorphism

Spec: specificity

SPG: superior parietal gyrus

SPM: Statistical Parametric Mapping

SN: salience network

SS: sagittal stratum

ST: stria terminalis

STG: superior temporal gyrus

SVM: support vector machine

T2D: type 2 diabetes

TBSS: tract based spatial statistics

TIV: total intracranial volume

TMP: temporal pole

TNF: tumor necrosis factor

UF: uncinate fasciculus

VAN: ventral attention network

VBM: voxel-based morphometry

VEGF: vascular endothelial growth factor

vPFC: ventral prefrontal cortex

vlPFC: ventrolateral prefrontal cortex

vmPFC: ventromedial prefrontal cortex

VS: ventral striatum

VTA: ventral tegmental area

WM: white matter

β-NGF: β-nerve growth factor

LIST OF FIGURES AND TABLES

Tables

Figures

Appendices

1. INTRODUCTION

Mood disorders, often referred to as affective disorders, represent mental health conditions marked by enduring and significant disturbances in a person's emotional state. This broad category primarily includes major depressive disorder (MDD), typified by sustained feelings of sadness and melancholy, and bipolar disorder (BD), defined by dramatic mood swings oscillating between depressive lows and manic highs (American Psychiatric Association, 2013).

These conditions impose a heavy toll on individuals, significantly affecting their cognition, emotions, behaviours, and overall quality of life. As some of the leading contributors to global disability, mood disorders pose a growing personal and socio-economic burden. Roughly one in three patients remain unresponsive to antidepressant treatments, and relapses are common, with around 44% and 70% of individuals experiencing recurrence within 1 and 5 years, respectively (Radua et al., 2017; Wittchen, 2012).

The exploration of multifaceted aspects of mood disorders may provide valuable insights that pave the way for potential advancements in diagnosis, treatment, and understanding of these disorders. This comprehensive understanding is crucial for devising more effective therapeutic interventions and improving the lives of individuals suffering from mood disorders.

1.1. Mood disorders

1.1.1. Major depressive disorder

MDD is the most common form of depressive disorders, and it ranks among the top ten causes of lifelong disability worldwide, particularly in high-income countries (Vos et al., 2017). According to the World Mental Health Survey, the prevalence of MDD across 18 countries is approximately 6%, with variations between different age groups. The median age of onset is around 25 years, and the risk of developing MDD increases up to the early 40s (Bromet et al., 2011). Notably, MDD affects women more frequently than men, with females having a twofold higher risk of developing the disorder.

The core features of MDD encompass one or more major depressive episodes lasting at least two weeks, characterized by persistent sadness, intense mental anguish, and feelings of despair, helplessness, and worthlessness. Anhedonia, the loss of interest and pleasure in daily activities, is also a defining symptom. Additionally, individuals commonly

experience fatigue and loss of energy, making even simple tasks seem overwhelmingly challenging. Impairments in thinking, concentration, and memory are also prevalent, with some cases exhibiting severe memory deficits (American Psychiatric Association, 2013; F. K. Goodwin & Jamison, 2007). Although many patients recover within one year, a depressive episode can last from 13 to 30 weeks on average. On average, a depressive episode lasts from 13 to 30 weeks, and approximately 70-90% of patients recover within one year. However, in outpatient settings, only 25% of patients achieve remission within 6 months, and more than half of patients still experience MDD after 2 years (Otte et al., 2016).

1.1.2. Bipolar disorder

BD is characterized by cycles of mania, hypomania, and depression, which significantly affects personal and social functioning (American Psychiatric Association, 2013). Mania can be defined as a state of elevated or irritable mood with increased activity, lasting at least one week. Symptoms include decreased sleep needs, rapid speech, and distractibility. Inappropriate and potentially harmful behaviours can also be present due to a mixture of increased activity, disinhibition, and grandiosity (American Psychiatric Association, 2013; F. K. Goodwin & Jamison, 2007). Hypomania is a milder form of mania without the severity of symptoms that cause impairment in daily functioning. Additionally, contrarily to mania, hypomania does not cause hospitalization and has a shorter duration (Benazzi, 2007). Depressive states in bipolar disorders are characterized by an inhibition of different aspects of mood, cognition, and behaviour, such as pessimistic mood and feeling of senselessness, significant reduction of mental activity, fatigue, loss of energy and impairment in volition and will. The severity of symptoms can vary from mild impairments in physical and mental activity to severe disruption of cognition associated to delusions, hallucinations and blurring of consciousness (American Psychiatric Association, 2013; F. K. Goodwin & Jamison, 2007).

BD can be classified as type I (BD-I) and type II (BD-II). In BD-I, the manic episode may have been preceded and followed by hypomanic or major depressive episodes, while BD-II is marked by hypomanic and depressive episodes without full-blown mania (American Psychiatric Association, 2013; F. K. Goodwin & Jamison, 2007). The worldwide prevalence of BD-I is estimated at approximately 0.6%, with higher rates in males. In contrast, BD-II manifests in about 0.4% of the population, showing a higher incidence

among females. BD-I manifests earlier with a mean age of onset at 18.4 years, compared to BD-II, which typically emerges at around 20 years of age (Merikangas et al., 2011). Bipolar disorder is characterized by recurrent mood episodes, which often emerge within the first year post initial diagnosis (Radua et al., 2017). Both BD-I and BD-II patients spend about half of the time in a symptomatic condition, with depressive symptoms being much more frequent than manic or hypomanic symptoms (Judd et al., 2002, 2003).

1.2. The problem of differential diagnosis

The differential diagnosis between BD and MDD represents a very complicated clinical task in mood disorders. About 60% of BD patients are initially misdiagnosed as being affected by MDD, and accurate BD diagnoses only come after 5-10 years, on average (F. K. Goodwin & Jamison, 2007; Hirschfeld, 2014). Misdiagnosis, though, causes important effects, such as selection of sub-optimal treatments, leading to potentially poor clinical outcomes and prognosis, as well as greater personal and healthcare costs (Hirschfeld et al., 2003). Currently, the most used criterion for the differentiation between MDD and BD is a positive history of manic or hypomanic episodes (American Psychiatric Association, 2013). However, the identification of previous manic or hypomanic episodes in depressed patients is extremely difficult.

The differential diagnosis between MDD and BD is complicated by the fact that the onset of BD is usually identified by a depressive episode, especially in BD-II (F. K. Goodwin & Jamison, 2007; Phillips & Kupfer, 2013). Another reason for the diagnostic complexity is the high prevalence of depressive episodes in BD (Phillips & Kupfer, 2013). BD patients are also more likely to seek treatment for depression than for mania, making the diagnosis even more challenging (Phillips & Kupfer, 2013). This is because depressive episodes cause subjective distress and functional impairment, whereas manic and hypomanic episodes lack subjective suffering due to the increase in energy, productivity, and creativity (Phillips & Kupfer, 2013).

The complexity in differential diagnosis is further exacerbated by the nuanced characterization of depressive mixed states. These states are delineated by the concurrent presentation of mania or hypomania within a depressive episode. Such manic or hypomanic manifestations may arise amid a depressive episode or manifest in the transitional phase between depression and mania or hypomania. Diagnostic criteria for mixed states necessitate the presence of a minimum of three manic or hypomanic

indicators, such as psychomotor agitation, distractibility, and crowded thoughts (American Psychiatric Association, 2013; Benazzi, 2007). Although mixed states can manifest also in MDD, such episodes are more frequently observed in BD, especially in patients with positive bipolar familiarity (Akiskal & Benazzi, 2003; Vázquez et al., 2018). Thus, the absence of distinct hypomanic or manic states in patients with a history of mixed depression further complicate the correct identification of mania and hypomania, and of BD in turn (Phillips & Kupfer, 2013).

Moreover, subthreshold symptoms of hypomania are common in unipolar depression. For instance, the longitudinal Early Developmental Stages of Psychopathology Study reported that around 40% of patients with a MDD diagnosis manifests subthreshold bipolar symptoms and they are more likely to convert into bipolar disorder during follow-up (Zimmermann et al., 2009). Similar results were reported by the National Comorbidity Survey Replication Study, which demonstrated that individuals with MDD and subthreshold hypomanic symptoms represent 39% of all unipolar patients (Angst et al., 2010). These findings highlight the importance of the accurate identification of subthreshold hypomanic symptoms for a correct diagnosis of BD.

1.3. Biological correlates of mood disorders

1.3.1. Biological correlates in MDD

Twin and family studies conducted over the last 40 years have elucidated the inherent genetic component involved in MDD, as evidenced by a twin heritability of 31%-42% (Sullivan et al., 2000). Candidate gene studies compared allele frequencies in selected genes between MDD cases and healthy controls (HC), yet they often led to controversial and uninformative results (Farrell et al., 2015; Sullivan et al., 2001). For example, 18 significant depressive disorder candidate genes (e.g., SLC6A5, BDNF, COMT, and HTR2A) were scrutinized, but no substantial support for any was found (Border et al., 2019). In another candidate gene study, a gene-environment association was suggested between SLC6A4/HTTLPR, early stress, and risk for MDD. However, subsequent studies, including one with similar methodology and setting, failed to replicate the results (Fergusson et al., 2011). The apparent absence of loci with major effects points to the existence of thousands of loci, each contributing minutely to MDD susceptibility (Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium et al., 2013), suggesting that MDD is a polygenic condition (McIntosh et al., 2019). This change of

perspective, in turn, paved the way to genome-wide association studies (GWASs) to understand the genetic components of MDD. Although 102 common genetic variants associated with MDD have been recognized (Howard et al., 2019), these variants account for only a small proportion of genetic contribution to MDD. However, sample size is extremely important in GWASs, since larger sample sizes increase the chances to find associated variants (McIntosh et al., 2019). The most recent GWAS examined data from 1.2 million individuals, discovering 178 genetic risk locations and 223 independently noteworthy single-nucleotide polymorphisms (SNPs) (Levey et al., 2021). The enlistment of such extensive cohorts was enabled by the employment of straightforward and economical approaches to identify cases, denoted as minimal phenotyping. Given the necessity of substantial samples to achieve robust statistical significance for genetic association, minimal phenotyping strategies has been employed, assuming that association would have been detectable in some of the contributing loci, even with a poorly measured phenotype. On the one hand, this led to the finding of hundreds of significant loci. On the other hand, though, minimal phenotyping led to a loss in specificity, with a large proportion of the identified loci not attributable to MDD (Flint, 2023).

Magnetic resonance imaging (MRI) has also been employed to unravel the neural underpinnings of MDD. A reduction in hippocampal grey matter (GM) volumes in MDD patients has been consistently reported by studies investigating volumetric changes in brain structures involved in the cortico-limbic network (Cole et al., 2011; Schmaal et al., 2016), also involving caudate, globus pallidus, and putamen, along with orbitofrontal (OFG) and anterior cingulate (ACC) cortices (Bora et al., 2012; Kempton et al., 2011). Structural alterations in the amygdala have also been extensively studied, yet leading to conflicting findings (J. P. Hamilton et al., 2008; Kronenberg et al., 2009; Saleh et al., 2012). It is noteworthy, though, that antidepressant medication might influence amygdala volume, with medicated patients exhibiting increased volumes and unmedicated patients showing decreased volumes (J. P. Hamilton et al., 2008). Integrity of white matter (WM) microstructure has also been investigated, showing alterations in tracts within the cortico-limbic network, mostly comprising anterior thalamic radiation (ATR) and cingulum bundle (CB) (W. Jiang et al., 2015; Liang et al., 2019).

Compared to HC, functional MRI (fMRI) studies have observed increased activations in the amygdala, dorsal ACC (dACC), and insula, along with reduced responses in the dorsolateral prefrontal cortex (dlPFC) and dorsal striatum in MDD patients, when processing negative stimuli (Frodl et al., 2010; Matthews et al., 2008). MDD patients also exhibit distinctive resting-state functional connectivity (rsFC) patterns across brain networks (Kaiser et al., 2015). Patients show an augmented representation of the global signal in the default mode network (DMN), reflecting abnormal connectivity within and between networks, potentially influencing the cognitive and emotional symptoms linked to depression (Scalabrini et al., 2020). Additionally, reduced rsFC in the fronto-parietal network has also been observed in these patients (Kaiser et al., 2015).Elevated levels of pro-inflammatory cytokines, such as interleukin (IL)-1β, IL-6, and tumor necrosis factor (TNF)-α, have also been found in individuals with MDD compared to non-depressed controls (Haapakoski et al., 2015; Maes et al., 1997; Mikova et al., 2001; A. H. Miller et al., 2009). Notably, cytokines, produced by immune cells like macrophages, are critical in managing body responses to infection and stress, and their imbalance can promote harmful inflammation (Gordon et al., 2014; Najjar et al., 2013). Findings suggest that increased cytokine levels may correlate with MDD severity and symptomatology, including sleep disturbances and anhedonia, potentially contributing to the development and persistence of depressive states (Motivala et al., 2005; Thomas et al., 2005). Longitudinal studies show that early-life elevated cytokine levels could be predictive of later development of depressive disorders, supporting a potential causal relationship (Au et al., 2015; Gimeno et al., 2009; Khandaker et al., 2014).

Additionally, both human and animal studies indicate that induced inflammation can lead to "sickness behaviour", which shares many symptoms with depression, such as social withdrawal and cognitive deficits. Interestingly, this reaction might be an evolved response intended to promote recovery from illness or infection (Dantzer et al., 2008; Merali et al., 2003; A. H. Miller et al., 2009; A. H. Miller & Raison, 2016).

Furthermore, research exploring anti-inflammatory treatments, like the cyclooxygenase (COX)-2 inhibitor celecoxib, in conjunction with traditional antidepressants has shown promising outcomes, underscoring the potential of targeting inflammation as a novel depression treatment strategy (Köhler-Forsberg et al., 2019; Müller et al., 2006). This

approach suggests that mitigating the pro-inflammatory state in patients with MDD could enhance recovery rates and offer a new therapeutic pathway.

1.3.2. Biological correlates in BD

Several genetic, neural, and inflammatory alterations have been found in BD patients. Twin-based studies have consistently reported high (60–85%) heritability of BD, indicating strong genetic components in its aetiology (Barnett & Smoller, 2009; Smoller & Finn, 2003). The initial endeavours to identify risk loci were based predominantly on a restricted assortment of genetic approaches, primarily genetic linkage analysis (Gordovez & McMahon, 2020; Grover et al., 2009). However, due to their ineffectiveness in deciphering intricate inheritance patterns, linkage methods failed to yield conclusive and consistent results in BD studies (Prathikanti & McMahon, 2001).

Despite some meta-analyses suggesting a modest contribution from a few well-studied candidates genes like SLC6A4 and DAOA (Allen et al., 2008; Detera-Wadleigh et al., 2007; Gatt et al., 2015; Hu et al., 2007; Kraft et al., 2007; Maheshwari et al., 2009; Schulze et al., 2005), the most substantial association evidence emerged from GWASs. The largest one by the Psychiatric Genomics Consortium (PGC) identified 64 independent genomic loci associated with BD (Mullins et al., 2021). The highest observed odds ratio for the identified SNPs at these loci stood at 1.15, corroborating the minor impacts of individual polymorphisms. This discovery facilitated the creation of the polygenic risk score (PRS), a method predicting risk based on genome-wide SNP profiles (Hara et al., 2023).

Current converging data emphasize at least three prominent genes. ANK3, situated on chromosome 10q21.2, emerged as one of the first genes linked to BD through GWASs (Durak et al., 2015; Hayashi et al., 2015; Mahon et al., 2011; Rueckert et al., 2013; E. N. Smith et al., 2009). Multiple studies have now established a significant connection between BD and SNPs in proximity to ANK3, with several of these SNPs influencing the expression of ANK3 (Hayashi et al., 2015; Lim et al., 2014; Lippard et al., 2017; Rueckert et al., 2013; Schulze et al., 2009). ANK3 is responsible for encoding ankyrin B, a protein crucial for axonal myelination, predominantly expressed in the brain among other tissues (Hannon et al., 2015). Additionally, CACNA1C, found on chromosome 12p13, has been linked to BD due to associations with genome-wide significant SNPs. Some of these SNPs are correlated with the expression of CACNA1C in various tissues, including the

nervous system, by encoding an L-type voltage-gated ion channel pivotal for neuronal development and synaptic signalling (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Dao et al., 2010; Ferreira et al., 2008; Gershon et al., 2014; Moskvina et al., 2009; Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011; Stahl et al., 2019). TRANK1, located on chromosome 3p22, has been found to be implicated due to significant associations with adjacent SNPs in BD and schizophrenia studies (D. T. Chen et al., 2013; Ikeda et al., 2018; X. Jiang et al., 2019; Mühleisen et al., 2014; Ruderfer et al., 2014). TRANK1 encodes a largely undefined protein, possibly contributing to the upkeep of the blood–brain barrier (Schiavone et al., 2017). Recent transcriptomic investigations propose that DCLK3, situated in the same 3p22 locus, may also play a contributory role in the risk for both BD and schizophrenia (Gandal, Zhang, et al., 2018; Hara et al., 2023).

Genetic foundations have also been scrutinized for symptoms overlapping between BD and other psychiatric conditions, as well as characteristics mainly observed in BD. The PGC GWAS has disclosed several genomic loci with genome-wide significant associations specifically with BD-I, and not BD-II, including ATP2B2, ATP2A2, CADM2, ANKHD1, GIT2, and B3GLCT (Mullins et al., 2021). Le-Niculescu et al. pinpointed genes like RLP3 and SLC6A4 as crucial pathways in mania (Le-Niculescu et al., 2021). GWASs have also revealed numerous loci related to mood-incongruent psychotic symptoms in severe BD patients (Goes et al., 2012), with recent analyses demonstrating significant associations between specific SNPs like rs9875793 in 3q26.1 and negative mood delusions in BD (Meier et al., 2012). Several genes linked to the circadian clock showed associations with the onset and relapse of BD, including CLOCK, ARNTL1, NPAS2, PER3, and NR1D1 (Etain et al., 2011; McCarthy et al., 2012). Moreover, a recent study utilizing PRSs revealed that insomnia PRS correlated with a heightened risk of BD-II compared to BD-I relative to HC, and hypersomnia PRS was more associated with BD-I than BD-II (K. J. S. Lewis et al., 2020).

However, the SNP-based heritability disclosed in the PGC3 BD GWAS was ~20%, which is much lower than the 65–80% reported in epidemiological studies, implying numerous unidentified genetic factors (Mullins et al., 2021). The Brainstorm consortium has shown significant correlations between BD and other psychiatric disorders, further supporting the notion of BD's clinical and genetic/biological heterogeneity (Brainstorm Consortium

et al., 2018; Gandal, Haney, et al., 2018). The PGC BD GWAS corroborated these correlations, revealing the genetic proximity of BD-I to schizophrenia and BD-II to MDD (Mullins et al., 2021).

GM alterations have been widely studied in BD patients, showing reduced volumes in cortical and subcortical regions related to fronto-limbic neural circuits (Phillips & Swartz, 2014; Teixeira et al., 2019). Among the subcortical regions, amygdala and hippocampus have been highlighted as having reduced volumes in BD patients compared to HC (Foland-Ross et al., 2012; Wijeratne et al., 2013). Despite evidence of volume reductions, other studies have identified increased volumes in these same regions (Lisy et al., 2011). However, a potential confounding variable that might account for the variability in the findings could be the extent of exposure to mood stabilizers like lithium, which is believed to enhance GM volume, potentially through neurogenesis (Zung et al., 2016). Volume alterations in BD patients have also been reported in cortical fronto-insular regions, including ACC, ventral prefrontal cortex (vPFC), superior temporal gyrus (STG), and insular structures (Bora et al., 2010; Foland-Ross et al., 2011; Hanford et al., 2016).

Several alterations in WM microstructure in limbic, frontal, parietal, and fronto-occipital systems have been observed in BD (Vai et al., 2019). Disrupted tracts encompass the fornix (FX), CB, corpus callosum (CC), and corona radiata (CR) (Benedetti et al., 2011; Emsell et al., 2013). Moreover, measures of interhemispheric integration and efficiency were significantly attenuated in the anterior section of the CC, compared to HC (Leow et al., 2013).

Functional abnormalities within fronto-limbic structures are believed to underlie the emotional and mood irregularities seen in BD (Phillips et al., 2008; Strakowski et al., 2005). This mechanism is characterized by an imbalance between overactivity of the emotional ventral system, comprising the amygdala, insula, ventral striatum (VS), ACC, and prefrontal cortex (PFC), and the underactivity of the dorsal system, which comprises the hippocampus, dACC, and dlPFC (Vai et al., 2019). Meta-analytic evidence has pointed out a reduced activation of the ventrolateral PFC (vlPFC) accompanied by increased activation in limbic structures (C.-H. Chen et al., 2011; Delvecchio et al., 2013). Additionally, task-based functional connectivity (FC) between dlPFC and amygdala has been found to be weaker in BD patients, compared to HC, during emotion-regulation tasks. Several rsFC alterations have also been reported between the PFC, ACC, and

mesolimbic regions, such as amygdala, thalamus, and insula (Anand et al., 2009; Anticevic et al., 2013; Chai et al., 2011; L. Chen et al., 2018; Chepenik et al., 2010; Ellard et al., 2018; M. Li et al., 2015; C.-H. Liu et al., 2012; Lois et al., 2014; Ongür et al., 2010). Intra-network connectivity is disrupted as well in the DMN and limbic networks (Doucet et al., 2017; Wang et al., 2017).

Research has delved deeply also into the role of inflammation and immune system dysregulation in BD, revealing intricate details through various studies and experiments. One theory suggests a significant role of impaired T cell regulation in the etiopathogenesis of BD and MDD. While both disorders exhibit T cell deficits, the timing differs. MDD patients show early impaired maturation of Th2 and Th17 cells, worsening with age, characterized by reduced T regulatory cells and heightened inflammatory monocyte activation. In contrast, BD patients have fewer regulatory T cells early in life, with levels normalizing as they age (Drexhage et al., 2011; Grosse et al., 2016).

Focusing on cytokines, several studies found increased IL-6 levels in BD patients irrespective of the disease phase (G. Anderson & Maes, 2015; Benedetti et al., 2020). Not only elevated IL-6, but also higher TNF-α levels and reduced anti-inflammatory IL-4 have been reported in BD patients compared to controls (Kim et al., 2007). Interestingly, mood stabilizers appeared to lower IL-6 levels after six weeks of treatment, pointing to their potential immune-regulating effects.

However, TNF-α research yields inconsistent results. Contrasting findings exist regarding TNF-α levels in BD patients. Elevated levels of TNF-α were found in both manic and depressed BD patients when compared to HC (O'Brien et al., 2006; Ortiz-Domínguez et al., 2007). However, other research did not observe these differences (Brietzke et al., 2009). Notably, a meta-analysis indicated higher levels of TNF-α and sTNF-R1 in manic individuals when contrasted with euthymic patients and HC, as well as increased TNF-α in depressed patients in comparison to the control group, pointing to the potential of TNF-α as a marker for different BD states (Munkholm et al., 2013).

C-reactive protein (CRP) levels have been observed to vary in different phases of BD, diminishing by 42% in depressive states compared to euthymic periods and 48% in comparison to manic episodes (Jacoby et al., 2016). This observation is reinforced by reports of heightened CRP in BD patients, especially during manic episodes, with a moderate decrease upon transitioning to euthymia (Fernandes et al., 2017). Moreover,

elevated baseline CRP in men experiencing depression was significantly linked with the onset of manic symptoms over a two-year span (Becking et al., 2013).

In terms of adaptive cytokines, the research presents divergent results. No substantial disparities were noted in IL-2 and INF-γ levels between BD patients and HC, whereas another study indicated elevated IL-2 levels in HC compared to both manic and depressed BD subjects (Kim et al., 2007; Ortiz-Domínguez et al., 2007). When considering chemokines, higher plasma levels of CCL11, CCL24, and CXCL10, along with diminished CXCL8, were documented in BD-I patients in contrast to controls, suggesting a possible chemokine imbalance as an inherent feature of BD due to their function in immune reactions and neurogenesis (Barbosa et al., 2013; Teixeira et al., 2018).

Furthermore, lithium therapy in euthymic BD patients was found to restore normal cytokine levels, potentially through the inhibition of GSK-3, which is instrumental in controlling cytokine creation and immune responses. This insight is crucial given the link between manic episodes and a rise in CRP and pro-inflammatory cytokines (Fernandes et al., 2016; van den Ameele et al., 2016).

1.4. Cognitive deficits in mood disorders

Beyond emotional disturbances, mood disorders also affect cognitive functioning (Douglas et al., 2018; Rock et al., 2014), which is considered as a core feature by the Research Domain Criteria (RDoC, Pan et al., 2019). Both MDD and BD patients frequently experience compromised performance in a wide span of cognitive domains, such as attention, memory, executive functions, and processing speed (Bortolato et al., 2016; Bourne et al., 2013; Rock et al., 2014). Meta-analytic evidence involving 644 individuals experiencing their first episode of depression in comparison to 570 HC, revealed deficiencies in psychomotor speed, attention, and executive functions, including attentional shifting, verbal fluency, and cognitive flexibility (R. S. C. Lee et al., 2012). Longitudinal studies have also shown that, compared to HC, mood disorder patients are at an increased risk of developing mild cognitive impairment and dementia, as a function of depressive episodes (Kessing & Andersen, 2017; Varghese et al., 2022). Crucially, cognitive dysfunction has been shown to be associated to poor response to treatment and a predictor of occupational and social functional impairment, with significant impacts on quality of life of the patients (Douglas et al., 2018; Pan et al., 2019; Poletti, Mazza, et al., 2021; Rock et al., 2014). Cognitive deficits also persist beyond acute depressive phases

(Rock et al., 2014), suggesting that they may represent a core feature of the disorder rather than being merely epiphenomenal to mood symptoms.

The precise neurobiological mechanisms driving cognitive impairments in depression remain incompletely understood. However, various levels of analysis, including genetics, molecular factors, cellular processes, and neural circuits, have all been implicated (Pan et al., 2019). The prevailing hypothesis suggests disruptions in the structure, function, and interconnections of brain circuits associated with cognitive control and function in MDD patients. Structural and functional alterations of the fronto-temporal and fronto-subcortical circuitry have been suggested to be key in the emergence of cognitive symptoms in depression, involving regions whose alterations play an important role in depression, such as hippocampus, amygdala, and ACC (Jiao et al., 2011; Pizzagalli, 2011). Executive functions can be affected by abnormalities in dorsal cognitive networks, comprising dACC, dlPFC, and hippocampus. Similarly, the ventral affective networks, which include the perigenual ACC, the amygdala, and the OFC, alongside the hippocampus, play a significant role in multiple cognitive tasks. These tasks range from determining salience, planning, working memory, and executive functions (Kheirbek & Hen, 2011).

Cognition is also supported by both interconnected and segregated mechanisms, depending on a balanced relationship between combined and separate cognitive structures. For instance, anti-correlation, meaning the selective alternating activation and deactivation of certain brain areas, is vital for optimal cognitive performance (Cha et al., 2014). Disruptions in the usual balance between key structures within the DMN might lead to the cognitive deficits and diminished cognitive proficiency seen in individuals with depression (J. P. Hamilton et al., 2011).

Furthermore, cognitive impairment in mood disorders has been linked to WM alterations in widespread set of tracts, comprising CR, CC, inferior (ILF) and superior (SLF) longitudinal fasciculi, inferior fronto-occipital fasciculus (IFOF), and CB (Melloni et al., 2019; Poletti et al., 2015). Additionally, alterations in some of these tracts have also been found to negatively mediate the effects of duration of illness on the cognitive outcome (Melloni et al., 2019).

Abnormalities in neural circuitry may also stem from imbalances in hormonal regulation, such as insulin resistance and glucocorticoid irregularities, dysregulation of neurotrophins

like brain-derived neurotrophic factor, heightened immunoinflammatory responses, and oxidative stress (Andreazza, 2012; M. Li et al., 2011; McAfoose & Baune, 2009; Ryan et al., 2012). Type 2 diabetes mellitus (T2D) and insulin resistance have been linked to cognitive impairments in mood disorders, through alterations in proapoptotic intracellular signalling cascades (Baker et al., 2011; McIntyre et al., 2009). Furthermore, activation of proinflammatory cytokines might contribute to the development of cognitive symptoms (Bortolato et al., 2015; Pan et al., 2019; Poletti, Mazza, et al., 2021).

High levels of CRP in patients have been linked to impairments in memory and attention (Chung et al., 2013). General cognitive dysfunction has been associated also to higher peripheral levels of IL-6, CD40 ligand, IL-1 receptor antagonist (IL-1Ra) (Barbosa et al., 2018; Doganavsargil-Baysal et al., 2013; Hope et al., 2015). TNF-α has been shown to have a positive relationship with inhibitory control (Barbosa et al., 2012) and a negative association with verbal memory (Doganavsargil-Baysal et al., 2013). Furthermore, TNF-α antagonists have been shown to mitigate cognitive impairment (Bortolato et al., 2015).

1.5. Precision psychiatry and machine learning techniques

Precision psychiatry represents an emerging approach, emphasizing the integration of several biological facets to tailor therapeutic interventions to the individual patient. Rather than utilizing the traditional one-size-fits-all model, precision psychiatry seeks to individualize treatment based on a deeper understanding of the underlying mechanisms and factors influencing psychiatric disorders (Fernandes et al., 2017; Insel & Cuthbert, 2015). Despite the efforts invested in defining the biological underpinnings of mood disorders have been key in the development of the field and in promoting the shift towards a personalized framework, the applied statistical approach limits the translational impact of the findings into the clinical practice (Nielsen et al., 2020; Vai et al., 2020). The cause resides in the application of average-group univariate statistics, which does not allow the definition of a generalizable predictive function (Vai et al., 2020). Conversely, machine learning (ML) techniques enable the prediction at the single-subject level, providing estimates of the algorithm generalisation ability in out-of-sample observations (Nielsen et al., 2020; Walter et al., 2019).

There are several methods to assess the generalisation ability of models (Varoquaux et al., 2017). A first strategy is the hold-out validation, in which the available data is split into two parts: a training set and a test set. The model is trained on the training set and

evaluated on the test set. This procedure allows to evaluate the model's predictive performance on observations it was not trained on. Another common strategy to assess the generalisation ability is the k-fold cross-validation (CV) technique, which consists in partitioning the dataset into k folds. The model is then iteratively trained on all the folds but one, which is left out as a test set. At each iteration, the model is trained on the joint $k-1$ folds (i.e., training set) and applied on the k^{th} left-out fold (i.e., test set). At the end of the process, all the folds have been used as a test set once, obtaining a predicted value for each observation. This allows to assess the generalization ability of the model on all the observations when they are not included in the training phase of the model but instead used as out-of-sample observations. Another CV scheme is the leave-one-out (LOO) CV, which is a special case of k-fold CV in which $k = n$, where n is the number of observations. In this case, at each iteration, one observation is left out as a test set, and the model is trained on the remaining samples. Despite the advantage of LOO CV in maximizing the sample size of the training set, it has been shown to be less reliable than k-fold CV (Varoquaux et al., 2017).

Another crucial aspect in precision psychiatry is the combination of several sources of information. To create predictive models that can generalize to previously unseen observations, the greatest amount of information possible is needed. For this reason, the framework of precision psychiatry requires the integration of features, information, and knowledge from multiple systems and at different scales (Fernandes et al., 2017). However, such an integration leads to the creation of large datasets, where observations are described in a very high dimensional space. Univariate statistics, though, are underpowered to deal with very high dimensional data, whereas ML is well equipped to handle such scenario (Breiman, 2001; Friedman et al., 2001). Nonetheless, when in a ML framework, high dimensional data still poses the threat of overfitting. Overfitting consists in extracting a function that can perfectly describe the training data, but cannot be generalised to unseen observations (Guyon & Elisseeff, 2003; Hua et al., 2009).

Some possible ways to mitigate this risk are supervised and unsupervised feature reduction methods, as well as embedded feature selection methods (e.g., elastic net penalized regression – EN), or kernel methods (e.g., support vector machine – SVM) (Mwangi et al., 2014). Additionally, though, these methods often require some free hyper-parameters to be set. The optimal value for the hyper-parameter is usually unknown,

requiring a tuning over a set of possible values. To find the hyper-parameters' values that best generalize to unseen data, a nested CV scheme is required. In nested CV, the training set is further iteratively split into training and test sets, creating an inner CV loop. With the application of this scheme, hyper-parameters can be optimized, and overfitting can be avoided, yet still enabling assessment of out-of-sample generalization ability of the model. For these reasons, ML procedures seem particularly suitable to overcome the limits of average-groups univariate statistics and to promote the implementation of tools that can be effectively used in clinical practice (Fernandes et al., 2017; Varoquaux et al., 2017).

1.6. Problem statement

Despite the availability of new interventions, the burden of mood disorders is still growing, with roughly one third of patients being unresponsive to antidepressant (Radua et al., 2017; Wittchen, 2012). An accurate and timely diagnosis is essential for a proper treatment and clinical course. However, the current diagnostic methods rely on the clinical assessment of symptoms, which are prone to misdiagnosis or selection of sub-optimal treatments, leading to potentially poor clinical outcomes as well as greater personal and healthcare costs (de Almeida & Phillips, 2013). This is particularly striking when considering the differential diagnosis between MDD and BD, which is based on a positive history of manic or hypomanic episodes (Han et al., 2019). In general, BD is characterised by a higher prevalence of depressive symptoms than hypomanic or manic ones, and its onset is usually identified by a depressive episode (Grande et al., 2016; Phillips & Kupfer, 2013). Due to this overlapping psychopathology, in particular in the early phases, approximately 60% of depressed BD patients are initially misdiagnosed as being affected by MDD and wait on average 5–10 years for a correct diagnosis (G. M. Goodwin, 2012; Hirschfeld et al., 2003).

Furthermore, mood disorders also affect cognitive functioning (Douglas et al., 2018; Rock et al., 2014). Both MDD and BD patients frequently experience compromised performance in a wide span of cognitive domains, such as attention, memory, executive functions, and processing speed (Bortolato et al., 2016; Bourne et al., 2013; Rock et al., 2014). Cognitive symptoms have been reported in acute episodes, in remission and euthymia (Cullen et al., 2016; Halvorsen et al., 2012; Poletti et al., 2014), significantly affecting the patients' outcome in psychosocial functioning (e.g., workplace

performance) and quality of life (McIntyre, Soczynska, et al., 2015; McIntyre, Xiao, et al., 2015). Additionally, cognitive dysfunction has been linked to poor response to treatment. Together, these aspects suggest that cognitive impairment represents a core dimension of mood disorders, yet lacking proper treatment (Poletti et al., 2014, 2017).

Consequently, to administer the optimal treatment and find novel therapeutic targets, it is crucial to identify reliable biomarkers capable of differentiating between MDD and BD patients and that can address the cognitive impairments in mood disorders. Previous studies have investigated these aspects, highlighting several abnormalities in genetics, inflammatory, as well as both structural and functional neuroimaging markers. However, most of these results were obtained using group-average statistics, limiting the translational impact of the findings.

2. AIM OF THE WORK

The aim of the present research project is twofold. The first objective is to investigate the biological substrate of mood disorders, and how it is differentially expressed in MDD and BD patients. The second objective is to identify the biomarkers capable of stratifying between cognitively impaired and cognitively intact mood disorder patients. To fulfil these objectives, we created ML predictive models, for the differentiation between depressed MDD and BD patients, as well as the identification of patients with cognitive deficits. The models were based on a wide amount of information, comprising voxel-based morphometry (VBM) measures, diffusion tensor imaging (DTI)-derived WM-integrity measures, resting-state fMRI (rsfMRI) features, peripheral inflammatory markers (PIMs), and PRS for several disorders and dimensions.

Importantly, all the methodological choices were oriented to minimize biases and maximize the generalizability of the results. State-of-the-art preprocessing pipelines were implemented to avoid information spillovers across subjects and to ensure the best standards for quality and replicability. Since MRI data were collected with two different scanners, a harmonization step was performed on neuroimaging features to mitigate potential discrepancies arising from the use of different equipment, which could compromise the consistency and reproducibility of downstream analyses and findings (Fortin et al., 2017, 2018; Leek et al., 2010; Yi et al., 2018). This harmonization was executed employing ComBat (Johnson et al., 2007), which has been proven to effectively remove the information related to the scanner (Fortin et al., 2017, 2018; Yi et al., 2018). All the features were then entered into ML predictive pipelines to obtain cross-validated predictions and estimates of the model's generalization accuracy. The ML pipeline was based on EN for vector-like data, whereas it was based on SVM for image- and matrix-like data. To concatenate multiple images or features a multiple kernel learning (MKL) algorithm was employed. Notably, we also investigated the features that relevantly contributed to the predictive models, thus integrating the predictive framework with an inferential approach.

The exploration of such a wide variety of aspects in an integrated and predictive framework is novel and it is fundamental to provide a comprehensive understanding of mood disorders with a stunning impact on patients' lives, bysensitivi bridging the gap between scientific knowledge and clinical practice.

3. RESULTS

3.1. Descriptive statistics

The sample was composed by 358 depressed patients (123 males; 235 females; mean age 48.17±10.35), including 139 MDD and 219 BD depressed patients (Table 1). MDD and BD samples did not significantly differ for age, sex, years of education, duration of illness, and medication load. Contrarily, a significant difference between MDD and BD patients was found for the number of episodes and the age of onset, with MDD patients showing fewer episodes and later onset.

Table 1. Descriptive statistics of the whole sample

Variable	Average ± s.d. / N			t/χ	p-value
	Total (N=358)	MDD (N=139)	BD (N=219)		
Age	48.17 ± 10.34	49.53 ± 10.08	47.36 ± 10.41	1.91	0.057
Sex	235 F / 123 M	95 F / 44 M	140 F / 79 M	0.55	0.457
Years of education	12.20 ± 3.78	12.55 ± 3.85	11.97 ± 3.71	1.41	0.158
Duration of illness	17.61 ± 10.82	17.11 ± 10.65	17.91 ± 10.91	-0.67	0.501
No. of episodes	10.00 ± 12.10	5.42 ± 5.86	12.94 ± 14.01	-5.98	<0.001*
Age of onset	30.58 ± 10.99	32.37 ± 12.18	29.45 ± 10.00	2.46	0.014*
Pharmacological load	4.58 ± 2.22	4.59 ± 2.21	4.57 ± 2.23	0.08	0.933

Average ± standard deviation is reported for continuous variables, whereas frequency was reported for categorical variables. Abbreviations: BD, bipolar disorder; F, female; M, male; MDD, major depressive disorder; s.d., standard deviation. * p<.05

The cognitive performance was assessed through the Brief Assessment of Cognition in Schizophrenia (BACS). For each cognitive domain, the raw scores were corrected for demographic variables, according to their effect on the specific domain. An equivalent score (ES) was then calculated on the corrected score and dichotomized as impaired and not impaired. Performance of the total sample for each cognitive domain of the BACS is reported in Table 2.

Table 2. Cognitive performance.

Domain	Raw score	Corrected score	Equivalent score
Verbal memory	35.88 ± 10.8	41.58 ± 10.21	2.08 ± 1.48
Working memory	17.01 ± 5.31	17.49 ± 5.25	1.60 ± 1.51
Psychomotor coordination	69.68 ± 18.16	72.13 ± 17.97	1.43 ± 1.50
Verbal fluency	43.70 ± 14.68	43.86 ± 14.33	2.04 ± 1.43
Processing speed	37.08 ± 12.83	39.82 ± 12.07	1.00 ± 1.29
Executive functions	14.72 ± 4.07	15.11 ± 3.98	2.12 ± 1.56

Average ± standard deviation is reported for raw, corrected, and equivalent scores, in each domain.

As a main target for cognitive impairment, a composite score was calculated across all the domains, capturing an overall global impairment (i.e., impaired vs. not impaired). Descriptive statistics highlighted that 32.74% of the patients had an overall global

cognitive impairment. The cognitively impaired group was significantly older, with fewer yeas of education, and lower pharmacological load (Table 3).

Table 3. Descriptive statistics of the subsample with BACS comparing cognitively intact and cognitively impaired patients at the composite score.

Variable	Average ± s.d. / N			t/χ	p-value
	Total (N=168)	Not impaired (N=113)	Impaired (N=55)		
Age	47.77 ± 10.77	46.08 ± 10.29	51.24 ± 10.91	-2.97	0.003*
Sex	103 F / 65 M	75 F / 38 M	28 F / 27 M	3.11	0.078
Diagnosis	79 BD / 89 MDD	51 BD / 62 MDD	28 BD / 27 MDD	0.29	0.590
Years of education	13.01 ± 3.66	13.56 ± 3.59	11.89 ± 3.55	2.82	0.005*
Duration of illness	17.89 ± 11.40	17.17 ± 11.00	19.36 ± 12.03	-1.17	0.244
No. of episodes	7.94 ± 8.66	7.48 ± 8.08	8.89 ± 9.69	-0.99	0.324
Age of onset	29.88 ± 11.22	28.91 ± 10.34	31.87 ± 12.59	-1.61	0.110
Pharmacological load	4.49 ± 2.01	4.17 ± 1.96	5.15 ± 1.97	-3.00	0.003*

Average ± standard deviation is reported for continuous variables, whereas frequency was reported for categorical variables. Abbreviations: BD, bipolar disorder; F, female; M, male; MDD, major depressive disorder; s.d., standard deviation. * p<.05

The descriptive statistics for all the single domains are reported in Appendix 1-6. Notably, 89.29% of the patients was impaired in at least on cognitive domain and 79.17% of the patients were characterized by impairment in more than one domain. The domain with the highest frequency of impairment was processing speed (73.81%), followed by working memory (57.74%), coordination (56.55%), verbal memory (41.07%), executive functions (37.50%), and verbal fluency (36.31%).

3.2. Combat assessment

To assess whether the ComBat algorithm effectively removed the information related to the MRI scanners from neuroimaging features, the accuracy of scanner-classification models was compared between pre- and post-correction (Figure 1A). Before correction, the models were very efficient in identifying the scanner with which observations were acquired, achieving a balanced accuracy (BA) in the 94.00-100.00% range. After correction, the ability of the models to differentiate between scanners dropped to a BA between 25.00% a 56.31%. Predictive accuracy of the rsfMRI features decreased by 42.00 percentage points for atlas-based connectivity (ABC), 44.00 points for dual regression components (DRCs) and seed-based connectivity (SBC), and 73.00 and 75.00 points for fractional amplitude of low-frequency fluctuations (fALFF) and regional homogeneity (ReHo), respectively. Considering DTI diffusivity measures, predictive accuracy decreased by 68.12 percentage points for axial diffusivity (AD), 66.06 points for

fractional anisotropy (FA), 63.81 points for mean diffusivity (MD), and 63.30 points for radial diffusivity (RD). VBM accuracy in scanner differentiation decreased by 43.69 percentage points. Complete description of the performance of the models is provided in Appendix 7.

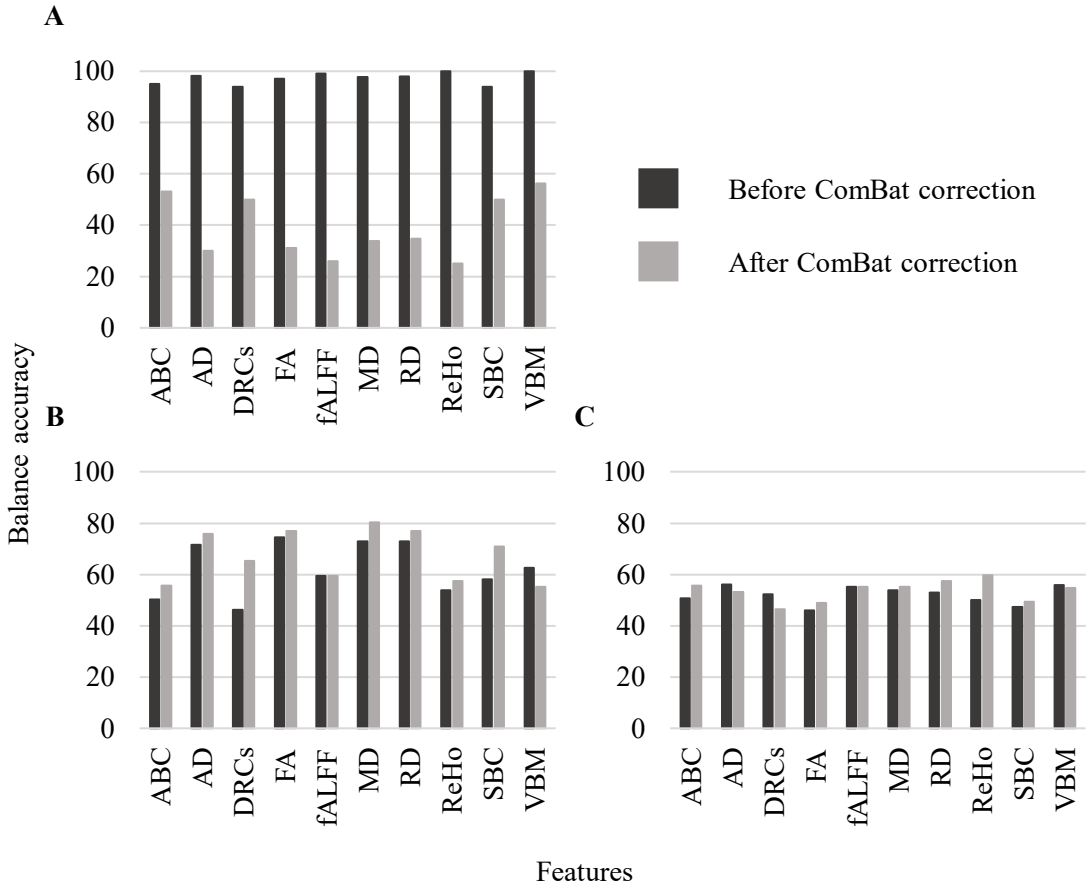


Figure 1. BA of the classification for (A) differentiation between the two scanners, (B) MDD vs. BD, and (C) cognitively intact vs. impaired. Results are reported for both pre- and post-ComBat correction, and for each MRI-derived feature.
Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; BA, balance accuracy; DRCs, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity, RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed-based connectivity; VBM, voxel-based morphometry.

To make sure that the correction process did not remove effects of interest, we also created pre- and post-correction models predicting the targets of interest (i.e., diagnosis and composite score at the BACS). No covariates were introduced in these models, in the attempt to isolate the effect of correction. In the models differentiating between MDD and BD (Appendix 8), only VBM showed a decrease in predictive performance, losing 7.25 percentage points of BA after ComBat correction. On the other features, the correction led to an improvement in the predictive accuracy, showing an increase of 5.52

percentage points for ABC, 19.11 points for DRCs, 0.13 points for fALFF, 3.40 points for ReHo, 12.75 points for SBC, and 4.33, 2.47, 7.31, and 3.90 points for AD, FA, MD, and RD respectively (Figure 1B).

Concerning the identification of cognitively impaired patients (Appendix 9), several features showed a growth in model's performance after the correction, with an increase of 5.00 percentage points for ABC, 9.76 points for ReHo, 1.91 points for SBC, 2.89 points for FA, 1.33 points for MD, and 4.55 points for RD. Contrarily, DRCs, AD, and VBM showed a reduction in predictive performance after correction of 5.95, 3.00, and 1.00 percentage points, respectively. The predictive accuracy of fALFF was the same between before and after correction (Figure 1C).

3.3. Diagnosis differentiation

3.3.1. Performance of the models

The results of the models differentiating between MDD and BD patients are summarised in Table 4. The best model was the MD-based model, with BA of 70.05%, sensitivity of 67.57%, specificity of 72.52%, negative predictive value (NPV) and positive predictive value (PPV) of 71.32% and 69.03% respectively, and 0.76 of area under the curve (AUC). Overall, all the DTI diffusivity measures reached an accuracy around 70%, except for FA that achieved a BA of 63.24%. Also the models based on SBC maps and PIMs reached a good performance, with BA of 66.22% and 62.39% respectively. The 1000-permutation test revealed that these models generated predictions significantly more accurate than a distribution of null models, whereas the other features could not significantly differentiate between diagnoses.

Table 4. Performance metrics for the models differentiating between MDD and BD patients.

Feature	Algorithm	BA	Sens	Spec	NPV	PPV	AUC	p-value	FDRq-values
MD	SVM	70.05	67.57	72.52	71.32	69.03	0.76	0.001	0.002*
RD	SVM	69.62	63.67	75.57	72.31	67.90	0.75	0.001	0.002*
AD	SVM	69.07	66.71	71.43	70.79	68.47	0.78	0.001	0.002*
SBC	MKL	66.22	69.36	63.08	65.47	68.05	0.71	0.002	0.003*
FA	SVM	63.24	61.62	64.86	64.56	62.61	0.73	0.001	0.002*
PIMs	EN	62.39	61.40	63.38	57.38	67.16	0.62	0.008	0.011*
fALFF	SVM	57.37	63.33	51.41	56.76	58.33	0.58	0.071	0.085
ReHo	SVM	56.22	64.23	48.21	54.71	58.24	0.61	0.098	0.103
VBM	SVM	52.27	62.73	41.82	52.50	51.63	0.53	0.224	0.209
Combined	MKL	51.19	59.05	43.33	52.22	48.96	0.51	0.343	0.288

PRS	EN	50.61	40.20	61.02	64.06	37.11	0.51	0.659	0.426
BACS	EN	50.50	80.77	20.22	47.01	54.55	0.50	0.769	0.461
DRCs	MKL	49.23	45.51	52.95	49.54	49.16	0.50	0.553	0.400
ABC	SVM	48.46	51.92	45	48.75	47.74	0.46	0.571	0.400

For each feature, accuracy metrics and p-value at the permutation test are reported. Models are sorted from the best to the worst model, considering BA. Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; AUC, area under the curve; BA, balance accuracy; BACS, brief assessment of cognition in schizophrenia; DRCs, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity; NPV, negative predictive value; PIMs, peripheral inflammatory markers; PPV, positive predictive value; PRS, polygenic risk score; RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed-based connectivity, Sens, sensitivity; Spec, specificity; VBM, voxel-based morphometry. * FDR corrected $p<0.05$

3.3.2. Feature importance of the DTI-based models for diagnosis differentiation

Feature importance analyses revealed that the DTI-based models comprised a widespread and mostly shared pattern across diffusivity measures, with BD patients showing average higher values in all the tracts and all the diffusivity measures.

The AD model (Figure 2A) relied on the contribution of the CC, the middle cerebellar peduncle (MCP), the pontine crossing tract (PCT), the bilateral CB, corticospinal tract (CST), internal (IC) and external (EC) capsule, sagittal stratum (SS), and superior corona radiata (SCR), the left superior cerebellar peduncle (SCP) and uncinate fasciculus (UF), and the right fronto-occipital fasciculus (FOF).

In the FA model (Figure 2B), the differentiating tracts were the FX, the genu of CC (GCC), the MCP and left cerebellar peduncle (CP), the bilateral CB, CST, medial lemniscus (LM), posterior thalamic radiation (PTR), superior fronto-occipital fasciculus (SFOF), SLF, and UF, and the right IC and EC, anterior CR (ACR), SS, and stria terminalis (ST).

The tracts that contributed to the prediction in the MD-based model were the GCC, the MCP and left CP, the bilateral IC, EC, CB, CST, the left LM, CR, and UF, and the right PTR, ST, and SFOF (Figure 2C).

The RD model revealed that the predictive tracts were the GCC, the FX, the MCP and bilateral inferior CP (ICP), the bilateral CB, ST, and UF, the right ACR, IC, CST, PTR, and SFOF, and left EC, LM, and SLF (Figure 2D).

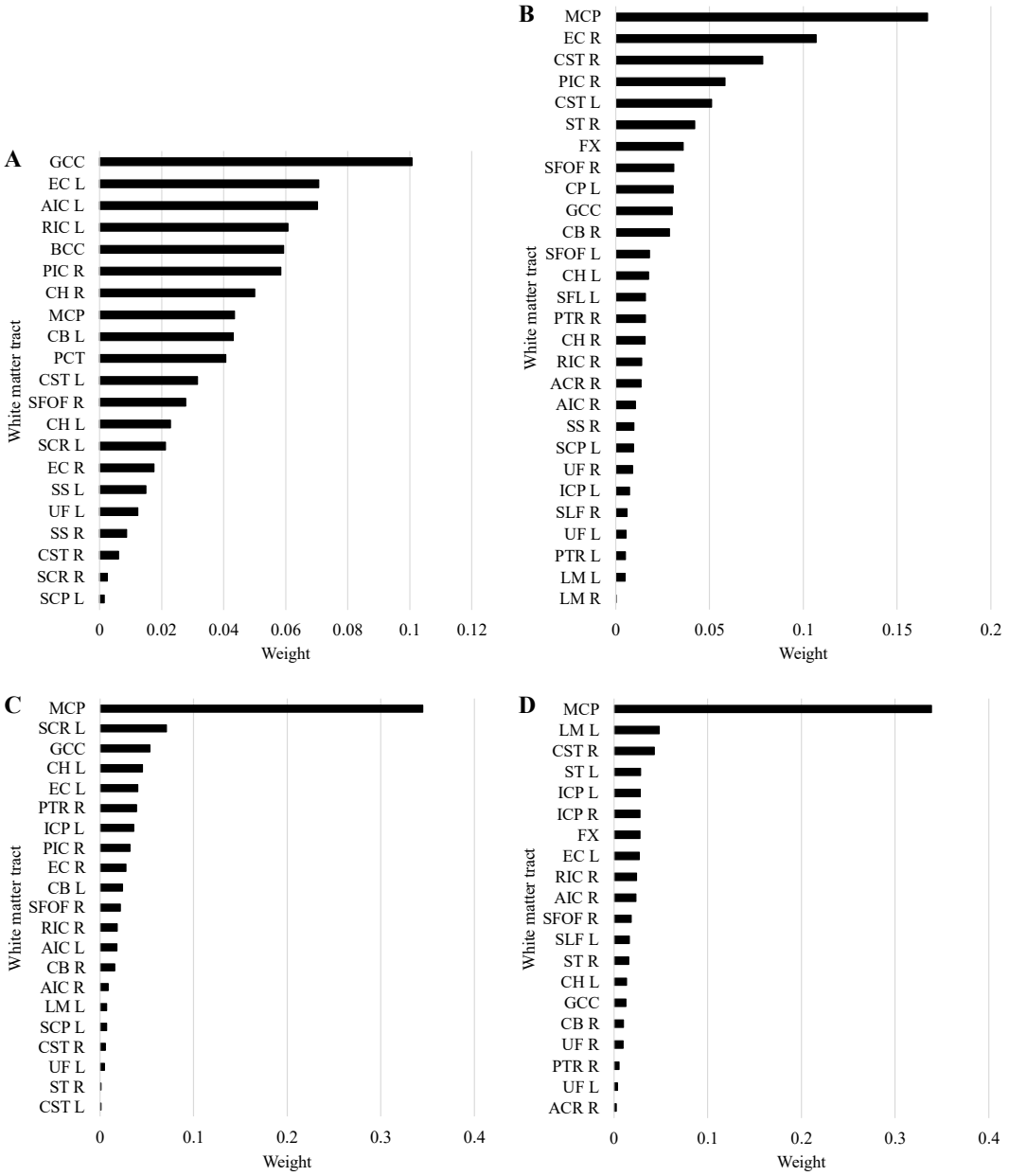


Figure 2. Median feature importance of the white matter tracts for the differentiation between MDD and BD patients for the A) AD-based model, B) FA-based model, C) MD-based model, and D) RD-based model. Only the tracts with median importance>0 are reported.
Abbreviations: ACR, anterior corona radiata; AD, axial diffusivity; AIC, anterior limb of internal capsule; ATR, anterior thalamic radiation; BCC, body of corpus callosum; BD, bipolar disorder; CB, cingulum bundle; CH, thalamic portion of cingulum; CP, cerebellar peduncle; CR, corona radiata; CST, corticospinal tract; EC, external capsule; FA, fractional anisotropy; FOF, fronto-occipital fasciculus; FX, fornix; GCC, genu of corpus callosum; IC, internal capsule; ICP, inferior cerebellar peduncle; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; L, left; LM, medial lemniscus; MCP, middle cerebellar peduncle; MD, mean diffusivity; MDD, major depressive disorder; PCT, pontine crossing tract; PIC, posterior limb of internal capsule; PTR, posterior thalamic radiation; R, right; RD, radial diffusivity; RIC, retrolenticular part of internal capsule; SCP, superior cerebellar peduncle; SCR, superior corona radiata; SFOF, superior fronto-occipital fasciculus; SLF, superior longitudinal fasciculus; SS, sagittal stratum; ST, stria terminalis; UF, uncinate fasciculus.

3.3.3. Feature importance of the SBC-based model for diagnosis differentiation

Feature importance analyses of the SBC model (Figure 2) showed that the predictive function relied on connectivity maps with periaqueductal grey (PAG), precuneus, the bilateral VS, the right inferior parietal lobule (IPL), rostral ACC (rACC), supramarginal gyrus (SMG), posterior insula (PI), and hippocampus, and left vlPFC, amygdala, and inferior temporal gyrus (ITG) as seed.

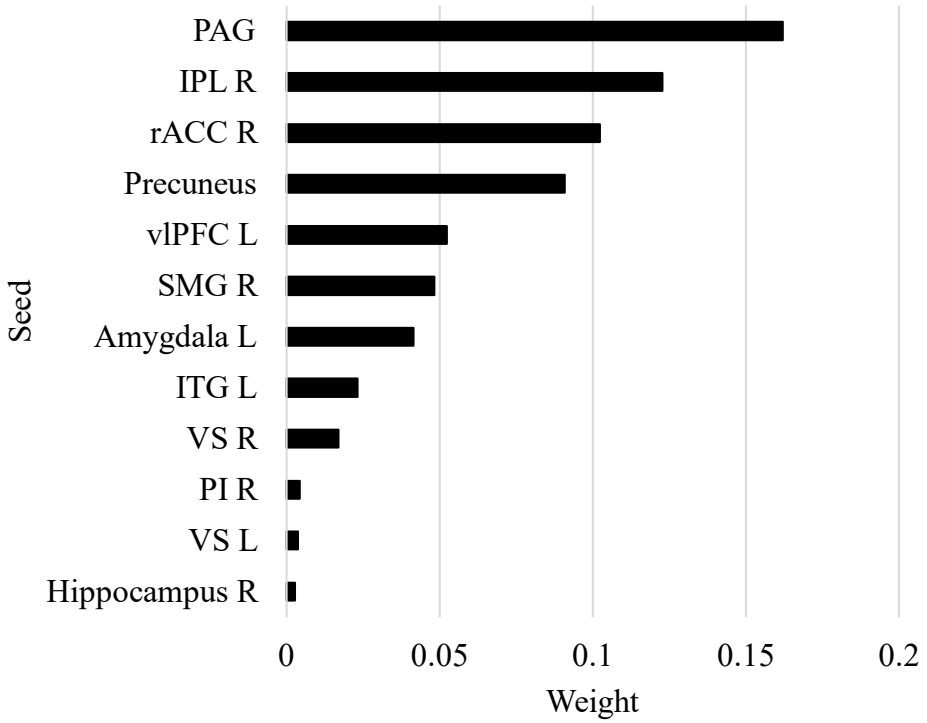


Figure 3. Median feature importance of the SBC maps for the differentiation between MDD and BD patients. Only the seeds with median importance>0 are reported.
Abbreviations: BD, bipolar disorder; IPL, inferior parietal lobule; ITG, inferior temporal gyrus; L, left; MDD, major depressive disorder; PAG, periaqueductal grey; PI, posterior insula; R, right; rACC, rostral anterior cingulate cortex; SBC, seed-based connectivity; SMG, supramarginal gyrus; vlPFC, ventrolateral prefrontal cortex, VS, ventral striatum.

To identify the connectivity patterns that significantly differentiated the diagnoses, additional models based on the single SBC maps were estimated (Table 5). When the SBC maps were considered separately for each seed, the models based on the FC map between the VS and the rest of the brain reached the best performance, with a BA of 70.06% for the left VS and 69.23% for the right VS. The SBC map seeded in the right PI was able to differentiate between MDD and BD with 66.60% of BA. Using left amygdala and right hippocampus as seeds for SBC maps the predictive BA was 64.29% and 62.18% respectively. The model based on PAG achieved a BA of 61.47% in differentiating the

two diagnostic groups. All these models performed significantly better than a distribution of null models, as assessed by the permutation test. Contrarily, the remaining models based on left ITG and vlPFC, right SMG, rACC, and IPL, and the precuneus did not reach significance threshold ($p<0.05$).

Table 5. Accuracy metrics of the models based on the single seeds SBC maps for the differentiation between MDD and BD patients.

Feature	BA	Sens	Spec	NPV	PPV	AUC	p-value
VS L	70.06	66.28	73.85	71.74	72.81	0.75	0.001*
VS R	69.23	79.23	59.23	67.10	73.86	0.76	0.001*
PI R	66.60	70.90	62.31	67.36	66.48	0.72	0.002*
Amygdala L	64.29	69.10	59.49	63.12	65.79	0.71	0.005*
Hippocampus R	62.18	64.74	59.62	62.05	63.61	0.70	0.008*
PAG	61.47	62.95	60.00	63.47	60.12	0.73	0.016*
ITG L	56.22	61.28	51.15	55.81	57.80	0.60	0.107
SMG R	55.71	58.21	53.21	58.37	57.19	0.57	0.115
rACC R	55.58	50.00	61.15	52.61	58.16	0.61	0.108
vlPFC L	52.44	48.85	56.03	54.19	51.80	0.56	0.281
Precuneus	52.31	59.74	44.87	51.99	52.77	0.53	0.257
IPL R	48.53	53.72	43.33	48.30	48.63	0.51	0.584

For each seed, accuracy metrics and p-value at the permutation test are reported for the differentiation between MDD and BD patients. All the models were estimated using SVM. Abbreviations: BD, bipolar disorder; IPL, inferior parietal lobule; ITG, inferior temporal gyrus; L, left; MDD, major depressive disorder; PAG, periaqueductal grey; PI, posterior insula; R, right; rACC, rostral anterior cingulate cortex; SBC, seed-based connectivity; Sens, sensitivity; SMG, supramarginal gyrus; Spec, specificity; vlPFC, ventrolateral prefrontal cortex, VS, ventral striatum. * $p<0.05$

3.3.4. Feature importance of the PIMs-based model for diagnosis differentiation

To identify the PIMs that significantly contribute the differentiation between MDD and BD patients, a 1000-bootstrap procedure was applied on the EN model (Figure 4). MDD was associated to higher values of SCGF-β, GM-CSF, RANTES, IL-16, M-CSF, and CCL13. Contrarily, higher levels of CCL8, CCL11, IL-4, and IFN-γ were associated to BD.

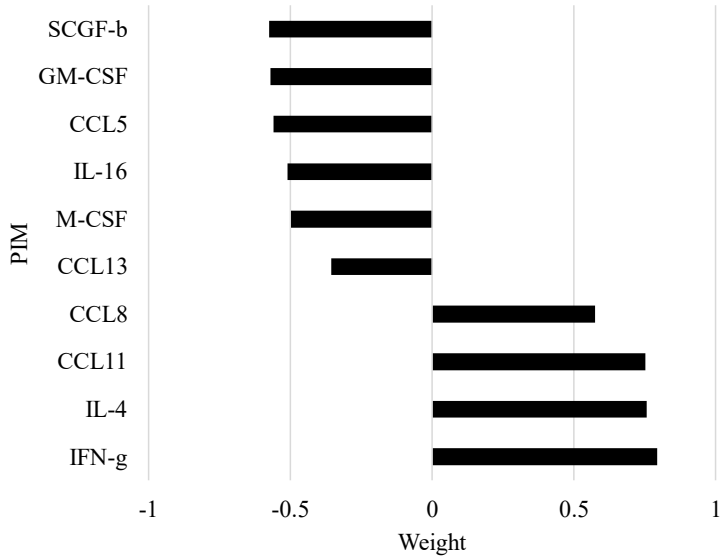


Figure 4. Mean contribution of the PIMs-based model for the differentiation between MDD and BD patients. Only the PIMs with a significant contribution at the bootstrap procedure were reported.
Abbreviations: CCL, C-C motif ligand; GM-CSF, granulocyte macrophage colony stimulating factor; IFN, interferon; IL, interleukin; M-CSF, macrophage colony stimulating factor; SCGF-β, stem cell growth factor β.

3.4. Prediction of cognitive deficit

3.4.1. Performance of the models

The results of the models for the identification of cognitively impaired patients at the composite score are summarised in Table 6. The best predictive performance was reached by the models based on DRCs and SBC with 68.33% of BA and 0.73 of AUC; ReHo with 66.67% of BA and 0.70 of AUC; and VBM with 64.00% of BA and 0.59 of AUC. Also, the RD-based models achieved a decent predictive accuracy with 60.44% of BA and 0.66 of AUC respectively. These models resulted to be significantly better than null models at the permutation test, except for the RD-based model that did not survive FDR correction for multiple comparisons.

Table 6. Performance metrics for the models identifying cognitively impaired patients at the composite score.

Feature	Algorithm	BA	Sens	Spec	PPV	NPV	AUC	p-value	FDRq-values
DRCs	MKL	68.33	56.67	80.00	72.67	66.51	0.73	0.005	0.019*
SBC	MKL	68.33	60.00	76.67	72.33	65.83	0.73	0.006	0.019*
ReHo	SVM	66.67	56.67	76.67	76.33	69.12	0.70	0.003	0.019*
VBM	SVM	64.00	70.00	58.00	66.76	68.67	0.59	0.008	0.019*
RD	SVM	60.44	50.89	70.00	66.24	58.47	0.66	0.035	0.066
FA	SVM	55.11	50.89	59.33	58.13	54.97	0.61	0.146	0.230

AD	SVM	54.11	54.89	53.33	54.22	56.56	0.61	0.230	0.311
MD	SVM	53.11	48.67	57.56	60.29	53.27	0.63	0.280	0.331
PRS	EN	51.00	61.22	40.78	32.97	68.85	0.51	0.537	0.507
Combined	MKL	50.67	36.00	65.33	46.67	50.86	0.67	0.424	0.445
PIMs	EN	50.62	61.54	39.71	28.07	72.97	0.51	0.836	0.637
fALFF	SVM	45.00	56.67	33.33	45.94	40.67	0.43	0.832	0.637
ABC	SVM	43.33	30.00	56.67	40.33	44.48	0.50	0.876	0.637

For each feature, accuracy metrics and p-value at the permutation test are reported. Models are sorted from the best to the worst model, considering BA. Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; AUC, area under the curve; BA, balance accuracy; DRCs, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity; NPV, negative predictive value; PPV, positive predictive value; PRS, polygenic risk score; RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed based connectivity, Sens, sensitivity; Spec, specificity; VBM, voxel-based morphometry. * FDR corrected p<0.05

3.4.2. Feature importance of the DRCs-based model for identification of cognitive impairment

The feature importance analyses of the DRCs-based model (Figure 5) revealed that the most differentiating components were represented by the ventral DMN (vDMN), primary visual network (PVN), posterior (PSN) and anterior (ASN) salience network, bilateral executive control network (ECN), and language network (LN).

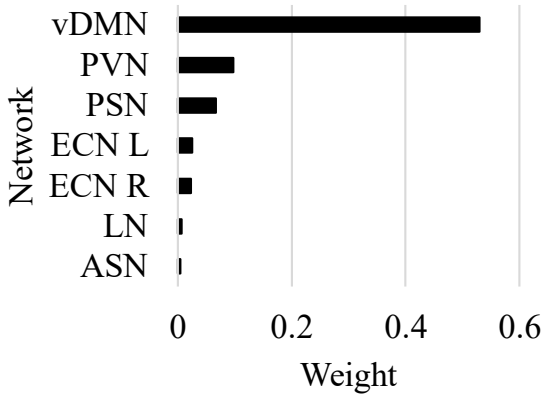


Figure 5. Median feature importance of the DRCs-based model for the identification of mood disorder patients with overall global impairment as measured by the composite score. Only the features with median importance>0 are reported.
Abbreviations: ASN, anterior salience network; DRCs, dual regression components; ECN, executive control network; L, left; LN, language network; PSN, posterior salience network; PVN, primary visual network; R, right; vDMN, ventral default mode network.

Additional single-component models were performed, to investigate the predictive performance of each DRC separately (Table 7). Only two components were able to significantly identify patients with cognitive impairment, referring to the vDMN and left ECN. The vDMN component reached 68% of BA and the left ECN achieved 61.67% of BA.

Table 7. Accuracy metrics of the models based on the single DRCs for the identification of patients with cognitive impairment.

Feature	BA	Sens	Spec	NPV	PPV	AUC	p-value
vDMN	68.33	56.67	80.00	70.48	68.33	0.71	0.003*
ECN L	61.67	56.67	66.67	70.72	59.06	0.67	0.033*
PVN	60.00	50.00	70.00	51.62	61.28	0.66	0.067
ASN	58.33	56.67	60.00	64.29	58.00	0.62	0.091
PSN	55.00	43.33	66.67	53.64	63.10	0.54	0.177
ECN R	51.67	43.33	60.00	55.50	49.13	0.48	0.364
LN	51.67	46.67	56.67	53.46	48.95	0.51	0.366

For each DRC, accuracy metrics and p-value at the permutation test are reported for the identification of cognitively impaired patients. All the models were estimated using SVM. Abbreviations: ASN, anterior salience network; AUC, area under the curve; BA, balance accuracy; ECN, executive control network; L, left; LN, language network; PPV, positive predictive value; PSN, posterior salience network; PVN, primary visual network; NPV, negative predictive value; R, right; Sens, sensitivity; Spec, specificity; vDMN, ventral default mode network. * $p<0.05$

3.4.3. Feature importance of the SBC-based model for identification of cognitive impairment

In the SBC model the maps contributing to prediction were seeded in the dorsomedial PFC (dmPFC), the PAG, the bilateral SMG, the left VS and middle temporal gyrus (MTG), and the right vlPFC (Figure 6).

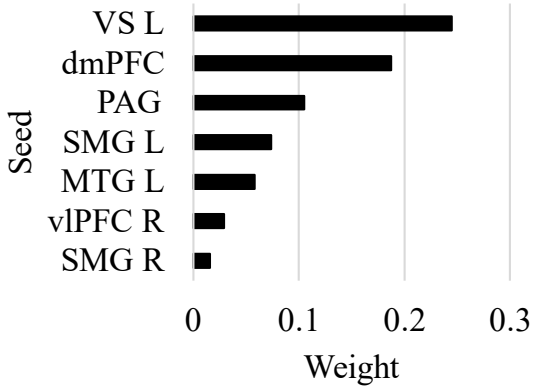


Figure 6. Median feature importance of the SBC-based model for the identification of mood disorder patients with overall global impairment as measured by the composite score. Only the features with median importance>0 are reported.
Abbreviations: dmPFC, dorsomedial prefrontal cortex; L, left; MTG, middle temporal gyrus; PAG, periaqueductal gray; R, right; SBC, seed-based connectivity; SMG, supramarginal gyrus; vlPFC, ventrolateral prefrontal cortex; VS, ventral striatum.

To investigate the connectivity patterns that can significantly identify patients with cognitive impairment, the SBC maps that contributed to prediction were entered in predictive models separately (Table 8). The predictive models that reached significance at the permutation test were based on the SBC maps seeded in the left VS (BA=66.67%) and MTG (BA=65.00%), and in the right SMG (BA=66.67%).

Table 8. Accuracy metrics of the models based on the single seeds SBC maps for the identification of patients with cognitive impairment.

Feature	BA	Sens	Spec	NPV	PPV	AUC	p-value
VS L	66.67	66.67	66.67	69.44	70.00	0.74	0.010*
SMG R	66.67	70.00	63.33	70.22	67.38	0.73	0.011*
MTG L	65.00	56.67	73.33	71.33	62.38	0.69	0.015*
dmPFC	58.33	40.00	76.67	62.00	56.43	0.70	0.108
vlPFC R	56.67	50.00	63.33	54.89	55.40	0.56	0.130
PAG	51.67	53.33	50.00	45.00	56.67	0.60	0.356
SMG L	46.67	40.00	53.33	46.57	46.57	0.43	0.704

For each seed, accuracy metrics and p-value at the permutation test are reported for the identification of cognitively impaired patients. All the models were estimated with SVM. Abbreviations: AUC, area under the curve; BA, balanced accuracy, dmPFC, dorsomedial prefrontal cortex; L, left; MTG, middle temporal gyrus; NPV, negative predictive value; PAG, periaqueductal grey; PPV, positive predictive value; R, right; Sens, sensitivity; SMG, supramarginal gyrus; Spec, specificity; vlPFC, ventrolateral prefrontal cortex; VS ventral striatum. * $p<0.05$

3.4.4. Feature importance of the ReHo-based model for identification of cognitive impairment

The ReHo-based model highlighted that the regions capable of identifying cognitively impaired patients included both midline and dorsal prefrontal and parietal regions, along with temporal structures. Specifically, the right middle frontal gyrus (MFG), posterior cingulate gyrus (PCG), and PoCG, and the left superior frontal gyrus (SFG) and STG were the regions included in the model, contributing to the prediction (Figure 7).

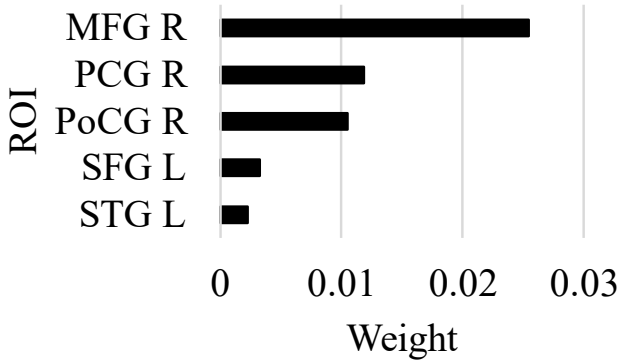


Figure 7. Median feature importance of the SBC-based model for the identification of mood disorder patients with overall global impairment as measured by the composite score. Only the features with median importance>0 are reported.
Abbreviations: L, left, MFG, middle frontal gyrus; PCG, posterior cingulate gyrus; PoCG, postcentral gyrus; R, right; ROI, region of interest; SBC, seed-based connectivity; SFG, superior frontal gyrus; STG, superior temporal gyrus.

3.4.5. Feature importance of the VBM-based model for identification of cognitive impairment

The VBM model relied on a widespread pattern pertaining to different systems (Figure 8). This pattern encompassed the bilateral opercular part of the inferior frontal gyrus (IFGPo), occipital fusiform gyrus (OFuG), and PI, the left amygdala, opercular part of the inferior frontal gyrus (IFGPt), middle cingulate gyrus (MCG), occipital pole (OCP), thalamus, and temporal pole (TMP), and the right angular gyrus (AG), caudate, frontal operculum (FO), ITG, medial frontal cortex (mFC), MFG, and SFG.

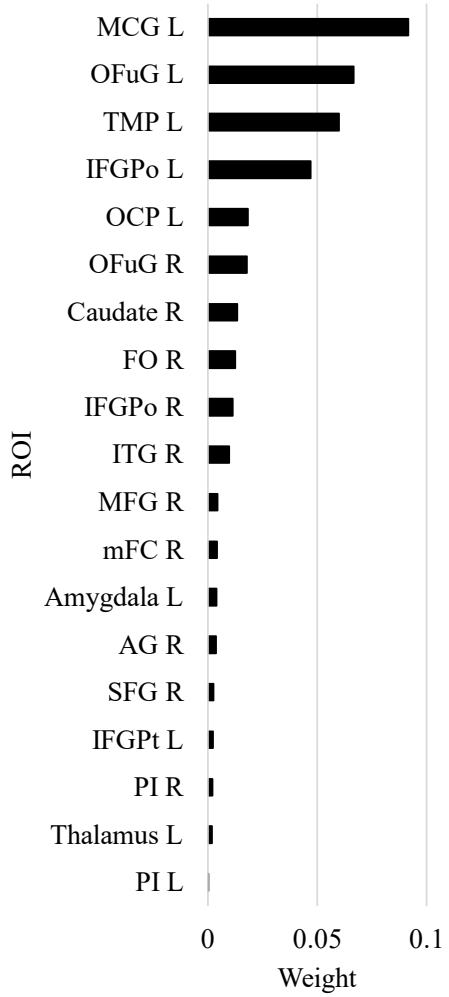


Figure 8. Median feature importance of the VBM-based model for the identification of mood disorder patients with overall global impairment as measured by the composite score. Only the features with median importance>0 are reported.
Abbreviations: AG, angular gyrus; IFGPo, opercular part of the inferior frontal gyrus; IFGPt, triangular part of the inferior frontal gyrus; ITG, inferior temporal gyrus; L, left; MCG, middle cingulate gyrus; mFC, medial frontal cortex; MFG, middle frontal gyrus; OFuG, occipital fusiform gyrus; PI, posterior insula; R, right; ROI, region of interest; SFG, superior frontal gyrus; TMP, temporal pole; VBM, voxel-based morphometry.

4. DISCUSSION

In this work, we created several predictive models for the differentiation between depressed MDD and BD patients, as well as for the identification of cognitive impairment in mood disorder patients. To fulfil this objective, we applied the ML framework on a wide panel of biological information, comprising structural and functional neuroimaging (i.e., VBM, DTI-derived measures, and rsfMRI-derived features), PIMs (i.e., cytokines and chemokines), and genetics (i.e., PRS for several disorders and dimensions). The application of ML techniques allowed us to make predictions at the single-subject level and assess model's performance on unseen data, which is particularly critical in the perspective of adopting a translational approach. In particular, we deployed state-of-the-art analysis pipelines that prevent performance inflation, promote robustness, and allow model's interpretation. The first strength of the work is the application of the Combat algorithm to remove the information related to the acquisition with different MRI scanners. As a second strength we removed the information related to nuisance variables that could intervene as confounders in the prediction. Notably, a wide set of confounders was regressed-out, also comprising non-linear effects of age, and the age-by-sex interaction effect. Third, we applied a nested CV scheme, avoiding data leakage between train and test sets, which increases the reliability of the estimated generalization capacity of the model.

4.1. Diagnosis differentiation

4.1.1. Models' performance

Concerning the differentiation between MDD and BD patients, the diffusivity measures derived from DTI images and the SBC maps were the features that led to a classification accuracy significantly higher than a distribution of null models. The models based on AD, RD, and MD were the best performing with BA higher than 70%, whereas the FA-based model was less accurate, with a BA of 64%. These results are in line with and improve the predictive performance of previous studies. Specifically, the AD, RD, and MD features were previously found to be able to classify the diagnostic groups with accuracy between 61% and 66%, whereas FA was found to correctly classify 56% and 68% of the cases (Deng et al., 2018; Vai et al., 2020).

The model based on SBC connectivity maps reached 66% of accuracy, with the significant models based on single-seed maps achieving predictive accuracy from 60% to

70%. Only one other study investigated the predictive performance of SBC for the differentiation between MDD and BD patients, reaching an overall accuracy of 91%, using ACC, insula, and amygdalae as seeds (H. Yu et al., 2020). It should be noted, though, that these results were obtained on a small sample (N<50) and using a LOO CV scheme, which might have led to performance inflation.

The other rsfMRI features did not significantly classify the diagnoses. Interestingly, previous studies obtained a predictive performance in the range 69-86% (M. Li et al., 2017; Rive et al., 2016; Shao et al., 2019), using degree of centrality (M. Li et al., 2017), dynamic communities of networks (Shao et al., 2019), or network connectivity (Rive et al., 2016) as features. Despite not being the same features that were used in the present work, information related to network connectivity should be present in the features that we extracted. However, the study employing network connectivity was the one that achieved the lowest predictive accuracy, thus confirming our results. Furthermore, all the studies were characterized by small sample sizes (N<100).

The PIMs also significantly differentiated between MDD and BD with 62% of accuracy. This accuracy is lower compared to that obtained in an analogous previous study, in which PIMs differentiated between MDD and BD with 90% of accuracy (Poletti, Vai, et al., 2021). Although it probably does not account for the whole difference in accuracy, it should be noted that in this work the effect of an extensive set of covariates was regressed-out. In the previous work, instead, fewer nuisance affects were considered and addressed by inserting them into the model.

The models based on VBM differentiated the two groups of patients with accuracy around chance level. Most of the previous studies reached higher accuracies using kernel methods on GM features, yet with very high variability across studies (i.e., accuracy range: 54-97%) (He et al., 2017; Redlich et al., 2014; Rive et al., 2016; Rubin-Falcone et al., 2018; Serpa et al., 2014). It should be noted that in all these studies, the implemented CV scheme was a LOO (Redlich et al., 2014; Rubin-Falcone et al., 2018; Serpa et al., 2014), which has been proven to be less reliable than the k-fold scheme implemented in the current work (Colombo et al., 2022; Varoquaux et al., 2017). Furthermore, the sample size of the previous studies was very small (N<60), and when testing the performance of the model on an external validation set the accuracy dropped by around ten percentage points (Redlich et al., 2014; Rubin-Falcone et al., 2018). Only one study achieved a very

high classification accuracy of 97.9% with a SVM on GM features (He et al., 2017). In this case, beyond the low sample size, a feature reduction procedure was carried-out outside the CV scheme, which could have further induced inflated performance on the test set, due to information spillover (Colombo et al., 2022; Mwangi et al., 2014). Lastly, a study differentiating MDD and BD adolescents from the ABCD cohort using a SVM with recursive feature elimination achieved around 80% accuracy with GM-related features (Y. Liu et al., 2022). The study was multicentric and implemented a leave-one-site-out CV embedding a 10-fold CV scheme. Importantly, though, none of these studies removed the effect of confounding variables (Rubin-Falcone et al., 2018). Notably, in one previous study from our group, VBM features were used to differentiate between MDD and BD patients, reaching 69.59% of accuracy (Vai et al., 2020). When combining GM and WM features with a MKL algorithm, the model reached 73.65% of accuracy, with VBM being the most predictive feature. Therefore, these results appear to be in contrast with the findings of the present work. However, despite the high similarity between the two studies, some differences should be considered. First, these previous models were trained on a smaller sample. Furthermore, the set of covariates that was used in the current work was more complete, including also non-linear and interaction effects. Lastly, in the present work, the features underwent a harmonization procedure using the ComBat algorithm. It is worth noting that the differentiating power of VBM decreased by 7 percentage points after harmonization, going from 62% of accuracy to 55% after correction, suggesting that the ComBat algorithm might have removed some biological information.

The models based on BACS and PRSs did not significantly differentiated between MDD and BD patients, reaching predictive accuracies around 50%. This suggests that the cognitive profile does not differ between the two diagnostic groups. Furthermore, the absence of significant predictive information in the PRSs, corroborate and support the presence of a shared genetic component between MDD and BD (Coleman et al., 2020). The model combining all the features did not reach significance, with accuracy around chance level. Although MKL can perform feature selection, a possible explanation of this negative result is that a majority of non-predictive features could introduce a high amount of noise, driving the model to make erroneous predictions.

4.1.2 Models' interpretation

The investigation of the DTI-based models revealed that the classification was based on widespread, yet subtle, patterns of WM alterations. When interpreting DTI results, though, the susceptibility of the diffusion tensor modelling to voxels containing multiple orientations should be considered (Jeurissen et al., 2013). Around 90% of WM voxels contain crossing-fibers, that is when tracts cross, kiss, branch, or merge (Basser et al., 2000; Jeurissen et al., 2013). In these voxels, the estimated fiber trajectories might not accurately represent the true pathways, resulting in biased estimates. There, anisotropy reflects a consensus average fiber direction (Basser et al., 2000; Pierpaoli & Basser, 1996). On the one hand, this can manifest as an increase in FA due to the crossing of tracts with higher intrinsic anisotropy. On the other hand, a decrease in FA can be observed due to averaging out of two crossing tracts (Basser et al., 2000; Jeurissen et al., 2013; Pierpaoli & Basser, 1996). Therefore, while the diffusion tensor offers reliable estimates of anisotropy in voxels with uniform fiber direction, these estimates intricately depend on the structure and architecture of the tissue in voxels with non-uniform fiber directions. However, the pattern found in this work included both coherently oriented WM tracts (e.g., CC) and more complex structures (e.g., IC), suggesting that these results are not driven by crossing-fibers artifacts, but rather by true anisotropy differences (Basser et al., 2000).

Peculiarly, in these tracts, BD patients were found to have average higher values across all the diffusivity measures. These results may appear paradoxical, as the various diffusivity measures reflect different, and sometimes opposite, undergoing biological processes.

AD quantifies water molecule displacement parallel to the orientation of axonal fibers, reflecting the preferential diffusion along the axons' main trajectory over the perpendicularly oriented myelin sheaths. This phenomenon of diffusion anisotropy is not solely attributable to the presence of myelin, but is also intricately tied to the axonal microarchitecture, encompassing elements such as microtubules and neurofilaments (Kinoshita et al., 1999). Preclinical models have demonstrated an inverse relationship between AD and axonal pathology, suggesting that a decrease in AD could serve as a surrogate indicator of axonal compromission (Boretius et al., 2012; Song et al., 2003; Sun et al., 2006). Contrarily, RD represents the diffusivity perpendicular to the axon. Also RD

is linked to myelin integrity, but its increase points to disrupted myelination (Song et al., 2002). MD is the mean molecular motion, independent of tissue directionality. Its increase is associated to an enlargement in the extracellular space (Schulte et al., 2005). FA measures the anisotropy level within a voxel, assessing the uniformity of direction in local tract structures, serving as an indicator of WM integrity. A decrease in FA points to disrupted water directionality, often representing structural tissue damage (Schulte et al., 2005).

Overall, an increase in AD and FA would be interpreted as an indication of WM integrity, whereas an increase in RD and MD values would suggest a disruption of axonal structure. Therefore, our results appear to be counterintuitive, since BD patients had higher levels across all the measures. However, in chronic diseases, when inflammation, axonal loss, axonal injury, and demyelination coexist, the increased isotropic diffusion seems to enhance both RD and AD (Winklewski et al., 2018). Higher levels of AD have also been reported in neurological developmental pathology that leads to loss of complexity of WM traits, or to cellular infiltration, gliosis and extracellular water changes caused by chronic inflammation (Aung et al., 2013; Haas et al., 2014). Furthermore, water diffusivity is measured on the single voxel, which has a volume of about $1mm^3$. Even if the estimation of the diffusion parameters is performed on an image which includes only voxels that are part of the WM skeleton, it is possible that water molecules of volumetric portions of voxels containing extra-neuritic space are included in the analysis. Therefore, WM damage could make water molecules flowing out of the injured tract, yet still diffusing in areas proximal to the tract. Since the flowing water molecules would diffuse in the same direction as inside the axon by inertia, this information could cause distorted anisotropy calculations (Churchill et al., 2018). Consequently, the concurrent increase in values of AD, FA, RD, and MD can be interpreted as being caused by an increase in free-water diffusion, reflecting abnormalities of both axon integrity and myelinating processes in BD patients.

Interestingly, for the FA, MD, and RD models the most differentiating tract was the MCP. The MCP is core in the cortico-ponto-cerebellar tract (Keser et al., 2015), which has been previously found to be impaired in BD patients (Hou et al., 2022).

The exploration of the SBC model highlighted the importance of FC maps seeded in the bilateral VS, right PI and hippocampus, left amygdala, and PAG. These regions are all

involved in processing emotions, rewards, stress responses, and pain perception. They are all part of an interconnected network that allows humans to experience and react to emotional stimuli, whether positive or negative. This includes the modulation of pleasure and pain, the processing and response to stressful stimuli, the formation and recall of emotional memories, and the regulation of mood and motivation.

The VS, including the nucleus accumbens, plays a pivotal role in reward anticipation and pleasure-related learning, integrating information about rewards and action plans (Berridge et al., 2009). It is also crucial in motivational processes, by associating stimuli with positive outcomes, thereby influencing behaviour aimed at achieving these outcomes (Sesack & Grace, 2010). Additionally, it is thought to be involved in decision-making, especially in situations involving cost-benefit analysis (Levita et al., 2009).

The insula is primarily involved in interoceptive awareness, subjective perception of the body's internal state, and the generation of conscious emotional feelings (Craig, 2009; Critchley & Harrison, 2013). However, the posterior portion of the insula is particularly crucial for the primary somatosensory processing of stimuli, especially those relating to pain and temperature (Augustine, 1996; Gehrlach et al., 2019).

The amygdala is known to be critical in emotional processing, memory consolidation, and decision-making (LeDoux, 2007; McGaugh, 2004). It plays a crucial role in fear conditioning and modulates the memory consolidation of arousing stimuli (McGaugh, 2004; Phelps & LeDoux, 2005). The amygdala is also thought to be involved in processing rewards and to motivate and reinforce behaviour (LeDoux, 2007; Phelps & LeDoux, 2005).

The hippocampus is well known to be crucial for memory consolidation and spatial navigation (Bird & Burgess, 2008; A. M. P. Miller et al., 2014). However, the ventral part of the hippocampus is implicated in emotion and stress responses, probably due to its connectivity with the amygdala (Bannerman et al., 2004; Fanselow & Dong, 2010; Strange et al., 2014).

The PAG is particularly important in both the descending and ascending modulation of pain, through its connections with prefrontal, cingulate, and insular cortices, amygdala, and hypothalamus. Specifically, the PAG is particularly dense of opioid μ-receptors, through which it probably mediates dopaminergic pathways and induces analgesia (C. Li et al., 2016; Vázquez-León et al., 2023; Zelena et al., 2018). It is involved in the

modulation of combative defence reactions and fight-or-flight responses (Vázquez-León et al., 2023; Weis et al., 2022; Zelena et al., 2018). Given its role in defensive behaviours, the PAG appears to be fundamental in the neurocircuitry of fear and anxiety (Weis et al., 2022). The PAG has also been found be activated during anticipation of a negative stimulus, suggesting its involvement in fear conditioning (Roy et al., 2014; Watson et al., 2016; Zelena et al., 2018).

Interestingly, all these regions are core components of the reward and aversion circuits, which play a key role in motivational control (Liang et al., 2022). The reward system is mainly composed by the projections of midbrain dopamine neurons of the ventral tegmental area (VTA) to the striatum, amygdala, hippocampus, and PFC (R. G. Lewis et al., 2021). It is responsible for mediating the physiological and cognitive processing of reward, which consists in associating stimuli to a positive outcome. This association leads to adjustments in the individual's behaviour, oriented to search for that particular positive stimulus (R. G. Lewis et al., 2021). Conversely, aversive stimuli can be associated to negative outcomes and experiences, such as pain or discomfort. This eventually leads to avoidance behaviours, reducing the likelihood to be exposed to similar potentially harmful stimuli in the future (W. Chen, 2022). Notably, the key brain regions of the aversion pathway include the PAG, amygdala, insula, lateral habenula, medial prefrontal cortex, ventral pallidum, and VTA (Liang et al., 2022).

Both the reward and the aversion circuits have been linked to anhedonia (Liang et al., 2022). Therefore, the observed differential connectivity patterns in these regions may mirror the phenomenological opposite polarity between the depressive and manic phases. This view is consistent with a spectral definition of anhedonia, rooted in the hyper- or hypo-activation of the reward system, which would reflect behaviours ranging from hyperhedonia to anhedonia, respectively (Rizvi et al., 2018). Furthermore, it is not completely clear whether anhedonia would represent a state- or trait-like quality, with evidence of persisting inter-episode anhedonia (Rizvi et al., 2018; Whitton & Pizzagalli, 2022). The combination of these two aspects, may support the presence of a trait-like spectrum of anhedonia, over which patients are distributed as a function of the state of the reward system. This interpretation would also be in line with the reward hypersensitivity model of BD, which posits that stable reward hypersensitivity would be responsible for mood swings, as a result of excessive deactivation or activation of the

reward system (Bart et al., 2021; Rizvi et al., 2018; Whitton & Pizzagalli, 2022). Overall, the different SBC patterns between MDD and BD in the core regions of the reward and aversive systems may reflect a differential sensitivity to reward between the two diagnostic groups.

Lastly, the investigation of the model based on PIMs highlighted a differential inflammatory profile between MDD and BD patients, which is in line with previous evidences (Bai et al., 2015; Brunoni et al., 2020; M.-H. Chen et al., 2019; Kim et al., 2002; Mota et al., 2013; Poletti, Vai, et al., 2021). Notably, consistently with the current results, higher levels of IFN-γ, IL-4 , and CCl11 have been previously reported in BD patients (Barbosa et al., 2013; Wollenhaupt-Aguiar et al., 2020). Furthermore, MDD patients have been previously found to have increased concentrations of GM-CSF, CCL5, IL-16, M-CSF, CCL13, in line with the results of the present work (Almulla et al., 2023; Maes et al., 2022; Powell et al., 2018; Schmidt et al., 2014). This pattern suggests that BD patients are associated to a more pro-inflammatory profile, with a main involvement of the monocyte/macrophage response. Instead, besides the inflammatory profile, MDD patients appear to be also characterized by the activation of regulatory mechanisms (Poletti, Vai, et al., 2021). In accordance with this view, up-to-date models for acute depression support anti-inflammatory properties of both GM-CSF and M-CSF (Hemmati et al., 2019; Ye et al., 2020).

4.2. Identification of cognitive impairment

Considering the models for the identification of mood disorder patients with a global cognitive impairment (i.e., composite score), the most predictive features were DRCs, SBC, ReHo, and VBM, achieving a predictive accuracy higher than 60% and reaching almost 70% of BA. These models resulted to be significantly different from a distribution of null models, indicating that are based on true information, rather than chance.

The identification of cognitively impaired patients relied on a widespread pattern of both functional and structural alterations. This pattern comprised regions that mostly cover the DMN, the salience network (SN), and the ECN. Furthermore, the strength and inter-network connectivity of the vDMN and the left ECN significantly identified patients with cognitive impairment.

The DMN has been linked to internally directed cognitive activities, such as self-monitoring and mind wandering, given its deactivation during goal-oriented tasks

compared to its baseline at rest (Anticevic et al., 2012; Raichle et al., 2001). The DMN is mainly constituted by ventromedial PFC (vmPFC), PCG, and regions in the medial temporal and parietal lobes (Andrews-Hanna et al., 2010; S. Lee et al., 2021; Raichle, 2015). More specifically, its dorsal component comprises the vmPFC, ACC, AG, OFG, SFG, PCG, MCG, thalamus, and hippocampus. The ventral component of the DMN, instead, involves the retrosplenial cortex, the PCG, MFG, SFG, AG, middle occipital gyrus, precuneus, and parahippocampal gyrus (Andrews-Hanna et al., 2010; S. Lee et al., 2021; Shirer et al., 2012).

The ECN has been associated to externally directed higher-order cognitive functions, such as attention, working memory, and decision-making (Bressler & Menon, 2010; Uddin, 2015). The ECN includes nodes in the left MFG, SFG, IFG, OFG, SMG, AG, ITG, MTG, precuneus, superior (SPG) and inferior (IPG) parietal gyri, thalamus and caudate (Shen et al., 2020; Shirer et al., 2012).

The SN is implicated in the detection, integration, and filtering of relevant interoceptive and emotional signals. It is mainly composed by a cingulate-frontal operculum system, which is anchored in the dACC and anterior insula (Seeley et al., 2007; Shirer et al., 2012). Notably, the SN also encompasses subcortical structures, such as the amygdala, VTA, VS (Menon, 2011).

Together, these networks are the most prominent macro-scale networks and their aberrant functional organization and dynamic cross-talk has been suggested to be implicated in several psychiatric and neurological disorders (Menon, 2011; Schimmelpfennig et al., 2023). During stimulus-driven information processing, the CEN and SN typically show increased activation, while the DMN is characterized by reduced activation. The opposite pattern is seen with self-referential processes. Notably, the SN is thought to play a key role in orchestrating the interaction between the CEN and the DMN, by guiding saliency (Menon, 2011; Schimmelpfennig et al., 2023; Seeley et al., 2007). The proper balancing between these networks is thought to be crucial for cognitive performance (Anticevic et al., 2012; Cha et al., 2014; Fox et al., 2005). Therefore, the results of this work may reflect an altered balancing between activation and suppression in these networks.

However, when considering the models predicting impairment in specific cognitive domains, only few of them reached more than 60% of BA, and these few models did not achieve statistical significance at the permutation test after correcting for multiple

comparisons. These results reveal that patients could be differentiated on a broad definition of cognitive impairment, but not on deficits in specific cognitive domains. It should be noted, though, that most of the patients were impaired in at least one cognitive domain, and frequently they concurrently manifested deficits over multiple domains. Therefore, when trying to identify patients impaired in a specific cognitive domain, the comparison was not necessarily against cognitively intact patients. Instead, the classification was mostly made between patients with cognitive deficit in a specific domain and patients who may exhibit deficits in disparate other cognitive domains. This potentially introduced some noise into the not-impaired class, thus reducing the predictive accuracy of the models, which might explain the negative results.

Nonetheless, these results still highlight a lack of specificity in the information embedded in these features, since they could not differentiate between specific cognitive domains. This lack of specificity is in line with the view that cognitive networks are highly redundant and pleiotropic: multiple nodes share similar roles and each node participate to several functions. On the one hand, such a structure provides considerable resilience to disruption. On the other hand, the failure of highly connected hubs can have a disruptive effect on the entire network (Millan et al., 2012).

Furthermore, these results indicate that patients do not stratify on fine-grained cognitive dimensions, but rather on coarse-grained cognitive functioning. In this view, these findings support the presence of two transdiagnostic subpopulations within mood disorders, differentiated by the presence or absence of a global cognitive impairment, related to a neurobiological substrate that include prefrontal, temporal, and sensorimotor cortices. Furthermore, given their high level of integration and abstraction, the involved regions are not only implicated in strictly cognitive functions, but also in emotional and social processing, such as emotional regulation and theory of mind. Consequently, there is the possibility that these broadly defined cognitive phenotypes might be an epiphenomenon overlapping with a better characterized phenotype. For instance, even though unrelatedly to cognitive performance, alterations in the DMN and in its balancing with other networks at rest have been previously found in depression, with both cognitive and emotional implications (Anticevic et al., 2012; J. P. Hamilton et al., 2011; Scalabrini et al., 2020). It is also worth noting that despite not having found subpopulations that reflect cognitive impairment in specific domains, patients may stratify at different levels

of granularity associated to the impairment in different combinations of cognitive dimensions.

Although DTI diffusivity, PIMs, and PRSs alterations have been previously linked to cognitive impairments in mood disorders, these features did not significantly identify patients with cognitive impairment in this work (Melloni et al., 2019; Millan et al., 2012; Poletti et al., 2015; Poletti, Mazza, et al., 2021). A possible explanation for these negative results might be that the present experiment is underpowered to detect these effects. For instance, the sample size could be insufficient to extract the relevant information needed to make predictions at the single subject level, failing to generalize to out-of-sample observations. However, in some specific domains, these features were able to identify cognitively impaired patients with predictive accuracies higher than 60%. These models were also significant at the permutation test, but they did not survive multiple comparisons correction. By increasing the number of permutations, the distribution of null models may be narrowed, possibly leading to higher inferential power. Lastly, it might be that other non-linear algorithms would achieve a better performance. Furthermore, the model combining all the features achieved an accuracy around 50%. Despite the ability of MKL to perform feature selection, removing non-predictive noisy features from the model, it might be that the amount of data and noise to be handled was too high in this case.

4.3. Limits of the study

Despite the innovations, the advancements, and the state-of-the-art methods of the present work, some limitations need to be addressed. First, the design of the present experiment was cross-sectional. The data on which the models were based were collected when BD were already diagnosed as such, as well as at the same time of cognitive assessment. These results should be validated in a longitudinal framework, collecting data at the onset, to predict the future switch to mania or the manifestation of cognitive deficits. For the same reason, patients were not drug-naïve, which may introduce some bias induced by the effects that life-time treatments can have on most of the measures used in this work (e.g., neuroimaging and inflammation). Additionally, the absence of ethnic variability in the sample limits the generalizability of the findings. Despite being higher than most of previous studies, the sample size was still small for machine learning analyses and is hardly sufficient to create a validated tool to be translated in real world applications.

However, the present work and its result can guide future similar studies that may exploit large datasets derived by international consortia, such as the Enhancing Neuro Imaging Genetics Through Meta Analysis Consortium (ENIGMA). Furthermore, future studies may include more sources of information, such as additional modalities and features.

Lastly, another possible limitation is the use of two different scanners for the acquisition of MRI data. This aspect is known to introduce artifacts, leading to higher variance and possibly introducing bias in the findings (Fortin et al., 2017, 2018; Leek et al., 2010; Yi et al., 2018). Consequently, these artifacts, also known as batch effects, should be accounted for. Currently, the ComBat algorithm has been proven to be extremely effective in removing such effects (Fortin et al., 2017, 2018). However, when the information related to the batch overlap with the true target signal, ComBat could also remove some target-related information, in the attempt to remove the batch effect. To prevent this scenario, a set of covariates was modelled in the ComBat algorithm, to protect their effect from erroneous removal. Using this strategy, though, in the case of shared information between batch and target effects, the performance of the correction may be suboptimal. In this work, although the correction worked properly, a difference in the predictive accuracy of the targets of interest was observed after correction. Despite most of the features being characterized by a slight increase in predictive performance, some models showed a more substantial increase, or even a decrease in accuracy. The decreased performance was probably due to an overlap between the batch and the target of interest. Supporting this view, the predictive accuracy of the VBM-based model was characterized by the most pronounced contraction, and concurrently VBM was the feature on which the correction was less effective. The increase in performance, instead, can have a twofold interpretation. On the one hand, the increase in performance can be interpreted as the successful removal of noisy information generated by the batch effect. On the other hand, there is the possibility that the increased performance may be caused by involuntary injection of a sham signal into the data. Furthermore, since ComBat correction is performed at the group level, observations may have influenced each other leading to information spillover, and performance inflation in turn.

4.4 Conclusion

Overall, we were able to differentiate between MDD and BD patients with accuracy around 70%, using ML models based on a comprehensive set of biological data, including

structural and functional neuroimaging, inflammatory markers, and genetics. With the same set of variables, we were also able to detect cognitively impaired patients with similar accuracy, namely 67%.

Results from diagnosis differentiation suggest that BD patients are characterized by a subtly lower integrity of WM over a widespread network, compared to MDD. MDD and BD patients were also found to have a different arrangement of functional connections of the reward system, probably reflecting the behavioural differences between depression and mania. Additionally, a differential inflammatory profile was found to differentiate between the diagnostic groups. The identification of patients with cognitive impairment, instead, relied on structural and functional alterations in regions belonging to the default mode, the executive control, and the salience networks, possibly defining novel therapeutic targets for subpopulations of mood disorder patients.

5. MATERIALS AND METHODS

5.1. Participants

The sample included 358 depressed patients (139 MDD and 219 BD) recruited at the IRCCS San Raffaele Scientific Institute, Milan, Italy. Inclusion criteria were: meeting the diagnostic criteria for MDD or BD with an ongoing depressive episode according to the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) criteria and having a score higher than 8 at the 21-Hamilton Depression Rating Scale (HDRS-21) (M. Hamilton, 1960), and being 18-65 years old. To reduce the possibility of uncorrected a priori labelling, MDD patients should also have a history of at least two previous depressive episodes. Participants were excluded if they had: a major medical and neurological disorder, pregnancy, mental retardation, or history of drug or alcohol abuse or dependency. After a complete description of the study, written informed consent was obtained. All procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. The study was approved by the local ethical committee.

5.2. Cognitive assessment

5.2.1. Battery description

The degree of cognitive impairment was assessed by administering the BACS to 168 subjects (89 MDD and 79 BD). BACS is a quick and affordable neuropsychological battery originally developed to assess schizophrenia patients, and it is the most frequently used tool in clinical trials (Bakkour et al., 2014). Six cognitive domains can be assessed by means of paper-and-pen tasks: verbal memory, working memory, coordination, processing speed, executive functions, and verbal fluency (Keefe et al., 2004).

Verbal memory was assessed using a list learning task, wherein patients were asked to recall as many words as possible from a previously presented list. To evaluate working memory, patients underwent a digit sequencing task that involved arranging clusters of numbers in ascending order from lowest to highest, with the length of the number sequences progressively increasing. The measurement of motor speed was conducted using a token motor task, where motor speed was represented by the patient's ability to place previously scattered tokens into a container within a 60 s timeframe. Verbal fluency was assessed with two tasks. In the first task, patients were given one minute to name as

many words as possible within a given category. In the second task, in two separate trials, patients had to generate as many words as possible that begin with a given letter in one minute. Symbol coding was used to measure the attention and speed of information processing domain, consisting in mapping the as many symbols as possible to numbers in 90 s. Executive functions were assessed using the Tower of London task, in which patients were instructed to determine the total number of moves required to make a picture match the arrangement of another picture, each displaying three differently coloured balls arranged on pegs in unique configurations.

BACS is characterized by speed and efficiency, as the assessment can be conducted in around 30 minutes, and provides a reliable, valid, and psychometrically stable assessment of cognitive ability in schizophrenia spectrum (Keefe et al., 2004). BACS was originally validated on a sample of 150 schizophrenic patients and 50 HC, on a standard neuropsychological battery including different tests such as Wechsler Memory Scale, Logical Memory subtest, and Wisconsin Card Sort Test. BACS has been shown to have good psychometric qualities with a good completion rate and test-retest reliability (intraclass correlation coefficient > 0.79) in schizophrenia assessment. The sensibility of the BACS is comparable to that obtained from other cognitive batteries and its overall score is highly correlated with the overall score obtained in the standard neuropsychological battery, for both patients and HC, thus proving to be a valuable and sensitive assessment tool (Keefe et al., 2004). The Italian version of the BACS was validated on a sample of 204 healthy adult subjects (Anselmetti et al., 2008). Since they found that the performance at the BACS was influenced by demographic characteristics such as age and education, these variables were considered to correct performance.

Although BACS was originally designed to assess cognitive performance in patients with schizophrenia, in the last decade, its application has also been extended to mood disorders, prompting sensibility in the detection of cognitive deficits also in MDD and BD (Cholet et al., 2014; Huang et al., 2020; Keefe et al., 2014). Specifically, cognitive impairment in mood disorders appeared to be midway within a continuum between cognitively intact HC and patients affected by schizophrenia with cognitive deficits.

5.2.2. Scoring procedure

For each subtest, the raw scores were recorded for each patient. Each raw score was then corrected for demographic variables for Italian sample according to their effect on the

specific domain. Specifically, both age and education were accounted for in raw scores in the verbal memory and motor speed domains. Verbal fluency, attention, and processing speed raw scores were corrected only for education and age. Raw scores at working memory were corrected for the square of both age and sex, whereas executive functions scores were controlled for the square of age and education (Anselmetti et al., 2008).

The corrected scores were used to calculate the ES for each cognitive domain (ES=0: pathological performance; ES=1: borderline performance; ES=2 or 3: intermediate performance; ES=4: near to the median value), allowing to directly compare cognitive performance either across domains and across subjects, net of the effect of age, sex, and education (Anselmetti et al., 2008). The ES method allows to integrate a qualitative and quantitative description of cognitive performance for each patient and enables retrospective ES calculation for available samples (Buonocore et al., 2021). The ES were then dichotomized into cognitively impaired (ES=0) and cognitively intact (ES≥1). Additionally, a composite score for each patient was estimated, reflecting a global impairment, by averaging the ESs across the six BACS domains (Palladini et al., 2022). Patients with a composite score ≤ 1 were assigned to the cognitively impaired group.

5.3. MRI data

5.3.1. Acquisition

T1-weightet, DTI, and fMRI images were acquired for 314 subjects (111 MDD and 203 BD) with two different 3.0 Tesla scanners. Acquisition of 196 subjects (50 MDD and 146 BD) was performed with Gyroscan Intera, Philips, Netherlands employing an 8 channels SENSE head coil (T1-weighted MPRAGE sequences: TR 25.00 ms, TE 4.6 ms, field of view FOV=230 mm,91 matrix=256×256, in-plane resolution 0.9×0.9 mm, yielding 220 transversal slices with a thickness of 0.8 mm). DTI sequences were also acquired for 171 subjects (46 MDD and 125 BD). SE EPI sequences (TR/TE=9000/58 ms, FoV (mm) 232(ap), 126 (fh), 240.00 (rl); acquisition matrix = 112×85; voxel acquisition 2.14×2.73×2.3; 55 contiguous, with in-plane voxel size 1.88×1.88 mm; SENSE acceleration factor=2; 1 b0 and 35 non-collinear directions of the diffusion gradients; b value=900 s/mm2) were used. A rsfMRI acquisition was also performed for 49 subjects (22 MDD and 27 BD), while instructed to keep their eyes closed. Gradient-echo EPI sequences were acquired. The scanning sessions comprised 100 sequential T2*-weighted volumes (interleaved ascending transverse slices covering whole brain), acquired using

an EPI pulse sequence (TR = 2000 ms; TE = 30 ms; flip angle = 77°; field of view = 220 mm; number of slices = 32; slice thickness = 4 mm; matrix size = 64 x 63 reconstructed up to 64 x 64 pixels). Two dummy scans before fMRI acquisition allowed obtaining longitudinal magnetization equilibrium. Total time acquisition was 3 min and 28 s.

Acquisition of 118 subjects (61 MDD and 57 BD) were instead performed with an Ingenia CX, Philips, The Netherlands using a 32-channel sensitivity encoding SENSE head coil (T1-weighted MPRAGE sequence: TR 8.00 ms, TE 3.7 ms, field of view FOV = 256 mm, matrix = 256 x 256, in-plane resolution 1 x 1 mm, yielding 182 transversal slices with a thickness of 1 mm). DTI sequences were also acquired for 111 subjects (56 MDD and 55 BD). SE EPI sequences (EPI factor= 43; TR/TE=5900/78 ms, FoV (mm) 232 (ap), 129 (fh), 240.00 (rl); acquisition matrix 112×85; 56 contiguous, 2.3-mm thick axial slices reconstructed with in-plane pixel size 1.88×1.88 mm; SENSE acceleration factor= 2; Multiband acceleration factor= 2; ten b0 and 96 non-collinear directions of the diffusion gradients: 60 b values=2855 s/mm2, 6 b values=700 s/mm2, 30 b values=1000 s/mm2) were acquired. Fat saturation was performed to avoid chemical shift artefacts. For rsfMRI acquisition 70 subjects (36 MDD and 34 BD) were instructed to keep their eyes closed. Gradient-echo EPI sequences were acquired. The scanning sessions comprised 200 sequential T2*-weighted volumes (interleaved ascending transverse slices covering whole brain, tilted 30° downward with respect to bicommissural line to reduce susceptibility artifacts in orbitofrontal region), acquired using an EPI pulse sequence (TR = 2000 ms; TE = 30 ms; flip angle = 85°; field of view = 192 mm; number of slices = 38; slice thickness = 3.7 mm; matrix size = 64 x 62 reconstructed up to 96 x 96 pixels). Six dummy scans before fMRI acquisition allowed obtaining longitudinal magnetization equilibrium. Total time acquisition was 6 min and 56 s.

5.3.2. Preprocessing

T1-weighted neuroanatomical images were processed using the Computational Anatomy Toolbox (CAT12) for statistical parametric mapping (SPM) toolbox (Gaser & Dahnke, 2016). T1 images were normalized in the Montreal Neurological Institute (MNI) space and segmented into GM, WM, and cerebrospinal fluid (CSF). Check of spatial alignment and sample homogeneity was performed to exclude outliers. GM maps were then smoothed with a 6 mm full width at half maximum (FWHM) Gaussian filter. The total intracranial volume (TIV) was then calculated.

DTI images were pre-processed using FMRIB Software Library (FSL) 6.0 tools (S. M. Smith et al., 2004, 2006; Woolrich et al., 2009). Images were corrected for the effects of eddy currents and head motion (Jenkinson et al., 2002; Jenkinson & Smith, 2001), and a brain mask was created using Brain Extraction Tool (BET) (S. M. Smith, 2002), which deletes non-brain tissues from the image. The images were also non-linearly registered to a standard template (FMRIB58-FA, FMRIB Centre University of Oxford, Department of Clinical Neurology, John Radcliffe Hospital Headington, Oxford, United Kingdom). A diffusion tensor model was fitted at each voxel to obtain voxel-wise maps of four diffusion indices: AD, RD, MD, and FA. The maps of all subjects were then merged into a common 4D image and skeletonized, as used in tract based spatial statistics (TBSS), in order to focus on the centers of all fibre bundles that are common to the participants (S. M. Smith et al., 2006).

Preprocessing was performed the Harmonized AnaLysis of Functional MRI pipeline (HALFpipe) version 1.2.1 (Waller et al., 2022). Functional images underwent motion correction, slice time correction, susceptibility distortion correction, coregistration, and spatial normalization to the MNI152NLin2009cAsym template. The preprocessed images were smoothed with a 6 mm FWHM kernel, and grand mean scaled. Either a Gaussian-weighted high-pass (width of 125 s) or a frequency-based band pass (low cut-off of 0.01 Hz and a high cut-off of 0.1 Hz) temporal filter was applied. Anatomical component correction (aCompCor) was performed to remove the effect of physiological nuisance regressors represented by the top five principal components of CSF signal (Muschelli et al., 2014). Several features were then extracted from the preprocessed and denoised data, comprising SBC, ABC, DRCs, ReHo, and fALFF. The SBC, ABC, and DRCs features were based on the Gaussian-weighted filter, whereas ReHo and fALFF were calculated using the frequency-based filter. SBC calculation generated a map for each seed (N=44), which represents the FC between that seed and all the other voxels. For ABC a pairwise functional connectivity matrix was obtained for each subject between all pairs of regions included, for a total of 434 regions of interest (ROIs): 400 ROIs matched to 17 large-scale resting-state networks from the Schaefer atlas (Schaefer et al., 2018), 17 subcortical ROIs from the Harvard-Oxford Atlas (Desikan et al., 2006) and 17 cerebellar ROIs from the Buckner 17-network atlas (Buckner et al., 2011). DRCs (Beckmann et al., 2009) were calculated on a set of 14 networks (Shirer et al., 2012) and represented the strength of

each network in the individual. For each subject, a voxel-wise map of ReHo and fALFF was also obtained: ReHo (Zang, 2004) is the local synchronization between a given voxel and its nearest neighbouring voxels, whereas fALFF (Q.-H. Zou et al., 2008) can be defined as the amplitude of the spontaneous neural activity.

5.3.3. Harmonization procedure

When samples are collected with different scanners, systematic differences in equipment or parameter configurations could generate technical artifacts (Fortin et al., 2017). Such non-biological effects are often referred to as batch effects (Leek et al., 2010), and are known to be sources of bias and variance in the acquired images (Fortin et al., 2018). Importantly, this source of variation across different scanners can negatively affect the consistency and reproducibility of the downstream analyses and findings (Fortin et al., 2017, 2018; Leek et al., 2010; Yi et al., 2018). Since in the present study samples were collected using two different scanners, a harmonization step was applied on the extracted features, prior to second-level analyses, to rule out the presence of batch effects. The harmonization was performed using ComBat (Johnson et al., 2007), a batch-effect correction tool which has been shown to be effective in removing unwanted inter-scanner variation on DTI (Fortin et al., 2017), VBM (Fortin et al., 2018), and rsfMRI (M. Yu et al., 2018) measures. The advantage of ComBat over the other traditionally employed methods resides in the use of empirical Bayes for parameter estimation, thus correcting for both additive and multiplicative batch effects (Fortin et al., 2017, 2018; M. Yu et al., 2018). The methods typically used to remove batch effects are based on residualization of the scanner-related information (Fortin et al., 2018). Let y_{ijv} be the feature value for scanner i, participant j, and voxel or ROI v. Let also X be an $n \times p$ matrix of p biological covariates of interest for n subjects and let Z be the $n \times k$ matrix representing the information about the scanner with which each observation has been acquired, with k being the number of scanners. Residualization methods fit a regression model:

$$y_{ijv} = \alpha_v + X_{ij}\beta_v + Z_{ij}\theta_v + \varepsilon_{ijv},$$

where α_v is the average feature value for voxel or ROI v, θ_v is the vector of k coefficients representing the scanner effect on the feature, β_v is the vector of k coefficients representing the effect of the biological covariates on the feature. Both θ_v and β_v are derived using an ordinary least square regression. The scanner effect is then removed form the original data by:

$$y_{ijv}^{Adj} = y_{ijv} - Z_{ij}\theta_v.$$

ComBat extends these residualization-based techniques using emipircal Bayes to estimate the parameters related to the scanner effects, improving the robustness and the reliability, especially with small sample sizes (Fortin et al., 2018). The ComBat model can be expressed as:

$$y_{ijv} = \alpha_v + X_{ij}\beta_v + Z_{ij}\theta_v + \delta_{iv}\varepsilon_{ijv},$$

where the δ_{iv} parameter describes the multiplicative scanner effect. ComBat can then be solved by:

$$y_{ijv}^{Adj} = \frac{y_{ijv} - \alpha_v - X_{ij}\beta_v - Z_{ij}\theta_v}{\delta_{iv}} + \alpha_v + X_{ij}\beta_v,$$

where the set of biological covariates X entered comprised age, sex, and diagnosis for all the features, except for the harmonization of the VBM maps where also the TIV was inserted. Note that in this context, the aim is to preserve the effect of the covariates on the features.

The performance of ComBat was assessed through the implementation of predictive models for the classification of the scanner with which each observation was acquired using the harmonized features. The classification accuracy can then be used as a proxy for harmonization quality: the lowest the accuracy, the lowest the residual information related to scanner in the feature, and the highest the harmonization quality.

5.4. Inflammatory markers

Bio-Plex Pro™ Human Cytokine 48-Plex and Bio-Plex Pro™ Human Chemokine 40-Plex assays (BIO-RAD) were used to detect plasma concentrations of immune analytes in 128 subjects (71 MDD and 57 BD). Plasma concentrations were detected through the bead-based Luminex system according to xMAP technology (Luminex 200™ system, Merck Millipore).

Analysed cytokines included: IL-1α, IL-1β, IL-1ra, IL-2, IL-2Rα, IL-3, IL-4, IL-5, IL-6, IL-7, IL-9, IL-10, IL-12 (p70), IL-12 (p40), IL-13, IL-15, IL-16, IL-17, IL-18, interferon (IFN)-α2, IFN-γ, TNF-α, TNF-β, hepatocyte growth factor (HGF), LIF; chemokines: C-C motif ligand (CCL) 21 (CCL21)/6Ckine, C-X-C motif chemokine ligand (CXCL)9/MIG, IL-8/CXCL8, BCA-1/CXCL13, CTACK/CCL27, ENA-78/CXCL5, Eotaxin/CCL11, Eotaxin-2/CCL24, Eotaxin-3/CCL26, Fractalkine/CX3CL1, GCP-2/CXCL6,GRO-α/CXCL1, GRO-β/CXCL2, I-309/CCL1, IP-10/CXCL10, I-

TAC/CXCL11, MCP-1/CCL2, MCP-2/CCL8, MCP-3/CCL7, MCP-4/CCL13, MDC/CCL22, MIP-1α/CCL3, MIP-1β, MIP-1δ/CCL15, MIP-3α/CCL20, MIP-3β/CCL19, MPIF-1/CCL23, SCYB16/CXCL16, SDF-1α+β/CXCL12, SDF-1α, TARC/CCL17, TECK/CCL25, RANTES/CCL5, TRAIL. Analysed growth factors consisted of granulocyte macrophage colony stimulating factor (GM-CSF), macrophage colony stimulating factor (M-CSF), basic fibroblast growth factor (basic FGF), granulocyte colony stimulating factor (G-CSF), β-nerve growth factor (β-NGF), platelet-derived growth factor-BB (PDGF-BB), vascular endothelial growth factor (VEGF), stem cell factor (SCF), stem cell growth factor β (SCGF-β); macrophage migration inhibitory factor (MIF).

These multiplexed sandwich immunoassays were developed from commercially available capture and detection antibodies and standard proteins, validated and approved by EDI-GMBH. Luminex experiments were performed according to the pre-optimized protocol provided by the manufacturer. Analyses were performed on observed concentrations (pg/ml) calculated using Belysa Immunoassay software (version 1.2).

Only cytokines with non-detected (i.e., missing) values < 10% were included in the analyses.

5.5. Genetic data

5.5.1. Genotyping

Blood samples were genotyped for 336 patients (126 MDD and 210 BD) using Infinium PsychArray 24 BeadChip (Illumina, Inc., San Diego), which is a cost-effective, high-density microarray specifically developed in collaboration with the Psychiatric Genomics Consortium for large-scale genetic studies focused on psychiatric predisposition and risk. Almost 600.000 genetic variants can be extracted with this microarray, including ~271.000 proven tag SNPs found on the Infinium Core-24 BeadChip, ~277,000 markers from the Infinium Exome-24 v1.1 BeadChip, and ~50,000 markers associated with common psychiatric disorders, such as schizophrenia, BD, autism-spectrum disorders (ASD), attention deficit hyperactivity disorder (ADHD), MDD, obsessive compulsive disorder, anorexia nervosa, and Tourette's syndrome (https://www.illumina.com/products/by-type/microarray-kits/infinium-psycharray.html).

5.5.2. Quality control

Quality control (QC) of genotyped data was performed following the guidelines provided by (C. A. Anderson et al., 2010), using PLINK 1.9 (Purcell et al., 2007). QC was performed separately for individuals and genetic variants. In the individuals' QC, subjects were discarded if the genotyped sex was different from the phenotypical one, the genotype rate was lower than 95%, or had an outlying autosomal heterozygosity (F_{het} > ±0.2). After QC on subjects, genetic variants with call rate lower than 98%, with minor allele frequency (MAF) lower than 1%, or extremely deviant from Hardy-Weinberg equilibrium ($p<10^{-8}$) were removed. Relatedness and population ancestry of the sample were also checked. Observations were removed if they had a degree of recent shared ancestry (i.e., identity by descendent, IBD) higher than 0.1875, which is the threshold corresponding to the halfway between third- and second-degree relatives. Lastly, subjects were excluded if the genotype distribution was deviant from the current population, that is if the first two components resulting from a principal component analysis were outlying more than 5 standard deviations from the mean.

5.5.3. Imputation

Genotype Imputation was performed via the Michigan Imputation Server using Minimac4, version 1.7.1 (https://imputationserver.sph.umich.edu), selecting the following options. The reference panel chosen was the latest version of the Haplotype Reference Consortium panel (HRC r1.1 2016), based on the GRCh37/hg19 array build. Imputation stability and reliability were assessed using the R2 metric, filtering out values under the 0.003 accuracy threshold. Eagle (version 2.4) was used for phasing. Post-imputation files contained approximately 14 million variants.

5.5.4. PRS calculation

PRSs for MDD, BD, schizophrenia, ADHD, anorexia nervosa, ASD, neuroticism, coronary artery disease (CAD), education, body mass index (BMI), CRP, and T2D were calculated. Each PRS was computed for each subject j as:

$$PRS_j = \sum_{i}^{N} \beta_i * dosage_{ij}$$

where β_i is the magnitude of association between the variant alleles at the i^{th} SNP with $i, 1 \dots N$ and the disorder derived from the latest GWAS, whereas $dosage_{ij}$ consists in the allele counts of variants associated with the disorder of interest. The association weights β_i were calculated with PRS-CS algorithm (source code available at

https://github.com/getian107/PRScs), a Python-based tool that infers variants' posterior effect sizes using the summary statistics derived from the GWAS listed in Table 7. PRSs were then mean centered and scaled to have zero-mean and unit variance.

Table 9. Main information of GWAS studies used as reference for calculating effect-sizes of variants associated with the psychiatric disorders of interest.

Disorder of interest	Authors	Total size of the sample	Number of cases	Number of controls	Ethnicity	PubMed ID
MDD	Wray et al. (2018)	500,199	170,756	329,443	European	29700475
BD	Mullins et al. (2021)	413,466	41,917	371,549	European	34002096
Schizophrenia	Trubetskoy et al. (2022)	320,404	76,755	243,649	European	35396580
ADHD	Demontis et al. (2018)	53,293	19,099	34,194	European	30478444
Anorexia	Watson et al. (2019)	72,517	16,992	55,525	European	31308545
Autism	Grove et al. (2019)	46,350	18,381	27,969	European	30804558
CAD	Nikpay et al. (2015)	184,305	60,801	123,504	European South Asian East Asian Hispanic African American	26343387
Education	Lee et al. (2018)	1,131,881	N.A.	N.A.	European	30038396
BMI	Locke et al. (2015)	322,154	N.A.	N.A.	European	25673413
CRP	Lightart et al. (2018)	204,402	N.A.	N.A.	European	30388399
T2D	Scott et al. (2017)	159,208	26,676	132,532	European	28566273
Neuroticism	De Moor et al. (2015)	63,661	N.A.	N.A.	European	25993607

Abbreviations: ADHD, attention and hyperactivity disorder; BD, bipolar disorder; BMI, body mass index; CAD, coronary artery disease; CRP, C-reactive protein; MDD, major depressive disorder; N.A., not available; T2D, type 2 diabetes.

5.6 Machine learning analyses

5.6.1. Algorithms

The ML analyses were performed with different algorithms based on the type of data. For vector data (i.e., PIMs, BACS scores, and PRSs) an EN penalized regression was employed. The EN penalization is the combination of the L_1 and L_2 regularizations which force a shrinkage of the coefficients toward the zero (Friedman et al., 2001; H. Zou & Hastie, 2005). This reduces the contribution of irrelevant, noisy, and redundant features, reducing overfitting in turn. The EN penalization can be defined as:

$$P_\alpha(\beta) = \sum_{j=1}^{p} \left(\frac{(1-\alpha)}{2} \beta_j^2 + \alpha \left| \beta_j \right| \right),$$

where α defines the trade-off between the L_1 and L_2 regularizations and β_j is the j^{th} regression coefficient for $j = 1, \dots, p$. The EN penalized regression can then be obtained by adding the EN penalization term $P_\alpha(\beta)$ to the regression loss function:

$$\left(\frac{1}{N} D(\beta_0, \beta) + \lambda P_\alpha(\beta)\right),$$

where N is the number of observations, D is the deviance of the model fit to the target (i.e., subject's diagnosis), β_0 is the intercept, β is a vector of p regression coefficients, and λ is the weighting factor that represents the strength of the regularization.

Since the number of BD patients was higher than that of MDD patients, class weighting was applied (King & Zeng, 2001) in the EN model to prevent from preferentially predicting the largest group. The weight of each class (w_c) was calculated to be inversely proportional to class frequencies, according to $w_c = \frac{n}{n_c s_c}$, where n_c is the number of classes, and s_c is the number of subjects in the class c (King & Zeng, 2001; H. Zou & Hastie, 2005).

For map- and matrix-like data a binary SVM was used as implemented in the Pattern Recognition for Neuroimaging Toolbox (PRoNTo - Schrouff et al., 2013). Let $X \in \mathbb{R}^{n,p}$ be the data matrix with n observations and p features, with x_i being the feature vector of the i^{th} observation and y_i its associated target class label. SVM aims to find the hyperplane that separates the observations of the two classes maximizing the margin, that is the distance between the hyperplane and the closest training observation of each class, named support vectors (Vapnik, 2000). The hyperplane can be defined by the dot product between a weight vector w and the feature vector x_i, resulting in the decision function:

$$f(x_i) = w \cdot x_i + b,$$

where b is a bias term. The decision function f is then solved by:

$$\min \frac{1}{2}\|w\|^2 + C \sum_i \xi_i \ \ subject\ to\ y_i(w \cdot x_i + b) \geq 1 - \xi_i \ \forall i \xi_i \geq 0 \ \forall i,$$

where $\sum_i \xi_i$ is an upper-bound on the number of classification errors, and C corresponds to the soft-margin parameter. The soft-margin parameter defines the trade-off between the classification performance and the maximization of the margin, that is the generalization ability. For high values of C, SVM tries to maximize the margin irrespective of the classification errors, whereas for low values of C the margin is narrower. However, it could be the case that data are not linearly separable in the original

feature space, but they may be linearly separable in a higher-dimensional space. It turns out, though, that the so called "kernel trick" can be used to efficiently perform calculations in the original feature space, yet implicitly projecting the data into a higher-dimensional space, where linear separation may be possible, thus avoiding computational overhead (Shawe-Taylor & Cristianini, 2004). For this reason, the feature vector x_i is replaced by the kernel function $K(x, x_i)$, that is this case was defined as the dot product:

$$K(x, x_i) = x \cdot x_i,$$

which represents a similarity matrix.

For the models in which multiple maps per subject were present (i.e., DRCs and SBC), a MKL algorithm was used for prediction. MKL is a SVM-based method that can linearly combine multiple kernels, rather than just using a single kernel function (Schrouff et al., 2018). One of the main advantages of MKL is the aim to find the best combination over a set of kernel functions, which is more efficient than combining different classifiers that use a single kernel (Schrouff et al., 2018). In MKL, the kernel function $K(x, x_i)$ can be considered as a linear combination of the basis M basis kernels:

$$K(x, x_i) = \sum_{m=1}^{M} d_m K_m(x, x_i)\,, with\ d_m \geq 0, \sum_{m=1}^{M} d_m = 1.$$

The decision function and its resolution are based on the primal SVM formulation, and therefore are:

$$f(x_i) = \sum_{m} w_m \cdot x_i + b,$$

solved by:

$$\min \frac{1}{2} \sum_{m} \frac{1}{d_m} \|w_m\|^2 + C \sum_{i} \xi_i$$

$$subject\ to\ y_i \left(\sum_{m} w_m \cdot x_i + b \right) \geq 1 - \xi_i\ \forall i \xi_i \geq 0\ \forall i \sum_{m} d_m = 1, d_m \geq 0\ \forall m.$$

5.6.2. Model estimation and statistical inference

Each feature was entered into the model separately, to classify between MDD and BD patients. Except for the models using the BACS as a predictor, the same models were also run to differentiate between the patients with a cognitive deficit and those without, both for the composite score and each cognitive domain separately.

The predictive performance of all the models was assessed through a 5-fold nested CV procedure (Pereira et al., 2009; Varoquaux et al., 2017). In the inner loop, the hyper-parameters of each algorithm were optimized. For EN, the α hyper-parameter was optimized over a set of 10 linearly spaced values in the range 0-1, whereas the λ hyper-parameter was optimized on 100 logarithmically spaced values between 10^{-5} and 10^{5}. For SVM and MKL, the soft-margin parameter C was optimized over a set of 6 values (0.01, 0.1, 1, 10, 100, 1000). In all the models, the effect of age, sex, education, medication load, number of previous episodes, duration of illness, and the interaction between age and sex was removed. Additionally, the square of age was also entered as a confounder in order to remove age-related non-linear effects (Alfaro-Almagro et al., 2021). For medication load, we categorized each medication into low-dose or high-dose groupings, scored as 0 (no medication), 1 (low dosage), or 2 (high dosage). We then combined all individual medication scores for each drug category in each individual participant to obtain a single composite score (Sackeim, 2001).

To avoid class imbalance, random subsampling was performed, preventing the algorithm to inflate the classification accuracy by systematically predicting the majority class. The predictions made by the model in the outer loop on the test set were finally used to calculate several metrics of performance, namely: BA, sensitivity, specificity, PPV, NPV, and AUC.

Each model underwent a 1000-permutation test to assess whether the classification accuracy was significantly different from chance level (Ojala & Garriga, 2009). Permutation test consisted in randomly shuffling the observations between the target classes, to break the relationship between the predictors and the target. At each iteration, a random classifier is trained, creating a distribution of null accuracy scores. Significance is defined as the number of times in which the random classifier performed better than the tested classifier. Significance threshold was set to $p<0.05$, applying Bonferroni correction for multiple comparisons within each target. Consequently, the threshold was $p<0.05/13=0.004$ for the models differentiating between MDD and BD, and $p<0.05/12=0.004$ for the models identifying cognitively impaired patients.

Additionally, for the models that resulted to be significantly different from null distribution, feature importance was calculated across CV folds to investigate the most predictive features. For the EN model, a 1000-bootstrap procedure was applied.

Contributing features were defined by 95% confidence intervals. For kernel methods, the most predictive features were defined as having median feature importance across folds >0. Feature importance was computed by applying an atlas-based parcellation to the data and running a MKL learning model, building one kernel for each parcel. For VBM the Harvard-Oxford atlas was used (S. M. Smith et al., 2004), for DTI features the Johns Hopkins University atlas was employed (Mori et al., 2008), whereas on fMRI features the Brainnetome atlas was applied (Fan et al., 2016). For vector-like features, a dummy atlas was purposefully created, indicating one each value as a separate parcel.

REFERENCES

Akiskal, H. S., & Benazzi, F. (2003). Family history validation of the bipolar nature of depressive mixed states. *Journal of Affective Disorders*, *73*(1–2), 113–122. https://doi.org/10.1016/s0165-0327(02)00330-0

Alfaro-Almagro, F., McCarthy, P., Afyouni, S., Andersson, J. L., Bastiani, M., Miller, K. L., Nichols, T. E., & Smith, S. M. (2021). Confound modelling in UK Biobank brain imaging. *NeuroImage*, *224*, 117002.

Allen, N. C., Bagade, S., McQueen, M. B., Ioannidis, J. P. A., Kavvoura, F. K., Khoury, M. J., Tanzi, R. E., & Bertram, L. (2008). Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: The SzGene database. *Nature Genetics*, *40*(7), 827–834. https://doi.org/10.1038/ng.171

Almulla, A. F., Algon, A. A. A., Tunvirachaisakul, C., Al-Hakeim, H. K., & Maes, M. (2023). *T helper-1 activation via interleukin-16 is a key phenomenon in the acute phase of severe, first-episode major depressive disorder and suicidal behaviors* (p. 2023.04.16.23288643). medRxiv. https://doi.org/10.1101/2023.04.16.23288643

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5^{TM}, 5th ed* (pp. xliv, 947). American Psychiatric Publishing, Inc. https://doi.org/10.1176/appi.books.9780890425596

Anand, A., Li, Y., Wang, Y., Lowe, M. J., & Dzemidzic, M. (2009). Resting state corticolimbic connectivity abnormalities in unmedicated bipolar disorder and unipolar depression. *Psychiatry research*, *171*(3), 189–198. https://doi.org/10.1016/j.pscychresns.2008.03.012

Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. *Nature protocols*, *5*(9), 1564–1573.

Anderson, G., & Maes, M. (2015). Bipolar disorder: Role of immune-inflammatory cytokines, oxidative and nitrosative stress and tryptophan catabolites. *Current Psychiatry Reports*, *17*(2), 8. https://doi.org/10.1007/s11920-014-0541-1

Andreazza, A. C. (2012). Combining redox-proteomics and epigenomics to explain the involvement of oxidative stress in psychiatric disorders. *Molecular bioSystems*, *8*(10), 2503–2512. https://doi.org/10.1039/c2mb25118c

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron*, *65*(4), 550–562. https://doi.org/10.1016/j.neuron.2010.02.005

Angst, J., Cui, L., Swendsen, J., Rothen, S., Cravchik, A., Kessler, R. C., & Merikangas, K. R. (2010). Major depressive disorder with subthreshold bipolarity in the National Comorbidity Survey Replication. *The American Journal of Psychiatry*, *167*(10), 1194–1201. https://doi.org/10.1176/appi.ajp.2010.09071011

Anselmetti, S., Poletti, S., Ermoli, E., Bechi, M., Cappa, S., Venneri, A., Smeraldi, E., & Cavallaro, R. (2008). The Brief Assessment of Cognition in Schizophrenia. Normative data for the Italian population. *Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, *29*(2), 85–92. https://doi.org/10.1007/s10072-008-0866-9

Anticevic, A., Brumbaugh, M. S., Winkler, A. M., Lombardo, L. E., Barrett, J., Corlett, P. R., Kober, H., Gruber, J., Repovs, G., Cole, M. W., Krystal, J. H., Pearlson,

G. D., & Glahn, D. C. (2013). Global Prefrontal and Fronto-amygdala Dysconnectivity in Bipolar I Disorder with Psychosis History. *Biological psychiatry*, *73*(6), 565–573. https://doi.org/10.1016/j.biopsych.2012.07.031

Anticevic, A., Cole, M. W., Murray, J. D., Corlett, P. R., Wang, X.-J., & Krystal, J. H. (2012). The role of default network deactivation in cognition and disease. *Trends in Cognitive Sciences*, *16*(12), 584–592. https://doi.org/10.1016/j.tics.2012.10.008

Au, B., Smith, K. J., Gariépy, G., & Schmitz, N. (2015). The longitudinal associations between C-reactive protein and depressive symptoms: Evidence from the English Longitudinal Study of Ageing (ELSA). *International Journal of Geriatric Psychiatry*, *30*(9), 976–984. https://doi.org/10.1002/gps.4250

Augustine, J. R. (1996). Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Research. Brain Research Reviews*, *22*(3), 229–244. https://doi.org/10.1016/s0165-0173(96)00011-2

Aung, W. Y., Mar, S., & Benzinger, T. L. (2013). Diffusion tensor MRI as a biomarker in axonal and myelin damage. *Imaging in Medicine*, *5*(5), 427–440. https://doi.org/10.2217/iim.13.49

Bai, Y.-M., Su, T.-P., Li, C.-T., Tsai, S.-J., Chen, M.-H., Tu, P.-C., & Chiou, W.-F. (2015). Comparison of pro-inflammatory cytokines among patients with bipolar disorder and unipolar depression and normal controls. *Bipolar Disorders*, *17*(3), 269–277. https://doi.org/10.1111/bdi.12259

Baker, L. D., Cross, D. J., Minoshima, S., Belongia, D., Watson, G. S., & Craft, S. (2011). Insulin resistance and Alzheimer-like reductions in regional cerebral glucose metabolism for cognitively normal adults with prediabetes or early type

2 diabetes. *Archives of Neurology*, *68*(1), 51–57. https://doi.org/10.1001/archneurol.2010.225

Bakkour, N., Samp, J., Akhras, K., El Hammi, E., Soussi, I., Zahra, F., Duru, G., Kooli, A., & Toumi, M. (2014). Systematic review of appropriate cognitive assessment instruments used in clinical trials of schizophrenia, major depressive disorder and bipolar disorder. *Psychiatry Research*, *216*(3), 291–302. https://doi.org/10.1016/j.psychres.2014.02.014

Bannerman, D. M., Rawlins, J. N. P., McHugh, S. B., Deacon, R. M. J., Yee, B. K., Bast, T., Zhang, W.-N., Pothuizen, H. H. J., & Feldon, J. (2004). Regional dissociations within the hippocampus—Memory and anxiety. *Neuroscience and Biobehavioral Reviews*, *28*(3), 273–283. https://doi.org/10.1016/j.neubiorev.2004.03.004

Barbosa, I. G., Ferreira, R. de A., Rocha, N. P., Mol, G. C., da Mata Chiaccjio Leite, F., Bauer, I. E., & Teixeira, A. L. (2018). Predictors of cognitive performance in bipolar disorder: The role of educational degree and inflammatory markers. *Journal of Psychiatric Research*, *106*, 31–37. https://doi.org/10.1016/j.jpsychires.2018.09.003

Barbosa, I. G., Rocha, N. P., Bauer, M. E., de Miranda, A. S., Huguet, R. B., Reis, H. J., Zunszain, P. A., Horowitz, M. A., Pariante, C. M., & Teixeira, A. L. (2013). Chemokines in bipolar disorder: Trait or state? *European Archives of Psychiatry and Clinical Neuroscience*, *263*(2), 159–165. https://doi.org/10.1007/s00406-012-0327-6

Barbosa, I. G., Rocha, N. P., Huguet, R. B., Ferreira, R. A., Salgado, J. V., Carvalho, L. A., Pariante, C. M., & Teixeira, A. L. (2012). Executive dysfunction in euthymic

bipolar disorder patients and its association with plasma biomarkers. *Journal of Affective Disorders*, *137*(1–3), 151–155. https://doi.org/10.1016/j.jad.2011.12.034

Barnett, J. H., & Smoller, J. W. (2009). The genetics of bipolar disorder. *Neuroscience*, *164*(1), 331–343. https://doi.org/10.1016/j.neuroscience.2009.03.080

Bart, C. P., Titone, M. K., Ng, T. H., Nusslock, R., & Alloy, L. B. (2021). Neural reward circuit dysfunction as a risk factor for bipolar spectrum disorders and substance use disorders: A review and integration. *Clinical Psychology Review*, *87*, 102035. https://doi.org/10.1016/j.cpr.2021.102035

Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., & Aldroubi, A. (2000). In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, *44*(4), 625–632. https://doi.org/10.1002/1522-2594(200010)44:4<625::aid-mrm17>3.0.co;2-o

Becking, K., Boschloo, L., Vogelzangs, N., Haarman, B. C. M., Riemersma-van der Lek, R., Penninx, B. W. J. H., & Schoevers, R. A. (2013). The association between immune activation and manic symptoms in patients with a depressive disorder. *Translational Psychiatry*, *3*(10), e314. https://doi.org/10.1038/tp.2013.87

Beckmann, C. F., Mackay, C. E., Filippini, N., & Smith, S. M. (2009). Group comparison of resting-state FMRI data using multi-subject ICA and dual regression. *Neuroimage*, *47*(Suppl 1), S148.

Benazzi, F. (2007). Bipolar disorder—Focus on bipolar II disorder and mixed depression. *Lancet (London, England)*, *369*(9565), 935–945. https://doi.org/10.1016/S0140-6736(07)60453-X

Benedetti, F., Absinta, M., Rocca, M. A., Radaelli, D., Poletti, S., Bernasconi, A., Dallaspezia, S., Pagani, E., Falini, A., Copetti, M., Colombo, C., Comi, G., Smeraldi, E., & Filippi, M. (2011). Tract-specific white matter structural disruption in patients with bipolar disorder. *Bipolar Disorders*, *13*(4), 414–424. https://doi.org/10.1111/j.1399-5618.2011.00938.x

Benedetti, F., Aggio, V., Pratesi, M. L., Greco, G., & Furlan, R. (2020). Neuroinflammation in Bipolar Depression. *Frontiers in Psychiatry*, *11*, 71. https://doi.org/10.3389/fpsyt.2020.00071

Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: «liking», «wanting», and learning. *Current Opinion in Pharmacology*, *9*(1), 65–73. https://doi.org/10.1016/j.coph.2008.12.014

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: Insights from spatial processing. *Nature Reviews Neuroscience*, *9*(3), Articolo 3. https://doi.org/10.1038/nrn2335

Bora, E., Fornito, A., Pantelis, C., & Yücel, M. (2012). Gray matter abnormalities in Major Depressive Disorder: A meta-analysis of voxel based morphometry studies. *Journal of Affective Disorders*, *138*(1–2), 9–18. https://doi.org/10.1016/j.jad.2011.03.049

Bora, E., Fornito, A., Yücel, M., & Pantelis, C. (2010). Voxelwise meta-analysis of gray matter abnormalities in bipolar disorder. *Biological Psychiatry*, *67*(11), 1097–1105. https://doi.org/10.1016/j.biopsych.2010.01.020

Border, R., Johnson, E. C., Evans, L. M., Smolen, A., Berley, N., Sullivan, P. F., & Keller, M. C. (2019). No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression Across Multiple Large

Samples. *The American Journal of Psychiatry*, *176*(5), 376–387. https://doi.org/10.1176/appi.ajp.2018.18070881

Boretius, S., Escher, A., Dallenga, T., Wrzos, C., Tammer, R., Brück, W., Nessler, S., Frahm, J., & Stadelmann, C. (2012). Assessment of lesion pathology in a new animal model of MS by multiparametric MRI and DTI. *NeuroImage*, *59*(3), 2678–2688. https://doi.org/10.1016/j.neuroimage.2011.08.051

Bortolato, B., Carvalho, A. F., Soczynska, J. K., Perini, G. I., & McIntyre, R. S. (2015). The Involvement of TNF-α in Cognitive Dysfunction Associated with Major Depressive Disorder: An Opportunity for Domain Specific Treatments. *Current Neuropharmacology*, *13*(5), 558–576. https://doi.org/10.2174/1570159x13666150630171433

Bortolato, B., Miskowiak, K. W., Köhler, C. A., Maes, M., Fernandes, B. S., Berk, M., & Carvalho, A. F. (2016). Cognitive remission: A novel objective for the treatment of major depression? *BMC Medicine*, *14*, 9. https://doi.org/10.1186/s12916-016-0560-3

Bourne, C., Aydemir, Ö., Balanzá-Martínez, V., Bora, E., Brissos, S., Cavanagh, J. T. O., Clark, L., Cubukcuoglu, Z., Dias, V. V., Dittmann, S., Ferrier, I. N., Fleck, D. E., Frangou, S., Gallagher, P., Jones, L., Kieseppä, T., Martínez-Aran, A., Melle, I., Moore, P. B., … Goodwin, G. M. (2013). Neuropsychological testing of cognitive impairment in euthymic bipolar disorder: An individual patient data meta-analysis. *Acta Psychiatrica Scandinavica*, *128*(3), 149–162. https://doi.org/10.1111/acps.12133

Brainstorm Consortium, Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., Duncan, L., Escott-Price, V., Falcone, G. J., Gormley, P., Malik, R.,

Patsopoulos, N. A., Ripke, S., Wei, Z., Yu, D., Lee, P. H., Turley, P., Grenier-Boley, B., Chouraki, V., … Murray, R. (2018). Analysis of shared heritability in common disorders of the brain. *Science (New York, N.Y.)*, *360*(6395), eaap8757. https://doi.org/10.1126/science.aap8757

Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, *16*(3), 199–215.

Bressler, S. L., & Menon, V. (2010). Large-scale brain networks in cognition: Emerging methods and principles. *Trends in Cognitive Sciences*, *14*(6), 277–290. https://doi.org/10.1016/j.tics.2010.04.004

Brietzke, E., Stertz, L., Fernandes, B. S., Kauer-Sant'anna, M., Mascarenhas, M., Escosteguy Vargas, A., Chies, J. A., & Kapczinski, F. (2009). Comparison of cytokine levels in depressed, manic and euthymic patients with bipolar disorder. *Journal of Affective Disorders*, *116*(3), 214–217. https://doi.org/10.1016/j.jad.2008.12.001

Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., de Girolamo, G., de Graaf, R., Demyttenaere, K., Hu, C., Iwata, N., Karam, A. N., Kaur, J., Kostyuchenko, S., Lépine, J.-P., Levinson, D., Matschinger, H., Mora, M. E. M., Browne, M. O., Posada-Villa, J., … Kessler, R. C. (2011). Cross-national epidemiology of DSM-IV major depressive episode. *BMC Medicine*, *9*(1), 90. https://doi.org/10.1186/1741-7015-9-90

Brunoni, A. R., Supasitthumrong, T., Teixeira, A. L., Vieira, E. L., Gattaz, W. F., Benseñor, I. M., Lotufo, P. A., Lafer, B., Berk, M., Carvalho, A. F., & Maes, M. (2020). Differences in the immune-inflammatory profiles of unipolar and bipolar

depression. *Journal of Affective Disorders*, *262*, 8–15. https://doi.org/10.1016/j.jad.2019.10.037

Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C., & Yeo, B. T. T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(5), 2322–2345. https://doi.org/10.1152/jn.00339.2011

Buonocore, M., Inguscio, E., Bosinelli, F., Bechi, M., Agostoni, G., Spangaro, M., Martini, F., Bianchi, L., Cocchi, F., Guglielmino, C., Repaci, F., Bosia, M., & Cavallaro, R. (2021). Disentangling Cognitive Heterogeneity in Psychotic Spectrum Disorders. *Asian Journal of Psychiatry*, *60*, 102651. https://doi.org/10.1016/j.ajp.2021.102651

Cha, D. S., De Michele, F., Soczynska, J. K., Woldeyohannes, H. O., Kaidanovich-Beilin, O., Carvalho, A. F., Malhi, G. S., Patel, H., Sim, K., Brietzke, E., Mansur, R., Dunlop, K. A. M., Alsuwaidan, M., Baskaran, A., Fagiolini, A., Reznikov, R., Kudlow, P. A., & McIntyre, R. S. (2014). The putative impact of metabolic health on default mode network activity and functional connectivity in neuropsychiatric disorders. *CNS & Neurological Disorders Drug Targets*, *13*(10), 1750–1758. https://doi.org/10.2174/1871527313666141130205024

Chai, X. J., Whitfield-Gabrieli, S., Shinn, A. K., Gabrieli, J. D. E., Nieto Castañón, A., McCarthy, J. M., Cohen, B. M., & Ongür, D. (2011). Abnormal medial prefrontal cortex resting-state connectivity in bipolar disorder and schizophrenia. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *36*(10), 2009–2017. https://doi.org/10.1038/npp.2011.88

Chen, C.-H., Suckling, J., Lennox, B. R., Ooi, C., & Bullmore, E. T. (2011). A quantitative meta-analysis of fMRI studies in bipolar disorder. *Bipolar Disorders*, *13*(1), 1–15. https://doi.org/10.1111/j.1399-5618.2011.00893.x

Chen, D. T., Jiang, X., Akula, N., Shugart, Y. Y., Wendland, J. R., Steele, C. J. M., Kassem, L., Park, J.-H., Chatterjee, N., Jamain, S., Cheng, A., Leboyer, M., Muglia, P., Schulze, T. G., Cichon, S., Nöthen, M. M., Rietschel, M., BiGS, McMahon, F. J., … Strauss, J. (2013). Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. *Molecular Psychiatry*, *18*(2), 195–205. https://doi.org/10.1038/mp.2011.157

Chen, L., Wang, Y., Niu, C., Zhong, S., Hu, H., Chen, P., Zhang, S., Chen, G., Deng, F., Lai, S., Wang, J., Huang, L., & Huang, R. (2018). Common and distinct abnormal frontal-limbic system structural and functional patterns in patients with major depression and bipolar disorder. *NeuroImage. Clinical*, *20*, 42–50. https://doi.org/10.1016/j.nicl.2018.07.002

Chen, M.-H., Chang, W.-C., Hsu, J.-W., Huang, K.-L., Tu, P.-C., Su, T.-P., Li, C.-T., Lin, W.-C., & Bai, Y.-M. (2019). Correlation of proinflammatory cytokines levels and reduced gray matter volumes between patients with bipolar disorder and unipolar depression. *Journal of Affective Disorders*, *245*, 8–15. https://doi.org/10.1016/j.jad.2018.10.106

Chen, W. (2022). Neural circuits provide insights into reward and aversion. *Frontiers in Neural Circuits*, *16*, 1002485. https://doi.org/10.3389/fncir.2022.1002485

Chepenik, L. G., Raffo, M., Hampson, M., Lacadie, C., Wang, F., Jones, M. M., Pittman, B., Skudlarski, P., & Blumberg, H. P. (2010). Functional connectivity

between ventral prefrontal cortex and amygdala at low frequency in the resting state in bipolar disorder. *Psychiatry research*, *182*(3), 207–210. https://doi.org/10.1016/j.pscychresns.2010.04.002

Cholet, J., Sauvaget, A., Vanelle, J.-M., Hommet, C., Mondon, K., Mamet, J.-P., & Camus, V. (2014). Using the Brief Assessment of Cognition in Schizophrenia (BACS) to assess cognitive impairment in older patients with schizophrenia and bipolar disorder. *Bipolar Disorders*, *16*(3), 326–336. https://doi.org/10.1111/bdi.12171

Chung, K.-H., Huang, S.-H., Wu, J.-Y., Chen, P.-H., Hsu, J.-L., & Tsai, S.-Y. (2013). The link between high-sensitivity C-reactive protein and orbitofrontal cortex in euthymic bipolar disorder. *Neuropsychobiology*, *68*(3), 168–173. https://doi.org/10.1159/000353613

Churchill, N. W., Caverzasi, E., Graham, S. J., Hutchison, M. G., & Schweizer, T. A. (2018). White matter during concussion recovery: Comparing diffusion tensor imaging (DTI) and neurite orientation dispersion and density imaging (NODDI). *Human Brain Mapping*, *40*(6), 1908–1918. https://doi.org/10.1002/hbm.24500

Cole, J., Costafreda, S. G., McGuffin, P., & Fu, C. H. Y. (2011). Hippocampal atrophy in first episode depression: A meta-analysis of magnetic resonance imaging studies. *Journal of Affective Disorders*, *134*(1–3), 483–487. https://doi.org/10.1016/j.jad.2011.05.057

Coleman, J. R. I., Gaspar, H. A., Bryois, J., Breen, G., Byrne, E. M., Forstner, A. J., Holmans, P. A., De Leeuw, C. A., Mattheisen, M., McQuillin, A., Whitehead Pavlides, J. M., Pers, T. H., Ripke, S., Stahl, E. A., Steinberg, S., Trubetskoy, V., Trzaskowski, M., Wang, Y., Abbott, L., … Wray, N. R. (2020). The

Genetics of the Mood Disorder Spectrum: Genome-wide Association Analyses of More Than 185,000 Cases and 439,000 Controls. *Biological Psychiatry*, *88*(2), 169–184. https://doi.org/10.1016/j.biopsych.2019.10.015

Colombo, F., Calesella, F., Mazza, M. G., Melloni, E. M. T., Morelli, M. J., Scotti, G. M., Benedetti, F., Bollettini, I., & Vai, B. (2022). Machine learning approaches for prediction of bipolar disorder based on biological, clinical and neuropsychological markers: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 104552.

Craig, A. D. (2009). How do you feel — now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*(1), Articolo 1. https://doi.org/10.1038/nrn2555

Critchley, H. D., & Harrison, N. A. (2013). Visceral influences on brain and behavior. *Neuron*, *77*(4), 624–638. https://doi.org/10.1016/j.neuron.2013.02.008

Cross-Disorder Group of the Psychiatric Genomics Consortium. (2013). Identification of risk loci with shared effects on five major psychiatric disorders: A genome-wide analysis. *Lancet (London, England)*, *381*(9875), 1371–1379. https://doi.org/10.1016/S0140-6736(12)62129-1

Cullen, B., Ward, J., Graham, N. A., Deary, I. J., Pell, J. P., Smith, D. J., & Evans, J. J. (2016). Prevalence and correlates of cognitive impairment in euthymic adults with bipolar disorder: A systematic review. *Journal of Affective Disorders*, *205*, 165–181. https://doi.org/10.1016/j.jad.2016.06.063

Dantzer, R., O'Connor, J. C., Freund, G. G., Johnson, R. W., & Kelley, K. W. (2008). From inflammation to sickness and depression: When the immune system

subjugates the brain. *Nature Reviews. Neuroscience*, *9*(1), 46–56. https://doi.org/10.1038/nrn2297

Dao, D. T., Mahon, P. B., Cai, X., Kovacsics, C. E., Blackwell, R. A., Arad, M., Shi, J., Zandi, P. P., O'Donnell, P., Bipolar Genome Study (BiGS) Consortium, Knowles, J. A., Weissman, M. M., Coryell, W., Scheftner, W. A., Lawson, W. B., Levinson, D. F., Thompson, S. M., Potash, J. B., & Gould, T. D. (2010). Mood disorder susceptibility gene CACNA1C modifies mood-related behaviors in mice and interacts with sex to influence behavior in mice and diagnosis in humans. *Biological Psychiatry*, *68*(9), 801–810. https://doi.org/10.1016/j.biopsych.2010.06.019

de Almeida, J. R. C., & Phillips, M. L. (2013). Distinguishing between unipolar depression and bipolar depression: Current and future clinical and neuroimaging perspectives. *Biological psychiatry*, *73*(2), 111–118.

Delvecchio, G., Sugranyes, G., & Frangou, S. (2013). Evidence of diagnostic specificity in the neural correlates of facial affect processing in bipolar disorder and schizophrenia: A meta-analysis of functional imaging studies. *Psychological Medicine*, *43*(3), 553–569. https://doi.org/10.1017/S0033291712001432

Deng, F., Wang, Y., Huang, H., Niu, M., Zhong, S., Zhao, L., Qi, Z., Wu, X., Sun, Y., Niu, C., He, Y., Huang, L., & Huang, R. (2018). Abnormal segments of right uncinate fasciculus and left anterior thalamic radiation in major and bipolar depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *81*, 340–349. https://doi.org/10.1016/j.pnpbp.2017.09.006

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., &

Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Detera-Wadleigh, S. D., Liu, C., Maheshwari, M., Cardona, I., Corona, W., Akula, N., Steele, C. J. M., Badner, J. A., Kundu, M., Kassem, L., Potash, J. B., Gibbs, R., Gershon, E. S., McMahon, F. J., & NIMH Genetics Initiative for Bipolar Disorder Consortium. (2007). Sequence variation in DOCK9 and heterogeneity in bipolar disorder. *Psychiatric Genetics*, *17*(5), 274–286. https://doi.org/10.1097/YPG.0b013e328133f352

Doganavsargil-Baysal, O., Cinemre, B., Aksoy, U. M., Akbas, H., Metin, O., Fettahoglu, C., Gokmen, Z., & Davran, F. (2013). Levels of TNF-α, soluble TNF receptors (sTNFR1, sTNFR2), and cognition in bipolar disorder. *Human Psychopharmacology: Clinical and Experimental*, *28*(2), 160–167. https://doi.org/10.1002/hup.2301

Doucet, G. E., Bassett, D. S., Yao, N., Glahn, D. C., & Frangou, S. (2017). The Role of Intrinsic Brain Functional Connectivity in Vulnerability and Resilience to Bipolar Disorder. *The American Journal of Psychiatry*, *174*(12), 1214–1222. https://doi.org/10.1176/appi.ajp.2017.17010095

Douglas, K. M., Gallagher, P., Robinson, L. J., Carter, J. D., McIntosh, V. V., Frampton, C. M., Watson, S., Young, A. H., Ferrier, I. N., & Porter, R. J. (2018). Prevalence of cognitive impairment in major depression and bipolar disorder. *Bipolar Disorders*, *20*(3), 260–274. https://doi.org/10.1111/bdi.12602

Drexhage, R. C., Hoogenboezem, T. H., Versnel, M. A., Berghout, A., Nolen, W. A., & Drexhage, H. A. (2011). The activation of monocyte and T cell networks in

patients with bipolar disorder. *Brain, Behavior, and Immunity*, *25*(6), 1206–1213. https://doi.org/10.1016/j.bbi.2011.03.013

Durak, O., de Anda, F. C., Singh, K. K., Leussis, M. P., Petryshen, T. L., Sklar, P., & Tsai, L.-H. (2015). Ankyrin-G regulates neurogenesis and Wnt signaling by altering the subcellular localization of β-catenin. *Molecular Psychiatry*, *20*(3), 388–397. https://doi.org/10.1038/mp.2014.42

Ellard, K. K., Zimmerman, J. P., Kaur, N., Van Dijk, K. R. A., Roffman, J. L., Nierenberg, A. A., Dougherty, D. D., Deckersbach, T., & Camprodon, J. A. (2018). Functional Connectivity Between Anterior Insula and Key Nodes of Frontoparietal Executive Control and Salience Networks Distinguish Bipolar Depression From Unipolar Depression and Healthy Control Subjects. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *3*(5), 473–484. https://doi.org/10.1016/j.bpsc.2018.01.013

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*(2), 194–201. https://doi.org/10.1016/j.biopsych.2012.09.023

Etain, B., Milhiet, V., Bellivier, F., & Leboyer, M. (2011). Genetics of circadian rhythms and mood spectrum disorders. *European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology*, *21 Suppl 4*, S676-682. https://doi.org/10.1016/j.euroneuro.2011.07.007

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., Fox, P. T., Eickhoff, S. B., Yu, C., & Jiang, T. (2016). The Human

Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cerebral Cortex (New York, N.Y.: 1991)*, *26*(8), 3508–3526. https://doi.org/10.1093/cercor/bhw157

Fanselow, M. S., & Dong, H.-W. (2010). Are The Dorsal and Ventral Hippocampus functionally distinct structures? *Neuron*, *65*(1), 7. https://doi.org/10.1016/j.neuron.2009.11.031

Farrell, M. S., Werge, T., Sklar, P., Owen, M. J., Ophoff, R. A., O'Donovan, M. C., Corvin, A., Cichon, S., & Sullivan, P. F. (2015). Evaluating historical candidate genes for schizophrenia. *Molecular Psychiatry*, *20*(5), 555–562. https://doi.org/10.1038/mp.2015.16

Fergusson, D. M., Horwood, L. J., Miller, A. L., & Kennedy, M. A. (2011). Life stress, 5-HTTLPR and mental disorder: Findings from a 30-year longitudinal study. *The British Journal of Psychiatry: The Journal of Mental Science*, *198*(2), 129–135. https://doi.org/10.1192/bjp.bp.110.085993

Fernandes, B. S., Steiner, J., Molendijk, M. L., Dodd, S., Nardin, P., Gonçalves, C.-A., Jacka, F., Köhler, C. A., Karmakar, C., Carvalho, A. F., & Berk, M. (2016). C-reactive protein concentrations across the mood spectrum in bipolar disorder: A systematic review and meta-analysis. *The Lancet. Psychiatry*, *3*(12), 1147–1156. https://doi.org/10.1016/S2215-0366(16)30370-4

Fernandes, B. S., Williams, L. M., Steiner, J., Leboyer, M., Carvalho, A. F., & Berk, M. (2017). The new field of 'precision psychiatry'. *BMC medicine*, *15*(1), 1–7.

Ferreira, M. A. R., O'Donovan, M. C., Meng, Y. A., Jones, I. R., Ruderfer, D. M., Jones, L., Fan, J., Kirov, G., Perlis, R. H., Green, E. K., Smoller, J. W., Grozeva, D., Stone, J., Nikolov, I., Chambert, K., Hamshere, M. L., Nimgaonkar, V. L.,

Moskvina, V., Thase, M. E., … Wellcome Trust Case Control Consortium. (2008). Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nature Genetics*, *40*(9), 1056–1058. https://doi.org/10.1038/ng.209

Flint, J. (2023). The genetic basis of major depressive disorder. *Molecular Psychiatry*, 1–12. https://doi.org/10.1038/s41380-023-01957-9

Foland-Ross, L. C., Brooks, J. O., Mintz, J., Bartzokis, G., Townsend, J., Thompson, P. M., & Altshuler, L. L. (2012). Mood-state effects on amygdala volume in bipolar disorder. *Journal of Affective Disorders*, *139*(3), 298–301. https://doi.org/10.1016/j.jad.2012.03.003

Foland-Ross, L. C., Thompson, P. M., Sugar, C. A., Madsen, S. K., Shen, J. K., Penfold, C., Ahlf, K., Rasser, P. E., Fischer, J., Yang, Y., Townsend, J., Bookheimer, S. Y., & Altshuler, L. L. (2011). Investigation of cortical thickness abnormalities in lithium-free adults with bipolar I disorder using cortical pattern matching. *The American Journal of Psychiatry*, *168*(5), 530–539. https://doi.org/10.1176/appi.ajp.2010.10060896

Fortin, J.-P., Cullen, N., Sheline, Y. I., Taylor, W. D., Aselcioglu, I., Cook, P. A., Adams, P., Cooper, C., Fava, M., & McGrath, P. J. (2018). Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage*, *167*, 104–120.

Fortin, J.-P., Parker, D., Tunç, B., Watanabe, T., Elliott, M. A., Ruparel, K., Roalf, D. R., Satterthwaite, T. D., Gur, R. C., & Gur, R. E. (2017). Harmonization of multi-site diffusion tensor imaging data. *Neuroimage*, *161*, 149–170.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, *102*(27), 9673–9678. https://doi.org/10.1073/pnas.0504136102

Friedman, J., Hastie, T., & Tibshirani, R. (2001). The elements of statistical learning. Vol. 1 Springer series in statistics. *New York*.

Frodl, T., Bokde, A. L. W., Scheuerecker, J., Lisiecka, D., Schoepf, V., Hampel, H., Möller, H.-J., Brückmann, H., Wiesmann, M., & Meisenzahl, E. (2010). Functional connectivity bias of the orbitofrontal cortex in drug-free patients with major depression. *Biological Psychiatry*, *67*(2), 161–167. https://doi.org/10.1016/j.biopsych.2009.08.022

Gandal, M. J., Haney, J. R., Parikshak, N. N., Leppa, V., Ramaswami, G., Hartl, C., Schork, A. J., Appadurai, V., Buil, A., Werge, T. M., Liu, C., White, K. P., CommonMind Consortium, PsychENCODE Consortium, iPSYCH-BROAD Working Group, Horvath, S., & Geschwind, D. H. (2018). Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science (New York, N.Y.)*, *359*(6376), 693–697. https://doi.org/10.1126/science.aad6469

Gandal, M. J., Zhang, P., Hadjimichael, E., Walker, R. L., Chen, C., Liu, S., Won, H., van Bakel, H., Varghese, M., Wang, Y., Shieh, A. W., Haney, J., Parhami, S., Belmont, J., Kim, M., Moran Losada, P., Khan, Z., Mleczko, J., Xia, Y., … Geschwind, D. H. (2018). Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science (New York, N.Y.)*, *362*(6420), eaat8127. https://doi.org/10.1126/science.aat8127

Gaser, C., & Dahnke, R. (2016). *CAT—A computational anatomy toolbox for the analysis of structural MRI data. HBM 2016*.

Gatt, J. M., Burton, K. L. O., Williams, L. M., & Schofield, P. R. (2015). Specific and common genes implicated across major mental disorders: A review of meta-analysis studies. *Journal of Psychiatric Research*, *60*, 1–13. https://doi.org/10.1016/j.jpsychires.2014.09.014

Gehrlach, D. A., Dolensek, N., Klein, A. S., Roy Chowdhury, R., Matthys, A., Junghänel, M., Gaitanos, T. N., Podgornik, A., Black, T. D., Reddy Vaka, N., Conzelmann, K.-K., & Gogolla, N. (2019). Aversive state processing in the posterior insular cortex. *Nature Neuroscience*, *22*(9), 1424–1437. https://doi.org/10.1038/s41593-019-0469-1

Gershon, E. S., Grennan, K., Busnello, J., Badner, J. A., Ovsiew, F., Memon, S., Alliey-Rodriguez, N., Cooper, J., Romanos, B., & Liu, C. (2014). A rare mutation of CACNA1C in a patient with bipolar disorder, and decreased gene expression associated with a bipolar-associated common SNP of CACNA1C in brain. *Molecular Psychiatry*, *19*(8), 890–894. https://doi.org/10.1038/mp.2013.107

Gimeno, D., Kivimäki, M., Brunner, E. J., Elovainio, M., De Vogli, R., Steptoe, A., Kumari, M., Lowe, G. D. O., Rumley, A., Marmot, M. G., & Ferrie, J. E. (2009). Associations of C-reactive protein and interleukin-6 with cognitive symptoms of depression: 12-year follow-up of the Whitehall II study. *Psychological Medicine*, *39*(3), 413–423. https://doi.org/10.1017/S0033291708003723

Goes, F. S., Hamshere, M. L., Seifuddin, F., Pirooznia, M., Belmonte-Mahon, P., Breuer, R., Schulze, T., Nöthen, M., Cichon, S., Rietschel, M., Holmans, P.,

Zandi, P. P., Bipolar Genome Study (BiGS), Craddock, N., & Potash, J. B. (2012). Genome-wide association of mood-incongruent psychotic bipolar disorder. *Translational Psychiatry*, *2*(10), e180. https://doi.org/10.1038/tp.2012.106

Goodwin, F. K., & Jamison, K. R. (2007). *Manic-depressive illness: Bipolar disorders and recurrent depression: Vol. Vol. 1*. New York: Oxford University Press. https://hsrc.himmelfarb.gwu.edu/books/214

Goodwin, G. M. (2012). Bipolar depression and treatment with antidepressants. *The British Journal of Psychiatry*, *200*(1), 5–6.

Gordon, S., Plüddemann, A., & Martinez Estrada, F. (2014). Macrophage heterogeneity in tissues: Phenotypic diversity and functions. *Immunological Reviews*, *262*(1), 36–55. https://doi.org/10.1111/imr.12223

Gordovez, F. J. A., & McMahon, F. J. (2020). The genetics of bipolar disorder. *Molecular Psychiatry*, *25*(3), 544–559. https://doi.org/10.1038/s41380-019-0634-7

Grande, I., Berk, M., Birmaher, B., & Vieta, E. (2016). Bipolar disorder. *The Lancet*, *387*(10027), 1561–1572.

Grosse, L., Hoogenboezem, T., Ambrée, O., Bellingrath, S., Jörgens, S., de Wit, H. J., Wijkhuijs, A. M., Arolt, V., & Drexhage, H. A. (2016). Deficiencies of the T and natural killer cell system in major depressive disorder: T regulatory cell defects are associated with inflammatory monocyte activation. *Brain, Behavior, and Immunity*, *54*, 38–44. https://doi.org/10.1016/j.bbi.2015.12.003

Grover, D., Verma, R., Goes, F. S., Mahon, P. L. B., Gershon, E. S., McMahon, F. J., Potash, J. B., NIMH Genetics Initiative Bipolar Disorder Collaborative, Bipolar

Disorder Phenome Group, Gershon, E. S., McMahon, F. J., & Potash, J. B. (2009). Family-based association of YWHAH in psychotic bipolar disorder. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics*, *150B*(7), 977–983. https://doi.org/10.1002/ajmg.b.30927

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, *3*(Mar), 1157–1182.

Haapakoski, R., Mathieu, J., Ebmeier, K. P., Alenius, H., & Kivimäki, M. (2015). Cumulative meta-analysis of interleukins 6 and 1β, tumour necrosis factor α and C-reactive protein in patients with major depressive disorder. *Brain, Behavior, and Immunity*, *49*, 206–215. https://doi.org/10.1016/j.bbi.2015.06.001

Haas, B. W., Barnea-Goraly, N., Sheau, K. E., Yamagata, B., Ullas, S., & Reiss, A. L. (2014). Altered microstructure within social-cognitive brain networks during childhood in Williams syndrome. *Cerebral Cortex (New York, N.Y.: 1991)*, *24*(10), 2796–2806. https://doi.org/10.1093/cercor/bht135

Halvorsen, M., Høifødt, R. S., Myrbakk, I. N., Wang, C. E. A., Sundet, K., Eisemann, M., & Waterloo, K. (2012). Cognitive function in unipolar major depression: A comparison of currently depressed, previously depressed, and never depressed individuals. *Journal of Clinical and Experimental Neuropsychology*, *34*(7), 782–790. https://doi.org/10.1080/13803395.2012.683853

Hamilton, J. P., Furman, D. J., Chang, C., Thomason, M. E., Dennis, E., & Gotlib, I. H. (2011). Default-mode and task-positive network activity in major depressive disorder: Implications for adaptive and maladaptive rumination. *Biological Psychiatry*, *70*(4), 327–333. https://doi.org/10.1016/j.biopsych.2011.02.003

Hamilton, J. P., Siemer, M., & Gotlib, I. H. (2008). Amygdala volume in major depressive disorder: A meta-analysis of magnetic resonance imaging studies. *Molecular Psychiatry*, *13*(11), 993–1000. https://doi.org/10.1038/mp.2008.57

Hamilton, M. (1960). A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry*, *23*(1), 56.

Han, K.-M., De Berardis, D., Fornaro, M., & Kim, Y.-K. (2019). Differentiating between bipolar and unipolar depression in functional and structural MRI studies. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *91*, 20–27.

Hanford, L. C., Nazarov, A., Hall, G. B., & Sassi, R. B. (2016). Cortical thickness in bipolar disorder: A systematic review. *Bipolar Disorders*, *18*(1), 4–18. https://doi.org/10.1111/bdi.12362

Hannon, E., Lunnon, K., Schalkwyk, L., & Mill, J. (2015). Interindividual methylomic variation across blood, cortex, and cerebellum: Implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*, *10*(11), 1024–1032. https://doi.org/10.1080/15592294.2015.1100786

Hara, T., Owada, Y., & Takata, A. (2023). Genetics of bipolar disorder: Insights into its complex architecture and biology from common and rare variants. *Journal of Human Genetics*, *68*(3), 183–191. https://doi.org/10.1038/s10038-022-01046-9

Hayashi, A., Le Gal, K., Södersten, K., Vizlin-Hodzic, D., Ågren, H., & Funa, K. (2015). Calcium-dependent intracellular signal pathways in primary cultured adipocytes and ANK3 gene variation in patients with bipolar disorder and healthy controls. *Molecular Psychiatry*, *20*(8), 931–940. https://doi.org/10.1038/mp.2014.104

He, H., Sui, J., Du, Y., Yu, Q., Lin, D., Drevets, W. C., Savitz, J. B., Yang, J., Victor, T. A., & Calhoun, V. D. (2017). Co-altered functional networks and brain structure in unmedicated patients with bipolar and major depressive disorders. *Brain Structure and Function*, *222*, 4051–4064.

Hemmati, S., Sadeghi, M. A., Mohammad Jafari, R., Yousefi-Manesh, H., & Dehpour, A. R. (2019). The antidepressant effects of GM-CSF are mediated by the reduction of TLR4/NF-κB-induced IDO expression. *Journal of Neuroinflammation*, *16*(1), 117. https://doi.org/10.1186/s12974-019-1509-1

Hirschfeld, R. M. (2014). Differential diagnosis of bipolar disorder and major depressive disorder. *Journal of Affective Disorders*, *169 Suppl 1*, S12-16. https://doi.org/10.1016/S0165-0327(14)70004-7

Hirschfeld, R. M., Lewis, L., & Vornik, L. A. (2003). Perceptions and impact of bipolar disorder: How far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder. *Journal of Clinical Psychiatry*, *64*(2), 161–174.

Hope, S., Hoseth, E., Dieset, I., Mørch, R. H., Aas, M., Aukrust, P., Djurovic, S., Melle, I., Ueland, T., Agartz, I., Ueland, T., Westlye, L. T., & Andreassen, O. A. (2015). Inflammatory markers are associated with general cognitive abilities in schizophrenia and bipolar disorder patients and healthy controls. *Schizophrenia Research*, *165*(2), 188–194. https://doi.org/10.1016/j.schres.2015.04.004

Hou, L., Lam, B. Y.-H., Wong, N. M. L., Lu, W., Zhang, R., Ning, Y., & Lin, K. (2022). Integrity of cerebellar tracts associated with the risk of bipolar disorder. *Translational Psychiatry*, *12*, 335. https://doi.org/10.1038/s41398-022-02097-4

Howard, D. M., Adams, M. J., Clarke, T.-K., Hafferty, J. D., Gibson, J., Shirali, M., Coleman, J. R. I., Hagenaars, S. P., Ward, J., Wigmore, E. M., Alloza, C., Shen, X., Barbu, M. C., Xu, E. Y., Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I. J., … McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, *22*(3), Articolo 3. https://doi.org/10.1038/s41593-018-0326-7

Hu, X.-Z., Rush, A. J., Charney, D., Wilson, A. F., Sorant, A. J. M., Papanicolaou, G. J., Fava, M., Trivedi, M. H., Wisniewski, S. R., Laje, G., Paddock, S., McMahon, F. J., Manji, H., & Lipsky, R. H. (2007). Association between a functional serotonin transporter promoter polymorphism and citalopram treatment in adult outpatients with major depression. *Archives of General Psychiatry*, *64*(7), 783–792. https://doi.org/10.1001/archpsyc.64.7.783

Hua, J., Tembe, W. D., & Dougherty, E. R. (2009). Performance of feature-selection methods in the classification of high-dimension data. *Pattern Recognition*, *42*(3), 409–424.

Huang, Y.-C., Lee, Y., Lee, C.-Y., Lin, P.-Y., Hung, C.-F., Lee, S.-Y., & Wang, L.-J. (2020). Defining cognitive and functional profiles in schizophrenia and affective disorders. *BMC Psychiatry*, *20*(1), 39. https://doi.org/10.1186/s12888-020-2459-y

Ikeda, M., Takahashi, A., Kamatani, Y., Okahisa, Y., Kunugi, H., Mori, N., Sasaki, T., Ohmori, T., Okamoto, Y., Kawasaki, H., Shimodera, S., Kato, T., Yoneda, H., Yoshimura, R., Iyo, M., Matsuda, K., Akiyama, M., Ashikawa, K., Kashiwase, K., … Iwata, N. (2018). A genome-wide association study identifies two novel

susceptibility loci and trans population polygenicity associated with bipolar disorder. *Molecular Psychiatry*, *23*(3), 639–647. https://doi.org/10.1038/mp.2016.259

Insel, T. R., & Cuthbert, B. N. (2015). Brain disorders? Precisely. *Science*, *348*(6234), 499–500. https://doi.org/10.1126/science.aab2358

Jacoby, A. S., Munkholm, K., Vinberg, M., Pedersen, B. K., & Kessing, L. V. (2016). Cytokines, brain-derived neurotrophic factor and C-reactive protein in bipolar I disorder—Results from a prospective study. *Journal of Affective Disorders*, *197*, 167–174. https://doi.org/10.1016/j.jad.2016.03.040

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, *17*(2), 825–841.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical image analysis*, *5*(2), 143–156.

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*(11), 2747–2766. https://doi.org/10.1002/hbm.22099

Jiang, W., Gong, G., Wu, F., Kong, L., Chen, K., Cui, W., Ren, L., Fan, G., Sun, W., Ma, H., Xu, K., Tang, Y., & Wang, F. (2015). The Papez Circuit in First-Episode, Treatment-Naive Adults with Major Depressive Disorder: Combined Atlas-Based Tract-Specific Quantification Analysis and Voxel-Based Analysis. *PLoS ONE*, *10*(5), e0126673. https://doi.org/10.1371/journal.pone.0126673

Jiang, X., Detera-Wadleigh, S. D., Akula, N., Mallon, B. S., Hou, L., Xiao, T., Felsenfeld, G., Gu, X., & McMahon, F. J. (2019). Sodium valproate rescues expression of TRANK1 in iPSC-derived neural cells that carry a genetic variant associated with serious mental illness. *Molecular Psychiatry*, *24*(4), 613–624. https://doi.org/10.1038/s41380-018-0207-1

Jiao, Q., Ding, J., Lu, G., Su, L., Zhang, Z., Wang, Z., Zhong, Y., Li, K., Ding, M., & Liu, Y. (2011). Increased Activity Imbalance in Fronto-Subcortical Circuits in Adolescents with Major Depression. *PLOS ONE*, *6*(9), e25159. https://doi.org/10.1371/journal.pone.0025159

Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, *8*(1), 118–127.

Judd, L. L., Akiskal, H. S., Schettler, P. J., Coryell, W., Endicott, J., Maser, J. D., Solomon, D. A., Leon, A. C., & Keller, M. B. (2003). A prospective investigation of the natural history of the long-term weekly symptomatic status of bipolar II disorder. *Archives of General Psychiatry*, *60*(3), 261–269. https://doi.org/10.1001/archpsyc.60.3.261

Judd, L. L., Akiskal, H. S., Schettler, P. J., Endicott, J., Maser, J., Solomon, D. A., Leon, A. C., Rice, J. A., & Keller, M. B. (2002). The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Archives of General Psychiatry*, *59*(6), 530–537. https://doi.org/10.1001/archpsyc.59.6.530

Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D., & Pizzagalli, D. A. (2015). Large-Scale Network Dysfunction in Major Depressive Disorder: A Meta-analysis of Resting-State Functional Connectivity. *JAMA Psychiatry*, *72*(6), 603–611. https://doi.org/10.1001/jamapsychiatry.2015.0071

Keefe, R. S. E., Fox, K. H., Davis, V. G., Kennel, C., Walker, T. M., Burdick, K. E., & Harvey, P. D. (2014). The Brief Assessment of Cognition In Affective Disorders (BAC-A):performance of patients with bipolar depression and healthy controls. *Journal of Affective Disorders*, *166*, 86–92. https://doi.org/10.1016/j.jad.2014.05.002

Keefe, R. S. E., Goldberg, T. E., Harvey, P. D., Gold, J. M., Poe, M. P., & Coughenour, L. (2004). The Brief Assessment of Cognition in Schizophrenia: Reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophrenia Research*, *68*(2–3), 283–297. https://doi.org/10.1016/j.schres.2003.09.011

Kempton, M. J., Salvador, Z., Munafò, M. R., Geddes, J. R., Simmons, A., Frangou, S., & Williams, S. C. R. (2011). Structural neuroimaging studies in major depressive disorder. Meta-analysis and comparison with bipolar disorder. *Archives of General Psychiatry*, *68*(7), 675–690. https://doi.org/10.1001/archgenpsychiatry.2011.60

Keser, Z., Hasan, K. M., Mwangi, B. I., Kamali, A., Ucisik-Keser, F. E., Riascos, R. F., Yozbatiran, N., Francisco, G. E., & Narayana, P. A. (2015). Diffusion tensor imaging of the human cerebellar pathways and their interplay with cerebral macrostructure. *Frontiers in Neuroanatomy*, *9*, 41. https://doi.org/10.3389/fnana.2015.00041

Kessing, L. V., & Andersen, P. K. (2017). Evidence for clinical progression of unipolar and bipolar disorders. *Acta Psychiatrica Scandinavica*, *135*(1), 51–64. https://doi.org/10.1111/acps.12667

Khandaker, G. M., Pearson, R. M., Zammit, S., Lewis, G., & Jones, P. B. (2014). Association of serum interleukin 6 and C-reactive protein in childhood with depression and psychosis in young adult life: A population-based longitudinal study. *JAMA Psychiatry*, *71*(10), 1121–1128. https://doi.org/10.1001/jamapsychiatry.2014.1332

Kheirbek, M. A., & Hen, R. (2011). Dorsal vs Ventral Hippocampal Neurogenesis: Implications for Cognition and Mood. *Neuropsychopharmacology*, *36*(1), Articolo 1. https://doi.org/10.1038/npp.2010.148

Kim, Y.-K., Jung, H.-G., Myint, A.-M., Kim, H., & Park, S.-H. (2007). Imbalance between pro-inflammatory and anti-inflammatory cytokines in bipolar disorder. *Journal of Affective Disorders*, *104*(1–3), 91–95. https://doi.org/10.1016/j.jad.2007.02.018

Kim, Y.-K., Suh, I.-B., Kim, H., Han, C.-S., Lim, C.-S., Choi, S.-H., & Licinio, J. (2002). The plasma levels of interleukin-12 in schizophrenia, major depression, and bipolar mania: Effects of psychotropic drugs. *Molecular Psychiatry*, *7*(10), 1107–1114. https://doi.org/10.1038/sj.mp.4001084

King, G., & Zeng, L. (2001). Logistic Regression in Rare Events Data. *Political Analysis*, *9*(2), 137–163. https://doi.org/10.1093/oxfordjournals.pan.a004868

Kinoshita, Y., Ohnishi, A., Kohshi, K., & Yokota, A. (1999). Apparent diffusion coefficient on rat brain and nerves intoxicated with methylmercury. *Environmental Research*, *80*(4), 348–354. https://doi.org/10.1006/enrs.1998.3935

Köhler-Forsberg, O., N Lydholm, C., Hjorthøj, C., Nordentoft, M., Mors, O., & Benros, M. E. (2019). Efficacy of anti-inflammatory treatment on major depressive

disorder or depressive symptoms: Meta-analysis of clinical trials. *Acta Psychiatrica Scandinavica*, *139*(5), 404–419. https://doi.org/10.1111/acps.13016

Kraft, J. B., Peters, E. J., Slager, S. L., Jenkins, G. D., Reinalda, M. S., McGrath, P. J., & Hamilton, S. P. (2007). Analysis of association between the serotonin transporter and antidepressant response in a large clinical sample. *Biological Psychiatry*, *61*(6), 734–742. https://doi.org/10.1016/j.biopsych.2006.07.017

Kronenberg, G., Tebartz van Elst, L., Regen, F., Deuschle, M., Heuser, I., & Colla, M. (2009). Reduced amygdala volume in newly admitted psychiatric in-patients with unipolar major depression. *Journal of Psychiatric Research*, *43*(13), 1112–1117. https://doi.org/10.1016/j.jpsychires.2009.03.007

LeDoux, J. (2007). The amygdala. *Current Biology*, *17*(20), R868–R874. https://doi.org/10.1016/j.cub.2007.08.005

Lee, R. S. C., Hermens, D. F., Porter, M. A., & Redoblado-Hodge, M. A. (2012). A meta-analysis of cognitive deficits in first-episode Major Depressive Disorder. *Journal of Affective Disorders*, *140*(2), 113–124. https://doi.org/10.1016/j.jad.2011.10.023

Lee, S., Parthasarathi, T., & Kable, J. W. (2021). The Ventral and Dorsal Default Mode Networks Are Dissociably Modulated by the Vividness and Valence of Imagined Events. *The Journal of Neuroscience*, *41*(24), 5243–5250. https://doi.org/10.1523/JNEUROSCI.1273-20.2021

Leek, J. T., Scharpf, R. B., Bravo, H. C., Simcha, D., Langmead, B., Johnson, W. E., Geman, D., Baggerly, K., & Irizarry, R. A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Reviews Genetics*, *11*(10), 733–739.

Le-Niculescu, H., Roseberry, K., Gill, S. S., Levey, D. F., Phalen, P. L., Mullen, J., Williams, A., Bhairo, S., Voegtline, T., Davis, H., Shekhar, A., Kurian, S. M., & Niculescu, A. B. (2021). Precision medicine for mood disorders: Objective assessment, risk prediction, pharmacogenomics, and repurposed drugs. *Molecular Psychiatry*, *26*(7), 2776–2804. https://doi.org/10.1038/s41380-021-01061-w

Leow, A., Ajilore, O., Zhan, L., Arienzo, D., GadElkarim, J., Zhang, A., Moody, T., Van Horn, J., Feusner, J., Kumar, A., Thompson, P., & Altshuler, L. (2013). Impaired inter-hemispheric integration in bipolar disorder revealed using brain network analyses. *Biological psychiatry*, *73*(2), 183–193. https://doi.org/10.1016/j.biopsych.2012.09.014

Levey, D. F., Stein, M. B., Wendt, F. R., Pathak, G. A., Zhou, H., Aslan, M., Quaden, R., Harrington, K. M., Nuñez, Y. Z., Overstreet, C., Radhakrishnan, K., Sanacora, G., McIntosh, A. M., Shi, J., Shringarpure, S. S., 23andMe Research Team, Million Veteran Program, Concato, J., Polimanti, R., & Gelernter, J. (2021). Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nature Neuroscience*, *24*(7), 954–963. https://doi.org/10.1038/s41593-021-00860-2

Levita, L., Hare, T. A., Voss, H. U., Glover, G., Ballon, D. J., & Casey, B. J. (2009). The bivalent side of the nucleus accumbens. *NeuroImage*, *44*(3), 1178–1187. https://doi.org/10.1016/j.neuroimage.2008.09.039

Lewis, K. J. S., Richards, A., Karlsson, R., Leonenko, G., Jones, S. E., Jones, H. J., Gordon-Smith, K., Forty, L., Escott-Price, V., Owen, M. J., Weedon, M. N.,

Jones, L., Craddock, N., Jones, I., Landén, M., O'Donovan, M. C., & Di Florio, A. (2020). Comparison of Genetic Liability for Sleep Traits Among Individuals With Bipolar Disorder I or II and Control Participants. *JAMA Psychiatry*, *77*(3), 303–310. https://doi.org/10.1001/jamapsychiatry.2019.4079

Lewis, R. G., Florio, E., Punzo, D., & Borrelli, E. (2021). The Brain's Reward System in Health and Disease. *Advances in experimental medicine and biology*, *1344*, 57–69. https://doi.org/10.1007/978-3-030-81147-1_4

Li, C., Sugam, J. A., Lowery-Gionta, E. G., McElligott, Z. A., McCall, N. M., Lopez, A. J., McKlveen, J. M., Pleil, K. E., & Kash, T. L. (2016). Mu Opioid Receptor Modulation of Dopamine Neurons in the Periaqueductal Gray/Dorsal Raphe: A Role in Regulation of Pain. *Neuropsychopharmacology*, *41*(8), Articolo 8. https://doi.org/10.1038/npp.2016.12

Li, M., Das, T., Deng, W., Wang, Q., Li, Y., Zhao, L., Ma, X., Wang, Y., Yu, H., Li, X., Meng, Y., Palaniyappan, L., & Li, T. (2017). Clinical utility of a short resting-state MRI scan in differentiating bipolar from unipolar depression. *Acta Psychiatrica Scandinavica*, *136*(3), 288–299. https://doi.org/10.1111/acps.12752

Li, M., Huang, C., Deng, W., Ma, X., Han, Y., Wang, Q., Li, Z., Guo, W., Li, Y., Jiang, L., Lei, W., Hu, X., Gong, Q., Merikangas, K. R., Palaniyappan, L., & Li, T. (2015). Contrasting and convergent patterns of amygdala connectivity in mania and depression: A resting-state study. *Journal of Affective Disorders*, *173*, 53–58. https://doi.org/10.1016/j.jad.2014.10.044

Li, M., Soczynska, J. K., & Kennedy, S. H. (2011). Inflammatory biomarkers in depression: An opportunity for novel therapeutic interventions. *Current Psychiatry Reports*, *13*(5), 316–320. https://doi.org/10.1007/s11920-011-0210-6

Liang, S., Wang, Q., Kong, X., Deng, W., Yang, X., Li, X., Zhang, Z., Zhang, J., Zhang, C., Li, X.-M., Ma, X., Shao, J., Greenshaw, A. J., & Li, T. (2019). White Matter Abnormalities in Major Depression Biotypes Identified by Diffusion Tensor Imaging. *Neuroscience Bulletin*, *35*(5), 867–876. https://doi.org/10.1007/s12264-019-00381-w

Liang, S., Wu, Y., Hanxiaoran, L., Greenshaw, A. J., & Li, T. (2022). Anhedonia in Depression and Schizophrenia: Brain Reward and Aversion Circuits. *Neuropsychiatric Disease and Treatment*, *18*, 1385–1396. https://doi.org/10.2147/NDT.S367839

Lim, C. H., Zain, S. M., Reynolds, G. P., Zain, M. A., Roffeei, S. N., Zainal, N. Z., Kanagasundram, S., & Mohamed, Z. (2014). Genetic association of LMAN2L gene in schizophrenia and bipolar disorder and its interaction with ANK3 gene polymorphism. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *54*, 157–162. https://doi.org/10.1016/j.pnpbp.2014.05.017

Lippard, E. T. C., Jensen, K. P., Wang, F., Johnston, J. a. Y., Spencer, L., Pittman, B., Gelernter, J., & Blumberg, H. P. (2017). Effects of ANK3 variation on gray and white matter in bipolar disorder. *Molecular Psychiatry*, *22*(9), 1345–1351. https://doi.org/10.1038/mp.2016.76

Lisy, M. E., Jarvis, K. B., DelBello, M. P., Mills, N. P., Weber, W. A., Fleck, D., Strakowski, S. M., & Adler, C. M. (2011). Progressive neurostructural changes in adolescent and adult patients with bipolar disorder. *Bipolar Disorders*, *13*(4), 396–405. https://doi.org/10.1111/j.1399-5618.2011.00927.x

Liu, C.-H., Ma, X., Li, F., Wang, Y.-J., Tie, C.-L., Li, S.-F., Chen, T.-L., Fan, T., Zhang, Y., Dong, J., Yao, L., Wu, X., & Wang, C.-Y. (2012). Regional

homogeneity within the default mode network in bipolar depression: A resting-state functional magnetic resonance imaging study. *PloS One*, *7*(11), e48181. https://doi.org/10.1371/journal.pone.0048181

Liu, Y., Chen, K., Luo, Y., Wu, J., Xiang, Q., Peng, L., Zhang, J., Zhao, W., Li, M., & Zhou, X. (2022). Distinguish bipolar and major depressive disorder in adolescents based on multimodal neuroimaging: Results from the Adolescent Brain Cognitive Development study®. *Digital Health*, *8*, 20552076221123705.

Lois, G., Linke, J., & Wessa, M. (2014). Altered functional connectivity between emotional and cognitive resting state networks in euthymic bipolar I disorder patients. *PloS One*, *9*(10), e107829. https://doi.org/10.1371/journal.pone.0107829

Maes, M., Bosmans, E., De Jongh, R., Kenis, G., Vandoolaeghe, E., & Neels, H. (1997). Increased serum IL-6 and IL-1 receptor antagonist concentrations in major depression and treatment resistant depression. *Cytokine*, *9*(11), 853–858. https://doi.org/10.1006/cyto.1997.0238

Maes, M., Rachayon, M., Jirakran, K., Sodsai, P., Klinchanhom, S., Gałecki, P., Sughondhabirom, A., & Basta-Kaim, A. (2022). The Immune Profile of Major Dysmood Disorder: Proof of Concept and Mechanism Using the Precision Nomothetic Psychiatry Approach. *Cells*, *11*(7), 1183. https://doi.org/10.3390/cells11071183

Maheshwari, M., Shi, J., Badner, J. A., Skol, A., Willour, V. L., Muzny, D. M., Wheeler, D. A., Gerald, F. R., Detera-Wadleigh, S., McMahon, F. J., Potash, J. B., Gershon, E. S., Liu, C., & Gibbs, R. A. (2009). Common and Rare Variants of DAOA in Bipolar Disorder. *American journal of medical genetics. Part B,*

Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics, *150B*(7), 960–966. https://doi.org/10.1002/ajmg.b.30925

Mahon, P. B., Pirooznia, M., Goes, F. S., Seifuddin, F., Steele, J., Lee, P. H., Huang, J., Hamshere, M., DePaulo, J. R., Kelsoe, J. R., Rietschel, M., Nöthen, M., Cichon, S., Gurling, H., Purcell, S., Smoller, J. W., Craddock, N., Schulze, ThomasG., McMahon, F. J., … Zandi, P. P. (2011). Genome-wide association analysis of age at onset and psychotic symptoms in bipolar disorder. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics*, *156B*(3), 370–378. https://doi.org/10.1002/ajmg.b.31172

Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke, S., Wray, N. R., Lewis, C. M., Hamilton, S. P., Weissman, M. M., Breen, G., Byrne, E. M., Blackwood, D. H. R., Boomsma, D. I., Cichon, S., Heath, A. C., Holsboer, F., Lucae, S., Madden, P. A. F., Martin, N. G., McGuffin, P., Muglia, P., Noethen, M. M., … Sullivan, P. F. (2013). A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry*, *18*(4), 497–511. https://doi.org/10.1038/mp.2012.21

Matthews, S. C., Strigo, I. A., Simmons, A. N., Yang, T. T., & Paulus, M. P. (2008). Decreased functional coupling of the amygdala and supragenual cingulate is related to increased depression in unmedicated individuals with current major depressive disorder. *Journal of Affective Disorders*, *111*(1), 13–20. https://doi.org/10.1016/j.jad.2008.05.022

McAfoose, J., & Baune, B. T. (2009). Evidence for a cytokine model of cognitive function. *Neuroscience and Biobehavioral Reviews*, *33*(3), 355–366. https://doi.org/10.1016/j.neubiorev.2008.10.005

McCarthy, M. J., Nievergelt, C. M., Kelsoe, J. R., & Welsh, D. K. (2012). A Survey of Genomic Studies Supports Association of Circadian Clock Genes with Bipolar Disorder Spectrum Illnesses and Lithium Response. *PLoS ONE*, *7*(2), e32091. https://doi.org/10.1371/journal.pone.0032091

McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual Review of Neuroscience*, *27*, 1–28. https://doi.org/10.1146/annurev.neuro.27.070203.144157

McIntosh, A. M., Sullivan, P. F., & Lewis, C. M. (2019). Uncovering the Genetic Architecture of Major Depression. *Neuron*, *102*(1), 91–103. https://doi.org/10.1016/j.neuron.2019.03.022

McIntyre, R. S., Rasgon, N. L., Kemp, D. E., Nguyen, H. T., Law, C. W. Y., Taylor, V. H., Woldeyohannes, H. O., Alsuwaidan, M. T., Soczynska, J. K., Kim, B., Lourenco, M. T., Kahn, L. S., & Goldstein, B. I. (2009). Metabolic syndrome and major depressive disorder: Co-occurrence and pathophysiologic overlap. *Current Diabetes Reports*, *9*(1), 51–59. https://doi.org/10.1007/s11892-009-0010-0

McIntyre, R. S., Soczynska, J. Z., Woldeyohannes, H. O., Alsuwaidan, M. T., Cha, D. S., Carvalho, A. F., Jerrell, J. M., Dale, R. M., Gallaugher, L. A., Muzina, D. J., & Kennedy, S. H. (2015). The impact of cognitive impairment on perceived workforce performance: Results from the International Mood Disorders

Collaborative Project. *Comprehensive Psychiatry*, *56*, 279–282. https://doi.org/10.1016/j.comppsych.2014.08.051

McIntyre, R. S., Xiao, H. X., Syeda, K., Vinberg, M., Carvalho, A. F., Mansur, R. B., Maruschak, N., & Cha, D. S. (2015). The prevalence, measurement, and treatment of the cognitive dimension/domain in major depressive disorder. *CNS Drugs*, *29*(7), 577–589. https://doi.org/10.1007/s40263-015-0263-x

Meier, S., Mattheisen, M., Vassos, E., Strohmaier, J., Treutlein, J., Josef, F., Breuer, R., Degenhardt, F., Mühleisen, T. W., Müller-Myhsok, B., Steffens, M., Schmael, C., McMahon, F. J., Bipolar Disorder Genome Study (BiGS) Consortium, Nöthen, M. M., Cichon, S., Schulze, T. G., Rietschel, M., Kelsoe, J. R., … Szelinger, S. (2012). Genome-wide significant association between a «negative mood delusions» dimension in bipolar disorder and genetic variation on chromosome 3q26.1. *Translational Psychiatry*, *2*(9), e165. https://doi.org/10.1038/tp.2012.81

Melloni, E. M. T., Poletti, S., Vai, B., Bollettini, I., Colombo, C., & Benedetti, F. (2019). Effects of illness duration on cognitive performances in bipolar depression are mediated by white matter microstructure. *Journal of Affective Disorders*, *249*, 175–182. https://doi.org/10.1016/j.jad.2019.02.015

Menon, V. (2011). Large-scale brain networks and psychopathology: A unifying triple network model. *Trends in Cognitive Sciences*, *15*(10), 483–506. https://doi.org/10.1016/j.tics.2011.08.003

Merali, Z., Brennan, K., Brau, P., & Anisman, H. (2003). Dissociating anorexia and anhedonia elicited by interleukin-1beta: Antidepressant and gender effects on

responding for «free chow» and «earned» sucrose intake. *Psychopharmacology*, *165*(4), 413–418. https://doi.org/10.1007/s00213-002-1273-1

Merikangas, K. R., Jin, R., He, J.-P., Kessler, R. C., Lee, S., Sampson, N. A., Viana, M. C., Andrade, L. H., Hu, C., Karam, E. G., Ladea, M., Medina-Mora, M. E., Ono, Y., Posada-Villa, J., Sagar, R., Wells, J. E., & Zarkov, Z. (2011). Prevalence and Correlates of Bipolar Spectrum Disorder in the World Mental Health Survey Initiative. *Archives of General Psychiatry*, *68*(3), 241–251. https://doi.org/10.1001/archgenpsychiatry.2011.12

Mikova, O., Yakimova, R., Bosmans, E., Kenis, G., & Maes, M. (2001). Increased serum tumor necrosis factor alpha concentrations in major depression and multiple sclerosis. *European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology*, *11*(3), 203–208. https://doi.org/10.1016/s0924-977x(01)00081-5

Millan, M. J., Agid, Y., Brüne, M., Bullmore, E. T., Carter, C. S., Clayton, N. S., Connor, R., Davis, S., Deakin, B., DeRubeis, R. J., Dubois, B., Geyer, M. A., Goodwin, G. M., Gorwood, P., Jay, T. M., Joëls, M., Mansuy, I. M., Meyer-Lindenberg, A., Murphy, D., … Young, L. J. (2012). Cognitive dysfunction in psychiatric disorders: Characteristics, causes and the quest for improved therapy. *Nature Reviews Drug Discovery*, *11*(2), Articolo 2. https://doi.org/10.1038/nrd3628

Miller, A. H., Maletic, V., & Raison, C. L. (2009). Inflammation and its discontents: The role of cytokines in the pathophysiology of major depression. *Biological Psychiatry*, *65*(9), 732–741. https://doi.org/10.1016/j.biopsych.2008.11.029

Miller, A. H., & Raison, C. L. (2016). The role of inflammation in depression: From evolutionary imperative to modern treatment target. *Nature Reviews. Immunology*, *16*(1), 22–34. https://doi.org/10.1038/nri.2015.5

Miller, A. M. P., Vedder, L. C., Law, L. M., & Smith, D. M. (2014). Cues, context, and long-term memory: The role of the retrosplenial cortex in spatial cognition. *Frontiers in Human Neuroscience*, *8*, 586. https://doi.org/10.3389/fnhum.2014.00586

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., Faria, A. V., Mahmood, A., & Woods, R. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage*, *40*(2), 570–582.

Moskvina, V., Craddock, N., Holmans, P., Nikolov, I., Pahwa, J. S., Green, E., Wellcome Trust Case Control Consortium, Owen, M. J., & O'Donovan, M. C. (2009). Gene-wide analyses of genome-wide association data sets: Evidence for multiple common risk alleles for schizophrenia and bipolar disorder and for overlap in genetic risk. *Molecular Psychiatry*, *14*(3), 252–260. https://doi.org/10.1038/mp.2008.133

Mota, R., Gazal, M., Acosta, B. A., de Leon, P. B., Jansen, K., Pinheiro, R. T., Souza, L. D., Silva, R. A., Oses, J. P., Quevedo, L., Lara, D. R., Ghisleni, G., & Kaster, M. P. (2013). Interleukin-1β is associated with depressive episode in major depression but not in bipolar disorder. *Journal of Psychiatric Research*, *47*(12), 2011–2014. https://doi.org/10.1016/j.jpsychires.2013.08.020

Motivala, S. J., Sarfatti, A., Olmos, L., & Irwin, M. R. (2005). Inflammatory markers and sleep disturbance in major depression. *Psychosomatic Medicine*, *67*(2), 187–194. https://doi.org/10.1097/01.psy.0000149259.72488.09

Mühleisen, T. W., Leber, M., Schulze, T. G., Strohmaier, J., Degenhardt, F., Treutlein, J., Mattheisen, M., Forstner, A. J., Schumacher, J., Breuer, R., Meier, S., Herms, S., Hoffmann, P., Lacour, A., Witt, S. H., Reif, A., Müller-Myhsok, B., Lucae, S., Maier, W., … Cichon, S. (2014). Genome-wide association study reveals two new risk loci for bipolar disorder. *Nature Communications*, *5*, 3339. https://doi.org/10.1038/ncomms4339

Müller, N., Schwarz, M. J., Dehning, S., Douhe, A., Cerovecki, A., Goldstein-Müller, B., Spellmann, I., Hetzel, G., Maino, K., Kleindienst, N., Möller, H.-J., Arolt, V., & Riedel, M. (2006). The cyclooxygenase-2 inhibitor celecoxib has therapeutic effects in major depression: Results of a double-blind, randomized, placebo controlled, add-on pilot study to reboxetine. *Molecular Psychiatry*, *11*(7), 680–684. https://doi.org/10.1038/sj.mp.4001805

Mullins, N., Forstner, A. J., O'Connell, K. S., Coombes, B., Coleman, J. R. I., Qiao, Z., Als, T. D., Bigdeli, T. B., Børte, S., Bryois, J., Charney, A. W., Drange, O. K., Gandal, M. J., Hagenaars, S. P., Ikeda, M., Kamitaki, N., Kim, M., Krebs, K., Panagiotaropoulou, G., … Andreassen, O. A. (2021). Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nature Genetics*, *53*(6), 817–829. https://doi.org/10.1038/s41588-021-00857-4

Munkholm, K., Vinberg, M., & Vedel Kessing, L. (2013). Cytokines in bipolar disorder: A systematic review and meta-analysis. *Journal of Affective Disorders*, *144*(1–2), 16–27. https://doi.org/10.1016/j.jad.2012.06.010

Muschelli, J., Nebel, M. B., Caffo, B. S., Barber, A. D., Pekar, J. J., & Mostofsky, S. H. (2014). Reduction of motion-related artifacts in resting state fMRI using

aCompCor. *NeuroImage*, *96*, 22–35. https://doi.org/10.1016/j.neuroimage.2014.03.028

Mwangi, B., Tian, T. S., & Soares, J. C. (2014). A review of feature reduction techniques in neuroimaging. *Neuroinformatics*, *12*, 229–244.

Najjar, S., Pearlman, D. M., Alper, K., Najjar, A., & Devinsky, O. (2013). Neuroinflammation and psychiatric illness. *Journal of Neuroinflammation*, *10*, 43. https://doi.org/10.1186/1742-2094-10-43

Nielsen, A. N., Barch, D. M., Petersen, S. E., Schlaggar, B. L., & Greene, D. J. (2020). Machine learning with neuroimaging: Evaluating its applications in psychiatry. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*(8), 791–798.

O'Brien, S. M., Scully, P., Scott, L. V., & Dinan, T. G. (2006). Cytokine profiles in bipolar affective disorder: Focus on acutely ill patients. *Journal of Affective Disorders*, *90*(2–3), 263–267. https://doi.org/10.1016/j.jad.2005.11.015

Ojala, M., & Garriga, G. C. (2009). Permutation Tests for Studying Classifier Performance. *2009 Ninth IEEE International Conference on Data Mining*, 908–913. https://doi.org/10.1109/ICDM.2009.108

Ongür, D., Lundy, M., Greenhouse, I., Shinn, A. K., Menon, V., Cohen, B. M., & Renshaw, P. F. (2010). Default mode network abnormalities in bipolar disorder and schizophrenia. *Psychiatry Research*, *183*(1), 59–68. https://doi.org/10.1016/j.pscychresns.2010.04.008

Ortiz-Domínguez, A., Hernández, M. E., Berlanga, C., Gutiérrez-Mora, D., Moreno, J., Heinze, G., & Pavón, L. (2007). Immune variations in bipolar disorder: Phasic

differences. *Bipolar Disorders*, *9*(6), 596–602. https://doi.org/10.1111/j.1399-5618.2007.00493.x

Otte, C., Gold, S. M., Penninx, B. W., Pariante, C. M., Etkin, A., Fava, M., Mohr, D. C., & Schatzberg, A. F. (2016). Major depressive disorder. *Nature Reviews Disease Primers*, *2*(1), Articolo 1. https://doi.org/10.1038/nrdp.2016.65

Palladini, M., Bravi, B., Colombo, F., Caselani, E., Di Pasquasio, C., D'Orsi, G., Rovere-Querini, P., Poletti, S., Benedetti, F., & Mazza, M. G. (2022). Cognitive remediation therapy for post-acute persistent cognitive deficits in COVID-19 survivors: A proof-of-concept study. *Neuropsychological Rehabilitation*, 1–18. https://doi.org/10.1080/09602011.2022.2075016

Pan, Z., Park, C., Brietzke, E., Zuckerman, H., Rong, C., Mansur, R. B., Fus, D., Subramaniapillai, M., Lee, Y., & McIntyre, R. S. (2019). Cognitive impairment in major depressive disorder. *CNS Spectrums*, *24*(1), 22–29. https://doi.org/10.1017/S1092852918001207

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage*, *45*(1), S199–S209.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron*, *48*(2), 175–187. https://doi.org/10.1016/j.neuron.2005.09.025

Phillips, M. L., & Kupfer, D. J. (2013). Bipolar disorder diagnosis: Challenges and future directions. *The Lancet*, *381*(9878), 1663–1671.

Phillips, M. L., Ladouceur, C. D., & Drevets, W. C. (2008). A neural model of voluntary and automatic emotion regulation: Implications for understanding the

pathophysiology and neurodevelopment of bipolar disorder. *Molecular Psychiatry*, *13*(9), 829, 833–857. https://doi.org/10.1038/mp.2008.65

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. *The American Journal of Psychiatry*, *171*(8), 829–843. https://doi.org/10.1176/appi.ajp.2014.13081008

Pierpaoli, C., & Basser, P. J. (1996). Toward a quantitative assessment of diffusion anisotropy. *Magnetic Resonance in Medicine*, *36*(6), 893–906. https://doi.org/10.1002/mrm.1910360612

Pizzagalli, D. A. (2011). Frontocingulate Dysfunction in Depression: Toward Biomarkers of Treatment Response. *Neuropsychopharmacology*, *36*(1), Articolo 1. https://doi.org/10.1038/npp.2010.166

Poletti, S., Aggio, V., Brioschi, S., Dallaspezia, S., Colombo, C., & Benedetti, F. (2017). Multidimensional cognitive impairment in unipolar and bipolar depression and the moderator effect of adverse childhood experiences. *Psychiatry and Clinical Neurosciences*, *71*(5), 309–317. https://doi.org/10.1111/pcn.12497

Poletti, S., Bollettini, I., Mazza, E., Locatelli, C., Radaelli, D., Vai, B., Smeraldi, E., Colombo, C., & Benedetti, F. (2015). Cognitive performances associate with measures of white matter integrity in bipolar disorder. *Journal of Affective Disorders*, *174*, 342–352. https://doi.org/10.1016/j.jad.2014.12.030

Poletti, S., Mazza, M. G., Calesella, F., Vai, B., Lorenzi, C., Manfredi, E., Colombo, C., Zanardi, R., & Benedetti, F. (2021). Circulating inflammatory markers impact

cognitive functions in bipolar depression. *Journal of Psychiatric Research*, *140*, 110–116. https://doi.org/10.1016/j.jpsychires.2021.05.071

Poletti, S., Papa, G. S., Locatelli, C., Colombo, C., & Benedetti, F. (2014). Neuropsychological deficits in bipolar depression persist after successful antidepressant treatment. *Journal of Affective Disorders*, *156*, 144–149. https://doi.org/10.1016/j.jad.2013.11.023

Poletti, S., Vai, B., Mazza, M. G., Zanardi, R., Lorenzi, C., Calesella, F., Cazzetta, S., Branchi, I., Colombo, C., Furlan, R., & Benedetti, F. (2021). A peripheral inflammatory signature discriminates bipolar from unipolar depression: A machine learning approach. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *105*, 110136. https://doi.org/10.1016/j.pnpbp.2020.110136

Powell, T. R., Gaspar, H., Chung, R., Keohane, A., Gunasinghe, C., Uher, R., Aitchison, K. J., Souery, D., Mors, O., Maier, W., Zobel, A., Rietschel, M., Henigsberg, N., Dernovšek, M. Z., Hauser, J., Frissa, S., Goodwin, L., Hotopf, M., Hatch, S. L., … Wang, H. (2018). *Assessing 42 inflammatory markers in 321 control subjects and 887 major depressive disorder cases: BMI and other confounders and overall predictive ability for current depression* (p. 327239). bioRxiv. https://doi.org/10.1101/327239

Prathikanti, S., & McMahon, F. J. (2001). Genome scans for susceptibility genes in bipolar affective disorder. *Annals of Medicine*, *33*(4), 257–262. https://doi.org/10.3109/07853890108998754

Psychiatric GWAS Consortium Bipolar Disorder Working Group. (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new

susceptibility locus near ODZ4. *Nature Genetics*, *43*(10), 977–983. https://doi.org/10.1038/ng.943

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., & Daly, M. J. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, *81*(3), 559–575.

Radua, J., Grunze, H., & Amann, B. L. (2017). Meta-Analysis of the Risk of Subsequent Mood Episodes in Bipolar Disorder. *Psychotherapy and Psychosomatics*, *86*(2), 90–98.

Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, *38*, 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(2), 676–682. https://doi.org/10.1073/pnas.98.2.676

Redlich, R., Almeida, J. R., Grotegerd, D., Opel, N., Kugel, H., Heindel, W., Arolt, V., Phillips, M. L., & Dannlowski, U. (2014). Brain morphometric biomarkers distinguishing unipolar and bipolar depression: A voxel-based morphometry–pattern classification approach. *JAMA psychiatry*, *71*(11), 1222–1230.

Rive, M. M., Redlich, R., Schmaal, L., Marquand, A. F., Dannlowski, U., Grotegerd, D., Veltman, D. J., Schene, A. H., & Ruhé, H. G. (2016). Distinguishing medication-free subjects with unipolar disorder from subjects with bipolar disorder: State matters. *Bipolar disorders*, *18*(7), 612–623.

Rizvi, S. J., Lambert, C., & Kennedy, S. (2018). Presentation and Neurobiology of Anhedonia in Mood Disorders: Commonalities and Distinctions. *Current Psychiatry Reports*, *20*(2), 13. https://doi.org/10.1007/s11920-018-0877-z

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. *Psychological Medicine*, *44*(10), 2029–2040. https://doi.org/10.1017/S0033291713002535

Roy, M., Shohamy, D., Daw, N., Jepma, M., Wimmer, G. E., & Wager, T. D. (2014). Representation of aversive prediction errors in the human periaqueductal gray. *Nature Neuroscience*, *17*(11), 1607–1612. https://doi.org/10.1038/nn.3832

Rubin-Falcone, H., Zanderigo, F., Thapa-Chhetry, B., Lan, M., Miller, J. M., Sublette, M. E., Oquendo, M. A., Hellerstein, D. J., McGrath, P. J., & Stewart, J. W. (2018). Pattern recognition of magnetic resonance imaging-based gray matter volume measurements classifies bipolar disorder and major depressive disorder. *Journal of affective disorders*, *227*, 498–505.

Ruderfer, D. M., Fanous, A. H., Ripke, S., McQuillin, A., Amdur, R. L., Schizophrenia Working Group of the Psychiatric Genomics Consortium, Bipolar Disorder Working Group of the Psychiatric Genomics Consortium, Cross-Disorder Working Group of the Psychiatric Genomics Consortium, Gejman, P. V., O'Donovan, M. C., Andreassen, O. A., Djurovic, S., Hultman, C. M., Kelsoe, J. R., Jamain, S., Landén, M., Leboyer, M., Nimgaonkar, V., Nurnberger, J., … Kendler, K. S. (2014). Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Molecular Psychiatry*, *19*(9), 1017–1024. https://doi.org/10.1038/mp.2013.138

Rueckert, E. H., Barker, D., Ruderfer, D., Bergen, S. E., O'Dushlaine, C., Luce, C. J., Sheridan, S. D., Theriault, K. M., Chambert, K., Moran, J., Purcell, S. M., Madison, J. M., Haggarty, S. J., & Sklar, P. (2013). Cis-acting regulation of brain-specific ANK3 gene expression by a genetic variant associated with bipolar disorder. *Molecular Psychiatry*, *18*(8), 922–929. https://doi.org/10.1038/mp.2012.104

Ryan, J. P., Sheu, L. K., Critchley, H. D., & Gianaros, P. J. (2012). A neural circuitry linking insulin resistance to depressed mood. *Psychosomatic Medicine*, *74*(5), 476–482. https://doi.org/10.1097/PSY.0b013e31824d0865

Sackeim, H. A. (2001). The definition and meaning of treatment-resistant depression. *Journal of Clinical Psychiatry*, *62*, 10–17.

Saleh, K., Carballedo, A., Lisiecka, D., Fagan, A. J., Connolly, G., Boyle, G., & Frodl, T. (2012). Impact of family history and depression on amygdala volume. *Psychiatry Research*, *203*(1), 24–30. https://doi.org/10.1016/j.pscychresns.2011.10.004

Scalabrini, A., Vai, B., Poletti, S., Damiani, S., Mucci, C., Colombo, C., Zanardi, R., Benedetti, F., & Northoff, G. (2020). All roads lead to the default-mode network—Global source of DMN abnormalities in major depressive disorder. *Neuropsychopharmacology*, *45*(12), Articolo 12. https://doi.org/10.1038/s41386-020-0785-x

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex*

(New York, N.Y.: 1991), *28*(9), 3095–3114. https://doi.org/10.1093/cercor/bhx179

Schiavone, S., Mhillaj, E., Neri, M., Morgese, M. G., Tucci, P., Bove, M., Valentino, M., Di Giovanni, G., Pomara, C., Turillazzi, E., Trabace, L., & Cuomo, V. (2017). Early Loss of Blood-Brain Barrier Integrity Precedes NOX2 Elevation in the Prefrontal Cortex of an Animal Model of Psychosis. *Molecular Neurobiology*, *54*(3), 2031–2044. https://doi.org/10.1007/s12035-016-9791-8

Schimmelpfennig, J., Topczewski, J., Zajkowski, W., & Jankowiak-Siuda, K. (2023). The role of the salience network in cognitive and affective deficits. *Frontiers in Human Neuroscience*, *17*, 1133367. https://doi.org/10.3389/fnhum.2023.1133367

Schmaal, L., Veltman, D. J., van Erp, T. G. M., Sämann, P. G., Frodl, T., Jahanshad, N., Loehrer, E., Tiemeier, H., Hofman, A., Niessen, W. J., Vernooij, M. W., Ikram, M. A., Wittfeld, K., Grabe, H. J., Block, A., Hegenscheid, K., Völzke, H., Hoehn, D., Czisch, M., … Hibar, D. P. (2016). Subcortical brain alterations in major depressive disorder: Findings from the ENIGMA Major Depressive Disorder working group. *Molecular Psychiatry*, *21*(6), 806–812. https://doi.org/10.1038/mp.2015.69

Schmidt, F. M., Lichtblau, N., Minkwitz, J., Chittka, T., Thormann, J., Kirkby, K. C., Sander, C., Mergl, R., Faßhauer, M., Stumvoll, M., Holdt, L. M., Teupser, D., Hegerl, U., & Himmerich, H. (2014). Cytokine levels in depressed and non-depressed subjects, and masking effects of obesity. *Journal of Psychiatric Research*, *55*, 29–34. https://doi.org/10.1016/j.jpsychires.2014.04.021

Schrouff, J., Monteiro, J. M., Portugal, L., Rosa, M. J., Phillips, C., & Mourão-Miranda, J. (2018). Embedding Anatomical or Functional Knowledge in Whole-Brain Multiple Kernel Learning Models. *Neuroinformatics*, *16*(1), 117–143. https://doi.org/10.1007/s12021-017-9347-8

Schrouff, J., Rosa, M. J., Rondina, J. M., Marquand, A. F., Chu, C., Ashburner, J., Phillips, C., Richiardi, J., & Mourão-Miranda, J. (2013). PRoNTo: Pattern Recognition for Neuroimaging Toolbox. *Neuroinformatics*, *11*(3), 319–337. https://doi.org/10.1007/s12021-013-9178-1

Schulte, T., Sullivan, E. V., Müller-Oehring, E. M., Adalsteinsson, E., & Pfefferbaum, A. (2005). Corpus callosal microstructural integrity influences interhemispheric processing: A diffusion tensor imaging study. *Cerebral Cortex (New York, N.Y.: 1991)*, *15*(9), 1384–1392. https://doi.org/10.1093/cercor/bhi020

Schulze, T. G., Detera-Wadleigh, S. D., Akula, N., Gupta, A., Kassem, L., Steele, J., Pearl, J., Strohmaier, J., Breuer, R., Schwarz, M., Propping, P., Nöthen, M. M., Cichon, S., Schumacher, J., NIMH Genetics Initiative Bipolar Disorder Consortium, Rietschel, M., & McMahon, F. J. (2009). Two variants in Ankyrin 3 (ANK3) are independent genetic risk factors for bipolar disorder. *Molecular Psychiatry*, *14*(5), 487–491. https://doi.org/10.1038/mp.2008.134

Schulze, T. G., Ohlraun, S., Czerski, P. M., Schumacher, J., Kassem, L., Deschner, M., Gross, M., Tullius, M., Heidmann, V., Kovalenko, S., Jamra, R. A., Becker, T., Leszczynska-Rodziewicz, A., Hauser, J., Illig, T., Klopp, N., Wellek, S., Cichon, S., Henn, F. A., … Rietschel, M. (2005). Genotype-phenotype studies in bipolar disorder showing association between the DAOA/G30 locus and persecutory delusions: A first step toward a molecular genetic classification of

psychiatric phenotypes. *The American Journal of Psychiatry*, *162*(11), 2101–2108. https://doi.org/10.1176/appi.ajp.162.11.2101

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., & Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(9), 2349–2356. https://doi.org/10.1523/JNEUROSCI.5587-06.2007

Serpa, M. H., Ou, Y., Schaufelberger, M. S., Doshi, J., Ferreira, L. K., Machado-Vieira, R., Menezes, P. R., Scazufca, M., Davatzikos, C., & Busatto, G. F. (2014). Neuroanatomical classification in a population-based sample of psychotic major depression and bipolar I disorder with 1 year of diagnostic stability. *BioMed research international*, *2014*.

Sesack, S. R., & Grace, A. A. (2010). Cortico-Basal Ganglia reward network: Microcircuitry. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *35*(1), 27–47. https://doi.org/10.1038/npp.2009.93

Shao, J., Dai, Z., Zhu, R., Wang, X., Tao, S., Bi, K., Tian, S., Wang, H., Sun, Y., Yao, Z., & Lu, Q. (2019). Early identification of bipolar from unipolar depression before manic episode: Evidence from dynamic rfMRI. *Bipolar Disorders*, *21*(8), 774–784. https://doi.org/10.1111/bdi.12819

Shawe-Taylor, J., & Cristianini, N. (2004). *Kernel Methods for Pattern Analysis*.

Shen, K.-K., Welton, T., Lyon, M., McCorkindale, A. N., Sutherland, G. T., Burnham, S., Fripp, J., Martins, R., & Grieve, S. M. (2020). Structural core of the

executive control network: A high angular resolution diffusion MRI study. *Human Brain Mapping*, *41*(5), 1226–1236. https://doi.org/10.1002/hbm.24870

Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., & Greicius, M. D. (2012). Decoding Subject-Driven Cognitive States with Whole-Brain Connectivity Patterns. *Cerebral Cortex*, *22*(1), 158–165. https://doi.org/10.1093/cercor/bhr099

Smith, E. N., Bloss, C. S., Badner, J. A., Barrett, T., Belmonte, P. L., Berrettini, W., Byerley, W., Coryell, W., Craig, D., Edenberg, H. J., Eskin, E., Foroud, T., Gershon, E., Greenwood, T. A., Hipolito, M., Koller, D. L., Lawson, W. B., Liu, C., Lohoff, F., … Kelsoe, J. R. (2009). Genome-wide association study of bipolar disorder in European American and African American individuals. *Molecular Psychiatry*, *14*(8), 755–763. https://doi.org/10.1038/mp.2009.43

Smith, S. M. (2002). Fast robust automated brain extraction. *Human brain mapping*, *17*(3), 143–155.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., Ciccarelli, O., Cader, M. Z., & Matthews, P. M. (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *Neuroimage*, *31*(4), 1487–1505.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., & Flitney, D. E. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*, S208–S219.

Smoller, J. W., & Finn, C. T. (2003). Family, twin, and adoption studies of bipolar disorder. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics*, *123C*(1), 48–58. https://doi.org/10.1002/ajmg.c.20013

Song, S.-K., Sun, S.-W., Ju, W.-K., Lin, S.-J., Cross, A. H., & Neufeld, A. H. (2003). Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. *NeuroImage*, *20*(3), 1714–1722. https://doi.org/10.1016/j.neuroimage.2003.07.005

Song, S.-K., Sun, S.-W., Ramsbottom, M. J., Chang, C., Russell, J., & Cross, A. H. (2002). Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *NeuroImage*, *17*(3), 1429–1436. https://doi.org/10.1006/nimg.2002.1267

Stahl, E. A., Breen, G., Forstner, A. J., McQuillin, A., Ripke, S., Trubetskoy, V., Mattheisen, M., Wang, Y., Coleman, J. R. I., Gaspar, H. A., de Leeuw, C. A., Steinberg, S., Pavlides, J. M. W., Trzaskowski, M., Byrne, E. M., Pers, T. H., Holmans, P. A., Richards, A. L., Abbott, L., … Bipolar Disorder Working Group of the Psychiatric Genomics Consortium. (2019). Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nature Genetics*, *51*(5), 793–803. https://doi.org/10.1038/s41588-019-0397-8

Strakowski, S. M., Delbello, M. P., & Adler, C. M. (2005). The functional neuroanatomy of bipolar disorder: A review of neuroimaging findings. *Molecular Psychiatry*, *10*(1), 105–116. https://doi.org/10.1038/sj.mp.4001585

Strange, B. A., Witter, M. P., Lein, E. S., & Moser, E. I. (2014). Functional organization of the hippocampal longitudinal axis. *Nature Reviews Neuroscience*, *15*(10), Articolo 10. https://doi.org/10.1038/nrn3785

Sullivan, P. F., Eaves, L. J., Kendler, K. S., & Neale, M. C. (2001). Genetic case-control association studies in neuropsychiatry. *Archives of General Psychiatry, 58*(11), 1015–1024. https://doi.org/10.1001/archpsyc.58.11.1015

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: Review and meta-analysis. *The American Journal of Psychiatry, 157*(10), 1552–1562. https://doi.org/10.1176/appi.ajp.157.10.1552

Sun, S.-W., Liang, H.-F., Trinkaus, K., Cross, A. H., Armstrong, R. C., & Song, S.-K. (2006). Noninvasive detection of cuprizone induced axonal damage and demyelination in the mouse corpus callosum. *Magnetic Resonance in Medicine, 55*(2), 302–308. https://doi.org/10.1002/mrm.20774

Teixeira, A. L., Colpo, G. D., Fries, G. R., Bauer, I. E., & Selvaraj, S. (2019). Biomarkers for bipolar disorder: Current status and challenges ahead. *Expert Review of Neurotherapeutics, 19*(1), 67–81. https://doi.org/10.1080/14737175.2019.1550361

Teixeira, A. L., Gama, C. S., Rocha, N. P., & Teixeira, M. M. (2018). Revisiting the Role of Eotaxin-1/CCL11 in Psychiatric Disorders. *Frontiers in Psychiatry, 9*, 241. https://doi.org/10.3389/fpsyt.2018.00241

Thomas, A. J., Davis, S., Morris, C., Jackson, E., Harrison, R., & O'Brien, J. T. (2005). Increase in interleukin-1beta in late-life depression. *The American Journal of Psychiatry, 162*(1), 175–177. https://doi.org/10.1176/appi.ajp.162.1.175

Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature Reviews Neuroscience, 16*(1), Articolo 1. https://doi.org/10.1038/nrn3857

Vai, B., Bertocchi, C., & Benedetti, F. (2019). Cortico-limbic connectivity as a possible biomarker for bipolar disorder: Where are we now? *Expert Review of Neurotherapeutics*, *19*(2), 159–172. https://doi.org/10.1080/14737175.2019.1562338

Vai, B., Parenti, L., Bollettini, I., Cara, C., Verga, C., Melloni, E., Mazza, E., Poletti, S., Colombo, C., & Benedetti, F. (2020). Predicting differential diagnosis between bipolar and unipolar depression with multiple kernel learning on multimodal structural neuroimaging. *European Neuropsychopharmacology*, *34*, 28–38.

van den Ameele, S., van Diermen, L., Staels, W., Coppens, V., Dumont, G., Sabbe, B., & Morrens, M. (2016). The effect of mood-stabilizing drugs on cytokine levels in bipolar disorder: A systematic review. *Journal of Affective Disorders*, *203*, 364–373. https://doi.org/10.1016/j.jad.2016.06.016

Vapnik, V. N. (2000). *The Nature of Statistical Learning Theory*. Springer. https://doi.org/10.1007/978-1-4757-3264-1

Varghese, S., Frey, B. N., Schneider, M. A., Kapczinski, F., & de Azevedo Cardoso, T. (2022). Functional and cognitive impairment in the first episode of depression: A systematic review. *Acta Psychiatrica Scandinavica*, *145*(2), 156–185. https://doi.org/10.1111/acps.13385

Varoquaux, G., Raamana, P. R., Engemann, D. A., Hoyos-Idrobo, A., Schwartz, Y., & Thirion, B. (2017). Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. *NeuroImage*, *145*, 166–179.

Vázquez, G. H., Lolich, M., Cabrera, C., Jokic, R., Kolar, D., Tondo, L., & Baldessarini, R. J. (2018). Mixed symptoms in major depressive and bipolar

disorders: A systematic review. *Journal of Affective Disorders*, *225*, 756–760. https://doi.org/10.1016/j.jad.2017.09.006

Vázquez-León, P., Miranda-Páez, A., Valencia-Flores, K., & Sánchez-Castillo, H. (2023). Defensive and Emotional Behavior Modulation by Serotonin in the Periaqueductal Gray. *Cellular and Molecular Neurobiology*, *43*(4), 1453–1468. https://doi.org/10.1007/s10571-022-01262-z

Vos, T., Abajobir, A. A., Abate, K. H., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abdulkader, R. S., Abdulle, A. M., Abebo, T. A., Abera, S. F., Aboyans, V., Abu-Raddad, L. J., Ackerman, I. N., Adamu, A. A., Adetokunboh, O., Afarideh, M., Afshin, A., Agarwal, S. K., Aggarwal, R., … Murray, C. J. L. (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, *390*(10100), 1211–1259. https://doi.org/10.1016/S0140-6736(17)32154-2

Waller, L., Erk, S., Pozzi, E., Toenders, Y. J., Haswell, C. C., Büttner, M., Thompson, P. M., Schmaal, L., Morey, R. A., Walter, H., & Veer, I. M. (2022). ENIGMA HALFpipe: Interactive, reproducible, and efficient analysis for resting-state and task-based fMRI data. *Human Brain Mapping*, *43*(9), 2727–2742. https://doi.org/10.1002/hbm.25829

Walter, M., Alizadeh, S., Jamalabadi, H., Lueken, U., Dannlowski, U., Walter, H., Olbrich, S., Colic, L., Kambeitz, J., Koutsouleris, N., Hahn, T., & Dwyer, D. B. (2019). Translational machine learning for psychiatric neuroimaging. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *91*, 113–121. https://doi.org/10.1016/j.pnpbp.2018.09.014

Wang, Y., Wang, J., Jia, Y., Zhong, S., Zhong, M., Sun, Y., Niu, M., Zhao, L., Zhao, L., Pan, J., Huang, L., & Huang, R. (2017). Topologically convergent and divergent functional connectivity patterns in unmedicated unipolar depression and bipolar disorder. *Translational Psychiatry*, *7*(7), e1165. https://doi.org/10.1038/tp.2017.117

Watson, T. C., Cerminara, N. L., Lumb, B. M., & Apps, R. (2016). Neural Correlates of Fear in the Periaqueductal Gray. *The Journal of Neuroscience*, *36*(50), 12707–12719. https://doi.org/10.1523/JNEUROSCI.1100-16.2016

Weis, C. N., Bennett, K. P., Huggins, A. A., Parisi, E. A., Gorka, S. M., & Larson, C. (2022). A 7-Tesla MRI study of the periaqueductal gray: Resting state and task activation under threat. *Social Cognitive and Affective Neuroscience*, *17*(2), 187–197. https://doi.org/10.1093/scan/nsab085

Whitton, A. E., & Pizzagalli, D. A. (2022). Anhedonia in Depression and Bipolar Disorder. *Current Topics in Behavioral Neurosciences*, *58*, 111–127. https://doi.org/10.1007/7854_2022_323

Wijeratne, C., Sachdev, S., Wen, W., Piguet, O., Lipnicki, D. M., Malhi, G. S., Mitchell, P. B., & Sachdev, P. S. (2013). Hippocampal and amygdala volumes in an older bipolar disorder sample. *International Psychogeriatrics*, *25*(1), 54–60. https://doi.org/10.1017/S1041610212001469

Winklewski, P. J., Sabisz, A., Naumczyk, P., Jodzio, K., Szurowska, E., & Szarmach, A. (2018). Understanding the Physiopathology Behind Axial and Radial Diffusivity Changes—What Do We Know? *Frontiers in Neurology*, *9*. https://www.frontiersin.org/articles/10.3389/fneur.2018.00092

Wittchen, H.-U. (2012). The Burden of Mood Disorders. *Science, 338*(6103), 15–15. https://doi.org/10.1126/science.1230817

Wollenhaupt-Aguiar, B., Librenza-Garcia, D., Bristot, G., Przybylski, L., Stertz, L., Kubiachi Burque, R., Ceresér, K. M., Spanemberg, L., Caldieraro, M. A., Frey, B. N., Fleck, M. P., Kauer-Sant'Anna, M., Passos, I. C., & Kapczinski, F. (2020). Differential biomarker signatures in unipolar and bipolar depression: A machine learning approach. *The Australian and New Zealand Journal of Psychiatry, 54*(4), 393–401. https://doi.org/10.1177/0004867419888027

Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., & Smith, S. M. (2009). Bayesian analysis of neuroimaging data in FSL. *Neuroimage, 45*(1), S173–S186.

Ye, T., Wang, D., Cai, Z., Tong, L., Chen, Z., Lu, J., Lu, X., Huang, C., & Yuan, X. (2020). Antidepressive properties of macrophage-colony stimulating factor in a mouse model of depression induced by chronic unpredictable stress. *Neuropharmacology*, *172*, 108132. https://doi.org/10.1016/j.neuropharm.2020.108132

Yi, H., Raman, A. T., Zhang, H., Allen, G. I., & Liu, Z. (2018). Detecting hidden batch factors through data-adaptive adjustment for biological effects. *Bioinformatics, 34*(7), 1141–1147.

Yu, H., Li, M.-L., Li, Y.-F., Li, X.-J., Meng, Y., Liang, S., Li, Z., Guo, W., Wang, Q., Deng, W., Ma, X., Coid, J., & Li, D. T. (2020). Anterior cingulate cortex, insula and amygdala seed-based whole brain resting-state functional connectivity differentiates bipolar from unipolar depression. *Journal of Affective Disorders, 274*, 38–47. https://doi.org/10.1016/j.jad.2020.05.005

Yu, M., Linn, K. A., Cook, P. A., Phillips, M. L., McInnis, M., Fava, M., Trivedi, M. H., Weissman, M. M., Shinohara, R. T., & Sheline, Y. I. (2018). Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fMRI data. *Human brain mapping*, *39*(11), 4213–4227.

Zang, Y. (2004). Ji ang T, Lu Y, He Y, Tian L. Regional homogeneity approach to fMRI data analysis. *Neuroimage*, *22*, 394–400.

Zelena, D., Menant, O., Andersson, F., & Chaillou, E. (2018). Periaqueductal gray and emotions: The complexity of the problem and the light at the end of the tunnel, the magnetic resonance imaging. *Endocrine Regulations*, *52*(4), 222–238. https://doi.org/10.2478/enr-2018-0027

Zimmermann, P., Brückl, T., Nocon, A., Pfister, H., Lieb, R., Wittchen, H.-U., Holsboer, F., & Angst, J. (2009). Heterogeneity of DSM-IV major depressive disorder as a consequence of subthreshold bipolarity. *Archives of General Psychiatry*, *66*(12), 1341–1352. https://doi.org/10.1001/archgenpsychiatry.2009.158

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, *67*(2), 301–320.

Zou, Q.-H., Zhu, C.-Z., Yang, Y., Zuo, X.-N., Long, X.-Y., Cao, Q.-J., Wang, Y.-F., & Zang, Y.-F. (2008). An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: Fractional ALFF. *Journal of neuroscience methods*, *172*(1), 137–141.

Zung, S., Souza-Duran, F. L., Soeiro-de-Souza, M. G., Uchida, R., Bottino, C. M., Busatto, G. F., & Vallada, H. (2016). The influence of lithium on hippocampal

volume in elderly bipolar patients: A study using voxel-based morphometry. *Translational Psychiatry*, *6*(6), e846. https://doi.org/10.1038/tp.2016.97

APPENDIX

Appendix 1. Descriptive statistics of the subsample with BACS scores, by the presence or absence of deficit at the coordination domain

Variable	Average ± s.d. / N			t/χ	p-value
	Total (N=168)	No deficit (N=73)	Impaired (N=95)		
Age	47.77 ± 10.77	45.96 ± 10.58	49.16 ± 10.71	-1.92	0.057
Sex	103 F / 65 M	44 F / 29 M	59 F / 36 M	0.01	0.935
Diagnosis	79 BD / 89 MDD	38 BD / 35 MDD	41 BD / 54 MDD	0.98	0.322
Years of education	13.01 ± 3.66	13.21 ± 3.61	12.86 ± 3.69	0.60	0.550
Duration of illness	17.89 ± 11.40	17.62 ± 11.37	18.09 ± 11.41	-0.27	0.789
No. of episodes	7.94 ± 8.66	7.75 ± 7.37	8.08 ± 9.54	-0.24	0.808
Age of onset	29.88 ± 11.22	28.34 ± 10.48	31.06 ± 11.61	-1.56	0.121
Pharmacological load	4.49 ± 2.01	3.86 ± 1.79	4.97 ± 2.04	-3.63	<0.001*

Average ± standard deviation is reported for continuous variables, whereas frequency was reported for categorical variables. Abbreviations: BD, bipolar disorder; F, female; M, male; MDD, major depressive disorder; s.d., standard deviation. * $p<.05$

Appendix 2. Descriptive statistics of the subsample with BACS scores, by the presence or absence of deficit at the executive functions domain

Variable	Average ± s.d. / N			t/χ	p-value
	Total (N=168)	No deficit (N=105)	Impaired (N=63)		
Age	47.77 ± 10.77	46.26 ± 10.68	50.29 ± 10.45	-2.37	0.019*
Sex	103 F / 65 M	61 F / 44 M	42 F / 21 M	0.88	0.347
Diagnosis	79 BD / 89 MDD	54 BD / 51 MDD	25 BD / 38 MDD	1.73	0.188
Years of education	13.01 ± 3.66	13.60 ± 3.46	12.03 ± 3.76	2.73	0.007*
Duration of illness	17.89 ± 11.40	17.10 ± 10.64	19.21 ± 12.44	-1.16	0.248
No. of episodes	7.94 ± 8.66	8.02 ± 7.53	7.81 ± 10.28	0.15	0.880
Age of onset	29.88 ± 11.22	29.16 ± 9.81	31.08 ± 13.15	-1.07	0.286
Pharmacological load	4.49 ± 2.01	4.43 ± 1.99	4.57 ± 2.05	-0.43	0.668

Average ± standard deviation is reported for continuous variables, whereas frequency was reported for categorical variables. Abbreviations: BD, bipolar disorder; F, female; M, male; MDD, major depressive disorder; s.d., standard deviation. * $p<.05$

Appendix 3. Descriptive statistics of the subsample with BACS scores, by the presence or absence of deficit at the fluency domain

Variable	Average ± s.d. / N			t/χ	p-value
	Total (N=168)	No deficit (N=107)	Impaired (N=61)		
Age	47.77 ± 10.77	45.43 ± 10.92	51.87 ± 9.16	-3.87	<0.001*
Sex	103 F / 65 M	75 F / 32 M	28 F / 33 M	8.59	0.003*
Diagnosis	79 BD / 89 MDD	51 BD / 56 MDD	28 BD / 33 MDD	0.00	0.953
Years of education	13.01 ± 3.66	13.17 ± 3.42	12.74 ± 4.03	0.73	0.466
Duration of illness	17.89 ± 11.40	16.36 ± 10.89	20.56 ± 11.76	-2.32	0.022*
No. of episodes	7.94 ± 8.66	7.74 ± 8.32	8.30 ± 9.23	-0.4	0.691
Age of onset	29.88 ± 11.22	29.07 ± 11.05	31.31 ± 11.37	-1.25	0.214
Pharmacological load	4.49 ± 2.01	4.18 ± 2.02	5.02 ± 1.89	-2.62	0.010*

Average ± standard deviation is reported for continuous variables, whereas frequency was reported for categorical variables. Abbreviations: BD, bipolar disorder; F, female; M, male; MDD, major depressive disorder; s.d., standard deviation. * p<.05

Appendix 4. Descriptive statistics of the subsample with BACS scores, by the presence or absence of deficit at the processing speed domain

Variable	Average ± s.d. / N			t/χ	p-value
	Total (N=168)	No deficit (N=44)	Imapired (N=124)		
Age	47.77 ± 10.77	40.57 ± 11.42	50.32 ± 9.27	-5.59	<0.001*
Sex	103 F / 65 M	32 F / 12 M	71 F / 53 M	2.66	0.103
Diagnosis	79 BD / 89 MDD	23 BD / 21 MDD	56 BD / 68 MDD	0.40	0.525
Years of education	13.01 ± 3.66	14.25 ± 3.30	12.57 ± 3.68	2.65	0.009*
Duration of illness	17.89 ± 11.40	14.68 ± 10.48	19.02 ± 11.49	-2.19	0.030*
No. of episodes	7.94 ± 8.66	8.39 ± 10.14	7.78 ± 8.07	0.40	0.693
Age of onset	29.88 ± 11.22	25.89 ± 8.70	31.30 ± 11.66	-2.80	0.006*
Pharmacological load	4.49 ± 2.01	3.84 ± 1.68	4.72 ± 2.07	-2.50	0.013*

Average ± standard deviation is reported for continuous variables, whereas frequency was reported for categorical variables. Abbreviations: BD, bipolar disorder; F, female; M, male; MDD, major depressive disorder; s.d., standard deviation. * p<.05

Appendix 5. Descriptive statistics of the subsample with BACS scores, by the presence or absence of deficit at the verbal memory domain

Variable	Average ± s.d. / N			t/χ	p-value
	Total (N=168)	No deficit (N=99)	Impaired (N=69)		
Age	47.77 ± 10.77	48.69 ± 9.74	46.45 ± 11.98	1.32	0.187
Sex	103 F / 65 M	71 F / 28 M	32 F / 37 M	9.96	0.002*
Diagnosis	79 BD / 89 MDD	46 BD / 53 MDD	33 BD / 36 MDD	0.00	0.987
Years of education	13.01 ± 3.66	13.16 ± 3.62	12.80 ± 3.70	0.63	0.528
Duration of illness	17.89 ± 11.40	19.36 ± 11.38	15.77 ± 11.08	2.02	0.045*
No. of episodes	7.94 ± 8.66	7.37 ± 8.02	8.75 ± 9.46	-1.01	0.313
Age of onset	29.88 ± 11.22	29.32 ± 10.96	30.68 ± 11.53	-0.77	0.443
Pharmacological load	4.49 ± 2.01	4.21 ± 2.03	4.88 ± 1.93	-2.13	0.035*

Average ± standard deviation is reported for continuous variables, whereas frequency was reported for categorical variables. Abbreviations: BD, bipolar disorder; F, female; M, male; MDD, major depressive disorder; s.d., standard deviation. * $p<.05$

Appendix 6. Descriptive statistics of the subsample with BACS scores, by the presence or absence of deficit at the working memory domain

Variable	Average ± s.d. / N			t/χ	p-value
	Total (N=168)	No deficit (N=71)	Imapired (N=97)		
Age	47.77 ± 10.77	46.42 ± 10.41	48.75 ± 10.92	-1.38	0.168
Sex	103 F / 65 M	44 F / 27 M	59 F / 38 M	0.00	1.000
Diagnosis	79 BD / 89 MDD	32 BD / 39 MDD	47 BD / 50 MDD	0.08	0.781
Years of education	13.01 ± 3.66	13.93 ± 3.66	12.34 ± 3.51	2.83	0.005*
Duration of illness	17.89 ± 11.40	18.62 ± 11.14	17.35 ± 11.55	0.71	0.479
No. of episodes	7.94 ± 8.66	8.97 ± 10.80	7.19 ± 6.59	1.32	0.189
Age of onset	29.88 ± 11.22	27.80 ± 10.47	31.40 ± 11.49	-2.07	0.040*
Pharmacological load	4.49 ± 2.01	4.04 ± 1.93	4.81 ± 2.01	-2.47	0.014*

Average ± standard deviation is reported for continuous variables, whereas frequency was reported for categorical variables. Abbreviations: BD, bipolar disorder; F, female; M, male; MDD, major depressive disorder; s.d., standard deviation. * $p<.05$

Appendix 7. Performance of scanner-classification models pre- and post-ComBat correction.

Data	Feature	BA	Sens	Spec	NPV	PPV	AUC
Before ComBat	ABC	95.00	96.00	94.00	94.18	96.00	0.98
	DRCs	94.00	94.00	94.00	94.67	94.67	0.99
	fALFF	99.00	100.00	98.00	98.18	100.00	1.00
	ReHo	100.00	100.00	100.00	100.00	100.00	1.00
	SBC	94.00	98.00	90.00	91.75	97.78	1.00
	AD	98.22	97.97	98.46	98.52	98.07	0.99
	FA	97.19	96.44	97.95	98.05	96.57	0.99
	MD	97.71	97.46	97.95	98.06	7.56	0.99
	RD	97.96	97.46	98.46	98.51	97.56	0.99
	VBM	100.00	100.00	100.00	100.00	100.00	1.00
After ComBat	ABC	53.00	74.00	32.00	52.20	54.76	0.55
	DRCs	50.00	60.00	40.00	50.00	50.00	0.20
	fALFF	26.00	14.00	38.00	13.10	28.31	0.13
	ReHo	25.00	8.00	42.00	6.62	25.79	0.03
	SBC	50.00	80.00	20.00	50.00	50.00	0.21
	AD	30.10	8.69	51.50	15.44	34.84	0.15
	FA	31.13	17.44	44.82	17.37	31.35	0.17
	MD	33.90	8.67	59.14	19.49	38.28	0.17
	RD	34.66	10.71	58.62	21.46	38.18	0.20
	VBM	56.31	54.94	57.69	59.40	56.35	0.48

For each feature, accuracy metrics are reported. Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; AUC, area under the curve; BA, balance accuracy; DR, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity; NPV, negative predictive value; PPV, positive predictive value; RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed-based connectivity, Sens, sensitivity; Spec, specificity; VBM, voxel-based morphometry.

Appendix 8. Performance of differentiation between MDD and BD patients before and after ComBat correction, for each MRI-derived feature.

Data	Feature	BA	Sens	Spec	NPV	PPV	AUC
Before ComBat	ABC	50.19	54.87	45.51	49.79	50.66	0.47
	DRCs	46.15	42.69	49.62	45.83	44.32	0.48
	fALFF	59.42	52.95	65.90	61.20	58.46	0.62
	ReHo	53.97	58.21	49.74	54.09	53.59	0.52
	SBC	58.21	44.10	72.31	61.67	57.38	0.60
	AD	71.57	75.71	67.43	70.85	76.85	0.83
	FA	74.48	71.67	77.29	76.68	76.77	0.85
	MD	73.02	76.62	69.43	71.90	75.55	0.84
	RD	73.05	74.67	71.43	72.58	74.37	0.85
	VBM	62.57	69.33	55.81	61.03	65.00	0.63
After ComBat	ABC	55.71	53.46	57.95	56.48	58.04	0.54
	DRCs	65.26	65.64	64.87	65.29	66.74	0.70
	fALFF	59.55	61.41	57.69	61.18	59.70	0.66
	ReHo	57.37	66.28	48.46	55.84	61.26	0.62
	SBC	70.96	77.69	64.23	69.24	74.78	0.77
	AD	75.90	77.52	74.29	76.01	77.19	0.88
	FA	76.95	71.67	82.24	80.24	75.77	0.87
	MD	80.33	83.33	77.33	78.92	82.35	0.88
	RD	76.95	77.48	76.43	76.54	78.25	0.87
	VBM	55.32	53.95	56.58	55.91	54.85	0.58

For each feature, accuracy metrics are reported. Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; AUC, area under the curve; BA, balance accuracy; DR, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity; NPV, negative predictive value; PPV, positive predictive value; RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed based connectivity, Sens, sensitivity, Spec, specificity; VBM, voxel-based morphometry.

Appendix 9. Performance of differentiation between cognitively intact and impaired patients at the composite score, before and after ComBat correction, for each MRI-derived feature.

Data	Feature	BA	Sens	Spec	NPV	PPV	AUC
Before ComBat	ABC	50.71	40.00	61.43	49.71	54.00	0.45
	DRCs	52.38	57.14	47.62	51.00	54.76	0.43
	fALFF	55.24	55.24	55.24	56.90	54.40	0.50
	ReHo	50.00	40.95	59.05	50.00	50.00	0.58
	SBC	47.38	31.43	63.33	50.30	46.17	0.52
	AD	56.22	55.78	56.67	59.02	58.52	0.58
	FA	46.00	57.56	34.44	44.16	45.07	0.42
	MD	53.89	53.56	54.22	54.50	55.94	0.61
	RD	52.89	61.56	44.22	53.00	54.77	0.57
	VBM	55.82	54.18	57.45	53.54	58.76	0.63
After ComBat	ABC	55.71	46.67	64.76	58.05	57.61	0.51
	DRCs	46.43	18.57	74.29	41.67	47.18	0.47
	fALFF	55.24	61.90	48.57	55.33	55.33	0.52
	ReHo	59.76	51.43	68.10	62.86	58.98	0.54
	SBC	49.29	48.10	50.48	51.81	47.43	0.51
	AD	53.22	55.56	50.89	53.86	53.93	0.56
	FA	48.89	64.22	33.56	48.44	49.43	0.47
	MD	55.22	62.00	48.44	55.27	56.95	0.60
	RD	57.44	68.44	46.44	57.43	57.83	0.59
	VBM	54.82	50.91	58.73	56.21	53.57	0.54

For each feature, accuracy metrics are reported. Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; AUC, area under the curve; BA, balance accuracy; DRCs, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity; NPV, negative predictive value; PPV, positive predictive value; RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed based connectivity, Sens, sensitivity; Spec, specificity; VBM, voxel-based morphometry.

Appendix 10. Performance metrics for the models identifying cognitively impaired patients at the coordination domain.

Feature	BA	Sens	Spec	PPV	NPV	AUC	p-value	FDRq-values
Combined	64.67	62.67	66.67	64.00	67.33	0.67	0.024	0.328
ReHo	57.86	70.00	45.71	56.13	62.24	0.58	0.084	0.573
fALFF	54.52	56.19	52.86	54.00	56.15	0.49	0.225	0.883
SBC	51.67	15.71	87.62	55.00	51.43	0.47	0.331	0.883
VBM	51.37	55.16	47.58	51.85	51.12	0.47	0.311	0.883
PRS	50.60	40.00	61.19	56.67	44.57	0.51	0.635	0.922
RD	50.58	61.03	40.13	50.01	52.16	0.49	0.388	0.883
PIMs	49.44	38.89	60.00	56.76	42.11	0.49	0.713	0.922
ABC	49.05	52.38	45.71	48.32	49.46	0.55	0.524	0.922
DRCs	47.62	42.86	52.38	47.65	47.78	0.50	0.644	0.922
FA	46.92	55.00	38.85	47.14	46.67	0.46	0.743	0.922
AD	41.79	48.21	35.38	43.39	39.00	0.42	0.987	1.036
MD	41.79	48.33	35.26	42.76	40.21	0.42	0.983	1.036

For each feature, accuracy metrics and p-value at the permutation test are reported. Models are sorted from the best to the worst model, considering BA. Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; AUC, area under the curve; BA, balance accuracy; DRCs, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity; NPV, negative predictive value; PPV, positive predictive value; PRS, polygenic risk score; RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed based connectivity; Sens, sensitivity, Spec, specificity; VBM, voxel-based morphometry.

Appendix 11. Performance metrics for the models identifying cognitively impaired patients at the executive functions domain.

Feature	BA	Sens	Spec	PPV	NPV	AUC	p-value	FDRq-values
ABC	59.52	75.24	43.81	57.04	65.33	0.72	0.088	0.400
RD	59.09	61.82	56.36	61.66	59.25	0.62	0.036	0.400
ReHo	59.05	51.90	66.19	65.83	66.80	0.64	0.060	0.400
FA	54.55	56.36	52.73	56.49	54.27	0.59	0.145	0.495
DRCs	50.24	46.67	53.81	58.00	49.80	0.56	0.416	0.766
SBC	50.24	71.90	28.57	49.68	53.17	0.46	0.413	0.766
PIMs	50.02	39.39	60.66	35.14	64.91	0.50	0.499	0.766
AD	49.09	56.36	41.82	49.54	48.53	0.55	0.531	0.766
MD	49.09	49.09	49.09	50.32	47.85	0.56	0.561	0.766
VBM	49.02	60.76	37.27	49.00	49.11	0.49	0.541	0.766
PRS	47.96	44.26	51.65	38.03	58.02	0.48	0.660	0.819
Combined	44.67	30.67	58.67	38.89	44.05	0.52	0.804	0.915
fALFF	40.48	49.52	31.43	39.56	39.88	0.45	0.957	1.005

For each feature, accuracy metrics and p-value at the permutation test are reported. Models are sorted from the best to the worst model, considering BA. Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; AUC, area under the curve; BA, balance accuracy; DRCs, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity; NPV, negative predictive value; PPV, positive predictive value; PRS, polygenic risk score; RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed based connectivity; Sens, sensitivity, Spec, specificity; VBM, voxel-based morphometry.

Appendix 12. Performance metrics for the models identifying cognitively impaired patients at the fluency domain.

Feature	BA	Sens	Spec	PPV	NPV	AUC	p-value	FDRq-values
DRCs	63.33	72.00	54.67	63.17	62.67	0.61	0.034	0.284
AD	60.00	56.00	64.00	61.52	59.02	0.60	0.042	0.284
SBC	58.00	54.67	61.33	65.33	63.67	0.62	0.116	0.284
RD	57.00	48.00	66.00	58.93	55.77	0.60	0.096	0.284
MD	56.00	46.00	66.00	58.03	55.29	0.64	0.125	0.284
VBM	54.45	58.73	50.18	52.83	56.66	0.54	0.172	0.335
ABC	53.67	52.00	55.33	54.57	52.86	0.59	0.262	0.447
PRS	50.82	21.43	80.21	38.71	63.64	0.51	0.077	0.284
PIMs	49.66	40.00	59.32	36.84	62.50	0.50	0.576	0.874
FA	48.00	42.00	54.00	50.03	44.81	0.48	0.647	0.883
ReHo	46.00	45.33	46.67	48.17	33.17	0.38	0.750	0.931
Combined	40.33	42.00	38.67	40.92	33.00	0.38	0.948	1.048
fALFF	30.67	34.67	26.67	32.14	20.00	0.32	0.998	1.048

For each feature, accuracy metrics and p-value at the permutation test are reported. Models are sorted from the best to the worst model, considering BA. Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; AUC, area under the curve; BA, balance accuracy; DRCs, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity; NPV, negative predictive value; PPV, positive predictive value; PRS, polygenic risk score; RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed based connectivity; Sens, sensitivity, Spec, specificity; VBM, voxel-based morphometry.

Appendix 13. Performance metrics for the models identifying cognitively impaired patients at the processing speed domain.

Feature	BA	Sens	Spec	PPV	NPV	AUC	p-value	FDRq-values
VBM	61.39	68.61	54.17	59.75	67.30	0.58	0.035	0.478
DRCs	58.00	64.00	52.00	58.13	67.50	0.50	0.111	0.758
Combined	54.50	38.00	71.00	46.67	56.67	0.54	0.258	0.928
MD	52.50	47.50	57.50	46.24	54.67	0.49	0.300	0.928
SBC	52.00	40.00	64.00	60.00	51.11	0.58	0.356	0.928
PRS	49.91	59.82	40.00	73.63	26.23	0.50	0.571	1.014
PIMs	49.86	79.71	20.00	73.33	26.32	0.50	0.408	0.928
RD	48.75	47.50	50.00	51.35	47.70	0.48	0.594	1.014
ABC	44.00	44.00	44.00	47.22	34.52	0.41	0.826	1.045
AD	41.25	50.00	32.50	43.00	37.38	0.39	0.967	1.045
fALFF	40.00	44.00	36.00	38.79	31.81	0.42	0.955	1.045
FA	40.00	50.00	30.00	40.62	38.67	0.40	0.977	1.045
ReHo	36.00	28.00	44.00	28.89	30.83	0.28	0.995	1.045

For each feature, accuracy metrics and p-value at the permutation test are reported. Models are sorted from the best to the worst model, considering BA. Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; AUC, area under the curve; BA, balance accuracy; DRCs, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity; NPV, negative predictive value; PPV, positive predictive value; PRS, polygenic risk score; RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed based connectivity; Sens, sensitivity, Spec, specificity; VBM, voxel-based morphometry.

Appendix 14. Performance metrics for the models identifying cognitively impaired patients at the verbal memory domain.

Feature	BA	Sens	Spec	PPV	NPV	AUC	p-value	FDRq-values
ReHo	59.52	65.71	53.33	56.41	65.56	0.60	0.048	0.573
MD	56.82	57.12	56.52	59.21	54.41	0.55	0.084	0.573
DRCs	54.29	52.86	55.71	54.29	54.29	0.58	0.205	0.700
AD	54.09	48.03	60.15	54.87	53.39	0.59	0.189	0.700
fALFF	51.9	56.67	47.14	43.27	63.33	0.49	0.336	0.764
SBC	51.9	45.24	58.57	49.93	54.21	0.54	0.291	0.764
PIMs	50.79	60.61	40.98	35.71	65.79	0.51	0.563	0.782
VBM	50.45	52.31	48.59	51.79	48.80	0.51	0.401	0.893
PRS	50.33	80.00	20.65	39.67	61.29	0.50	0.915	0.893
ABC	48.10	46.19	50.00	46.59	47.50	0.48	0.589	0.961
Combined	46.67	50.00	43.33	45.24	46.89	0.44	0.723	0.961
FA	45.68	41.36	50.00	42.11	46.44	0.41	0.842	0.961
RD	44.70	48.33	41.06	45.64	42.41	0.46	0.915	0.961

For each feature, accuracy metrics and p-value at the permutation test are reported. Models are sorted from the best to the worst model, considering BA. Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; AUC, area under the curve; BA, balance accuracy; DRCs, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity; NPV, negative predictive value; PPV, positive predictive value; PRS, polygenic risk score; RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed based connectivity; Sens, sensitivity, Spec, specificity; VBM, voxel-based morphometry.

Appendix 15. Performance metrics for the models identifying cognitively impaired patients at the working memory domain.

Feature	BA	Sens	Spec	PPV	NPV	AUC	p-value	FDRq-values
VBM	62.63	61.03	64.23	63.58	62.66	0.62	0.013	0.177
ABC	60.89	57.50	64.29	66.16	64.10	0.64	0.049	0.334
PRS	56.23	58.62	53.85	62.96	49.30	0.56	0.124	0.498
fALFF	53.93	60.36	47.50	53.00	57.99	0.50	0.227	0.498
PIMs	50.66	80.39	20.93	54.67	47.37	0.51	0.146	0.620
Combined	50.00	33.33	66.67	49.05	50.57	0.42	0.473	0.922
MD	49.77	45.30	54.24	48.81	50.34	0.50	0.460	0.922
DRCs	45.71	25.71	65.71	54.64	45.37	0.49	0.781	1.043
SBC	45.71	51.43	40.00	45.49	43.33	0.42	0.787	1.043
AD	45.61	48.64	42.58	44.00	46.68	0.49	0.840	1.043
RD	43.86	44.09	43.64	42.84	44.17	0.36	0.925	1.043
ReHo	43.57	70.71	16.43	44.09	39.58	0.44	0.892	1.043
FA	40.38	41.82	38.94	38.65	41.17	0.35	0.993	1.043

For each feature, accuracy metrics and p-value at the permutation test are reported. Models are sorted from the best to the worst model, considering BA. Abbreviations: ABC, atlas-based connectivity; AD, axial diffusivity; AUC, area under the curve; BA, balance accuracy; DRCs, dual regression components; FA, fractional anisotropy; fALFF, fractional amplitude of low-frequency fluctuations; MD, mean diffusivity; NPV, negative predictive value; PPV, positive predictive value; PRS, polygenic risk score; RD, radial diffusivity; ReHo, regional homogeneity; SBC, seed based connectivity; Sens, sensitivity, Spec, specificity; VBM, voxel-based morphometry.