BMC Psychiatry

SYSTEMATIC REVIEW Open Access

Neuroimaging and neurophysiologic biomarkers for diagnosis and prognosis of depressive disorders, bipolar disorder, anxiety disorders, obsessive compulsive disorder, posttraumatic stress disorder, and substance use disorder: an evidence map

Catherine Sowerby[1*], Adrienne Landsteiner[1], Kristen Ullman[1], Maylen Anthony[1], Caleb Kalinowski[1], Michele R. Spoont[1,2,3], Scott Sponheim[3,4], Kelvin Lim[3,5], Jose V. Pardo[3,6], Timothy J. Wilt[1,2,7] and Wei Duan-Porter[1,2*]

Abstract

Background Advancements in precision medicine, particularly the use of neuroimaging and neurophysiologic techniques, may improve diagnosis, prognosis, and treatment of mental health disorders. Recent efforts to develop large neuroimaging datasets have yielded promising results for identifying mental health biomarkers. This scoping review identifies and characterizes studies of neuroimaging and neurophysiologic techniques used to address a variety of mental health disorders.

Methods We searched MEDLINE and Embase (January 2010-September 2023). Eligible studies examined neuroimaging and neurophysiologic techniques (e.g., magnetic resonance imaging [MRI] or electroencephalogram [EEG]) for diagnosis, prognosis, and/or treatment response for eligible mental health disorders. From eligible studies, we abstracted information on populations, clinical settings, imaging techniques, study designs, outcomes, and analytic approaches.

Results From 58,824 unique search results, we identified 441 eligible primary studies and 27 systematic reviews addressing mental health disorders. Most studies focused on depressive disorders ($k=320$ primary studies [17 systematic reviews]); fewer examined bipolar disorders ($k=61$ [3]), posttraumatic stress disorder (PTSD; $k=39$ [2]), obsessive compulsive disorder (OCD; $k=26$ [1]), anxiety disorders ($k=22$ [3]), or substance use disorders (SUD; $k=25$ [0]). Three-quarters of primary studies used MRI-based techniques and 20% employed EEG. Two-thirds of studies focused on diagnosis (nearly all cross-sectional); the remaining studies mostly addressed symptom response to

*Correspondence:
Catherine Sowerby
Catherine.Sowerby@va.gov
Wei Duan-Porter
wei.duanporter@va.gov

Full list of author information is available at the end of the article

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

various treatments, including antidepressants and psychotherapy. Most primary studies were small ($N < 100$; $k = 263$), and generally included y oung and middle-aged adults; only 5 focused on older adults (sample mean age ≥ 65). Studies were most commonly conducted in China ($k = 181$), the United States ($k = 83$), or Canada ($k = 22$).

Conclusions Although many eligible studies evaluated MRI or EEG for diagnosis and/or treatment response for depressive disorders, most were small and cross-sectional. There was less existing evidence examining other neuroimaging techniques or focusing on other mental health disorders (PTSD, OCD, anxiety disorders, or SUD). Given these evidence gaps, it is likely premature to implement neuroimaging and neurophysiologic tests in clinical settings. To determine clinical utility, future research should use large samples in longitudinal designs and investigate a broader set of disorders.

Trial registration https://doi.org/10.17605/OSF.IO/5PHG2. Clinical trial number: not applicable.

Keywords Mental health, Medical neurobiology, Psychiatry, Neuroscience, Medical sciences

Background

In the past several decades, there have been significant advancements in precision medicine, specifically the use of biomarkers in diagnosis, prognosis, and treatment of medical conditions. Precision medicine aims to personalize care based on more detailed knowledge of individual variations in disease mechanism and prognosis [1]. For treatment of mental health disorders, precision medicine has primarily focused on identifying drug metabolic pathways and other genomic factors in treatment response [2]. Other approaches use neuroimaging techniques and neuropsychological measurements to assess brain structure, function, and activity [3].

Recent efforts to systematically collect and examine large neuroimaging datasets have yielded promising results for identifying biomarkers for individuals with psychiatric disorders. Several ongoing large-scale, population-based studies also aim to advance precision medicine, including the United Kingdom (UK) Biobank database [4, 5]. Despite advances in the development and availability of neuroimaging and neurophysiologic assessment techniques, substantial challenges exist in application to clinical practice [2, 3, 6]. Larger studies and newer paradigms are likely needed for translation into clinical settings, particularly due to clinical heterogeneity and symptom overlap among psychiatric disorders [6].

In response to congressional mandates to advance and, if appropriate, implement precision medicine for mental health treatment of United States (US) veterans, the Department of Veterans Affairs (VA) convened a workgroup to identify targets for future research and clinical demonstration projects. As a critical first step to understanding the state of research on using neuroimaging and neurophysiologic biomarkers to support diagnosis and/or treatment for psychiatric disorders, the VA Evidence Synthesis Program (ESP) conducted an evidence map. An evidence map is well suited to characterize the current evidence base, particularly when it may consist largely of exploratory studies [7, 8]. An evidence map is also the most appropriate method for meeting the broad scope of assessing the state of the science across numerous techniques and disorders [7, 8].

To facilitate future research and investment in precision medicine approaches to mental health diagnosis and treatment, we summarize the main results of the VA ESP evidence map evaluating the use of neuroimaging and neurophysiologic biomarkers in the diagnosis and treatment of depressive disorders, bipolar disorders, anxiety disorders, obsessive-compulsive disorder (OCD), posttraumatic stress disorder (PTSD), and substance use disorders (SUD) [9]. We provide descriptive information about the existing studies and highlight weaknesses and gaps in the evidence, as determined by the volume and characteristics of studies.

Methods

The protocol (developed *a priori)* is registered with Open Science Framework (S1 File; https://doi.org/10.17605/OSF.IO/5PHG2). Clinical trial number: not applicable.

Scope and key question

The main goals of the VA ESP report were to guide research policy and potential clinical demonstration projects of neuroimaging and neurophysiologic tests for a variety of mental health disorders and traumatic brain injury (TBI) [9]. As an initial step towards understanding the evidence base for this broad scope, we conducted an evidence map to describe the characteristics of the existing published literature. In this paper, we focus on results addressing diagnosis and prognosis of mental health disorders:

> *Key Question: What are the quantity, distribution, and characteristics of evidence assessing the accuracy and utility of neuroimaging and neurophysiologic biomarkers in the diagnosis and clinical management of the following disorders: depressive disorders, bipolar disorder, anxiety disorders, OCD, PTSD, and SUD?*

Data sources and searches

We searched MEDLINE and Embase from January 2010 to September 2023 (search strategy in S2 File). We elected to search for articles published after 2010 because a key goal of this review was to understand whether there were more recent advances within the evidence base on neuroimaging and neurophysiologic techniques for psychiatric disorders. We also searched websites for VA ESP and Agency for Healthcare Research and Quality Evidence-based Practice Centers to identify relevant reviews.

Study selection

Eligible English-language studies and reviews included adults with at least one disorder of interest, employed at least one test of interest (e.g., magnetic resonance imaging [MRI], including functional MRI [fMRI], and evoked potentials and electroencephalogram [EEG]), and assessed diagnostic accuracy compared with gold standard assessments, clinical prognosis, and/or treatment response. Exclusion criteria were pediatric populations, mental health symptoms or cognitive functioning only in neurodegenerative disorders or intracranial injury, and exclusive use of cross-sectional data for prognosis or treatment response. Detailed eligibility criteria are provided in Table 1.

Abstracts were uploaded in DistillerSR [10] and screened with the assistance of machine-learning tools (DistillerSR Artificial Intelligence System [DAISY]) in three separate phases involving a variable number of human reviewers. In the first phase, two human reviewers were required to agree on exclusion (abstracts could be included by single human reviewer); approximately 12,000 abstracts were reviewed, and this phase also served as training for DAISY. The second phase occurred when the rate of inclusion had fallen below 5% and DAISY predicted likelihood of inclusion was ≤0.4 for all remaining articles. During this phase, a single human reviewer screened about 7,000 abstracts, with DAISY serving as the second reviewer; DAISY predicted likelihood of inclusion was 0.2–0.3 for these articles. Finally, no human reviewer screened the remaining abstracts with DAISY predicted likelihood < 0.2 (25,912 abstracts); before this point, the inclusion rate had been 0.066% for the last 1,526 abstracts screened by a human reviewer. We conservatively estimate that an additional 17 abstracts may have been missed for Full-text review, due to stopping of single human review of abstracts in this last phase. Full-text review was primarily conducted by a single human reviewer and DAISY acting as the second reviewer, after two initial pilot rounds to ensure common understanding and operationalization of eligibility criteria. Approximately 50% of articles included at full-text review were also evaluated again by a second human reviewer. The use of machine-learning tools in screening and selection of studies has been documented to have good sensitivity, while enabling more efficient processing of a large amount of published evidence [11–13].

Data abstraction

We abstracted the following data: population characteristics; neuroimaging tests and/or EEG; clinical setting; comparator for diagnostic accuracy, diagnostic, prognostic, and/or treatment response outcomes; study design; and analytic methods (e.g., statistical models and diagnostic standards). Data from ~50% of articles were verified by a second reviewer.

Quality assessment and summary of results

As appropriate for evidence maps, we did not conduct formal quality or risk of bias assessments of individual studies. We present results organized by conditions of interest, focusing on characteristics of study populations, outcomes, and study designs (including analytic methods). We also note methodological limitations and evidence gaps.

Results

Overview

From 58,824 unique search results, we identified 441 eligible primary studies and 27 systematic reviews on mental health conditions (Fig. 1). Most articles focused on depressive disorders (*k*=320primary studies and 17 systematic reviews [14–30]). Fewer primary studies addressed bipolar disorder, anxiety disorders, OCD, PTSD, or SUD. Systematic reviews also addressed anxiety disorders (*k*=3) [31–33], bipolar disorders (*k*=4) [34–37], PTSD (*k*=2) [38, 39], or OCD (*k*=1) [40] (S4 File). Most systematic reviews employed MRI-based techniques (*k*=16) [15, 17, 18, 23–26, 28, 31, 33, 35–37, 39, 41, 42] or EEG (*k*=5) [14, 21, 27, 29, 30], or included multiple types of neuroimaging data (*k*=7) [19, 20, 22, 32, 34, 38, 40].

Figure 2 depicts the distribution of primary studies and Table 2 summarizes the population characteristics and methods. Citations and detailed characteristics are provided in S5 File. Below, we first describe primary studies focusing on depressive disorders, and then more briefly present evidence on other mental health disorders.

Depressive disorders

Two thirds of the primary studies focusing on depressive disorders employed structural and/or functional MRI (*k*=212) (Table 2). Remaining studies used EEG (*k*=100) or other neuroimaging techniques. Diagnostic accuracy was the most common outcome assessed (*k*=201, 63%) and most studies were small with total *N*<100 (*k*=187, 58%). Only nine studies had total *N*>1000 (see Figure 3). Study populations consisted of predominantly young and middle-aged adults, with substantial proportions of

Table 1 Eligibility criteria for selection of studies

	Inclusion criteria	Exclusion criteria
Population	$N \geq 30$, ≥ 18 years of age or older, with the following conditions: • Depression • Anxiety (including OCD, phobias and panic disorders) • Posttraumatic stress disorder (PTSD) • Substance use disorder (SUD) • Bipolar disorder • Traumatic brain injury (TBI)	• Pediatric populations or mixed pediatric and adult populations without stratified data • Non-human studies • Stroke patients (CVA) • Multiple sclerosis • Intracranial hemorrhage (intracranial etiology, e.g., burst aneurysm) • Neurodegenerative conditions (e.g., dementia, Parkinson's disease) except if assessed as outcomes or subsets of eligible conditions
Neuroimaging or Neurophysiologic Technique	• Magnetic resonance imaging (MRI) • Functional magnetic resonance imaging (fMRI) • Diffusion tensor imaging (DTI) • Perfusion weighted imaging (PWI) • Magnetic resonance spectroscopy (MRS) • Positron emission tomography (PET) • Single photon emission computed tomography (SPECT) • Arterial spin labeling (ASL) • Magnetoencephalography (MEG) • Evoked potentials and electroencephalogram (EEG) • Paired pulse transcranial magnetic stimulation (ppTMS)	NA
Outcomes	Diagnostic accuracy compared with: • Validated structured clinical interviews (e.g., MINI, SCID-5, WHO WMH-CIDI) • Validated clinician reported instruments of symptoms (e.g., CAPS, HDRS, HAM-A) • Patient-reported measures of mental health symptoms (e.g., PCL-5, PHQ-9, HADS, BDI, GAD-7, AUDIT) • Measures of cognition, other psychiatric symptoms (e.g., delusions, hallucinations) Prognosis and treatment response: • Change in symptoms, cognition, functioning (e.g., SF-36, WHODAS) • Sobriety/abstinence or reduction in substance use (SUD only) • Recurrence or relapse (study must define criteria and use validated measures) • Sensitivity (vs. lack of response) to treatment • Self-harm behaviors or suicide risk • Adverse events and side effects	Only between-group differences (e.g., in neuroimaging findings) or cross-sectional correlations with symptom severity
Timing	Published 2010 or later	Earlier than 2010
Setting	Any	
Study design	Observational studies, trials, and systematic reviews	Study protocols, case reports, abstracts, editorials; for prognosis and treatment response outcomes, cross-sectional studies

Abbreviations. *ASL* Arterial spin labeling, *AUDIT* Alcohol screening questionnaire, *BDI* Beck Depression Inventory, *CAPS* Clinically Administered Posttraumatic Stress Disorder Scale, *CVA* Cerebrovascular accident, *DTI* Diffusion tensor imaging, *EEG* Evoked potentials and electroencephalogram, *fMRI* Functional magnetic resonance imaging, *GAD-7* General Anxiety Disorder-7, *HADS* Hospital Anxiety and Depression Scale, *HAM-A* Hamilton Anxiety Rating Scale, *HDRS* Hamilton Depression Rating Scale, *MEG* Magnetoencephalography, *MINI* Mini International Neuropsychiatric Interview, *MRI* Magnetic resonance imaging, *MRS* Magnetic resonance spectroscopy, *OCD* Obsessive compulsive disorder, *PCL-5* Posttraumatic Stress Disorder Checklist for Diagnostic and Statistical Manual of Mental Disorders, 5th Edition, *PET* Positron emission tomography, *PHQ-9* Patient Health Questionnaire-9, *ppTMS* Paired pulse transcranial magnetic stimulation, *PTSD* Posttraumatic Stress Disorder, *PWI* Perfusion weighted imaging, *SCID-5* Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition, *SF-36* 36-Item Short Form Survey, *SPECT* Single photon emission computed tomography, *SUD* Substance use disorder, *TBI* Traumatic brain injury, *WHO* World Health Organization, *WHODAS* World Health Organization Disability Assessment Schedule, *WMH-CIDI* World Mental Health Composite International Diagnostic Interview

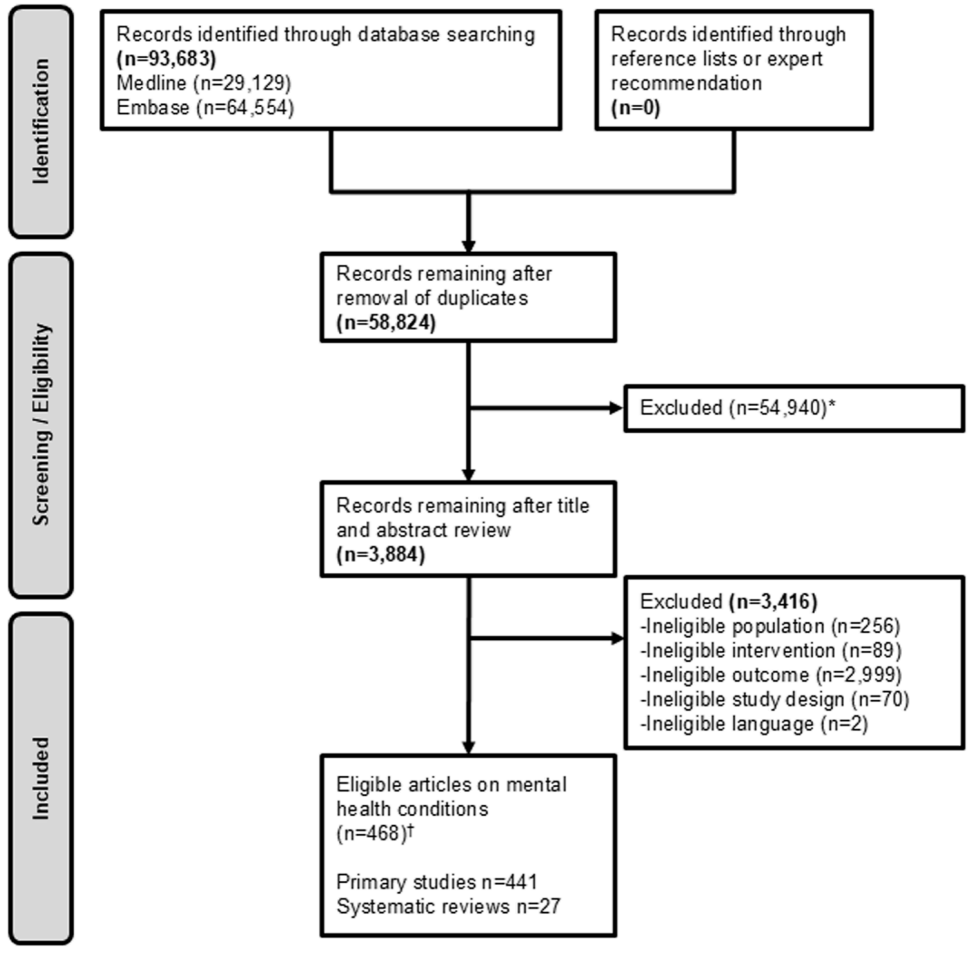

Fig. 1 Search & selection of eligible studies. *n=27,915 excluded with probability score <0.2 by machine-learning algorithm DAISY on the DistillerSR platform (Evidence Partners). †There were 9 other eligible articles on traumatic brain injury that were included in the original VA ESP report [9]

women. Studies were conducted in various countries, most commonly in China (*k*=131).

Primary studies evaluating diagnosis

Table 3 summarizes characteristics of 201 studies evaluating diagnosis of depressive disorders. Of 142 studies using MRI-based techniques, the largest proportion used structural MRI (k=64), fMRI (k=78; 67 resting and 11 task-specific), or both (k=5). A few used other MRI-based techniques like DTI (k=8) and ASL (k=2). Most MRI studies were cross-sectional (k=133). More than half the studies had N<100 (k=122) and only 11 had N>1000. Most included healthy controls (k=193), while a fifth also included participants with bipolar disorder (k=32). A quarter of studies focused on medication-free participants (k=54), variably defined as those who were medication naive or those not on medications at the time of the study. Few included participants during their first episode of depression (k=22). The most common diagnostic standards were clinician assessments (eg, Hamilton Depression Rating Scale [HAM-D; k=112]) or structured interviews (eg, Structured Clinical Interview for DSM [SCID; k=111] and Mini-International Neuropsychiatric Interview [MINI; k=22]). Fewer studies used patient-reported measures such as the Beck Depression Inventory (BDI; k=25).

Fifty-five studies employed EEG and nearly all were cross-sectional. Most were also very small, with only nine having N≥100 (range 30–400). All but eight included healthy controls (k=47); only two studies focused on participants in their first episode of depression. Clinician assessments (HAM-D and Montgomery-Asberg Depression Rating Scale [MADRS]) were the most frequently used diagnostic standard (k=28).

Eleven studies used other neuroimaging techniques, including magnetoencephalography (MEG; k=7), positron emission tomography (PET; k=2), and single-photon emission computerized tomography (SPECT; k=1). All but one were cross-sectional and most were small (N=41-108), with the single exception of the SPECT study (N=4,541). MEG studies were conducted in China or Taiwan, while PET and SPECT studies occurred in

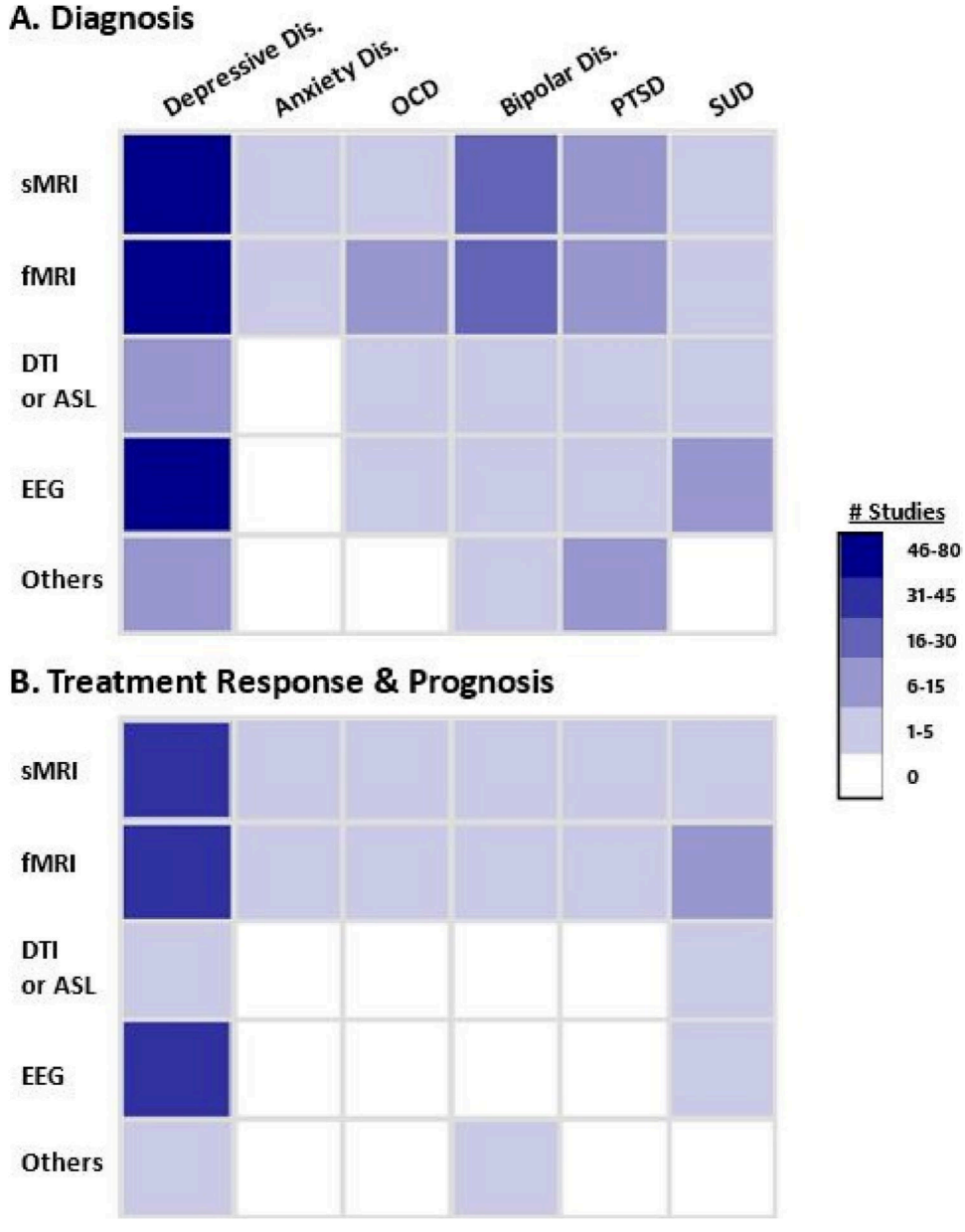

Fig. 2 Number of primary studies using neuroimaging or neurophysiologic data to evaluate (**A**) diagnosis or (**B**) treatment response and prognosis for various mental health disorders. *Abbreviations*. ASL= arterial spin labeling; dis=disorder; DTI=diffusion tensor imaging; EEG=electroencephalogram; fMRI=functional magnetic resonance imaging; MRI=magnetic resonance imaging; OCD=obsessive compulsive disorder; PTSD=posttraumatic stress disorder; sMRI=structural magnetic resonance imaging; SUD=substance use disorder. Notes. "Others" category includes positron emission tomography (PET), single-photon emission computerized tomography (SPECT), and magnetoencephalography (MEG)

the US. MEG and PET studies included healthy controls, while the SPECT study compared data of individuals with cognitive disorders, depression, or both. Four MEG studies focused on medication-free patients, all of which used clinician assessments (HAM-D) or structured interviews (SCID or MINI) as the diagnostic standard.

Primary studies evaluating treatment response & prognosis

Eighty-nine studies evaluated association of neuroimaging or neurophysiologic changes with treatment response and most of these used MRI-based techniques (k=57; see Table 4 for summary characteristics of studies evaluating treatment response). Twenty studies used structural MRI, 31 used fMRI (15 resting, 14 task-specific, and two both resting and task), two used both, and five studies used DTI. Most studies evaluated response to antidepressants (k=58), while others examined psychotherapy, electroconvulsive therapy (ECT), repetitive transcranial magnetic stimulation (rTMS), transcranial direct current

Table 2 Summary characteristics of primary studies

Study characteristics	Depressive disorders (k = 320)	Anxiety disorders[a] (k = 22)	OCD (k = 26)	Bipolar disorders (k = 61)	PTSD (k = 39)	SUD (k = 25)
Imaging technique						
MRI-based techniques:						
sMRI	94	7	8	30	9	6
fMRI	118	11	16	28	15	11
DTI or ASL	14	——	1	5	1	2
EEG	100	4	2	1	11	9
Others[b]	18	——	——	4	6	——
Outcomes						
Diagnosis	201	16	22	53	31	14
Prognosis	9	1	——	1	1	9
Treatment response[c]	119	7	4	7	7	2
Study design & methods						
Cross-sectional	178	13	17	49	25	11
Cohort/longitudinal observational	126	9	9	12	12	13
Randomized controlled trial (RCT)	16	——	——	——	2	1
Used machine learning	181	12	10	38	18	7
Models validated	216	15	15	36	18	12
Country						
US/Canada	68	7	1	14	20	12
China	131	10	20	23	8	4
UK/Europe	44	2	1	12	2	5
Others[d]	72	3	3	11	6	3
NR	5	——	1	1	3	1
Sample sizes (total N)[e]						
30–99	187	17	14	28	21	17
100–200	82	5	10	25	10	6
201–500	33	——	2	6	6	1
501–1000	5	——	——	——	——	——
> 1,000	9	——	——	1	1	1
NR	4	——	——	1	1	——
Age (mean or median, years)						
18–25	17	7	4	8	2	——
26–44	231	14	14	44	26	15
45–64	35	——	——	4	5	6
> 65	6	——	——	1	——	——
NR	28	5	1	2	6	4
% Women						
0–15	2	——	——	——	11	10
16–40	16	6	10	4	5	10
41–70	226	17	8	48	14	2
> 70	36	2	——	4	5	——
Not reported	38	3	1	4	4	3
Other groups included						
Medication free	94	13	12	4	8	4
First episode	24	——	——	2	——	——
Treatment resistant	30	——	2	5	——	——
Healthy controls	219	14	21	43	27	14

Abbreviations. *ASL* Arterial spin labeling, *DTI* Diffusion tensor imaging, *EEG* Electroencephalogram, *MRI* Magnetic resonance imaging, *OCD* Obsessive compulsive disorder, *PTSD* Posttraumatic stress disorder, *RCT* Randomized controlled trial, *SUD* Substance use disorder, *UK* United Kingdom, *US* United States

[a] Includes general anxiety disorder, panic disorder, and social anxiety disorder

[b] Includes positron emission tomography (PET), single-photon emission computerized tomography (SPECT), and magnetoencephalography (MEG)

[c] For SUD, this was abstinence vs. relapse after or during treatment

[d] Includes other countries not included in categories above, as well as studies done in multiple countries

[e] Also includes healthy controls if among participants

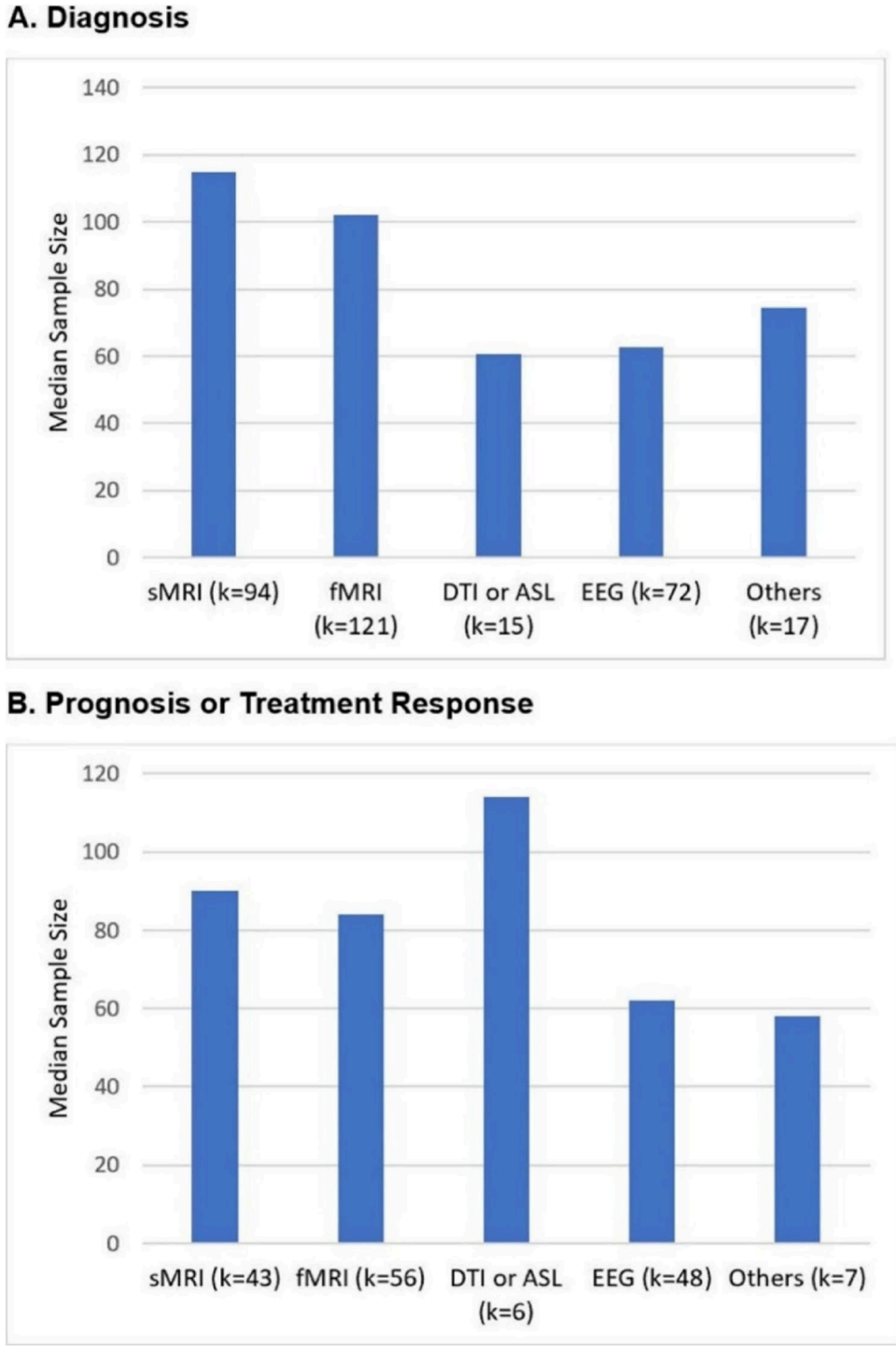

Fig. 3 Median sample size of included studies evaluating (**A**) diagnosis or (**B**) treatment response or prognosis for depressive disorders. *Abbreviations*. ASL = arterial spin labeling; DTI = diffusion tensor imaging; EEG = electroencephalogram; fMRI = functional magentic resonance imaging; sMRI = structural magnetic resonance imaging. *Notes*. "Others" category includes positron emission tomography (PET), single-photon emission computerized tomography (SPECT), and magnetoencephalography (MEG). Number of included studies is indicated for each type of imaging or neurophysiologic data

stimulation (tDCS), theta burst stimulation (TBS), and inpatient multi-modal treatment.

Five studies evaluated other neuroimaging and neurophysiologic techniques in response to antidepressants (MEG $k = 1$ and PET $k = 2$), or rTMS (SPECT $k = 2$, conducted by a single research group). Sample sizes were very small ($N = 33–58$). The SPECT studies also included individuals with bipolar disorder.

Table 3 Summary characteristics of primary studies evaluating diagnosis of depressive disorders

Study characteristics	sMRI (k = 64)	fMRI (k = 78)	ASL or DTI (k = 10)	EEG (k = 55)	Other techniques[a] (k = 11)
Depression subgroups					
Medication free	16	27	2	5	4
First episode	7	12	--	2	1
Treatment resistant	1	3	--	——	——
Other groups included					
Healthy controls	57	71	8	47	10
Bipolar disorder patients	16	10	3	1	2
Country					
US/Canada	9	7	4	6	3
China	35	51	5	20	6
UK/Europe	10	9	1	9	——
Other[b]	9	10	——	18	2
Not reported	1	1	——	2	——
Sample sizes (total N)					
30–99	26	36	7	46	7
100–249	26	23	1	6	2
250–499	6	8	1	3	——
500–999	1	5	——	——	——
> 1,000	4	5	1	——	1
Not reported	1	1	——	——	1
Diagnostic accuracy standard					
Clinician interviews	61	67	9	35	11
Clinician assessments	58	75	8	36	8
Patient-reported outcomes	14	31	1	41	3
Study design					
Cross-sectional	55	69	9	49	10
Cohort/longitudinal	9	9	1	5	1
RCT	——	——	——	1	——
Analytic methods					
Sensitivity/specificity	55	69	9	48	8
Machine learning	42	58	7	34	4
Models validated	46	60	9	42	8

Abbreviations. *ASL* Arterial spin labeling, *DTI* Diffusion tensor imaging, *EEG* Electroencephalogram, *fMRI* Functional magnetic resonance imaging, *RCT* Randomized controlled trial, *sMRI* Structural magnetic resonance imaging (structural or functional), *UK* United Kingdom, *US* United States

[a]Includes positron emission tomography (PET), single-photon emission computerized tomography (SPECT), and magnetoencephalography (MEG)

[b]Includes other countries not included in categories above, as well as studies done in multiple countries

Finally, two studies used MRI to evaluate prognosis in depressive disorders. One evaluated the association between changes in hippocampal volumes and symptoms over two years for US adults 60 years or older at baseline, diagnosed with depression ($N = 72$) or not ($N = 90$). The second study used task-specific fMRI and machine learning methods to classify symptom trajectories for 118 young and middle-aged Dutch adults (mean age 36–43 years) over two years.

Other conditions

Anxiety disorders

Four studies evaluated diagnosis of anxiety disorders, all using either structural MRI ($k = 1$) or fMRI ($k = 3$; one resting and two task-specific). These addressed general anxiety disorder ($k = 3$), social anxiety disorder ($k = 1$), and panic disorder ($k = 1$). Studies were small ($N = 40$–93) and included young adults with substantial representation of women. All included healthy controls and used the SCID and/or the Hamilton Anxiety Rating Scale (HAM-A) as gold standards. Three studies were cross-sectional and the last was a longitudinal cohort. All assessed model accuracy and three conducted model validation; two used machine learning methods

Six studies evaluated structural MRI ($k = 2$) and fMRI ($k = 4$; all task-specific) in response to psychotherapy ($k = 4$), antidepressants ($k = 1$), and a computer-based behavioral intervention ($k = 1$). These addressed general anxiety disorder ($k = 4$), social anxiety disorder ($k = 1$), and panic disorder ($k = 2$). All but one study had total

Table 4 Summary of characteristics of included studies evaluating treatment response in depressive disorders

Study characteristics	Response to treatments				
	Antidepressants (k = 71)	Psychotherapies (k = 5)	ECT (k = 10)	rTMS (k = 25)	Other[a] (k = 10)
Imaging technique					
sMRI	21	——	8	1	1
fMRI	23	4	3	6	5
DTI	4	——	——	——	1
EEG	23	1	——	17	4
Others[b]	5	——	——	2	——
Depression subgroups					
Medication free	36	2	——	4	3
Treatment resistant	5	——	7	14	1
Other groups included					
Healthy controls	33	3	2	5	1
Bipolar disorder	——	——	4	2	1
Country					
US/Canada	23	5	5	8	2
China	21	——	1	1	2
UK/Europe	11	——	1	5	2
Other[c]	16	——	3	10	4
NR	——	——	——	1	——
Sample sizes (total N)					
30–99	39	3	8	21	5
100–249	27	2	2	3	4
250–499	5	——	——	——	1
500–999	——	——	——	——	——
> 1,000	——	——	——	1	——
Study design					
Cohort/longitudinal	60	5	10	23	9
RCT	11	——	——	2	1
Analytic methods					
ROC (or sensitivity/specificity)	61	3	8	24	10
Machine learning	22	——	8	11	4
Models validated	36	——	9	16	5

Abbreviations. *DTI* Diffusion tensor imaging, *ECT* Electroconvulsive therapy, *MRI* Magnetic resonance imaging (structural or functional), *NR* Not reported, *ROC* Receiver operating curve, *rTMS* Repetitive transcranial magnetic stimulation, *UK* United Kingdom, *US* United States

[a]Includes 1 study on transcranial direct current stimulation (tDCS), 1 study on theta burst stimulation (TBS) vs. repetitive transcranial magnetic stimulation (rTMS), and 1 using multi-modal inpatient treatment

[b]Includes positron emission tomography (PET), single-photon emission computerized tomography (SPECT), and magnetoencephalography (MEG)

[c]Includes other countries not included in categories above, as well as studies done in multiple countries

$N < 100$ (range = 34–135). Participants were young adults, and women were well represented (four studies had more than 50%). All but one undertook model validation and two employed machine learning methods

Obsessive compulsive disorder (OCD)

Seventeen studies evaluated diagnosis of OCD, with four using structural MRI, 11 using resting-state fMRI, and one each with DTI and EEG. All studies included young adults and had $N < 200$. Most included healthy controls ($k = 14$) and half had 16–40% women participants ($k = 10$). The most commonly used diagnostic standard was the Yale Brown Obsessive Compulsive Scale (Y-BOCS; $k = 15$), though 11 also used SCID. Nearly all studies evaluated model accuracy ($k = 14$), while seven undertook model validation and five used machine learning.

Only two studies evaluated treatment response in OCD. A Korean study employed baseline structural MRI and machine learning methods to predict response after antidepressants for 4 months in 56 participants. One US study used machine learning methods to examine the association of resting-state fMRI changes with symptoms after four weeks of cognitive behavioral therapy for 42 US adults.

Bipolar disorders

Forty-one studies evaluated diagnosis for bipolar disorders; nearly all used MRI-based techniques ($k = 20$

structural MRI, $k = 18$ fMRI [12 resting and six task-specific], $k = 3$ DTI, and $k = 2$ ASL). One study evaluated EEG for diagnosis. Half of studies were small with total sample sizes less than 100 ($k = 23$); only three studies had more than 250 participants (range 251–441). More than half of studies also included participants with unipolar depression ($k = 27$) and healthy controls ($k = 23$). Most were cross-sectional ($k = 31$), while three were longitudinal (to confirm symptoms and diagnosis over 1–2 years). Sixteen studies used structured clinical interviews (MINI and/or SCID) and standardized clinician assessments (Young Mania Rating Scale [YMRS]) as the diagnostic standard. Another 18 studies used only structured interviews, and three used only YMRS. About half of diagnostic studies used machine learning methods ($k = 25$) and undertook model validation ($k = 24$).

All four prognostic studies also included participants with depressive disorders, and thus were discussed above in the section on depressive disorders studies. Briefly, these studies all used MRI-based techniques to evaluate outcomes after ECT, were small ($N = 33$–122), and included middle-aged adults (mean age range 39–56). Three studies used machine learning and validated models.

Posttraumatic stress disorder (PTSD)

Thirty studies focused on PTSD, with most evaluating diagnosis ($k = 24$). Fourteen diagnostic studies employed MRI-based techniques, five used EEG, and six used other techniques ($k = 2$ SPECT, $k = 1$ PET, and $k = 3$ MEG [one of these also used MRI]). Most diagnostic studies were small, with the majority having $N < 100$ ($k = 15$). The remaining sample sizes were 116–217 ($k = 8$) and 2,137 for one large database study. Six studies included participants who also had TBI. Of the 24 studies evaluating diagnosis, 11 used structured interviews (SCID) and 12 used clinician assessments (Clinician Administered PTSD Scale [CAPS]) as the diagnostic standard. Some also used patient-reported outcome measures such as the PTSD Checklist (PCL; $k = 7$). Nearly all studies were cross-sectional in design, with only two being cohort studies. All studies evaluated model accuracy, 10 undertook model validation, and nine used machine learning methods.

Six prognostic studies all used either structural MRI ($k = 2$) or fMRI ($k = 4$; two resting and two task-specific). Sample sizes ranged from 53–135, with most having $N < 100$ ($k = 5$). One study included alcohol use disorder (AUD). All studies included relatively young adults and reported on predictive models for response to psychotherapy treatment, using CAPS to assess PTSD severity. One study included both psychotherapy and TMS. One study used machine learning methods, and three assessed model accuracy and performed model validation.

Substance use disorders (SUD)

Twenty studies addressed SUD, with the majority focusing on AUD ($k = 12$, 60%). Remaining studies focused on cocaine use disorder ($k = 3$), opioid use disorder ($k = 2$, and methamphetamine use disorder ($k = 3$). About half evaluated diagnosis ($k = 9$), while the rest estimated prediction of relapse ($k = 6$) or treatment response ($k = 5$). Of the nine studies that evaluated diagnostic accuracy, three used MRI-based techniques and six used EEG. Most were small with $N < 100$ ($k = 6$) and all included healthy controls. Six used structured interviews (SCID) and three used patient-reported outcomes (Alcohol Use Disorders Identification Test [AUDIT]) as the diagnostic standard. All studies were cross-sectional and evaluated model accuracy, five undertook model validation, and six used machine learning methods.

Of the 11 prognostic studies, two used structural MRI, six used fMRI (4 resting and 2 task-specific), and one used both in addition to ASL. The remaining two studies used EEG. Most studies were small with $N < 100$ ($k = 8$); one study had a total sample size greater than 1000 ($N = 1{,}376$). All study samples had less than 40% women and a mean age less than 55 years; only one study reported race. Seven studies assessed model accuracy, three undertook model validation, and five used machine learning methods.

Discussion

Our evidence map found that the majority of existing studies used MRI-based techniques to evaluate depressive disorders. Anxiety disorders had the smallest number of studies. Most studies were small, focused on diagnosis, and cross-sectional. Even longitudinal studies often lacked important data on history of exposures, symptoms, and treatments. Very few included only participants with their first episode or new onset of symptoms. Additionally, no existing study evaluated prediction of adverse events, despite this being an important factor in patient and clinician decisions to stop or switch therapies. These methodological limitations have been previously noted and likely contribute to poor replicability and validity of precision medicine studies in mental health [6, 43, 44]. Whereas most existing studies included less than 100 participants, current estimates are that thousands of individuals are needed to provide stable and valid results regarding important associations between neuroimaging biomarkers and complex clinical phenotypes [6].

A recent analysis of neural heterogeneity showed that extreme regional variation of brain deviations characterizes mental illness [45], which may explain the well-established clinical heterogeneity of psychiatric disorders [46, 47]. Despite the high degree of neural heterogeneity between individual cases, case-control studies often rely on group mean comparisons that ignore this variation

[46–48]. Longitudinal and person-specific data on brain networks, symptoms, and exposures are likely needed to further clarify clinical phenotypes, along with consideration of transdiagnostic dimensional approaches [45, 49–51].

Other challenges include high cost and technical difficulty of assessments and differing measures of symptoms and treatment response [6, 52–54]. Furthermore, to account for changes in brain structure and functioning over time, studies should use comparisons with age-standardized findings developed from large populations [43], instead of small samples of age-matched controls. Recent efforts to systematically collect and examine large neuroimaging datasets have yielded more promising results [6, 43]. Thus, future work may produce insights that improve diagnosis and treatment outcomes in mental health.

In order to use precision medicine approaches to improve mental health treatment, the evidence base needs to be greatly expanded. Specifically, we need better evidence from larger, longitudinal studies, such as the Adolescent Brain Cognitive Development (ABCD) study in the US [55], the UK Biobank [5], and the Human Connectome Project [56]. These large-scale studies require substantial upfront and ongoing investments, as well as fundamental changes in research organization and incentives to promote cooperation and data sharing [53, 54, 57]. It may be possible to build on existing large studies that collect neuroimaging data by adding more detailed assessments of mental health symptoms and relevant exposures, or vice versa. Alternatively, it may be possible to pool results from smaller studies if they used standardized common data models, study protocols, and biomarker assessment techniques in clearly defined clinical populations.

We sought to identify and describe the evidence for a broad range of neuroimaging and neurophysiologic biomarkers in the diagnosis, prognosis, and/or treatment response for a variety of mental health conditions. Although this evidence map makes an important contribution by describing the current evidence base and existing gaps, there are several limitations. As typical for evidence maps, we did not formally assess the quality of included studies or extract detailed results on the analytic models (e.g., control variables or performance metrics). Additionally, we limited our search to English-language studies published in 2010 or later, and thus, our map does not include evidence published in non-English languages or before 2010. Finally, we employed machine learning techniques in identification of relevant studies; it is possible that we missed some eligible studies.

Conclusions

The evidence base on neuroimaging and neurophysiologic biomarkers for mental health disorders is large but predominantly focused on using MRI to evaluate depressive disorders. Existing studies had substantial methodological limitations, including small sample sizes and cross-sectional designs. These concerns indicate that it is likely premature to apply neuroimaging and neurophysiologic tests to evaluate and treat mental health conditions in clinical settings. Future research is needed to more rigorously evaluate the utility of these biomarkers in the diagnosis and management of mental health conditions across a wide spectrum of patient populations and clinical settings.

Abbreviations

ABCD	Adolescent Brain Cognitive Development
ASL	arterial spin labeling
AUD	Alcohol use disorder
AUDIT	Alcohol Use Disorders Identification Test
BDI	Beck Depression Inventory
CAPS	Clinician Administered Posttraumatic Stress Disorder Scale
CVA	Cerebrovascular accident
DAISY	DistillerSR's Artificial Intelligence System
DTI	diffusion tensor imaging
EEG	electroencephalogram
ESP	Evidence Synthesis Program
fMRI	functional magnetic resonance imaging
GAD-7	General Anxiety Disorder-7
HAM-A	Hamilton Anxiety Rating Scale
HAM-D/HDRS	Hamilton Depression Rating Scale
k	Number of studies
MADRS	Montgomery-Asberg Depression Rating Scale
MEG	Magnetoencephalography
MINI	Mini-International Neuropsychiatric Interview
MRI	Magnetic resonance imaging
MRS	Magnetic resonance spectroscopy
N	Sample size
OCD	obsessive compulsive disorder
PCL	Posttraumatic Stress Disorder Checklist
PET	Positron emission tomography
PHQ-9	Patient Health Questionnaire-9
ppTMS	Paired pulse transcranial magnetic stimulation
PTSD	Posttraumatic stress disorder
PWI	Perfusion weighted imaging
RCT	Randomized controlled trial
ROC	Receiver operating curve
rTMS	Repetitive transcranial magnetic stimulation
SCID	Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders (DSM)
SF-36	36-Item Short Form Survey
sMRI	structural magnetic resonance imaging
SPECT	Single photon emission computed tomography
SUD	substance use disorder
TBI	Traumatic brain injury
TBS	Theta burst stimulation
tDCS	Transcranial direct current stimulation
TMS	Transcranial magnetic stimulation
UK	United Kingdom
US	United States
VA	Department of Veterans Affairs
WHO	World Health Organization
WHODAS	World Health Organization Disability Assessment Schedule
WMH-CIDI	World Mental Health Composite International Diagnostic Intervie
Y-BOCS	Yale Brown Obsessive Compulsive Scale
YMRS	Young Mania Rating Scale

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12888-025-07429-4.

Supplementary Material 1. S1 File. Protocol. S2 File. Search Strategy. S3 File. AI Methods Description. S4 File. Included Systematic Reviews by Mental Health Condition and Outcome. S5 File. Detailed Characteristics of Included Primary Studies. S6 File. Supplementary Materials References

Acknowledgements
The authors are grateful to the Operational Partners of the original VA report, who contributed to the development and approval of the final project scope and provided feedback on the draft report: Stuart W. Hoffman, PhD; Sumitra Muralidhar, PhD; Vetisha L. McClair, PhD; Clifford Smith, PhD, ABPP; Emily Hartwell, PhD; Wendy Tenhula, PhD.

Authors' contributions
Conception/design: CS, WD-P, AL, KU, MS, SS, KL, JVP, and TJW. Acquisition, analysis, or interpretation of data: CS, WD-P, AL, KU, CK, MA, and TJW. First draft of the manuscript: CS and WD-P. Manuscript drafting, revision, approval: CS, WD-P, AL, KU, CK, MA, MS, SS, KL, JVP, and TJW. Overall guarantors: CS and WD-P. All authors read and approved the final manuscript.

Funding
This work was supported by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development, Evidence Synthesis Program, Grant #09–009.

Data availability
All data generated or analyzed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Minneapolis VA Health Care System, Center for Care Delivery and Outcomes Research, Minneapolis, MN, USA
[2]University of Minnesota School of Medicine, Department of Medicine, Minneapolis, MN, USA
[3]University of Minnesota School of Medicine, Department of Psychiatry and Behavioral Sciences, Minneapolis, MN, USA
[4]Minneapolis VA Health Care System, Division of Psychology, Minneapolis, MN, USA
[5]Minneapolis VA Health Care System, Geriatric Research Education and Clinical Center, Minneapolis, MN, USA
[6]Minneapolis VA Health Care System, Cognitive Neuroimaging Unit, Minneapolis, MN, USA
[7]University of Minnesota School of Public Health, Division of Health Policy and Management, Minneapolis, MN, USA

Received: 10 February 2025 / Accepted: 9 September 2025
Published online: 06 May 2026

References

1. National Academy of Medicine. In: Finkelman EM, McGinnis JM, McClellan MB, Dzau VJ, editors. Vital Directions for Health & Health Care: An Initiative of the National Academy of Medicine. Washington, DC: National Academies Press (US) Copyright 2023 by the National Academy of Sciences. All rights reserved. 2017. pp 249–268
2. Fernandes BS, Williams LM, Steiner J, Leboyer M, Carvalho AF, Berk M. The new field of 'precision psychiatry'. BMC Med. 2017;15(1):80.
3. Williams LM. Precision psychiatry: a neural circuit taxonomy for depression and anxiety. Lancet Psychiatry. 2016;3(5):472–80.
4. Littlejohns TJ, Holliday J, Gibson LM, Garratt S, Oesingmann N, Alfaro-Almagro F, et al. The UK biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions. Nat Commun. 2020;11(1):2624.
5. Biobank UK. [Available from: https://www.ukbiobank.ac.uk/.
6. Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, et al. Reproducible brain-wide association studies require thousands of individuals. Nature. 2022;603(7902):654–60.
7. Bragge P, Clavisi O, Turner T, Tavender E, Collie A, Gruen RL. The global evidence mapping initiative: scoping research in broad topic areas. BMC Med Res Methodol. 2011;11:92.
8. Miake-Lye IM, Hempel S, Shanman R, Shekelle PG. What is an evidence map? A systematic review of published evidence maps and their definitions, methods, and products. Syst Rev. 2016;5:28.
9. Ullman K, Landsteiner A, Anthony M, Sowerby K, Kalinowski C, Duan-Porter W, et al. Neuroimaging and neurophysiologic biomarkers for mental health: an evidence map. Washington: Evidence Synthesis Program, Health Services Research and Development Service; 2022.
10. DistillerSR. Version 2.35. DistillerSR Inc. 2023. Accessed Mar 2022-Dec 2023. https://www.distillersr.com/
11. Gates A, Guitard S, Pillay J, Elliott SA, Dyson MP, Newton AS, et al. Performance and usability of machine learning for screening in systematic reviews: a comparative evaluation of three tools. Syst Rev. 2019;8(1):278.
12. Hamel C, Kelly SE, Thavorn K, Rice DB, Wells GA, Hutton B. An evaluation of distillersr's machine learning-based prioritization tool for title/abstract screening - impact on reviewer-relevant outcomes. BMC Med Res Methodol. 2020;20(1):256.
13. Shekelle PG, Shetty K, Newberry S, Maglione M, Motala A. Machine learning versus standard techniques for updating searches for systematic reviews: a diagnostic accuracy study. Ann Intern Med. 2017;167(3):213–5.
14. Bruun CF, Arnbjerg CJ, Kessing LV. Electroencephalographic parameters differentiating melancholic depression, non-melancholic depression, and healthy controls. Syst Rev Front Psychiatry. 2021;12:648713.
15. Cohen SE, Zantvoord JB, Wezenberg BN, Bockting CLH, van Wingen GA. Magnetic resonance imaging for individual prediction of treatment response in major depressive disorder: a systematic review and meta-analysis. Transl Psychiatry. 2021;11(1):168.
16. De Crescenzo F, Ciliberto M, Menghini D, Treglia G, Ebmeier KP, Janiri L. Is (18)F-FDG-PET suitable to predict clinical response to the treatment of geriatric depression? A systematic review of PET studies. Aging Ment Health. 2017;21(9):889–94.
17. Dichter GS, Gibbs D, Smoski MJ. A systematic review of relations between resting-state functional-MRI and treatment response in major depressive disorder. J Affect Disord. 2015;172:8–17.
18. Enneking V, Leehr EJ, Dannlowski U, Redlich R. Brain structural effects of treatments for depression and biomarkers of response: a systematic review of neuroimaging studies. Psychol Med. 2020;50(2):187–209.
19. Fu CH, Steiner H, Costafreda SG. Predictive neural biomarkers of clinical response in depression: a meta-analysis of functional and structural neuroimaging studies of pharmacological and psychological therapies. Neurobiol Dis. 2013;52:75–83.
20. Gillett G, Tomlinson A, Efthimiou O, Cipriani A. Predicting treatment effects in unipolar depression: a meta-review. Pharmacol Ther. 2020;212:107557.
21. Khosla A, Khandnor P, Chand T. Automated diagnosis of depression from EEG signals using traditional and deep learning approaches: a comparative analysis. Biocybern Biomed Eng. 2022;42(1):108–42.
22. Levy A, Taib S, Arbus C, Peran P, Sauvaget A, Schmitt L, et al. Neuroimaging biomarkers at baseline predict electroconvulsive therapy overall clinical response in depression: a systematic review. J ECT. 2019;35(2):77–83.
23. Long Z, Du L, Zhao J, Wu S, Zheng Q, Lei X. Prediction on treatment improvement in depression with resting state connectivity: a coordinate-based meta-analysis. J Affect Disord. 2020;276:62–8.
24. Masse-Sibille C, Djamila B, Julie G, Emmanuel H, Pierre V, Gilles C. Predictors of response and remission to antidepressants in geriatric depression: a systematic review. J Geriatr Psychiatry Neurol. 2018;31(6):283–302.

25. Scheepens DS, van Waarde JA, Lok A, de Vries G, Denys D, van Wingen GA. The link between structural and functional brain abnormalities in depression: a systematic review of multimodal neuroimaging studies. Front Psychiatry. 2020;11:485.
26. Siegel-Ramsay JE, Bertocci MA, Wu B, Phillips ML, Strakowski SM, Almeida JRC. Distinguishing between depression in bipolar disorder and unipolar depression using magnetic resonance imaging: a systematic review. Bipolar Disord. 2022;24(5):474–498.
27. Simon L, Blay M, Galvao F, Brunelin J. Using EEG to predict clinical response to electroconvulsive therapy in patients with major depression: a comprehensive review. Front Psychiatry. 2021;12:643710.
28. Sinha P, Joshi H, Ithal D. Resting state functional connectivity of brain with electroconvulsive therapy in depression: meta-analysis to understand its mechanisms. Front Hum Neurosci. 2020;14:616054.
29. van der Vinne N, Vollebregt MA, van Putten M, Arns M. Frontal alpha asymmetry as a diagnostic marker in depression: fact or fiction? A meta-analysis. Neuroimage Clin. 2017;16:79–87.
30. Widge AS, Bilge MT, Montana R, Chang W, Rodriguez CI, Deckersbach T, et al. Electroencephalographic biomarkers for treatment response prediction in major depressive illness: a meta-analysis. Am J Psychiatry. 2019;176(1):44–56.
31. Qing X, Gu L, Li D. Abnormalities of localized connectivity in Obsessive-compulsive disorder: a voxel-wise meta-analysis. Front Hum Neurosci. 2021;15:739175.
32. Santos VA, Carvalho DD, Van Ameringen M, Nardi AE, Freire RC. Neuroimaging findings as predictors of treatment outcome of psychotherapy in anxiety disorders. Prog Neuropsychopharmacol Biol Psychiatry. 2019;91:60–71.
33. Xu J, Van Dam NT, Feng C, Luo Y, Ai H, Gu R, et al. Anxious brain networks: a coordinate-based activation likelihood estimation meta-analysis of resting-state functional connectivity studies in anxiety. Neurosci Biobehav Rev. 2019;96:21–30.
34. Hozer F, Houenou J. Can neuroimaging disentangle bipolar disorder? J Affect Disord. 2016;195:199–214.
35. Librenza-Garcia D, Kotzian BJ, Yang J, Mwangi B, Cao B, Pereira Lima LN, et al. The impact of machine learning techniques in the study of bipolar disorder: a systematic review. Neurosci Biobehav Rev. 2017;80:538–54.
36. Seeberg I, Kjaerstad HL, Miskowiak KW. Neural and behavioral predictors of treatment efficacy on mood symptoms and cognition in mood disorders: a systematic review. Front Psychiatry. 2018;9(JUL):337.
37. Whalley HC, Papmeyer M, Sprooten E, Lawrie SM, Sussmann JE, McIntosh AM. Review of functional magnetic resonance imaging studies comparing bipolar disorder and schizophrenia. Bipolar Disord. 2012;14(4):411–31.
38. Colvonen PJ, Glassman LH, Crocker LD, Buttner MM, Orff H, Schiehser DM, et al. Pretreatment biomarkers predicting PTSD psychotherapy outcomes: a systematic review. Neurosci Biobehav Rev. 2017;75:140–56.
39. Nelson MD, Tumpap AM. Posttraumatic stress disorder symptom severity is associated with left hippocampal volume reduction: a meta-analytic study. CNS Spectr. 2017;22(4):363–72.
40. Fullana MA, Abramovitch A, Via E, Lopez-Sola C, Goldberg X, Reina N, et al. Diagnostic biomarkers for obsessive-compulsive disorder: a reasonable quest or ignis fatuus? Neurosci Biobehav Rev. 2020;118:504–13.
41. Haghbayan H, Boutin A, Laflamme M, Lauzier F, Shemilt M, Moore L, et al. The prognostic value of MRI in moderate and severe traumatic brain injury: A systematic review and Meta-Analysis. Crit Care Med. 2017;45(12):e1280–8.
42. Hulkower MB, Poliak DB, Rosenbaum SB, Zimmerman ME, Lipton ML. A decade of DTI in traumatic brain injury: 10 years and 100 articles later. AJNR Am J Neuroradiol. 2013;34(11):2064–74.
43. Bethlehem RAI, Seidlitz J, White SR, Vogel JW, Anderson KM, Adamson C, et al. Brain charts for the human lifespan. Nature. 2022;604(7906):525–33.
44. Esterman M, Stumps A, Jagger-Rickels A, Rothlein D, DeGutis J, Fortenbaugh F, et al. Evaluating the evidence for a neuroimaging subtype of posttraumatic stress disorder. Sci Transl Med. 2020. https://doi.org/10.1126/scitranslmed.aaz9343.
45. Segal A, Parkes L, Aquino K, Kia SM, Wolfers T, Franke B et al. Regional, circuit and network heterogeneity of brain abnormalities in psychiatric disorders. Nat Neurosci. 2023;26:1613–1629.
46. Fried EI, Nesse RM. Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR*D study. J Affect Disord. 2015;172:96–102.
47. Young G, Lareau C, Pierre B. One quintillion ways to have PTSD comorbidity: recommendations for the disordered DSM-5. Psychol Injury Law. 2014;7(1):61–74.
48. Wolfers T, Doan NT, Kaufmann T, Alnæs D, Moberget T, Agartz I, et al. Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. JAMA Psychiatr. 2018;75(11):1146–55.
49. Dalgleish T, Black M, Johnston D, Bevan A. Transdiagnostic approaches to mental health problems: current status and future directions. J Consult Clin Psychol. 2020;88(3):179–95.
50. Gillan CM, Seow TXF. Carving out new transdiagnostic dimensions for research in mental health. Biol Psychiatry Cogn Neurosci Neuroimaging. 2020;5(10):932–4.
51. Yucel M, Oldenhof E, Ahmed SH, Belin D, Billieux J, Bowden-Jones H, et al. A transdiagnostic dimensional approach towards a neuropsychological assessment for addiction: an international Delphi consensus study. Addiction. 2019;114(6):1095–109.
52. Benkarim O, Paquola C, Park BY, Kebets V, Hong SJ, Vos de Wael R, et al. Population heterogeneity in clinical cohorts affects the predictive accuracy of brain imaging. PLoS Biol. 2022;20(4):e3001627.
53. Specht K. Current challenges in translational and clinical fMRI and future directions. Front Psychiatry. 2019;10:924.
54. White T, Blok E, Calhoun VD. Data sharing and privacy issues in neuroimaging research: opportunities, obstacles, challenges, and monsters under the bed. Hum Brain Mapp. 2022;43(1):278–91.
55. ABCD Research Consortium. Adolescent Brain Cognitive Development (ABCD) Study [Available from: https://abcdstudy.org/about/.
56. Elam JS, Glasser MF, Harms MP, Sotiropoulos SN, Andersson JLR, Burgess GC, et al. The human connectome project: A retrospective. NeuroImage. 2021;244:118543.
57. Kozlov M. Your brain expands and shrinks over time - these charts show how. Nature. 2022;604(7905):230–1.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.