DOI:
10.1111/bdi.13247

Document Version
Early version, also known as pre-print

Miskowiak Kamilla Woznica (Orcid ID: 0000-0003-2572-1384)
Yalin Nefize (Orcid ID: 0000-0003-4298-5074)
Burdick Katherine (Orcid ID: 0000-0003-4417-4988)
Bowie Christopher (Orcid ID: 0000-0002-1983-8861)
Carvalho Andre F (Orcid ID: 0000-0002-2500-5671)
Dols Annemiek (Orcid ID: 0000-0003-1964-0318)
Douglas Katie (Orcid ID: 0000-0002-5344-2959)
Kessing Lars (Orcid ID: 0000-0001-9377-9436)
Lafer Beny (Orcid ID: 0000-0002-6132-9999)
Lewandowski Kathryn Eve (Orcid ID: 0000-0003-1477-6187)
McIntyre Roger S (Orcid ID: 0000-0003-4733-2523)
Porter Richard (Orcid ID: 0000-0002-8695-3966)
Schaffer Ayal (Orcid ID: 0000-0001-6220-5042)
Vieta Eduard (Orcid ID: 0000-0003-1673-8248)
Stokes Paul (Orcid ID: 0000-0002-2274-8330)

Title: Can magnetic resonance imaging enhance the assessment of potential new treatments for cognitive impairment in mood disorders? A systematic review and position paper by the International Society for Bipolar Disorders Targeting Cognition Task Force

Running head: Neuroimaging to evaluate pro-cognitive treatments

Authors: Kamilla W. Miskowiak[1,2], Nefize Yalin[3], Ida Seeberg[1], Katherine E. Burdick[4,5], Vicent Balanzá-Martínez[6], Caterina del Mar Bonnin[7], Christopher R. Bowie[8], Andre F. Carvalho[9], Annemieke Dols[10], Katie Douglas[11], Peter Gallagher[12], Gregor Hasler[13], Lars V. Kessing[1,14], Beny Lafer[15], Kathryn E. Lewandowski[4,16], Carlos López-Jaramillo[17], Anabel Martinez-Aran[7], Roger S. McIntyre[18], Richard J. Porter[11], Scot E. Purdon[19], Ayal Schaffer[20], Tomiki Sumiyoshi[21], Ivan J. Torres[22], Tamsyn E. Van Rheenen[23,24], Lakshmi N. Yatham[22], Allan H. Young[3], Eduard Vieta[7], Paul R.A. Stokes[3]

[1]Copenhagen Affective disorder Research Centre (CADIC), Psychiatric Centre Copenhagen, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; [2]Department of Psychology, University of Copenhagen, Copenhagen, Denmark; [3]Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; [4]Department of Psychiatry, Harvard Medical School, Boston, MA, USA; [5]Department of Psychiatry, Brigham and Women's Hospital, Boston, MA, USA; [6]Teaching Unit of Psychiatry and Psychological Medicine, Department of Medicine, University of Valencia, CIBERSAM, Valencia, Spain.[7]Clinical Institute of Neuroscience, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Spain; [8]Department of Psychology, Queen's University, Kingston, Canada; [9]IMPACT Strategic Research Centre (Innovation in Mental and Physical Health and Clinical Treatment), Deakin University, Geelong, Vic., Australia; [10]Department of Old Age Psychiatry, GGZ in Geest, Amsterdam UMC, location VUmc, Amsterdam Neuroscience, Amsterdam Public Health research institute, Amsterdam, The Netherlands; [11]Department of Psychological Medicine, University of Otago, Christchurch, New Zealand; [12]Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle-upon-Tyne, UK; [13]Psychiatry Research Unit, University of Fribourg, Fribourg, Switzerland; [14]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark; [15]Bipolar Disorder Research Program, Institute of Psychiatry, Hospital das Clinicas, Faculdade de Medicina, Universidade de São Paulo, São Paulo, SP, Brazil; [16]McLean Hospital, Schizophrenia and Bipolar Disorder Program, Belmont, MA, US; [17]Research Group in Psychiatry, Department of Psychiatry, Universidad de Antioquia, Medellín, Colombia; [18]Mood Disorders Psychopharmacology Unit, Brain and Cognition Discovery Foundation, University of Toronto, Toronto, Canada; [19]Department of Psychiatry, University of Alberta, Edmonton, Canada; [20]Department of Psychiatry, University of Toronto; [21]Department of Preventive Intervention for Psychiatric Disorders, National Institute of Mental Health, National Center of

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/bdi.13247

Neurology and Psychiatry, Tokyo, Japan; [22]Department of Psychiatry, University of British Columbia, Vancouver, Canada; [20]Department of Psychiatry, University of British Columbia, Vancouver, Canada; [23]Melbourne Neuropsychiatry Centre, Department of Psychiatry, University of Melbourne, Carlton, Australia; [24]Centre for Mental Health, Faculty of Health, Arts and Design, Swinburne University, Australia.

Corresponding author**:** Kamilla Woznica Miskowiak, Neurocognition and Emotion in Mood disorder (NEAD) Group, CADIC, Psychiatric Centre Copenhagen, Copenhagen University Hospital, Blegdamsvej 9, DK-2100 Copenhagen, Tel: 0045 38647087, Email: Kamilla.miskowiak@regionh.dk

Word count: 6,172

Figures / tables: 1/5

Supplementary material: Search profile, PRISMA 2020 Checklist

ABSTRACT

Background: Developing treatments for cognitive impairment is key to improving the functioning of people with mood disorders. Neuroimaging may assist in identifying brain-based efficacy markers. This systematic review and position paper by the International Society for Bipolar Disorders Targeting Cognition Task Force examines the evidence from neuroimaging studies of pro-cognitive interventions.

Methods: We included magnetic resonance imaging (MRI) studies of candidate interventions in people with mood disorders or healthy individuals, following the procedures of the Preferred Reporting Items for Systematic reviews and Meta-Analysis 2020 statement. Searches were conducted on PubMed/MEDLINE, PsycInfo, EMBASE, Cochrane Library and Clinicaltrials.gov from inception to 30th April 2021. Two independent authors reviewed the studies using the National Heart, Lung, Blood Institutes of Health Quality Assessment Tool for Controlled Intervention Studies and quality of neuroimaging methodology assessment checklist.

Results: We identified 26 studies (N=702). Six investigated cognitive remediation or pharmacological treatments in mood disorders (N=190). In healthy individuals, 14 studies investigated pharmacological interventions (N=319), two cognitive training (N=73) and four neuromodulatory treatments (N=120). Methodologies were mostly rated as 'fair'. 77% studies investigated effects with task-based fMRI. Findings varied but most consistently involved treatment-associated cognitive control network (CCN) activity increases with cognitive improvements, or CCN activity decreases with no cognitive change, and increased functional connectivity. In mood disorders, treatment-related default mode network suppression occurred.

Conclusions: Modulation of CCN and DMN activity is a putative efficacy biomarker. Methodological recommendations are to pre-declare intended analyses and use task-based fMRI, paradigms probing the CCN, longitudinal assessments, mock scanning and out-of-scanner tests.

Funding: No funding

INTRODUCTION

Cognitive impairment in mood disorders has emerged as a new treatment priority to improve functioning and quality of life in people with bipolar disorders (BD) or major depressive disorder (MDD). However, very few treatment options demonstrate efficacy in treating persistent cognitive impairment during remitted phases of these disorders and currently there are no pharmacological agents approved by regulatory authorities for the treatment of cognitive deficits in people with mood disorders. The search for new treatments is impeded to some extent by our limited insight into the neurocircuitry characteristics underlying cognitive impairment in people with mood disorders and, hence, a lack of identified neurocircuitry targets for pro-cognitive interventions. Consequently, it is unclear whether candidate treatments may effectively correct neurocircuitry abnormalities underlying cognitive impairments.

Lack of neurocircuitry targets is a common methodological challenge in treatment development for central nervous system (CNS) disorders. In particular, high failure rates of about 85% for CNS drug trials are likely related to candidate compounds not effectively engaging key neural circuitry dysfunctions[1,2]. While CNS treatment development typically relies on animal models, the discovery of therapeutic-like effects of compounds in these models has turned out to be a poor predictor of efficacy in human populations[3]. The consequent abandonment of investment in this field by many large pharmaceutical companies, the so-called 'death of CNS drug development', has led to a large unmet clinical need for new CNS treatments[4], including pro-cognitive interventions.

The Food and Drug Administration (FDA) Critical Path Initiative -the national US strategy for transforming development, evaluation and manufacturing of FDA-regulated medical products- highlighted neuroimaging in human populations as a key tool to accelerate the screening and selection of new candidate CNS treatments[5]. Indeed, neuroimaging may provide a powerful new way of assessing the potential of candidate treatments for cognitive impairment in mood disorders. The International Society for Bipolar Disorders (ISBD) Targeting Cognition Task Force has previously emphasized the key role that neuroimaging assessments can have in pro-cognitive intervention trials to investigate the neural correlates of potential pro-cognitive efficacy of candidate treatments[6]. Specifically, assessing the effects of new treatments on the brain, for example using magnetic resonance imaging (MRI), may support the identification of changes in cognition-relevant function and structure, which could then inform the potential effectiveness of the proposed treatment.

In a recent systematic review of functional MRI (fMRI) studies in mood disorders[7], aberrant (mainly *hypo*-) blood-oxygen-level-dependent (BOLD) activity in the medial and dorsal prefrontal cortex (DPFC) and parietal cortex cognitive control network (CCN) and *hyper*-activity (i.e., failure to suppress) in the default mode network (DMN) were identified as the most consistent neural correlates of cognitive impairment across domains[7]. However, the findings regarding the *direction* of the aberrant activity and its exact locations varied across studies, which is likely due to the cognitive heterogeneity among people with mood disorders and methodological differences between studies[8-11]. In keeping with this, recent fMRI studies that compared neuronal activity between *cognitively impaired* (with global cognition scores ≥1 SD below the mean of healthy controls) and *cognitively normal* people with BD found that cognitive impairments were associated with working memory(WM)-related *hypo*-activity in the DPFC and parietal regions coupled with *hyper*-activity in the DMN[12,13]. At a structural level, cognitively impaired BD individuals were found to display lower cerebral white matter volume and greater DPFC volume or grey matter, respectively, than cognitively normal individuals and healthy participants[14,15]. In MDD, extensive white matter dysfunction has also been observed in globally cognitively impaired individuals relative to those with either selective impairments or normal cognition[16]. However, other structural studies in BD revealed no differences between impaired and normal neurocognitive subgroups[12,17], suggesting that task-based fMRI could be a more sensitive assay of cognitive impairments[12,18]. Together, the emerging findings highlight modulation of aberrant fMRI BOLD responses in the DPFC and DMN and, possibly, structural change in white matter and DPFC as putative neuronal targets for pro-cognitive interventions in mood disorders.

The aims of the present systematic review were twofold. The first aim was to examine the findings and evaluate the quality of evidence from functional and structural MRI studies of the neuronal effects of candidate pro-cognitive treatments in people with mood disorders in full or partial remission, as well as in healthy individuals for whom neuronal changes were not confounded by potential clinical symptoms and concomitant medication. This focus differs from that in a prior review[7], which delineated the neural underpinnings of cognitive impairments and examined the neural basis of *direct* or *indirect* cognitive improvement with pro-cognitive interventions or symptom reduction in remitted or symptomatic people with mood disorders, respectively.

The second aim was to provide consensus-based methodological guidance by the ISBD Targeting Cognition Task Force for neuroimaging assessments in pro-cognitive intervention trials and direction for putative neurocircuitry targets with the greatest potential to serve as 'surrogate' endpoints.

METHODS AND MATERIALS

The protocol of this systematic review was registered with PROSPERO (ID CRD42020218099) and the review process has been conducted in concordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Updated 2020 guidelines [19].

Study Identification

A comprehensive systematic search was performed on the PubMed/MEDLINE, PsycInfo, EMBASE and Cochrane Library databases as well as Clinicaltrials.gov from inception to 30th April 2021. The search profile included four elements "Mood disorder and or healthy volunteers", "Cognition", "Intervention" and "Neuroimaging" with each of their combinations and alternative keywords in the respective databases (please see S1 in Supplementary material).

Inclusion and Exclusion Criteria

Studies that aimed to evaluate the potential of new pro-cognitive treatments using MRI in healthy adult participants (aged between 18 and 65) and adult participants (aged between 18 and 65) with mood disorders who were on average in full or partial remission as diagnosed using International Classification of Disease (ICD) 10 or Diagnostic and Statistical Manual of Mental Disorders (DSM) 5 criteria, or earlier versions of ICD and DSM, were included in this systematic review. Studies were included if they were written in English, had cognitive function as the primary focus and, accordingly, included 'non-emotional' cognition tests as outcome measure, involved a control group, reported cognitive data acquired inside and/or outside the scanner, and had MRI assessments at baseline and after treatment completion. Randomized controlled trials, open-label trials or relevant experimental trials were eligible for inclusion. In contrast, references that were conference abstracts, case reports, meta-analyses, reviews, theses, or were not published and/or not peer reviewed were excluded. Also, ongoing electroconvulsive therapy in parallel with the treatment of interest was an exclusion criterion.

Study Selection and Data Extraction

In accordance with the inclusion and exclusion criteria, two authors (NY and IS) performed a primary title/abstract screening followed by a secondary full-text screening for potentially eligible articles in an independent manner. Reference lists of eligible studies and relevant review articles were also screened for additional studies. All searches were re-run prior to the final analysis and any further studies identified were retrieved for inclusion. Agreement between the two authors was high; 100% for primary screening and 87% for secondary screening. Discrepancies were discussed with the senior researchers (PRS, KWM) and consensus was reached in all cases.

The following information was extracted from the included studies: year of publication, participant details (including number, mean age, sex distribution, diagnosis, psychiatric comorbidities, and medication status), intervention details (dose, duration, frequency), cognitive parameters (cognitive tasks used in the study including inside or outside the scanner), MRI related parameters (MRI technique and type of measure, acquisition, pre-processing, statistical analysis), changes in MRI, cognitive task and mood scale measures as well as presence and frequency of adverse events that were reported in the findings. Data extraction was performed by one researcher (NY) and another researcher (IS) checked the extracted data.

Outcome Measures

Primary outcomes of the included studies were changes in MRI measures and cognitive performance in response to interventions compared with a placebo/ sham/ active control/ treatment as usual (TAU) control condition. MRI measures of effect were defined as the blood-oxygen-level dependent (BOLD) signal for fMRI, functional connectivity (FC), regional homogeneity, amplitude of low frequency fluctuations, clustering co-efficient, characteristic path length or node degree for resting state functional MRI (rsfMRI), gray/white matter volume, arterial spin labelling (ASL), cortical thickness or surface area for structural MRI (sMRI), mean/radial/axial diffusivity or fractional anisotropy for diffusion tensor imaging (DTI) and metabolite levels for magnetic resonance spectroscopy (MRS). Cognitive performance measures involved accuracy and response times on neurocognitive tests of psychomotor speed, attention, vigilance, learning and memory, and executive functions

including reasoning, problem solving or verbal fluency. Secondary outcome measures included the presence and frequency of psychiatric symptoms and change in mood scale scores associated with the intervention compared with control conditions.

Quality Assessment

Quality assessment was conducted with two different checklists, one of which focused on the general study methodology and the other on neuroimaging methodology. For general study methodology quality, the National Heart, Lung, Blood Institutes of Health (NHBIH) Quality Assessment Tool for Controlled Intervention Studies was used[20]. For the quality assessment of neuroimaging methodology, quality assessment tools used in previous meta-analyses[21-26] were modified and a new checklist was created (see S2 in Supplementary material). For both quality assessment tools, studies that fulfilled 75% or higher of the stated criteria were categorized as good-quality studies, whereas studies that fulfilled 50% to 75% of the stated criteria were categorized as fair quality, and below 50% were categorized as poor quality. Each study was assessed by one researcher (NY) using abovementioned checklists and the scores were semi-independently evaluated by another researcher (IS). Discrepancies in scores were discussed between two researchers and resolved by consensus in all cases. The PRISMA 2020 checklist was completed (Supplementary material).

RESULTS

The study identification and selection process in accordance with our inclusion and exclusion criteria is summarized in the PRISMA 2020 flowchart (Figure 1). The initial search identified 4073 references, from which 3980 remained after removing duplicates. Out of these, 3878 were excluded in primary title/abstract screening and 102 papers were found eligible for full-text screening. Secondary full-text screening resulted in 26 papers that met inclusion criteria and were included in the review.

Quality Assessment

Nine of 26 studies were classified as good-quality studies according to the Quality Assessment Tool for Neuroimaging Methodology in Individual Studies criteria[28-32,50-52,55] (see S3 in supplementary material and

individual summary tables, right columns), 14 were categorized as fair-quality studies[27,35-37,39,40,42,45-49,53,54], and three studies[38,41,44] were classified as poor-quality studies.

Regarding *general study* methodology, nine studies were classified as good quality according to NHBIH Quality Assessment Tool for Controlled Intervention Studies criteria [28-30,32,38-41,49,51] (see S4 in supplementary material), 17 studies were classified as fair quality and none as poor-quality studies (see S4 in supplementary material).

For the completeness of the review, findings of all studies are presented in brief, however with a greater focus on the good-quality studies.

Studies in mood disorders

Six of the 26 pro-cognitive studies were conducted in people with BD or MDD, with individual group sizes of 13-35 and an average age range of 30-43 years among groups[27-32] (see Table 1). The methodological quality of these studies was rated as good (scores ≥75%) for five studies and as fair (≥50%) for one study (Table 1). Five of the six pro-cognitive studies used task-based fMRI with the administration of n-back WM or picture encoding tasks within the scanner and also involved cognitive tests tapping into WM and executive functions or memory outside the scanner[27-29,31,32], while one structural MRI study involved administration of Rey Auditory Verbal Learning Test (RAVLT) outside the scanner[30]. Of these, four studies examined neuronal changes in response to pharmacological interventions. Three studies explored the neuronal effects of eight weekly intravenous high-dose erythropoietin (EPO; 40,000 IE) (n=30-35) vs. placebo (saline; n=26-34) administrations[29,30,33], while one study investigated the effects of two weeks of daily vortioxetine (20 mg; n=24 MDD and n=24 HC) vs. placebo (n=24 MDD and n=24 HC) treatment[32].The final two studies examined the neuronal changes in response to cognitive remediation (CR) interventions; one examined the effects of 12 weeks of weekly, group-based CR combined with computerized training between sessions (n=13) vs. TAU (n=14)[27] and the other study explored the *early* effects after two weeks of biweekly group-based Action-Based CR (ABCR) combined with daily computerized training (n=26) vs. weekly therapist-led conversation group meetings (n=19).

The EPO studies in people with mood disorders identified treatment-related improvements in n-back WM accuracy, picture recall, and RAVLT total recall compared to saline across BD and MDD groups, with no

differences in treatment effects between mood disorder groups[28-30]. These cognitive improvements were accompanied by, and correlated with, EPO-related *increases* in activity within the dorsolateral PFC (DLPFC), dorsomedial PFC (DMPFC), medial temporal and superior parietal regions[28,29], *reduced* activity in the hippocampus, a predefined region of the DMN[29], and increased volume in a subregion of the left hippocampus[28-30]. In contrast, vortioxetine treatment was accompanied by *reduced* activity in the DLPFC and occipito-parietal nodes of the WM network. Vortioxetine also reduced DMN activity in the hippocampus. These effects occurred in the absence of changes in WM performance across MDD and HC[32].

The study of weekly CR over 12 weeks revealed no significant changes in cognitive performance or fMRI BOLD activity during n-back WM or picture encoding[27]. In contrast, the study of ABCR revealed early increases in DLPFC and DMPFC activity after two weeks of treatment vs. control group sessions, which correlated with and predicted subsequent improvement in executive function measured with the One-Touch Stockings of Cambridge (OTS; CANTAB) after 10 weeks of treatment [31].

Studies in healthy individuals

Twenty of the identified studies were conducted in healthy individuals[35-55] with individual group sizes of n=4-47 and mean age ranges of 20-33 years (see Tables 2-4). The methodological quality was "fair" (scores ≥50%) in 13 of these studies, good (scores ≥75%) in four studies and poor (scores<50%) in three studies (Tables 2-4). Of these, 15 studies used task-based fMRI with the administration of varied paradigms including n-back WM tasks, a continuous performance task, a novelty detection task, a variable attentional control task, a psychomotor vigilance task, a delayed-match-to-sample task, a spatial paired associates learning task and a semantic fluency task [35-37,39-42,44,45,49-53,55]. Of the remaining five studies that only involved administration of cognitive tasks outside the scanner, one used ASL with prose recall, n-back WM and digit span tasks[49], one used resting state fMRI with an episodic memory task[46], one used MRS with an associative learning task[54], and two used real-time fMRI with attentional score scale, continuous performance test, digit span and letter memory tasks[47,48].

Fourteen studies investigated the effects of pharmacological compounds[35-38,40-42,44,49-54], two examined the effects of different cognitive training interventions[39,55] and four assessed neuromodulation techniques[45-48].

Pharmacological intervention studies in healthy individuals

The 14 pharmacological intervention studies in HC, of which only three were of good quality, are reported in Table 2. Four studies investigated compounds that mainly target the serotonin or norepinephrine systems[44,53,54], four explored treatments mainly targeting dopaminergic systems[35,42,50,52], one investigated a compound targeting glutamatergic neurotransmission[36], one examined a compound targeting the immune response[51] and four investigated compounds with other mechanisms of action[38,40,41,49].

Compounds targeting the serotonin and norepinephrine systems

Three studies investigated the effects of selective serotonin reuptake inhibitors (SSRIs) or serotonin and norepinephrine reuptake inhibitors (SNRIs)[44,53,54]. In one study, where participants received three weeks of escitalopram (10 mg/day) (n=16) or placebo (n=20), changes in MRS measures of glutamatergic turnover (Glx/tCr) and associate learning[54] were investigated. The study revealed escitalopram-related reduction in hippocampal Glx/tCr ratios in the absence of cognitive change[54]. Of the other two smaller studies, one found *increases* in medial frontal cortex activation alongside improvement in semantic fluency[44] with one week of the SNRI venlafaxine (75 mg/day) treatment (n=4, venlafaxine or placebo groups) and another found *decreased* WM-related activation in inferior frontal gyrus and *increases* in the thalamus, caudate and anterior cingulate activations with SSRI escitalopram (10 mg/day) treatment (n=10; cross-over study)[53].

A study of a single dose study of atomoxetine (60 mg), an SNRI used as a non-stimulant treatment of attention deficit hyperactivity disorder (ADHD) vs. placebo (n=19; cross-over study) found an acute *increase* in PFC, temporal and occipital response and in FC between anterior insula and fronto-parietal network during high-load WM[37]. While atomoxetine did not produce cognitive improvement, the treatment-related increase in insula-DLPFC connectivity during WM correlated with WM performance increase over time.

Interventions targeting the dopaminergic system

Two studies, both rated as good methodological quality, investigated the effects of one week of treatment with tolcapone (day 1: 300 mg/day; days 2-7: 600 mg/day), a selective chatechol-O-methyltransferase (COMT) inhibitor, that is thought to increase dopaminergic neurotransmission in the PFC and is approved for treatment of

Parkinson's disease[50,52]. The first study found that tolcapone vs. placebo (n=34; cross-over study) *reduced* WM-related activity in the DLPFC in the absence of change in WM performance during fMRI, and additionally improved executive function and WM test performance *outside* the scanner[52]. The other study found a similar activity *reduction* after tolcapone vs. placebo (n=20; cross-over study) within the dorsal cingulate cortex during an attention control task, in the absence of changes in cognitive performance levels[50].

A third study investigated the effects of PF-06412562 (3 mg or 15 mg twice daily; n=27 per group), a D1/D5 receptor agonist vs. placebo (n=22) for five to seven days[35] and found no changes in MRI or cognitive measures.

Finally, a single study investigated the acute effects of a single dose of a combination of the nonselective dopamine agonist L-dopa/carbidopa (100 mg/25 mg) and the D2-receptor antagonist haloperidol (2 mg) for indirect D1-receptor stimulation in comparison with placebo (n=12, cross-over study)[42]. The study revealed a treatment-related *decrease* in occipital and temporal responses and *increased* FC between DLPFC and salience network during WM, in the absence of changes in cognitive performance.

Interventions targeting the glutamatergic system

One study investigated the effect of *add-on* treatment with phenytoin, an antiepileptic medication that primarily acts at the glutamatergic synapse by inhibiting voltage-gated sodium channels, to hydrocortisone treatment (known to produce verbal memory deficits). This was compared with hydrocortisone alone, phenytoin alone and placebo (n=15, cross-over study)[36]. Phenytoin counteracted verbal memory decline associated with hydrocortisone but this was accompanied by no change in neuronal activity during a picture encoding with the combined treatment vs. hydrocortisone [36].

Interventions targeting the hypothalamic–pituitary–adrenal (HPA) axis

One study, rated as good methodological quality, investigated the effects of mifepristone, a glucocorticoid and progesterone receptor antagonist, on brain activity based on evidence for negative effects of hypercortisolaemia on cognitive functions and evidence for promising cognitive benefits of mifepristone in BD[56]. The study revealed that a single dose of mifepristone (600 mg) vs. placebo (n=20; cross-over study) decreased fusiform, angular and precuneal cortices activations, in the absence of change in cognitive performance[51].

Other pharmacological interventions

One study[40] investigated the effects of a single dose of methylene blue (280 mg), an inhibitor of nitric oxide synthase with putative effects on neuroplasticity, that has been found to improve hypotension and improve residual mood and anxiety symptoms in BD[58]. The study revealed that methylene blue (n=13) *increased* activity in PFC and occipito-parietal cortices during a short-term memory task and in the insula during a sustained attention task relative to placebo (n=13), in the absence of changes in cognitive performance[40].

A study of the acute effects of a single dose of (2.75 g) green tea extract vs. placebo (n=12; cross-over study) on task performance found increased FC between the right superior parietal lobule and right middle frontal gyrus, that was correlated with a trend-level increase in cognitive performance[49].

A third study, with a poor methodological quality rating, investigated the effects of a single dose of cannabidiol (600 mg)vs placebo (n=15; cross-over study) on cerebral blood flow measured by ASL and memory function based on emerging evidence that cannabidiol may improve aspects of memory function[57]. The study found that while cannabidiol did not improve memory performance, it did *increase* orbitofrontal cortex blood flow and that this correlated with memory improvement over time[38].

Finally, a study, with a poor methodological quality rating, investigated the effects of betahistine, a Meniere's disease medication that is a H3-receptor antagonist/H1-receptor agonist and is purported to improve cognitive function because it increases histamine neurotransmission. This study found no change in MRI or cognitive performance measures after betahistine vs. placebo treatment (n=16; cross-over study)[41].

Behavioral intervention studies

The two studies of cognitive training interventions are displayed in Table 3. One study found that practicing a WM task three times weekly over six weeks (n=15) relative to a no-practice condition (n=14) was associated with *increased* WM-specific activity in the ventrolateral PFC, and was accompanied by improvement in WM performance[39]. The other study investigated the effects of computerized training of processing speed over five days (n=23) in comparison to no intervention (n=21)[55]. This study revealed training-specific gray matter increase in the precentral gyrus and decrease in left superior temporal and occipital regions, increased neuronal activity in

the left perisylvian region during simple cognitive processes (but not WM), and greater resting state functional connectivity between the left perisylvian area and occipital regions, which accompanied performance improvements within psychomotor speed [55].

Neuromodulation and neurofeedback intervention studies

The neuromodulation studies, all rated as fair methodological quality, are displayed in Table 4. One study of transcranial magnetic stimulation (TMS) revealed significant improvements in memory retrieval with theta-burst stimulation vs. beta-frequency stimulation and sham (n=24; cross-over study)[46] but no change in functional connectivity in the hippocampal-cortical network was observed. However, memory improvements correlated with increase in hippocampal connectivity with posterior cingulate, mPFC and precuneus[46]. The other TMS study found that high frequency repetitive TMS vs. sham (n=39; cross-over study) reduced FC between right DLPFC and left hippocampus during WM performance (but not during resting state) in the absence of change in cognitive performance[45].

The remaining two studies used neurofeedback with real-time fMRI (rtfMRI), as pro-cognitive interventions. One study of the effects of two rtfMRI training vs. sham sessions seven days apart (n=15) vs. no training (n=15) showed that neurofeedback aiming to upregulate left DLPFC activity was accompanied by *increased* functional connectivity between the DMN, salience network (SN) and CCN in the absence of performance change[48]. Notably, the greater functional connectivity between SN-DMN, SN-CCN and within the CEN correlated with improvements in WM performance. The other study found that real-time neurofeedback delivered over five sessions (n=18) vs. sham (n=9) produced auditory cortex deactivation to noise stimulation in the absence of changes in cognitive performance on attention control tests[47].

DISCUSSION

This systematic review by the ISBD Targeting Cognition Task Force provides a comprehensive overview of the evidence for alterations in structural and functional MRI measures associated with pro-cognitive interventions in people with mood disorders and healthy individuals. We identified 26 studies, of which most investigated the effects of pharmacological interventions in healthy people, and only six investigated the effects of interventions

in people with BD or MDD in full or partial remission. The quality of the neuroimaging methodology and general methodology was rated as 'good' in approximately one-third of the studies; in five (83%) studies in mood disorders but only in four (20%) healthy participant studies, which were mostly rated as 'fair'. A common methodological limitation was small sample size, with only 11 studies (42%) including more than 20 participants per treatment arm.

Convergent treatment-related neurocircuitry changes across people with mood disorders and healthy individuals

Treatment-related changes in neural activity varied across the 20 task-based fMRI studies with respect to the direction and location of the signal change, likely because of differences between studies in fMRI paradigms, interventions, and participant characteristics. Six fMRI studies of either people with mood disorders[28,29,31], all rated as good methodological quality, *or* healthy individuals[39,44,55], rated as mixed quality, found that treatment-related brain activation *increases* in CCN regions, including the DPFC and occipito-parietal cortex, were accompanied by and/or correlated with *improved cognitive performance*. Five fMRI studies, of which four were rated as good quality, found treatment-related DPFC and occipito-parietal *activity reductions* that were accompanied by *no cognitive change* in cognitively normal people with MDD[32] *or* healthy individuals [50-53]. Finally, treatment-related FC increase was observed in five (25%) studies [37,42,48,49,55], of which four were rated as good quality. This emerging evidence for convergent treatment-related neurocircuitry changes across mood disorders and healthy individuals is interesting because of the generally good quality ratings of these studies.

In contrast, too few studies involved structural MRI, MRS and resting state fMRI to make any inferences regarding possible consistent patterns of change. Nevertheless, the structural hippocampal increase that accompanied EPO-related cognitive improvement[30] is consistent with observations from pro-cognitive interventions in other neuropsychiatric disorders[59], potentially through increase in neurogenesis. Further studies are warranted to investigate whether structural hippocampal changes are common structural imaging biomarkers of treatment effects on cognition in mood disorders. Notably such structural changes are more likely to be observed with longer-term treatment, for example at least several weeks of treatment, whereas treatment-related fMRI BOLD changes were observable already following acute treatment administration[37,42,49,51].

Treatment-related increase in cognitive capacity or neural efficiency – two sides of the same coin?

Insufficient statistical power could potentially explain the absence of performance change in the five studies reporting treatment-related reductions in task-related brain activations. Indeed, power calculations for previous cognition trials in mood disorders showed that group sizes of 22-26 participants are typically needed for detection of a differential cognitive change between groups with moderate effect sizes[60,61]. Nevertheless, only one of the five fMRI studies [53] involved a sample size smaller than 20, rendering this explanation unlikely. A more probable reason was that four of these studies used cross-over rather than parallel-group designs. Although within-group cross-over designs may be powerful in detecting change, a problem with this approach is that learning effects can occur with repeated cognitive testing, particularly between the first and second test sessions[62]. This could have masked *treatment-related* cognitive changes in cross-over studies, which involve twice as many cognitive assessments as parallel-group studies. Indeed, five of the six studies that showed treatment-related performance improvements used a parallel-group design.

Methodological differences notwithstanding, the identified fMRI studies, predominantly rated as good quality, provide emerging evidence for activity changes during task performance in the DPFC and associated CCN regions and, possibly, in the DMN in response to different interventions across mood disorders and healthy participants. These emerging findings are consistent with a recently proposed neurocircuitry-based biomarker model for pro-cognitive effects in mood disorders[7,63]. The model is based on the observation of a *bell-shaped dose-response association* between task difficulty and DPFC BOLD response and proposes that the dose-response curve in mood disorders and schizophrenia is *shifted towards the left,* such that peak BOLD response in the DPFC occurs at lower task difficulty levels than in healthy individuals [7,64]. The model posits that dorsal PFC *hypo*-activity may be a marker of lower *cognitive capacity* (i.e., impaired performance), while *hyper*-activity may be a marker of less *cortical efficiency* (i.e., need for more PFC resources to maintain task performance). Building on this, pro-cognitive interventions are proposed to shift the dose-response curve towards normality. This provides testable hypotheses regarding the *direction* of treatment-related neurocircuitry change, depending on whether cognitive performance is improved by the intervention or remains stable. Specifically, key next-step hypotheses to test are that biomarkers for pro-cognitive effects involve either: (i) reversal of pre-treatment DPFC *hypo-activity* in people who show treatment-related performance improvement, or (ii) reduction of pre-treatment DPFC *hyper-activity* in people who show no performance change (i.e., increased neural efficiency).

Interestingly, two meta-analyses of neural activity changes in response to CR in people with schizophrenia found that *increase* in task-related DPFC activity was the most reliable marker of cognitive improvement[65,66]. In contrast, less consistent CR related activity change was found in other cognition-relevant regions[65,66]. Together, these findings provide preliminary evidence for modulation of DPFC –and possibly DMN– activity as a putative transdiagnostic biomarker model for pro-cognitive effects, which deserves further investigation.

Challenges in using fMRI to assess potential pro-cognitive effects

Although fMRI shows real promise to better assess the potential of pro-cognitive interventions, there are limitations of using fMRI in this way. First, fMRI task test-retest reliability varies within individuals and across brain regions studied[67,68]. For example, the n-back WM task shows fair-to-good within-participant reliability depending on which brain region is examined (ICCs=0.44-0.57)[68], whereas the reliability of an episodic memory recall task is worse (ICC=0.36)[69]. The reliability of task-related brain activations needs to be considered when interpreting the findings of cognitive intervention fMRI studies. This is because large intra-individual variations may limit the statistical power for detection of a treatment-related effect on the fMRI signal, particularly when using a repeated-measures design. fMRI BOLD response also only provides an *indirect* measure of neuronal activity and the understanding of its biological basis is still emerging[70]. Nevertheless, the neurocircuitry biomarker model is at a *systems level* in the brain. As such, fMRI can provide insights into the final common downstream effects on brain function of different behavioural and biological pro-cognitive interventions with distinct molecular and cellular mechanisms. A limitation of the proposed biomarker model discussed above is that it is only suited for examining neurocircuitry target engagement by interventions purported to improve aspects of cognition that involve executive control, WM and/or sustained attention. If the targeted cognitive domain is episodic memory or emotional cognition for example, treatment-related changes would be expected in other neurocircuitries. Further, the interpretation of treatment-related change in fMRI activations will also depend on the specific brain regions interrogated.

Methodological recommendations for neuroimaging in intervention trials

It is noteworthy that only six MRI studies of pro-cognitive interventions involved people with mood disorders. In contrast, a brief search indicates that more than 30 MRI studies have investigated the neuronal correlates of pro-cognitive (predominantly CR) interventions in people with schizophrenia[65]. While observed treatment-related changes in neurocircuitry activity were comparable across these diagnostic groups, the discrepancy in the numbers of studies highlights the need for further studies to gain deeper insight into the neurocircuitry biomarkers of pro-cognitive effects in mood disorders. The MRI design of such studies may be optimized in the following ways (see also Table 5);

First, based on the identified studies in this review, we suggest that the priority is given to task-related fMRI assessments given the fMRI BOLD response is a sensitive assay of treatment effects that is measurable before observable behavioral improvements emerge (e.g., [34]) and that task-related fMRI provides a neural substrate of how the brain works while performing cognitive tests of interest (i.e., in a controlled context). Because of the scarcity of resting state and structural MRI interventions studies, further structural imaging and resting state imaging is needed to clarify the value of these techniques as biomarkers for treatments targeting cognition.

Second, fMRI paradigms that engage the CCN –the strongest determinant of fluid intelligence[71]– seem best suited for examining neural activity changes to treatments that target aspects of cognition involving elements of executive control, WM and/or sustained attention. This includes n-back WM tasks, which not only activate the CCN but also show aberrant CCN and DMN activity in mood disorders[7,12,13] and treatment-related activity change[28,29,31,32,72]. Importantly, the task should fit the cognitive deficit being targeted; hence, if the targeted domain is episodic memory or emotional cognition, tasks that probe these aspects of cognition and engage the associated temporal, limbic and ventral PFC neurocircuitries would be best suited.

Third, given the choices of fMRI analysis packages and regions of interest (ROIs), we suggest that researchers specify in a published study protocol before study commencement that they will use standardized fMRI analysis methods using task-relevant ROIs that produced robust activations in previous studies (e.g. CCN regions including dPFC for n-back WM). This can reduce the risk of selective reporting and increase comparability between studies.

A fourth recommendation is to *ideally* conduct longitudinal fMRI before and either after treatment/control or, alternatively, at *an early point* in treatment, if the purpose is to investigate whether change

in brain activation is instrumental for subsequent cognitive improvements. However, given the high costs of fMRI, an alternative approach for studies with a randomized controlled design could be to conduct a *single fMRI assessment* at treatment completion or early in treatment. This would enable comparison of matched intervention groups, for whom any differences in neuronal activity would be presumed to reflect the intervention.

Fifth, based on our experience of fMRI research, we recommend that participants are acclimatized to the scanner environment in a mock scanner. This can help participants familiarize themselves with the MRI environment before their real scanning session and will likely reduce variability in the fMRI BOLD response due to effects of anxiety and stress. Commercial mock scanners may involve a significant cost and a less costly alternative could be to support participants by familiarizing themselves with lying in an MRI scanner environment by using immersive virtual reality including replicating the noise of the scanner.

Sixth, when practically possible, the inclusion of a pre-study cognitive test session is recommended in cross-over studies to minimize learning effects during the active trial participation. This may improve the signal-to-noise ratio and, thereby, potentially statistical power for detection of treatment-related behavioral change.

Seventh, we recommend inclusion of additional neurocognitive tests *outside* the scanner. This is because fMRI paradigms are generally optimized for the detection of differences in brain activation but not behavioral performance[73], with the latter being particularly important outcomes given their strong association with daily functioning. Additional more difficult out-of-scanner tests will ensure optimal sensitivity to treatment-related change in cognitive performance. Analysis of the associations between changes in task-related brain activation and out-of-scanner cognitive performance will aid interpretations of the observed changes in neuronal activity and provide insight into the *functional relevance* of treatment effects on neurocircuitry activity.

Finally, and more broadly, we endorse fMRI intervention studies in healthy participants to provide proof-of-concept without any confounding effects of mood symptoms and medication before evaluating candidate interventions in individuals with mood disorders. While the effects on neurocircuitry activity in healthy individuals may differ from that in mood disorders due to different cognitive performance levels (so called ceiling effects), fMRI studies in healthy populations may still provide insights into which task-relevant brain networks are key treatment targets that warrant further investigation in people with mood disorders. Based on this, studies including fully or partially remitted (rather than symptomatic) people with mood disorders are recommended to minimize pseudospecificity and confounding effects of clinical symptoms, in line with our previous

recommendation for pro-cognitive interventions trials in mood disorders[6]. Studies may also consider narrowing phenotypes by including pre-screening for cognitive impairment in people with mood disorders, as previously recommended[6], and attending to illness stages (e.g., focusing on either early or later stages). Ultimately, such attempts to reduce variance in the characteristics of participants may further improve the signal-to-noise ratio in these studies.

Limitations

A limitation of the review was that it involved no quantitative analysis of fMRI related to brain activations or changes in structural MRI measures in response to the candidate interventions. The rationale for this *a priori* decision was the differences in treatment modalities and schedules (acute and long-term treatments), participant groups (mood disorders and healthy individuals), MRI techniques (structural, functional, MRS and ASL) and, for the fMRI studies, differences in task-based paradigms and consistent reporting of quantitative comparable fMRI measures. Additionally, even for fMRI-based studies, effect sizes are rarely reported, making quantitative comparisons very challenging. A strength of our review is that it provides a comprehensive overview of the state-of-the art findings from MRI studies of putative pro-cognitive interventions in mood disorders and healthy individuals as well as evaluations of the quality of the evidence. This has enabled evaluation of a putative neurocircuitry-based biomarker model for pro-cognitive effects and provided a basis for methodological recommendations to guide future work.

Conclusions

In conclusion, there is a pressing need for more effective treatments targeting persistent cognitive impairment in mood disorders, but preclinical biomarker models provide poor predictive validity of efficacy in humans. In this systematic review by the ISBD Targeting Cognition Task, we identified 26 MRI investigations of neuronal target engagement by behavioral, pharmacological and neuromodulatory interventions in people with mood disorders and healthy individuals. Treatment-related change in task-based fMRI BOLD activity was investigated in most studies, while studies using resting state fMRI or structural MRI were scarce. Treatment-related activity *increases* in DPFC and associated regions were accompanied by and/or correlated with cognitive improvements (indicating

enhanced cognitive capacity), while activity *reduction* in regions of this network was observed in the absence of performance changes (suggesting greater neural efficiency). Another common finding was the increase in FC within the task relevant networks. Together, the findings provide emerging evidence to support the testable hypotheses that modulation of dorsal PFC and DMN activity may be a neurocircuitry-based biomarker model for pro-cognitive effects. Nevertheless, findings were not uniform and the putative neurocircuitry-based biomarkers for pro-cognitive effects should therefore be regarded as preliminary. The quality of the neuroimaging – and general methodology of studies was mostly rated as 'fair', suggesting a need for methodological guidance. Based on this, methodological recommendations were proposed for future pro-cognitive trials with fMRI outcome measures (summarized in Table 5). Following these suggestions by the ISBD Targeting Cognition Task Force may aid detection of neurocircuitry target engagement. Further studies are warranted to examine whether the identified MRI, and particularly fMRI, measures constitute biomarkers that can inform go no-go decisions in future treatment development strategies.

Acknowledgements

The authors thank the International Society for Bipolar Disorders executives and staff for their support of the ISBD Targeting Cognition Task Force. This work is non-funded. KWM holds a five-year Lundbeck Foundation Fellowship (grant no. R215-2015-4121). TVR was supported by an NHMRC Early Career Fellowship (GNT1088785) and a Dame Kate Campbell Fellowship from the University of Melbourne. KD holds a Sir Charles Hercus Health Research Fellowship from the Health Research Council of New Zealand (grant no. 19/082). EV thanks the support of the Spanish Ministry of Science and Innovation (PI15/00283, PI18/00805) integrated into the Plan Nacional de I+D+I and co-financed by the ISCIII-Subdirección General de Evaluación and the Fondo Europeo de Desarrollo Regional (FEDER); the Instituto de Salud Carlos III; the CIBER of Mental Health (CIBERSAM); the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement (2017 SGR 1365), the CERCA Programme, and the Departament de Salut de la Generalitat de Catalunya for the PERIS grant SLT006/17/00357.

Competing Interests

KWM has received consultancy fees from Janssen and Lundbeck in the past three years. RP uses computer software at no cost for research – provided by SBT-pro and has received support for travel to educational meetings from Servier and Lundbeck. AS has received advisory or speaking fees from AbbVie, Janssen, Lundbeck, Otsuka, and Sunovion in the past three years. KD uses computer software at no cost for research provided by SBT-pro. LVK has received consultancy fees from Lundbeck and Teva in the past three years. EV has received grants and served as consultant, advisor or CME speaker for the following entities: AB-Biotics, Abbott, Allergan, Angelini, Dainippon Sumitomo Pharma, Ferrer, Gedeon Richter, Janssen, Lundbeck, Otsuka, Sage, Sanofi-Aventis, Sunovion, Takeda, all them unrelated to the present work. VBM has been a consultant, advisor or Continuing Medical Education (CME) speaker over the last three years for the following companies: Angelini, Lundbeck, Nutrición Médica, and Otsuka. AY has conducted paid lectures and advisory boards for Allergan, AstraZeneca, Bionomics, BrainCells Inc., Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Janssen, Lundbeck, Novartis, Otsuka Pharmaceutical Co., Pharmaceutica, Pfizer, Roche, Sanofi-Aventis, Servier Laboratories, Sunovion and Wyeth. He was lead Investigator for the EMBOLDEN Study (AstraZeneca), BCI Neuroplasticity Study and Aripiprazole Mania Study, and has been involved in investigator-initiated studies for AstraZeneca, Eli Lilly and Wyeth. IJT has received has served as consultant for Lundbeck Canada, Sumitomo Dainippon and Community Living British Columbia. LNY has been on speaker/advisory boards for, or has received research grants from Alkermes, Allergan, AstraZeneca, Bristol Myers Squibb, CANMAT, CIHR, Dainippon Sumitomo Pharma, GSK, Janssen, Lilly, Lundbeck, Merck, Otsuka, Pfizer, Sanofi, Sunovion and Teva. RM has received personal fees from Lundbeck, Janssen, Purdue, Pfizer, Otsuka, Allergan, Takeda, Neurocrine, Sunovion, Minerva, Intra-Cellular, Abbvie, and Eisai and is a shareholder in the 420 Company and CEO of Champignon. AD, AC, BL, IS, MBJ, CRB, CMB, KEL, PG, CLJ, AMA, SEP, TS and TVR report no conflict of interest. PRS reports non-financial support from Janssen Research and Development LLC, personal fees

and non-financial support from Frontiers in Psychiatry, personal fees from Allergan, a grant from H Lundbeck, grants and non-financial support from Corcept Therapeutics, outside the submitted work.

Data availability statement: Not applicable

REFERENCES

1. Borsook D, Becerra L, Fava M. Use of functional imaging across clinical phases in CNS drug development. *Transl Psychiatry.* 2013;3:e282.
2. Nathan PJ, Phan KL, Harmer CJ, Mehta MA, Bullmore ET. Increasing pharmacological knowledge about human neurological and psychiatric disorders through functional neuroimaging and its application in drug discovery. *Curr Opin Pharmacol.* 2014;14:54-61.
3. Millan MJ, Agid Y, Brune M, et al. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. *Nat Rev Drug Discov.* 2012;11(2):141-168.
4. Cutler R. Worldwide Clinical Trials. *http://www wwctrials com.* 2011. http://www.wwctrials.com.
5. FDA Critical Path Initiative. *FDA Critical Path Initiative* 2017; http://www.fda.gov/oc/initiatives/criticalpath/.
6. Miskowiak KW, Burdick KE, Martinez-Aran A, et al. Methodological recommendations for cognition trials in bipolar disorder by the International Society for Bipolar Disorders Targeting Cognition Task Force. *Bipolar Disord.* 2017;19(8):614-626.
7. Miskowiak KW, Petersen CS. Neuronal underpinnings of cognitive impairment and -improvement in mood disorders. *CNS Spectr.* 2019;24(1):30-53.
8. Burdick KE, Russo M, Frangou S, et al. Empirical evidence for discrete neurocognitive subgroups in bipolar disorder: clinical implications. *Psychol Med.* 2014;44(14):3083-3096.
9. Jensen JH, Knorr U, Vinberg M, Kessing LV, Miskowiak KW. Discrete neurocognitive subgroups in fully or partially remitted bipolar disorder: Associations with functional abilities. *J Affect Disord.* 2016;205:378-386.
10. Pu S, Noda T, Setoyama S, Nakagome K. Empirical evidence for discrete neurocognitive subgroups in patients with non-psychotic major depressive disorder: clinical implications. *Psychol Med.* 2018;48(16):2717-2729.
11. Van Rheenen TE, Lewandowski KE, Tan EJ, et al. Characterizing cognitive heterogeneity on the schizophrenia-bipolar disorder spectrum. *Psychol Med.* 2017;47(10):1848-1864.
12. Alonso-Lana S, Goikolea JM, Bonnin CM, et al. Structural and Functional Brain Correlates of Cognitive Impairment in Euthymic Patients with Bipolar Disorder. *PLoS One.* 2016;11(7):e0158867.
13. Zarp Petersen J, Varo C, Skovsen CF, et al. Neuronal underpinnings of cognitive impairment in bipolar disorder: A large data-driven functional magnetic resonance imaging study. *Bipolar Disord.* 2021.
14. Chakrabarty T, Torres IJ, Su WW, Sawatzky R, Keramatian K, Yatham LN. Cognitive subgroups in first episode bipolar I disorder: Relation to clinical and brain volumetric variables. *Acta Psychiatr Scand.* 2021;143(2):151-161.
15. Macoveanu J, Freeman KO, Kjaerstad HL, Knudsen GM, Kessing LV, Miskowiak KW. Structural brain abnormalities associated with cognitive impairments in bipolar disorder. *Acta Psychiatr Scand.* 2021.
16. Liang S, Wang Q, Kong X, et al. White Matter Abnormalities in Major Depression Biotypes Identified by Diffusion Tensor Imaging. *Neurosci Bull.* 2019;35(5):867-876.

17. Karantonis JA, Rossell SL, Carruthers SP, et al. Brain morphology does not clearly map to cognition in individuals on the bipolar-schizophrenia-spectrum: a cross-diagnostic study of cognitive subgroups. *J Affect Disord.* 2021;281:776-785.

18. Karantonis JA, Carruthers SP, Rossell SL, et al. A Systematic Review of Cognition-Brain Morphology Relationships on the Schizophrenia-Bipolar Disorder Spectrum. *Schizophr Bull.* 2021.

19. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev.* 2015;4:1.

20. NIH National Heart Lung and Blood Institute. https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools.

21. Poldrack RA, Fletcher PC, Henson RN, Worsley KJ, Brett M, Nichols TE. Guidelines for reporting an fMRI study. *Neuroimage.* 2008;40(2):409-414.

22. Iwabuchi SJ, Krishnadas R, Li C, Auer DP, Radua J, Palaniyappan L. Localized connectivity in depression: a meta-analysis of resting state functional imaging studies. *Neurosci Biobehav Rev.* 2015;51:77-86.

23. Laura G, Silvia T, Nikolaos P, Patrizia P. The Role of fMRI in the Assessment of Neuroplasticity in MS: A Systematic Review. *Neural plasticity.* 2018;2018:3419871.

24. Das TK, Javadzadeh A, Dey A, et al. Antioxidant defense in schizophrenia and bipolar disorder: A meta-analysis of MRS studies of anterior cingulate glutathione. *Progress in neuro-psychopharmacology & biological psychiatry.* 2019;91:94-102.

25. Wang W, Zhao Y, Hu X, et al. Conjoint and dissociated structural and functional abnormalities in first-episode drug-naive patients with major depressive disorder: a multimodal meta-analysis. *Sci Rep.* 2017;7(1):10401.

26. Jiang J, Zhao YJ, Hu XY, et al. Microstructural brain abnormalities in medication-free patients with major depressive disorder: a systematic review and meta-analysis of diffusion tensor imaging. *J Psychiatry Neurosci.* 2017;42(3):150-163.

27. Macoveanu J, Demant KM, Vinberg M, Siebner HR, Kessing LV, Miskowiak KW. Towards a biomarker model for cognitive improvement: No change in memory-related prefrontal engagement following a negative cognitive remediation trial in bipolar disorder. *J Psychopharmacol.* 2018;32(10):1075-1085.

28. Miskowiak KW, Macoveanu J, Vinberg M, et al. Effects of erythropoietin on memory-relevant neurocircuitry activity and recall in mood disorders. *Acta Psychiatr Scand.* 2016;134(3):249-259.

29. Miskowiak KW, Vinberg M, Glerup L, et al. Neural correlates of improved executive function following erythropoietin treatment in mood disorders. *Psychol Med.* 2016;46(8):1679-1691.

30. Miskowiak KW, Vinberg M, Macoveanu J, et al. Effects of Erythropoietin on Hippocampal Volume and Memory in Mood Disorders. *Biol Psychiatry.* 2015;78(4):270-277.

31. Ott CV, Macoveanu J, Bowie CR, et al. Change in prefrontal activity and executive functions after action-based cognitive remediation in bipolar disorder: a randomized controlled trial. *Neuropsychopharmacology.* 2020.

32. Smith J, Browning M, Conen S, et al. Vortioxetine reduces BOLD signal during performance of the N-back working memory task: a randomised neuroimaging trial in remitted depressed patients and healthy controls. *Mol Psychiatry.* 2018;23(5):1127-1133.

33. Miskowiak KW, Macoveanu J, Vinberg M, et al. Effects of erythropoietin on memory-relevant neurocircuitry activity and recall in mood disorders. *Acta Psychiatrica Scandinavica.* 2016;134(3):249-259.

34. Ott CV, Macoveanu J, Bowie CR, et al. Change in prefrontal activity and executive functions after action-based cognitive remediation in bipolar disorder: a randomized controlled trial. *Neuropsychopharmacology.* 2021;46(6):1113-1121.

35. Balice-Gordon R, Honey GD, Chatham C, et al. A Neurofunctional Domains Approach to Evaluate D1/D5 Dopamine Receptor Partial Agonism on Cognition and Motivation in Healthy Volunteers With Low Working Memory Capacity. *The international journal of neuropsychopharmacology.* 2020;23(5):287-299.

36. Brown ES, Lu H, Denniston D, et al. A randomized, placebo-controlled proof-of-concept, crossover trial of phenytoin for hydrocortisone-induced declarative memory changes. *J Affect Disord.* 2013;150(2):551-558.

37. Hernaus D, Casales Santa MM, Offermann JS, Van Amelsvoort T. Noradrenaline transporter blockade increases fronto-parietal functional connectivity relevant for working memory. *Eur Neuropsychopharmacol.* 2017;27(4):399-410.

38. Bloomfield MAP, Green SF, Hindocha C, et al. The effects of acute cannabidiol on cerebral blood flow and its relationship to memory: An arterial spin labelling magnetic resonance imaging study. *J Psychopharmacol.* 2020;34(9):981-989.

39. Jolles DD, Grol MJ, Van Buchem MA, Rombouts SA, Crone EA. Practice effects in the brain: Changes in cerebral activation after working memory practice depend on task demands. *Neuroimage.* 2010;52(2):658-668.

40. Rodriguez P, Zhou W, Barrett DW, et al. Multimodal Randomized Functional MR Imaging of the Effects of Methylene Blue in the Human Brain. *Radiology.* 2016;281(2):516-526.

41. van Ruitenbeek P, Mehta MA. Potential enhancing effects of histamine H(1) agonism/H(3) antagonism on working memory assessed by performance and bold response in healthy volunteers. *Br J Pharmacol.* 2013;170(1):144-155.

42. van Ruitenbeek P, Hernaus D, Mehta MA. A proof-of-principle study of the effect of combined haloperidol and levodopa administration on working memory-related brain activation in humans. *Human psychopharmacology.* 2018;33(5):e2675.

43. Takeuchi H, Taki Y, Sassa Y, et al. The Associations between Regional Gray Matter Structural Changes and Changes of Cognitive Performance in Control Groups of Intervention Studies. *Frontiers in human neuroscience.* 2015;9:681.

44. Xie Q, Liu Y, Li CY, et al. The modulation of venlafaxine on cortical activation of language area in healthy subjects with fMRI study. *Psychopharmacology (Berl).* 2012;223(4):417-425.

45. Bilek E, Schafer A, Ochs E, et al. Application of high-frequency repetitive transcranial magnetic stimulation to the DLPFC alters human prefrontal-hippocampal functional interaction. *J Neurosci.* 2013;33(16):7050-7056.

46. Hermiller MS, VanHaerents S, Raij T, Voss JL. Frequency-specific noninvasive modulation of memory retrieval and its relationship with hippocampal network connectivity. *Hippocampus.* 2019;29(7):595-609.

47. Sherwood MS, Parker JG, Diller EE, et al. Self-directed down-regulation of auditory cortex activity mediated by real-time fMRI neurofeedback augments attentional processes, resting cerebral perfusion, and auditory activation. *Neuroimage.* 2019;195:475-489.

48. Zhang Q, Zhang G, Yao L, Zhao X. Impact of real-time fMRI working memory feedback training on the interactions between three core brain networks. *Front Behav Neurosci.* 2015;9:244.

49. Schmidt A, Hammann F, Wolnerhanssen B, et al. Green tea extract enhances parieto-frontal connectivity during working memory processing. *Psychopharmacology (Berl).* 2014;231(19):3879-3888.

50. Magalona SC, Rasetti R, Chen J, et al. Effect of tolcapone on brain activity during a variable attentional control task: a double-blind, placebo-controlled, counter-balanced trial in healthy volunteers. *CNS Drugs.* 2013;27(8):663-673.

51. Yalin N, Kempton MJ, Mazibuko N, Mehta MA, Young AH, Stokes PR. Mifepristone enhances the neural efficiency of human visuospatial memory encoding and recall. *Psychoneuroendocrinology.* 2020;125:105116.

52. Apud JA, Mattay V, Chen J, et al. Tolcapone improves cognition and cortical information processing in normal human subjects. *Neuropsychopharmacology.* 2007;32(5):1011-1020.

53. Rose EJ, Simonotto E, Spencer EP, Ebmeier KP. The effects of escitalopram on working memory and brain activity in healthy adults during performance of the n-back task. *Psychopharmacology (Berl).* 2006;185(3):339-347.

54. Spurny B, Vanicek T, Seiger R, et al. Effects of SSRI treatment on GABA and glutamate levels in an associative relearning paradigm. *Neuroimage.* 2021;232:117913.

55. Takeuchi H, Taki Y, Hashizume H, et al. Effects of training of processing speed on neural systems. *J Neurosci.* 2011;31(34):12139-12148.

56. Watson S, Gallagher P, Porter RJ, et al. A randomized trial to examine the effect of mifepristone on neuropsychological performance and mood in patients with bipolar depression. *Biol Psychiatry.* 2012;72(11):943-949.

57. Batalla A, Bos J, Postma A, Bossong MG. The Impact of Cannabidiol on Human Brain Function: A Systematic Review. *Front Pharmacol.* 2020;11:618184.

58. Alda M, McKinnon M, Blagdon R, et al. Methylene blue treatment for residual symptoms of bipolar disorder: randomised crossover study. *Br J Psychiatry.* 2017;210(1):54-60.

59. Ott CV, Johnson CB, Macoveanu J, Miskowiak K. Structural changes in the hippocampus as a biomarker for cognitive improvements in neuropsychiatric disorders: A systematic review. *Eur Neuropsychopharmacol.* 2019;29(3):319-329.

60. Miskowiak KW, Ehrenreich H, Christensen EM, Kessing LV, Vinberg M. Recombinant human erythropoietin to target cognitive dysfunction in bipolar disorder: a double-blind, randomized, placebo-controlled phase 2 trial. *J Clin Psychiatry.* 2014;75(12):1347-1355.

61. Ott CV, Vinberg M, Kessing LV, Bowie CR, Forman JL, Miskowiak KW. Effect of Action-Based Cognitive Remediation on cognitive impairment in patients with remitted bipolar disorder: A randomized controlled trial. *Bipolar Disord.* 2020.

62. Bartels C, Wegrzyn M, Wiedl A, Ackermann V, Ehrenreich H. Practice effects in healthy adults: a longitudinal study on frequent repetitive cognitive testing. *BMC Neurosci.* 2010;11:118.

63. Petersen CS, Miskowiak KW. Toward a transdiagnostic neurocircuitry-based biomarker model for pro-cognitive effects: challenges, opportunities, and next steps. *CNS Spectr.* 2020:1-5.

64. Van Snellenberg JX, Torres IJ, Thornton AE. Functional neuroimaging of working memory in schizophrenia: task performance as a moderating variable. *Neuropsychology.* 2006;20(5):497-510.

65. Mothersill D, Donohoe G. Neural Effects of Cognitive Training in Schizophrenia: A Systematic Review and Activation Likelihood Estimation Meta-analysis. *Biol Psychiatry Cogn Neurosci Neuroimaging.* 2019;4(8):688-696.

66. Ramsay IS, MacDonald AW, III. Brain Correlates of Cognitive Remediation in Schizophrenia: Activation Likelihood Analysis Shows Preliminary Evidence of Neural Target Engagement. *Schizophr Bull.* 2015.

67. Zandbelt BB, Gladwin TE, Raemaekers M, et al. Within-subject variation in BOLD-fMRI signal changes across repeated measurements: quantification and implications for sample size. *Neuroimage.* 2008;42(1):196-206.

68. Plichta MM, Schwarz AJ, Grimm O, et al. Test-retest reliability of evoked BOLD signals from a cognitive-emotive fMRI test battery. *Neuroimage.* 2012;60(3):1746-1758.

69. Holiga S, Sambataro F, Luzy C, et al. Test-retest reliability of task-based and resting-state blood oxygen level dependence and cerebral blood flow measures. *PLoS One.* 2018;13(11):e0206583.

70. Howarth C, Mishra A, Hall CN. More than just summed neuronal activity: how multiple cell types shape the BOLD response. *Philosophical transactions of the Royal Society of London Series B, Biological sciences.* 2021;376(1815):20190630.

71. Barbey AK. Network Neuroscience Theory of Human Intelligence. *Trends in cognitive sciences.* 2018;22(1):8-20.

72. Higgins JP, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0.* The Cochrane Collaboration; 2011.

73. Price CJ, Friston KJ. Scanning patients with tasks they can perform. *Human brain mapping.* 1999;8(2-3):102-108.

Table 1. Studies investigating the pro-cognitive effects of pharmacological and behavioural interventions using magnetic resonance imaging in mood disorders.

Author	Study Design	Comparison (Intervention/Control)	Group	Age (mean±SD/ median)	Gender (M%)	Cognitive Outcome Measures	MRI Technique &Strength	Main Findings	Quality rating
Behavioural intervention studies									
Macoveanu et al 2018	Single blind parallel group RCT	12-week cognitive remediation + standard treatment	13 BD	34.4±6.5	46%	Picture encoding task n-back task	Task fMRI 3T	Cognitive remediation has no effect on behavioural measures for both tasks (p>0.05). Cognitive remediation has no significant effect on the neural responses of DLPFC and hippocampus for both tasks (p>0.05).	71% (fair)
		Standard treatment	14 BD	30.1±6.8	43%				
Ott et al 2020	Single blind parallel group RCT	2 weeks action based cognitive remediation, consisting of twice weekly group-based therapy and daily computerised training	26 BD	36 (23)	23%	Spatial n-back task One touch stockings of Cambridge*	Task fMRI 3T	Action based cognitive remediation has no effect on behavioural measures in both tasks (p>0.05). Action based cognitive remediation was associated with an increase in DLPFC during both high load and general working memory in comparison to control group (p=0.02). DLPFC activity increase predicted improved executive functions (measured with one touch stockings of Cambridge)	80% (good)
		2 weeks unstructured, therapist-led weekly conversation group	19 BD	38 (22)	21%				

								after treatment completion (10 weeks). In whole brain analysis, enhanced activity in middle frontal, inferior frontal and middle temporal gyrus was seen in action based cognitive remediation group compared to control group (p=0.015, 0.049, 0.0003).	
Pharmacological intervention studies									
Miskowiak et al 2015	Double blind parallel group RCT	8-week EPO (40 000 IU)	19 BD+16 TRD	40±10	31%	RAVLT*	Structural MRI 3T	Shape analysis revealed volume increase in CA1-3/subiculum region of left hippocampus in EPO versus saline independent of changes in mood symptoms and across all patients (p=0.02, FWE corrected). EPO improved RAVLT total recall compared to saline independent of changes in mood symptoms (p=0.04). Volume increase in CA1-3/subiculum region was the only predictor of verbal memory improvement (p=0.006) among variables including age, gender, diagnosis and changes in mood symptoms.	85% (good)
		Saline (Sodium Chloride 0.9%)	17 BD+17 TRD	43±12	32%				

Miskowiak et al 2016a	Double blind parallel group RCT	8-week EPO (40 000 IU)	18 BD+14 UD	40±11	34%	Picture encoding task	Task fMRI 3T	EPO improved picture recall compared to saline (p=0.01) without any differences in BD versus TRD (p=0.046, 0.04). EPO increased encoding related activity in DLPFC and left medial temporal and superior parietal regions, but not in hippocampus without any differences in BD versus TRD (p<0.05, FWE corrected). EPO-related neural activity change correlated with improvement of picture recall.	85% (good)
		Saline (Sodium Chloride 0.9%)	16 BD+14 UD	42±12	33%				
Miskowiak et al 2016b	Double blind parallel group RCT	8-week EPO (40 000 IU)	16 BD+14 UD	39±10	33%	Spatial n-back task	Task fMRI 3T	EPO improved working memory accuracy compared to saline (p=0.045) without any differences in BD versus TRD (p=0.046, 0.04). EPO increased WM load-related activity in the right SFG and enhanced WM load-related deactivation of the left hippocampus (p<0.05, FWE corrected). EPO-related neural activity change correlated with WM improvement.	85% (good)
			16 BD+10 UD	40±12	42%				
Smith et al 2018	Double blind parallel group	14 days vortioxetine (20 mg)	24 UD 24 HC	33.1±9.0 34.5±8.9	33% 37%	Verbal n-back task	Task fMRI 3T	Vortioxetine reduced n-back related activity in right DLPFC, left hippocampus, right insula,	80% (good)

	RCT	Placebo	24 UD 24 HC	38.1±8.8 33.8±9.1	54% 54%			posterior parietal cortex, fusiform and lingual gyrus in both UD and HC ($p<0.05$, FWE corrected).	

(BD: Bipolar Disorder, CA: cornu ammonis, DLPFC: Dorsolateral prefrontal cortex EPO: Erythropoietin, fMRI: Functional magnetic resonance imaging, HC: healthy controls, IU: International units, M:Male, MRI: Magnetic resonance imaging, RAVLT: Rey Auditory verbal learning test, RCT: Randomised controlled trial, SD: standard deviation UD: Unipolar depression, T:Tesla, TRD: Treatment resistant depression)

**Conducted outside the scanner*

Table 2. Studies investigating the pro-cognitive effects of pharmacological interventions using magnetic resonance imaging in healthy controls

Author	Study Design	Comparison (Intervention/Control)	Group (n)	Age (mean±SD)	Gender (M%)	Cognitive Outcome Measures	MRI Technique &Strength	Main Findings	Quality rating
Alda et al									
Apud et al 2007	Double blind cross-over RCT	1st day-300 mg/day tolcapone 2nd-7thday- 600 mg/day tolcapone	34	N/A, inclusion criteria: 18-55	51%	n-back task	Task fMRI 3T	No significant difference between tolcapone and placebo on accuracy and response time (p>0.05). Tolcapone was associated with decreased activation in bilateral DLPFC compared to placebo (p<0.02, p<0.03, corrected).	77% (good)
		7 day-placebo							
Balice-Gordon et al. 2020	Double blind parallel group RCT	1 week PF-06412562 6 mg/day (D1 and D5 partial agonist)	27	32±N/A	100%	n-back task AX-continuous performance task	Task fMRI 3T	No significant effect of both doses of PF-06412562 on both tasks(p>0.05).	54% (fair)
		1 week PF-06412562 30 mg/day (D1 and D5 partial agonist)	27						
		Placebo	22						
Bloomfield et al. 2020	Double blind cross-over RCT	Single dose Cannabidiol 600 mg	15	24±5	40%	Prose recall task * n-back task * Digit span task*	ASL 3T	Cannabidiol increased hippocampus CBF (p=0.004). No main effect of the cannabidiol on task performances (p>0.05). Cannabidiol induced increase in OFC CBF was associated with decreased RT in 2- back task (p=0.005).	46% (poor)
		Placebo							

Brown et al 2013	Double blind cross-over RCT	Placebo (3.5 days) + Hydrocortisone (160 mg/day, 2.5 days) Phenytoin (400 gr/day, 3.5 day) + Placebo (2.5 days) Phenytoin (400 gr/day, 3.5 day) + Hydrocortisone (160 mg/day, 2.5 days) Placebo (3.5 day) + Placebo (2.5 day)	15	25±8	40%	RAVLT* Novelty detection task	Task fMRI 3T	RAVLT total score is better in phenytoin+ hydrocortisone than hydrocortisone alone (p=0.01). Novelty detection task related hippocampal activation was reduced with hydrocortisone alone, phenytoin alone and hydrocortisone + phenytoin compared to placebo (p=0.02, <0.01, <0.01). Combination of hydrocortisone and phenytoin was associated with lower activation in para-hippocampus in novelty detection task (p<0.01). No significant correlations between changes in brain activation and changes in RAVLT (p>0.05).	69% (fair)
Hernaus et al. 2017	Double blind cross-over RCT	Single dose Atomoxetine 60 mg	19	N/A, inclusion criteria: 18-30 years	100%	n-back task	Task fMRI 3T	No significant difference between atomoxetine and placebo on accuracy and response time (p>0.05). Atomoxetine increased activity in middle temporal, middle cingulate and fronto-orbital cortices on 2 back vs. X back compared to placebo (p<0.05, FWE corrected). Atomoxetine increased activity in precentral and occipital cortices	54% (fair)

		Placebo						during 3 back vs. X back compared to placebo ($p<0.05$, FWE corrected). Atomoxetine increased FC between anterior insula and fronto-parietal network during 3 back vs. X back ($p<0.05$, FWE corrected). Atomoxetine induced increases in insula-DLPFC FC negatively correlated reaction time variability and working memory capacity ($p=0.02$, $p=0.02$).	
Magalona et al 2013	Double blind cross-over RCT	Tolcapone 300 mg/day (1st day) Tolcapone 600/day (2nd-7th day) Placebo	20	33±9	55%	Variable attentional control task	Task fMRI 3T	No significant effect of tolcapone on accuracy and reaction time ($p>0.05$). There was increased left IPC activation with tolcapone compared to placebo when all task conditions considered together ($p=0.013$, FWE corrected). There was a trend towards to decreased activation in dorsal cingulate with tolcapone compared to placebo when all task conditions considered together ($p=0.057$, FWE corrected). There was a significant positive correlation between COMT enzyme activity and activity in dorsal cingulate ($p=0.016$, 0.013).	77% (good)

Rodriguez et al. 2016	Double blind parallel group RCT	Single dose 280 mg methylene blue	13	29±10	31%	Psychomotor vigilance task Delayed match to sample task	Task fMRI 3T	No significant differences in behavioural measures for both tasks between methylene blue and placebo (p>0.05). Methylene blue increased the activity in bilateral anterior and posterior insula compared to placebo during attention phase of the psychomotor vigilance task (p=0.01-0.008). Methylene blue increased the activity in prefrontal, parietal and occipital cortices compared to placebo during all phases of delayed match to sample task (p=0.03-0.0003).	62% (fair)
		Food colorant	13	31±10	38%				
Rose et al 2006	Non-randomised open label crossover study	Escitalopram 10 mg for 7 days	10	25±N/A	70%	n-back task	Task fMRI 1.5T	No significant effect of escitalopram on accuracy and reaction time (p>0.05). No significant difference between escitalopram and medication-free condition on brain activity in whole-brain analysis (p>0.05, corrected). In region of interest analysis, while escitalopram increased activation in thalamus. anterior cingulate gyrus and caudate, it decreased activation in inferior frontal gyrus in comparison to medication free condition (<0.05, corrected).	69% (fair)
		Medication free for 7 days							
van Ruitenbeek et al 2013	Double blind cross-over RCT	Single dose Betahistine 96 mg/day	16	25±1	50%	n-back task Spatial paired associates learning task	Task fMRI 1.5T	No significant effect of Betahistine on both tasks (p>0.05).	38% (poor)
		Placebo							

van Ruitenbeek et al 2018	Double blind cross-over RCT	Single dose haloperidol 2m g +L-dopa 100 mg/ carbidopa 25 mg Placebo (Ascorbic acid)	12	25±5	100%	n-back task	Task fMRI 3T	No significant difference between combined haloperidol/L-dopa/carbidopa and placebo on accuracy and response time (p>0.05). Combined haloperidol/L-dopa/carbidopa decreased activation in occipital and temporal cortices compared to placebo during 2 back vs. 0 back condition (p<0.005, FWE corrected). Combined haloperidol/L-dopa/carbidopa increased connectivity between DLPFC and areas within the salience network (p<0.05, FWE corrected).	57% (fair)
Schmidt et al 2014	Double blind cross-over RCT	Single dose milk whey-based soft drink with green tea extract of 2,75 g/l Placebo (Milk whey-based soft drink without green tea extract)	12	24±3	100%	n-back task	Task fMRI 3T	Green tea extract was associated with a trend toward to an improved task performance compared to control drink (p=0.066). Green tea extract increased the connectivity from the right SPL to the right MFG compared to control drink (p=0.03, uncorrected). The effect of green tea extract on task performance was significantly correlated with greater right SPL to the right MFG connectivity (p=0.035).	54% (fair)
Spurny et al 2021	Double blind parallel group RCT	21 days of escitalopram 10 mg/day 21 days of placebo	16 20	N/A, inclusion criteria: 18-65 years	N/A N/A	Character associative relearning task*	MRS 3T	No significant differences in associative learning effects under escitalopram compared to placebo on brain gamma aminobutyric acid (p=0.349) and glutamate + glutamine (p=0.118) concentrations in relation to total creatine in	69% (fair)

								hippocampus, insula, putamen, pallidum, thalamus.	
Xie et al 2012	Double blind cross-over RCT	1 week venlafaxine 75 mg/day followed by 1 week placebo	4	29±4	100%	Animal & tool naming task	Task fMRI 1.5 T	The scores on the animal and tool naming test were higher on venlafaxine compared to placebo ($p<0.01$). Venlafaxine administration resulted in hyperactivation in medial frontal cortex (BA9/44) compared to placebo ($p<0.05$, FWE corrected). Difference in activation level between venlafaxine and placebo in left medial frontal cortex was positively correlated with the difference between venlafaxine and placebo in naming test score for both animals and tools ($p<0.001$).	31% (poor)
		1 week placebo followed by 1 week venlafaxine 75 mg/day	4	28±3	100%				
Yalin et al 2021	Double blind cross-over RCT	Single dose 600 mg mifepristone	20	27±8	100%	n-back task Spatial paired associates learning task	Task fMRI 3T	No significant differences in behavioral measures for both tasks between mifepristone and placebo ($p>0.05$). Mifepristone administration was significantly associated with decreased fusiform cortex activations in encoding blocks ($p=0.007$, $p=0.031$) and decreased angular and precuneal cortices activations in the recall block ($p=0.017$, $p=0.009$) of spatial paired associates learning task. Mifepristone administration did not significantly affect fMRI brain activations in the n-back task ($p>0.05$).	85% (good)
		Placebo							

(ASL: Arterial spin labelling, CBF: Cerebral Blood Flow, COMT: Catechol-O-methyltransferase, D: Dopamine, DLPFC: Dorsolateral prefrontal cortex, FC: Functional connectivity, FWE: Family wise error, fMRI: Functional magnetic resonance imaging, IPC: Inferior parietal cortex, M: Male, MFG: Middle frontal gyrus ,MRI: Magnetic resonance imaging, MRS: Magnetic resonance spectroscopy) N/A: Not available, OFC: Orbitofrontal cortex, RAVLT: Rey auditory verbal learning test, RCT: Randomised controlled trial, SD: standard deviation, SPL: Superior parietal lobule ,T: Tesla)

**Conducted outside the scanner*

Table 3. Studies investigating the pro-cognitive effects of behavioural interventions using magnetic resonance imaging in healthy controls

Author	Study Design	Comparison (Intervention/Control)	Group (n)	Age (mean±SD)	Gender (M%)	Cognitive Outcome Measures	MRI Technique &Strength	Main Findings	Quality rating
Jolles et al 2010	Non-randomised open label parallel group study	Practice of the task 3 times per week for 6 weeks	15	22±2	47%	Verbal working memory task	Task fMRI 3T	Decreased response time and increased accuracy were found in practice group compared to control group ($p=0.001$, $p<0.01$). Stronger activation increases in practice group compared to control group for load 5 manipulation trials relative to maintenance trials in VLPFC ($p<0.01$). For maintenance trials, increased activation was found in medial prefrontal cortex, paracingulate cortex, frontal lobe and lingual cortex in only practice group over time ($p<0.05$, cluster corrected). For manipulation trials, increased activation was found in striatum in only practice group over time $p<0.05$, cluster corrected).	54% (fair)
		No intervention in 6 weeks	14	22±2	43%				
Takeuchi et al 2011	Randomised open label parallel group study	Computerized processing speed training for 5 days within 6-day period	23	22±2	48%	n-back task	Task fMRI 3T	For 0-back condition, larger increases in functional activity from pre to post measures in rolandic operculum and temporal gyrus was found in processing speed training group compared to control group ($p<0.001$, corrected).	77% (good)
		No intervention in 6 days	21	21±2	57%				

								For 2-back condition, no regions showed a statistically significant change in functional activity between groups (p>0.05).	

(ACC: Anterior cingulate cortex, FWE: Family wise error, fMRI: Functional magnetic resonance imaging, IPC: Inferior parietal cortex, M: Male, MFG: Middle frontal gyrus, MRI: Magnetic resonance imaging, rGMD: Regional grey matter density, SD: standard deviation,T: Tesla, VLPFC: Ventrolateral prefrontal cortex)

**Conducted outside the scanner*

Table 4. Studies investigating the pro-cognitive effect of neuromodulation interventions using magnetic resonance imaging in healthy controls

Author	Study Design	Comparison (Intervention/Control)	Targeted Brain area	Group (n)	Age (mean±SD)	Gender (M%)	Cognitive Outcome Measures	MRI Technique &Strength	Main Findings	Quality rating
Bilek et al. 2013	Single blind cross-over RCT	High frequency repetitive TMS Sham TMS	Right DLPFC	39	25±2	70%	n-back task	Task fMRI 3T	No significant difference between repetitive TMS and placebo on task performance (p>0.69). Repetitive TMS reduced FC between right DLPFC and left hippocampus compared to sham TMS in 2 back vs. 0 back (p=0.04, FWE corrected).	69% (fair)
Hermiller et al. 2018	Single blind cross-over RCT	Theta-burst TMS Intermittent theta-burst TMS	Individualized parietal cortex location per participant	24	24±3	42%	Episodic memory task*	Resting state fMRI 3T	Memory retrieval was significantly better on theta-burst compared to sham (p=0.01) and beta-frequency stimulation (p=0.03). Theta-burst induced changes in memory retrieval compared to sham stimulation were	69% (fair)

		Beta-frequency TMS							positively correlated with increased hippocampal connectivity with posterior cingulate and medial frontal cortex ($p<0.05$). Theta-burst induced changes in memory retrieval compared to beta frequency stimulation were positively correlated with increased hippocampal connectivity with precuneal cortex ($p<0.05$).	
		Sham TMS								
Sherwood et al 2019	Single blind parallel group RCT	Real neurofeedback (5 session)	Bilateral auditory cortex	18	23±1	61%	Attentional score scale* Continuous performance test*	Real time fMRI 3T	No difference in continuous performance test or attentional score scale in real neurofeedback group compared to sham neurofeedback ($p>0.05$). Auditory cortex deactivation was greater in real neurofeedback group compared to sham neurofeedback ($p=0.029$). The changes in cognitive tests were not correlated to changes in auditory	62% (fair)
		Sham neurofeedback (5 session)		9	24±3	44%				

									cortex activation within groups (p>0.05).	
Zhang et al 2015	Parallel group RCT (Blinding information N/A)	Real neurofeedback (2 session)	DLPFC	15	22±4	53%	Digit span task* Letter memory task*	Real time fMRI 3T	FC of fourteen ROIs pairs (part of DMN, SE, CEN network) were increased in real neurofeedback group compared to four ROIs pairs (part of DMN and SE network) in sham neurofeedback (p<0.05) Improvements in letter memory and digit span task were correlated with increased FC between SN-DMN, SN-CEN, CEN-CEN in real feedback group (p<0.05).	54% (fair)
		Sham neurofeedback (2 session)		15	22±3	53%				

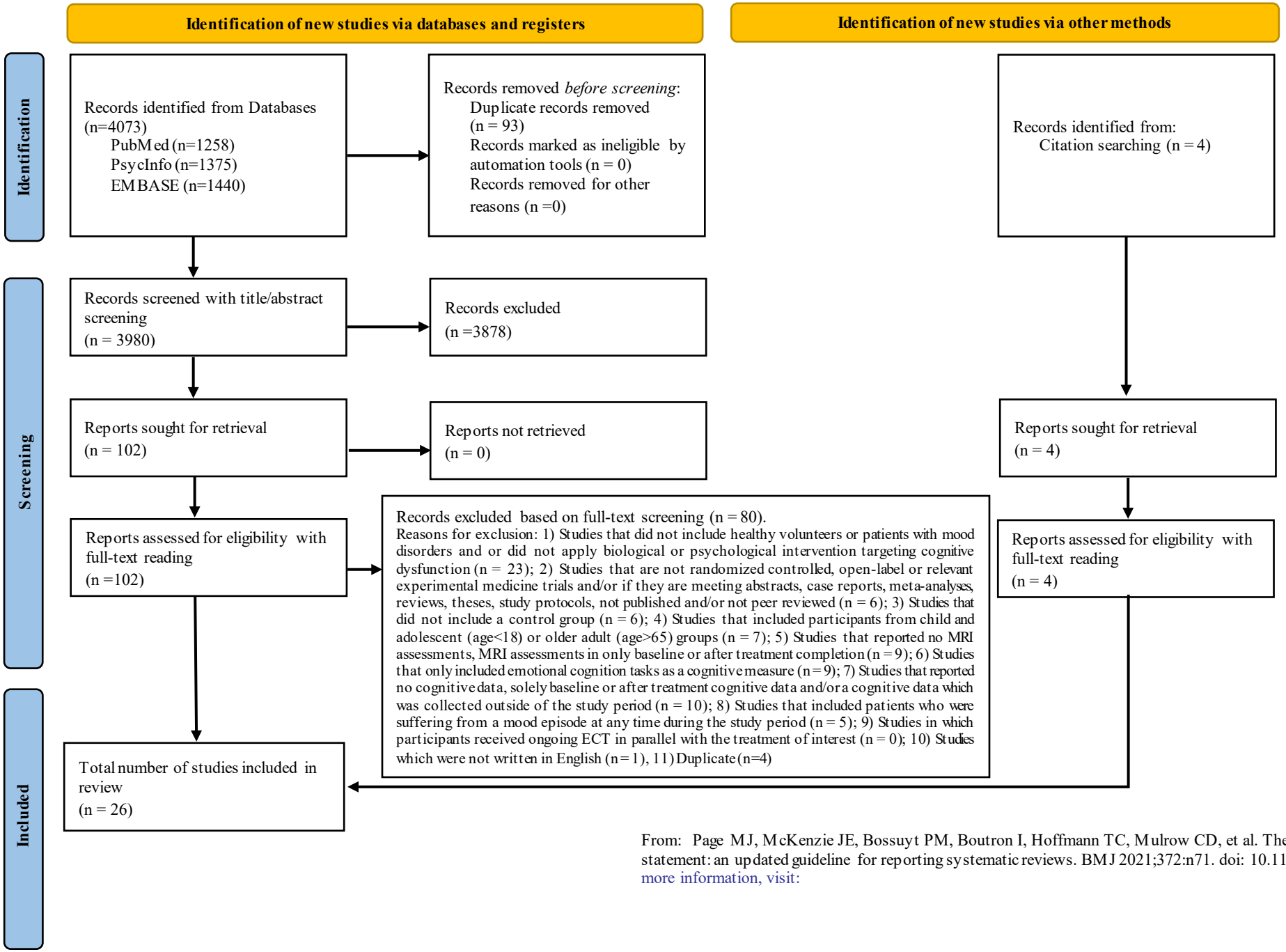
(CEN: Central executive network DMN: Default mode network, FC: Functional connectivity, FWE: Family wise error, DLPFC: Dorsolateral prefrontal cortex, fMRI: Functional magnetic resonance imaging, N/A: Not available, RCT: Randomised controlled trial, ROI: Region of interest SN: Salience network, T: Tesla, TMS: Transcranial magnetic stimulation)

**Conducted outside the scanner*

Table 5. Suggestions for pro-cognitive intervention studies with MRI as an outcome measure.

(i)	Implement task-related fMRI, for which the extant evidence is most robust
(ii)	Use fMRI tasks that reflect the cognitive deficit(s) the intervention is purported to improve
(iii)	Specify planned MRI analysis approach in published study protocol prior to study start
(iv)	Ideally, conduct longitudinal MRI assessments
(v)	When possible, use a mock scanner before the first fMRI scan
(vi)	Include a pre-study cognitive test session in cross-over studies
(vii)	Conduct additional neurocognitive testing outside the scanner
(viii)	Conduct a healthy participant study to obtain proof-of-concept before testing the intervention in partially or fully remitted individuals with mood disorders

Figure 1. PRISMA flowchart



From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: