

Review

# Methodological and Quality Flaws in the Use of Artificial Intelligence in Mental Health Research: Systematic Review

Roberto Tornero-Costa<sup>1\*</sup>, MSc; Antonio Martinez-Millana<sup>1\*</sup>, PhD; Natasha Azzopardi-Muscat<sup>2\*</sup>, PhD; Ledia Lazzeri<sup>2\*</sup>, MD; Vicente Traver<sup>1\*</sup>, PhD; David Novillo-Ortiz<sup>2\*</sup>, PhD

<sup>1</sup>Instituto Universitario de Investigación de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas, Universitat Politècnica de València, Valencia, Spain

<sup>2</sup>Division of Country Health Policies and Systems, World Health Organization, Regional Office for Europe, Copenhagen, Denmark

\*all authors contributed equally

**Corresponding Author:**

David Novillo-Ortiz, PhD

Division of Country Health Policies and Systems

World Health Organization, Regional Office for Europe

Marmorej 51

Copenhagen, 2100

Denmark

Phone: 45 45 33 7198

Email: [dnovillo@who.int](mailto:dnovillo@who.int)

## Abstract

**Background:** Artificial intelligence (AI) is giving rise to a revolution in medicine and health care. Mental health conditions are highly prevalent in many countries, and the COVID-19 pandemic has increased the risk of further erosion of the mental well-being in the population. Therefore, it is relevant to assess the current status of the application of AI toward mental health research to inform about trends, gaps, opportunities, and challenges.

**Objective:** This study aims to perform a systematic overview of AI applications in mental health in terms of methodologies, data, outcomes, performance, and quality.

**Methods:** A systematic search in PubMed, Scopus, IEEE Xplore, and Cochrane databases was conducted to collect records of use cases of AI for mental health disorder studies from January 2016 to November 2021. Records were screened for eligibility if they were a practical implementation of AI in clinical trials involving mental health conditions. Records of AI study cases were evaluated and categorized by the International Classification of Diseases 11th Revision (ICD-11). Data related to trial settings, collection methodology, features, outcomes, and model development and evaluation were extracted following the CHARMS (Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies) guideline. Further, evaluation of risk of bias is provided.

**Results:** A total of 429 nonduplicated records were retrieved from the databases and 129 were included for a full assessment—18 of which were manually added. The distribution of AI applications in mental health was found unbalanced between ICD-11 mental health categories. Predominant categories were Depressive disorders (n=70) and Schizophrenia or other primary psychotic disorders (n=26). Most interventions were based on randomized controlled trials (n=62), followed by prospective cohorts (n=24) among observational studies. AI was typically applied to evaluate quality of treatments (n=44) or stratify patients into subgroups and clusters (n=31). Models usually applied a combination of questionnaires and scales to assess symptom severity using electronic health records (n=49) as well as medical images (n=33). Quality assessment revealed important flaws in the process of AI application and data preprocessing pipelines. One-third of the studies (n=56) did not report any preprocessing or data preparation. One-fifth of the models were developed by comparing several methods (n=35) without assessing their suitability in advance and a small proportion reported external validation (n=21). Only 1 paper reported a second assessment of a previous AI model. Risk of bias and transparent reporting yielded low scores due to a poor reporting of the strategy for adjusting hyperparameters, coefficients, and the explainability of the models. International collaboration was anecdotal (n=17) and data and developed models mostly remained private (n=126).

**Conclusions:** These significant shortcomings, alongside the lack of information to ensure reproducibility and transparency, are indicative of the challenges that AI in mental health needs to face before contributing to a solid base for knowledge generation and for being a support tool in mental health management.

(*JMIR Ment Health* 2023;10:e42045) doi: [10.2196/42045](https://doi.org/10.2196/42045)

## KEYWORDS

artificial intelligence; mental health; health research; review methodology; systematic review; research methodology; research quality; trial methodology

## Introduction

Mental health represents a vital element of individual and collective well-being, but stressful or adverse living, working, or economic conditions and social inequalities, violence, and conflict can put it at risk. The COVID-19 pandemic has demonstrated how vulnerable mental health can be. Mental health conditions represent one of the leading causes of suffering and disability in the European Region. In 2021, over 150 million people in the WHO (World Health Organization) European Region lived with a mental health condition, and only 1 in 3 people with depression receive the care they need. To address these gaps in mental health services and support, many of which have been exacerbated by the pandemic, WHO/Europe launched a new Pan-European Mental Health Coalition [1]. Mental health is a top priority for the WHO and is a flagship initiative of the European Programme of Work 2020-2025 [2].

Artificial intelligence (AI) has been increasingly used to provide methods and tools for improved diagnosis and treatment of diseases since 2010. AI is defined as the reproducibility of human-like reasoning and pattern extraction to solve problems [3]. AI involves a variety of methods that expand traditional statistical techniques. AI can find patterns that support decision making and hypotheses validation. AI offers a new scope of powerful tools to automate tasks, support clinicians, and deepen understanding of the causes of complex disorders. AI's presence and potential in health care are rapidly increasing in recent years. AI models need to be fed with the adequate data to be integrated in the clinical workflow and ensuring data quality is crucial [4]. Digitized data in health care are available in a range of formats, including structured data such as electronic health records or medical images, and nonstructured schemas, such as clinical handwritten notes [5].

Because of the possibilities AI offers, policymakers may gain insight into more efficient strategies to promote health and into the current state of mental disorders. However, AI often involves a complex use of statistics, mathematical approaches, and high-dimensional data that could lead to bias, inaccurate

interpretation of results, and overoptimism of AI performance if it is not adequately handled [6]. Further, several lacking areas cause concern: transparent reporting in AI models that undermine replicability, potential ethical concerns, validation of generalizability, and positive collaboration in the research community [7,8].

The goals of this review are to map the applications of AI techniques in mental health research, reveal the prominent mental health aspects in this framework, and to assess the methodological quality of the recent scientific literature and evolution of this field in the last 5 years. Systematic reviews and meta-analyses (PRISMA [Preferred Reporting Items for Systematic Reviews and Meta-Analyses] 2020 statement) [9] will be used to design the search strategy and to funnel selection in this systematic overview.

## Methods

### Search Strategy

A systematic literature search was conducted on clinical trials on mental health disorders involving AI techniques using 4 electronic databases: PubMed, Scopus, IEEE Xplore, and Cochrane (Table 1). Search string queries are detailed in Appendix S1 in [Multimedia Appendix 1](#).

### Inclusion and Exclusion Criteria

We specified 3 inclusion criteria for screening. Records were included if they reported a clinical trial (either interventional or observational), were related to mental health disorders, and featured an application of AI. For the final eligibility assessment, exclusion criteria were defined to constrain the review: the reported AI case is not applied for a mental health outcome (ie, applying tools to improve image quality), the record was not published in English, or the report was not published in the last 5 years to review the specific application of these techniques in clinical mental health research. These criteria were designed to evaluate the researching lines in mental health disorders in the last few years, which include the democratization of frameworks and tools for AI application.

**Table 1.** Databases consulted and filters related to our search criteria applied in the search engines.

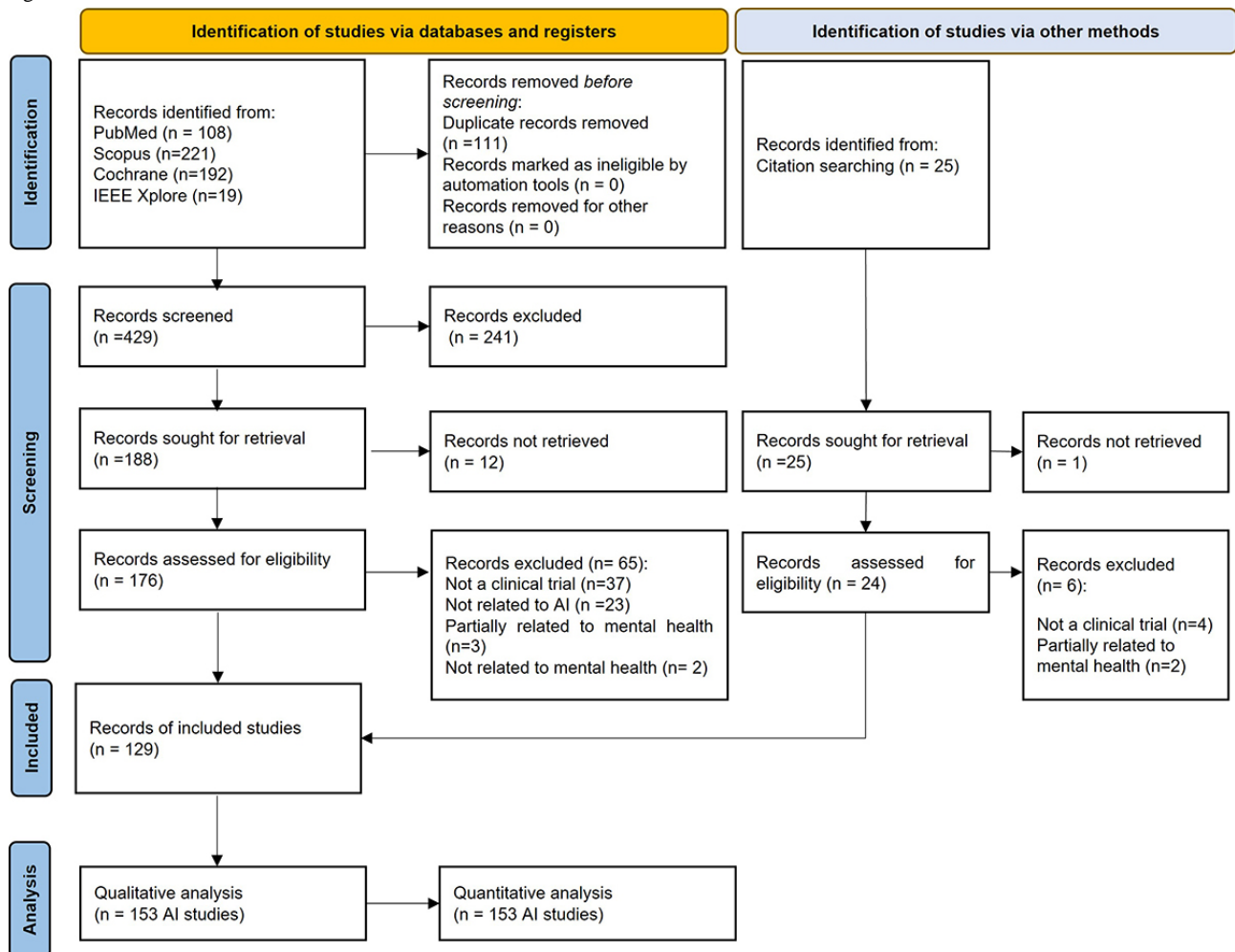
Database	Filters in search engine
PubMed	<ul style="list-style-type: none"> <li>• Article type: Clinical trial</li> <li>• Language: English</li> <li>• Publication date: 5 years</li> </ul>
Scopus	<ul style="list-style-type: none"> <li>• Document type: Article</li> <li>• Language: English</li> </ul>
IEEE Xplore	<ul style="list-style-type: none"> <li>• Range years: 2016-2021</li> </ul>
Cochrane	<ul style="list-style-type: none"> <li>• Type: Trials</li> <li>• Publication date: 2016-2021</li> <li>• Language: English</li> </ul>

**Selection Process**

Figure 1 shows the flow diagram of the selection process. Records from the scientific literature were identified in the 4 databases defined in Table 1. The resulting data sets were combined in a Microsoft Excel spreadsheet (.xlsx), rearranged by DOI (digital object identifier), and checked for possible erroneous entries. Duplicated records were assessed by comparing DOI names and titles of the publication. A simple code in R 4.2.0 win32 (R Foundation for Statistical Computing)

was used to find and tag records whose DOI name and title were already found in the database. The results were manually reviewed to correct minor errors due to misspellings of DOI or the title in the record database. The eligibility criteria for inclusion were then manually evaluated by the title and abstract of each record, and selected records were sought for retrieval. Retrieved records were fully screened and were dismissed if they did not meet the inclusion criteria or met the exclusion criteria. Finally, data and details were extracted for included AI studies.

**Figure 1.** Selection process: PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram. AI: artificial intelligence.



## Data Extraction

For the included AI applications, 11 categories and 35 data indicators are reported. These indicators were adapted from the CHARMS (Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies) guideline [10]. In addition, publication-related data such as author(s), title, journal, year of publication, and DOI were extracted for the analysis.

Categories were designed to evaluate the goal of the clinical trial and AI, the accessibility and quality of the development process of the data, how the AI model was designed and developed, results, and the reported discussion. Not only are the categories for data extraction designed for detailing the AI models, but also they evaluate the differences and linkages between trial design, data collection, and AI implementation. For further details, see Appendix S2 in [Multimedia Appendix 1](#).

## Quality Assessment

Risk of bias was evaluated by combining the Cochrane tool for randomized controlled trials [11] and PROBAST (Prediction Model Risk of Bias Assessment Tool) guidelines [12]. The Cochrane Handbook for Systematic Reviews of Interventions [13] accounted for the trial design and whether masking and blindness were applied or should have been. The PROBAST guidelines accounted for the suitability of the methodology for collecting the data, candidate predictors, and outcome definition for the AI model as well as how the AI model was applied and analyzed. Both guidelines were considered together to evaluate possible biased relations between trial design and AI applicability. Details of the methodology are provided in Appendix S3 in [Multimedia Appendix 1](#).

## Results

### Articles Identified From the Database Searching

The search identified 540 records, all published in English. Excluding 111 duplicates, 429 articles were screened according to the eligibility criteria based on the title and abstract. The screening process concluded with 241 records excluded for not meeting the inclusion criteria. Of those, 188 records were sought for retrieval, with 12 found impossible to retrieve. Thus, 176 records were assessed for eligibility and 65 were excluded due to not being a clinical trial (n=37), not related to AI (n=23), partially related to mental health (n=3), or not related to mental health (n=2). Furthermore, limitations of search queries were minimized as much as possible by manually adding a selection of AI studies that were not found in the search (n=25). Records from this selection were also screened and sought for retrieval and eventually 18 studies were included. Ultimately, 129 records were included in the analysis. A record could involve 2 or more different cases of AI use for a different outcome, from now on

referred to as an AI study. A total of 153 AI studies or AI applications were analyzed. [Table 2](#) summarizes the most important information extracted for this systematic review. Details on the final analysis for each study can be found in [Multimedia Appendix 2](#) (see also [14-142]).

Most used private data (n=142), and a small fraction used public data (n=10) or a mix (n=1). Most studies aimed to develop a new model (n=152); only 1 study aimed to validate a current model. No AI study was intended to update a previously developed AI model. Concerning mental health categories based on the International Classification of Diseases 11th Revision (ICD-11), nearly one-half of the studies (77/153, 50.3%) related to mood disorders, which combines the Depressive disorders (n=70) and Bipolar or related disorders (n=4) categories; 3 other studies used data from patients within both categories, labeled as "mood episodes." The second most common category was Schizophrenia or other primary psychotic disorders (n=26), and the third was Disorders specifically associated with stress (n=12). Some studies included participants with different mental disorders (n=10).

Other categories were Anxiety or fear-related disorders (n=7); Secondary mental or behavioral syndromes associated with disorders or diseases classified elsewhere (n=5); Disorders due to substance use (n=5); Neurocognitive and dementia disorders (n=4); Neurodevelopmental disorders (n=1); Obsessive-compulsive or related disorders (n=1); Feeding or eating disorders (n=1); Bodily distress disorder (n=1); Personality disorders (n=1); and Mental or behavioral disorders associated with pregnancy, childbirth, or the puerperium, without psychotic symptoms (n=1).

Only 28.1% (43/153) of studies used original data collected within the study, while 71.9% (110/153) of studies retrieved data from databases or were a secondary analysis of clinical trials not designed for that purpose. The most common type of trial design was randomized clinical trial (n=62), followed by prospective cohort study designs (n=30) and nonrandomized clinical trial designs (n=15). Further, we found longitudinal naturalistic studies (n=15), cross-sectional designs (n=14), case-control designs (n=9), and case reports (n=2). Two reports of AI cases used a mix of trial designs and 4 did not report this or the references were unclear. [Figure 2](#) shows the distribution of study design based on the prospective or retrospective collection of data.

Not all studies reported enough details to evaluate recruitment of participants (n=17). Almost one-half of the studies collected data from different locations (n=75), whereas the rest only reported 1 location (n=61). Of the multisite studies, only one-third used international collection (17/153, 11.1%). Only 13 of the 43 (30%) prospective collection studies followed a multisite collection method (n=13), and only 1 study was international.

**Table 2.** Key summary of artificial intelligence studies (N=153) analyzed (n=129 articles).

Mental health disorder (ICD-11 <sup>a</sup> ) section number: category	Artificial intelligence model family	Data type
6A0: Neurodevelopmental disorders	<ul style="list-style-type: none"> <li>Regression [14]</li> </ul>	<ul style="list-style-type: none"> <li>Mixed<sup>b</sup> [14]</li> </ul>
6A2: Schizophrenia or other primary psychotic disorders	<ul style="list-style-type: none"> <li>Competing models<sup>c</sup> [15-17]</li> <li>Ensembled models [18,19]</li> <li>Regression [20-23]</li> <li>Statistical clustering [24]</li> <li>SVM<sup>d</sup> [25-28]</li> <li>Trees [29-31]</li> <li>Regression and statistical clustering [32]</li> <li>Regression and hierarchical clustering [33]</li> </ul>	<ul style="list-style-type: none"> <li>Genomic data [29]</li> <li>Medical image [15-19,21-23,27,28,33]</li> <li>Mixed [20,26,30,32]</li> <li>Questionnaires and scales [24,25,31]</li> </ul>
6A6: Bipolar or related disorders	<ul style="list-style-type: none"> <li>Bayesian [34]</li> <li>Manifold [35]</li> <li>Regression [36]</li> <li>SVM [37]</li> </ul>	<ul style="list-style-type: none"> <li>Medical image [35,37]</li> <li>Mixed [36]</li> <li>Questionnaires and scales [34]</li> </ul>
6A7: Depressive disorders	<ul style="list-style-type: none"> <li>Bayesian [38-40]</li> <li>Competing models [41-53]</li> <li>Deep learning [54]</li> <li>Ensembled models [55-57]</li> <li>Hierarchical clustering [58,59]</li> <li>Markov model [60]</li> <li>Mixture model [61]</li> <li>Mixture model, regression, and trees [62]</li> <li>Regression [63-74]</li> <li>Relevance vector machine [75]</li> <li>Statistical learning [76,77]</li> <li>SVM [78-86]</li> <li>Trees [87-93]</li> <li>Trees and hierarchical clustering [94]</li> <li>Trees and statistical learning [95]</li> </ul>	<ul style="list-style-type: none"> <li>Audio recording [64]</li> <li>Biomarkers [51,54,95]</li> <li>Biosignal [41]</li> <li>I<sup>e</sup> [80,81,87,93]</li> <li>Genomic data [44,70,74]</li> <li>Medical image [38,40,48,67-69,75,82,85]</li> <li>Mixed [42,43,45,46,49,50,52,53,55-57,59,61,63,65,66,71-73,77,79,83,86,88-92,94]</li> <li>Questionnaires and scales [47,58,60,62,76]</li> <li>Text [39]</li> <li>Video image [78,84]</li> </ul>
6A6, 6A7: Mood episodes	<ul style="list-style-type: none"> <li>Competing models [96]</li> <li>SVM [97]</li> <li>Regression [98]</li> </ul>	<ul style="list-style-type: none"> <li>Medical image [97,98]</li> <li>Questionnaires and scales [96]</li> </ul>
6B0: Anxiety or fear-related disorders	<ul style="list-style-type: none"> <li>Competing models [99-101]</li> <li>Regression [102]</li> <li>SVM [103]</li> <li>Trees [104]</li> </ul>	<ul style="list-style-type: none"> <li>Biomarkers [101]</li> <li>Biosignal [100]</li> <li>Mixed [99,103,104]</li> <li>Text [102]</li> </ul>
6B2: Obsessive-compulsive or related disorders	<ul style="list-style-type: none"> <li>Competing models [105]</li> </ul>	<ul style="list-style-type: none"> <li>I [105]</li> </ul>
6B4: Disorders specifically associated with stress	<ul style="list-style-type: none"> <li>Competing models [106]</li> <li>Ensembled models [107,108]</li> <li>Hierarchical clustering [109]</li> <li>Mixture model and regression [110]</li> <li>SVM [111-115]</li> <li>Trees [116]</li> </ul>	<ul style="list-style-type: none"> <li>Audio recording [116]</li> <li>Biosignal [115]</li> <li>Medical image [112-114]</li> <li>Mixed [106-108,110,111]</li> <li>Questionnaires and scales [109]</li> </ul>
6B8: Feeding or eating disorders	<ul style="list-style-type: none"> <li>Competing models [117]</li> </ul>	<ul style="list-style-type: none"> <li>Mixed [117]</li> </ul>
6C2: Bodily distress disorder	<ul style="list-style-type: none"> <li>Regression [118]</li> </ul>	<ul style="list-style-type: none"> <li>Mixed [118]</li> </ul>
6C4: Disorders due to substance use	<ul style="list-style-type: none"> <li>Competing models [119]</li> <li>Regression [120]</li> <li>Trees [121,122]</li> </ul>	<ul style="list-style-type: none"> <li>Medical image [120]</li> <li>Mixed [119,121,122]</li> </ul>

Mental health disorder (ICD-11 <sup>a</sup> ) section number: category	Artificial intelligence model family	Data type
6D1: Personality disorders	<ul style="list-style-type: none"> <li>Trees [123]</li> </ul>	<ul style="list-style-type: none"> <li>Mixed [123]</li> </ul>
6D7, 6D8: Neurocognitive disorders and dementia	<ul style="list-style-type: none"> <li>Competing models [124]</li> <li>Ensembled models [125,126]</li> <li>Trees [127]</li> </ul>	<ul style="list-style-type: none"> <li>Mixed [124-127]</li> </ul>
6E2: Mental or behavioral disorders associated with pregnancy, childbirth, or the puerperium, without psychotic symptoms	<ul style="list-style-type: none"> <li>Competing models [128]</li> </ul>	<ul style="list-style-type: none"> <li>Mixed [128]</li> </ul>
6E6: Secondary mental or behavioral syndromes associated with disorders or diseases classified elsewhere	<ul style="list-style-type: none"> <li>Bayesian [129]</li> <li>Competing models [130]</li> <li>Regression [131]</li> <li>Trees [132,133]</li> </ul>	<ul style="list-style-type: none"> <li>Mixed [130-133]</li> <li>Questionnaires and scales [129]</li> </ul>
Combination of some ICD-11 categories in mental health	<ul style="list-style-type: none"> <li>Bayesian [134]</li> <li>Ensembled models [135-137]</li> <li>Regression [138,139]</li> <li>Regression and support vector machines [140]</li> <li>Trees [141]</li> </ul>	<ul style="list-style-type: none"> <li>Biosignal [140]</li> <li>Mixed [135-137,139]</li> <li>Questionnaires and scales [134,138,141]</li> </ul>
Unspecified <sup>f</sup>	<ul style="list-style-type: none"> <li>Competing models [142]</li> </ul>	<ul style="list-style-type: none"> <li>Mixed [142]</li> </ul>

<sup>a</sup>ICD-11: International Classification of Diseases 11th Revision.

<sup>b</sup>Mixed: combination of type of data and predictors.

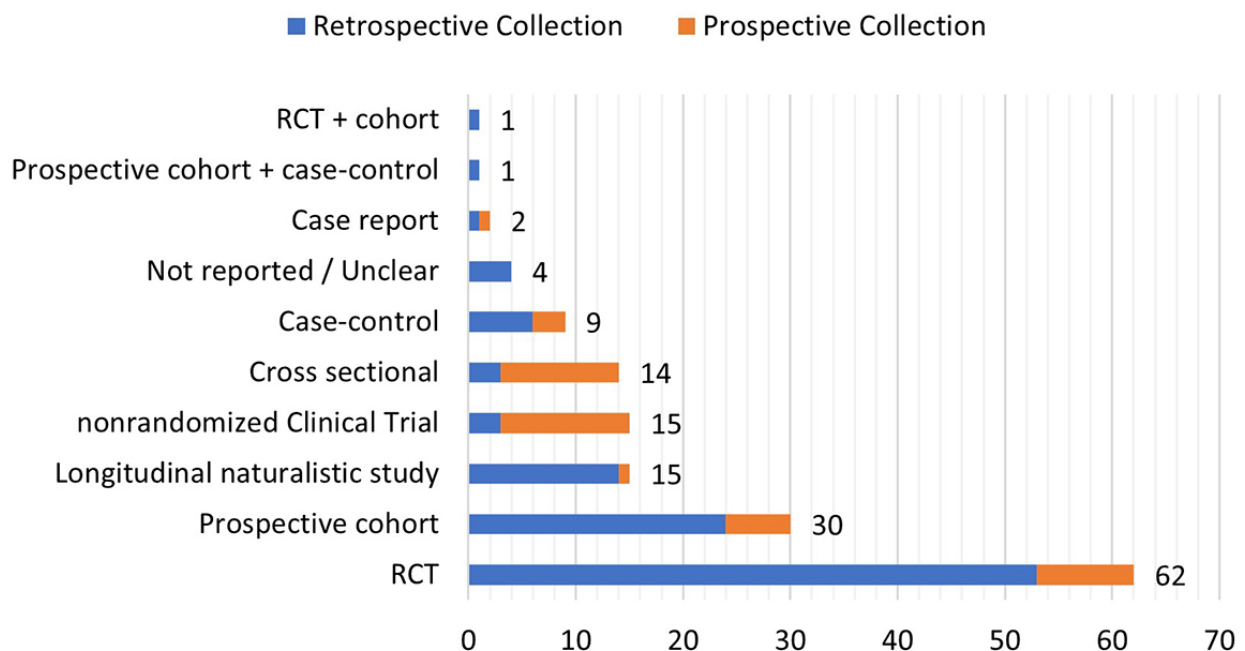
<sup>c</sup>Competing models: the study was designed for evaluate several types of artificial intelligence model families without assessing a priori adequacy.

<sup>d</sup>SVM: support vector machine.

<sup>e</sup>I: electronic health records.

<sup>f</sup>Unspecified: The outcome of the study is “mental health problems,” therefore, it could not be classified in any specific category.

**Figure 2.** Count of trial designs where data were retrieved. Orange specifies only studies with their own designed trial. RCT: randomized clinical trial.



### AI Applications

Studies were categorized according to the intended use of the AI models in the research. The most common category was studies for evaluating treatments, Treatment quality (n=44),

followed by Subgroups/patterns identification (n=31), Predictor identification (n=28), Prognosis (n=23), Diagnose (n=20), and Forecasting symptoms (n=7).

Most Treatment quality applications used retrospective data from previous randomized clinical trials (n=28) and in these studies the clinical arms were treated as different cohorts to compare AI outcomes and performance (n=28). The same results can be found in Predictor identification, where close to one-half of the studies collected data from previous randomized clinical trials to compare different clinical arms (n=13). In the Subgroups/patterns identification category, studies collected data from a balanced mix of study designs, while in Prognosis the most common method was prospective cohort studies. In the Diagnose and Forecasting symptoms studies, none of the categories stood out in particular. More detailed results are presented in Appendix S4 in [Multimedia Appendix 1](#).

**Figure 3** presents a dashboard that summarizes the AI model results regarding candidate predictors, preprocessing pipeline, AI techniques, and validation. For candidate predictors, many studies used a combination of data (n=73). The most individually common category was Medical image (n=33), which relates to medical imaging analysis (ie, region of interest or voxel-based morphometry), and the second was Questionnaires and scales (n=20), defined as any self-reported or interview-reported scale for symptom severity, conditions, or actual mood. The third was Biosignal collection (n=11), such as electroencephalography or electrocardiography and related analyses. Other data categories were Biomarkers (n=5), Genomic data (n=3), Electronic health records or I (n=2), Text (n=2), Video image analysis (n=2), and Audio recording (n=2). I refers to historical, demographic, and clinical information collected in hospitals and specialty care sites. Text refers to any data that are used for natural language processing analysis, such as written text or speech. Audio recording was introduced as the analysis of audio and voice features unrelated to language processing. The Mixed category (n=73 studies) combined data from I and different questionnaires and symptom scales (49/73, 67%); the remaining studies included other categories such as Biomarkers (n=7), Medical image (n=4), Genomic data (n=3), Biosignal (n=3), and Text (n=2). Medical image was also combined with Genomic data (n=1) and Biomarkers (n=1). Besides, Biomarkers were combined with Questionnaires and scales (n=2) and with Biosignals (n=1).

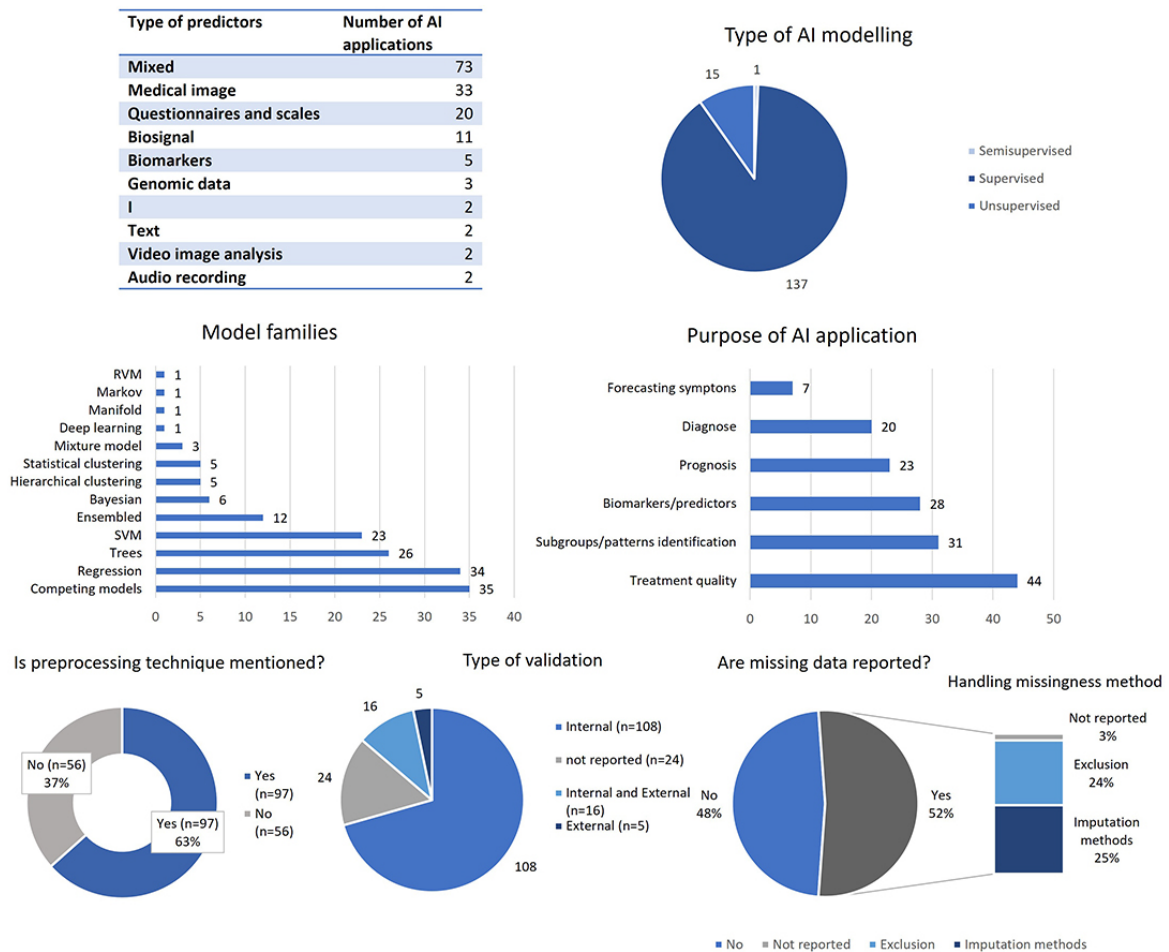
When evaluating data quality, methods to assess data suitability, and preprocessing pipelines, only 12/153 studies (7.8%)

considered the statistical power of the sample size; 37.3% (57/153) of studies used a sample size of 150 or less to train the AI models. Only 13.7% (21/153) reported external validation (n=5) or reported both internal and external validation (n=16). The rest of the AI studies used only internal validation (n=108) or did not report the validation method (n=24). Only 38.6% (59/153) of studies reported a method to assess significance of their performance results, while the majority did not detail any (n=94).

AI studies used supervised learning, semisupervised learning, and unsupervised learning methods. No reinforcement learning algorithms were found. Regarding AI algorithms, the most popular family of techniques was regression (n=34), followed by trees (n=26) and support vector machine (n=23), which constitutes most AI studies. Other algorithms were Bayesian (n=6), statistical clustering (n=5), hierarchical clustering (n=5), mixture model (n=3), deep learning (n=1), manifold (n=1), Markov model (n=1), and relevance vector machine (n=1). In some cases, an Ensembled model was designed with the inclusion of different types of AI algorithms (n=12). Another category, Competing models (n=35), refers to the AI studies that did not predefine a specific AI technique or algorithm based on features of the data and instead applied different techniques with the intention to retain the model algorithm with the best performance to their outcome definition. These 35 studies used 144 AI techniques in total.

Regarding preprocessing methods, only 63.4% (97/153) of studies reported whether they applied any preprocessing technique to data or that preprocessing was not needed, while the rest did not report any (n=56). Regarding data gaps, only 52.3% (80/153) of studies reported or mentioned if there were some missing data in samples or not, while 47.7% (73/153) did not. Of the studies reporting missing data, 2.6% (4/153) did not report any method to handle missing data bias, whereas 24.2% (37/153) opted for excluding the samples and 25.5% (39/153) chose to impute the missing values from the data distribution by different imputation methods. Of these, only 2 studies detailed the type of missingness. The proportion of reporting missing data was similar for both retrospective and prospective data collection.

**Figure 3.** Dashboard and descriptive analytics on AI developing and preprocessing pipeline. AI: artificial intelligence; EHR: electronic health record; RVM: relevance vector machine; SVM: support vector machine.



### Risk of Bias and Transparency

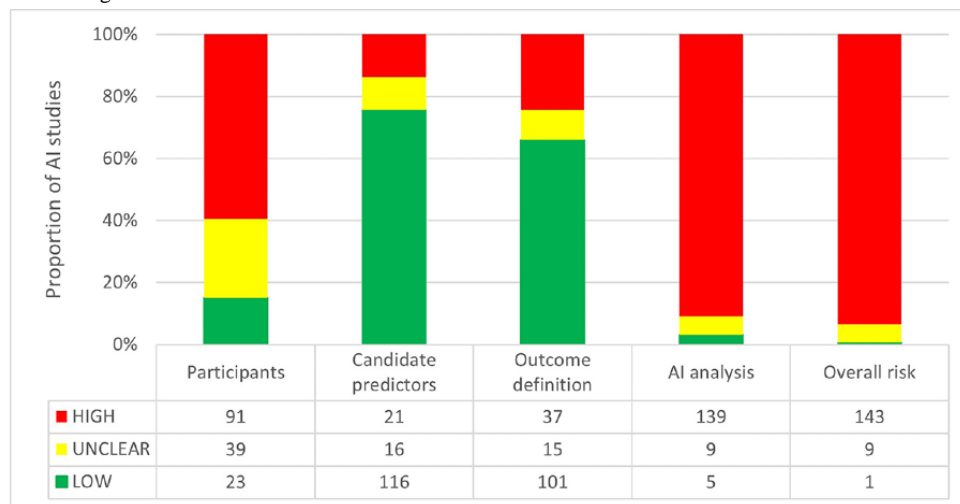
The risk of introducing bias is defined in Appendix S3 in Multimedia Appendix 1. Only 23 studies were found unlikely to introduce bias due to the trial design and evaluation of participants, whereas in the majority of studies the risk was high or unclear. In most cases, the risk of bias due to participants and the trial features was introduced by bias in the distribution of participants—that is, inclusion and exclusion criteria, loss of follow-up, and participants withdrawal—or the sample was not enough to be considered a good example of the target population. The definition and collection of the candidate predictors were mostly a low risk of introducing bias (n=16), with a few studies possibly introducing bias (n=21) or having an unclear risk of bias (n=16). Results for the outcome definition in the AI model are similar, with most studies evaluated as having a low risk of bias (n=101). Some studies were categorized as unclear (n=15) or high risk (n=37) due to unclear definitions of outcomes or combining data set of different populations whose outcomes were evaluated with different methods. The most important risk of bias was found when applying AI algorithms and their evaluation. Only a few studies were evaluated as unlikely to introduce some bias (n=5) and

the vast majority of AI analysis introduced a high risk of bias (n=139), while 9 could not be assessed properly (n=9). The main issues for bias in the AI analysis were not appropriately preprocessing and arranging the data for the specifications of the applied AI model, a bad handling of missingness, or an insufficient validation of the performance to account for overfitting and optimism (Figure 4). Appendix S5 in Multimedia Appendix 1 shows a stratified analysis of the risk of bias based on disorders, study designs, and outcome.

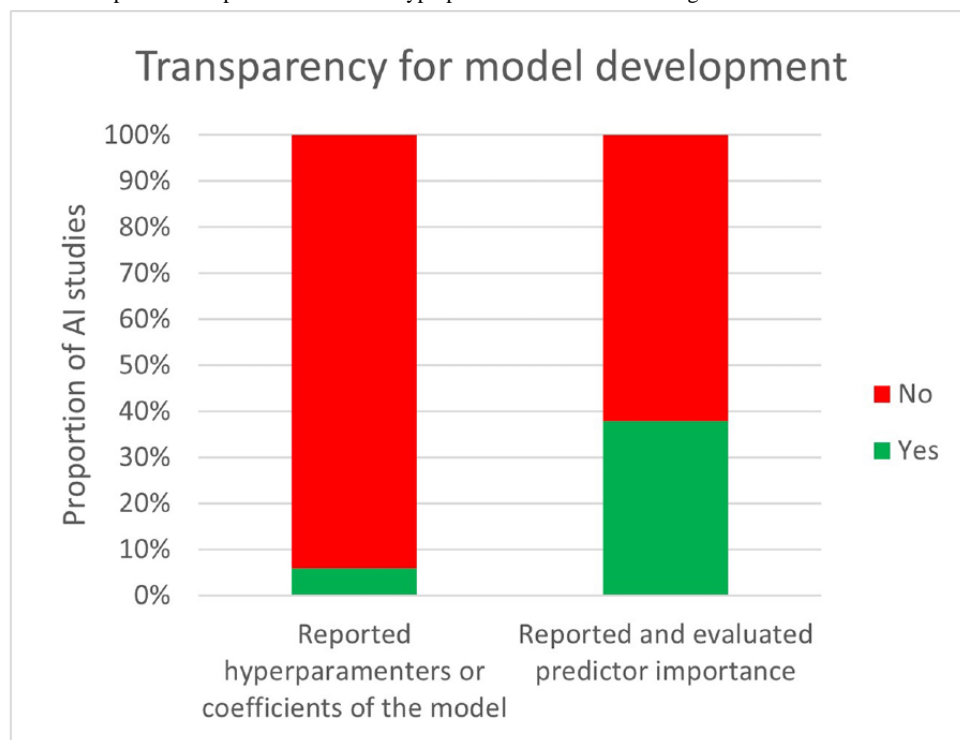
Overall, only 1 AI study could be assessed as a low risk of being biased. The most contributing categories to the overall risk were Participants and AI analysis. Most studies were likely unbiased about the definition and collection of predictors and the outcome but they failed to apply these data later in the model—bad data engineering or bad validation of the models—or the trial design had some flaws from the beginning. It is worth mentioning that only 9 of the 153 models reported any hyperparameter tuning or coefficients of the models and most of them reported basic trees models coefficients and decision rules. Only 58 studies mentioned or reported predictor importance and less than one-half reported the ranking and evaluated the methodology to test it (Figure 5).



**Figure 4.** Analysis of the risk of bias following PROBAST (Prediction Model Risk of Bias Assessment Tool) categories as defined in [Multimedia Appendix 1](#). AI: artificial intelligence.



**Figure 5.** Reporting of candidate predictor importance as well as hyperparameters for model tuning and coefficients of models. AI: artificial intelligence.



## Discussion

### Principal Findings

This overview summarizes the development and status of AI applications in mental health care. The review focused on the period from 2016 to 2021 to understand the latest advancements of these applications in mental health research, including aspects related to the methodological quality, the risk of bias, and the transparency. Results may be limited by the keywords applied in the search queries. This analysis is only representative for the records retrieved in this search. However, the samples analyzed may be sufficient to judge the quality and current status of this field. Significant methodological flaws were found involving the statistical processes of AI applications and data preprocessing. Many studies overlook evaluating or reporting

on data quality and how it fits the model. Some studies applied several AI techniques, here aggregated as “compelling models” studies, to select the most efficient technique without assessing their suitability for the problem they face, which may lead to overoptimism. Further, missing data management and the recruitment of participants are rarely reported, making it difficult to account for the risk of model overfitting. Preprocessing pipelines are not sufficiently reported, which hampers the reproducibility of the study or the adaptation of the AI techniques to the specific type of study. The use of reporting guidelines, such as the CONSORT-AI (AI version of the Consolidated Standards of Reporting Trials) for clinical trial reports involving AI [143], the SPIRIT-AI (AI version of the Standard Protocol Items: Recommendations for Interventional Trials) for clinical trial protocols involving AI [144], or the MI-CLAIM (Minimum Information About Clinical Artificial

Intelligence Modeling) checklist on minimum information for clinical AI modeling [145], would be very useful to ensure that the basic information about the study design and implementation is reported in these types of studies.

Some predictive models are being updated for AI reporting, such as TRIPOD-AI (AI version of the Transparent Reporting of a Multivariable Prediction Model of Individual Prognosis or Diagnosis) and PROBAST-AI (ie, the AI version of the PROBAST) [146]. However, these guidelines are rarely followed in the reviewed records. They lack transparent reporting on AI model features such as coefficients, hyperparameters, or predictor selection. Encouraging transparent reporting should be prioritized, as it would benefit second external validations and provide better accountability for reported models.

The incorporation of AI in mental health research is unbalanced between ICD-11 mental health disorders. Most research focusses on depressive disorders, where it combines severity questionnaires and scales with electronic health record and psychotic disorders using medical image data. External validation is very uncommon. Conducting suitable trial designs for the intended AI outcomes is understandably difficult in terms of money, time, and resources. Thus, it is common to apply data collected retrospectively. However, the original trial designs do not fit the specifications for AI development and most studies do not assess the appropriateness of these data. Notably, many authors may not understand the need to ensure an optimal preprocessing pipeline. In these cases, authors are aware of the poor performance of the models, but the proposed approach for improvement is suggested directly from a trial perspective rather than from assessing possible statistical bias or mistakes in model development, which could save cost and time over designing new studies.

## Challenges

AI studies were analyzed to identify challenges and opportunities involving the use of AI in mental health. Typically, AI studies reported insufficient samples to ensure model generalizability [68,84,103]. Several authors reported bias because of the difficulty in adapting typical trial designs to an AI context. For example, some authors detail the constraining boundaries for selecting participants in randomized clinical trials as a limiting factor, which reduces the sample size and could overlook confounders [68,90]. Most randomized clinical trials noted possible variance between the collected data and the real-world data. However, observational studies can also introduce bias in AI models if the imbalance between cohorts is not adequately addressed [84,128]. In these studies, the variety in features such as prescribed medication could introduce confounders and bias that are difficult to manage [94]. Furthermore, in long-term studies, lack of follow-up or other conditions leading to a decrease of patients is an important limitation, mostly for prognosis studies or predictive evolution of condition severity [30,58]. These issues are worse for retrospective collection of data, where trial designs tend to diverge from the problematics of AI. Besides, some authors are aware of bias due to gaps, but most did not properly evaluate this risk.

A noticeable lack of internationalization was detected. Many studies focused only on local data, which contributes to small sample sizes and poor generalizability [115,127]. Encouraging partnerships and collaborations across countries and centers should be a priority, as it could facilitate external validation [71]. Some authors mention difficulties reconciling clinical practices with AI study requirements, usually due to ethics problems related to clinical practice in patients that can overlook confounders, that is, making it difficult to apply placebo controls in some interventions [82,115].

Another challenge is the explainability of complex AI models, which could make researchers reluctant to adopt techniques that map high-order interactions or “black-box” algorithms [81,122]. Researchers prefer simpler algorithms. The few studies that reported model coefficients and some explanation used decision trees. Another challenge is that contradictory findings could occur among studies [85].

Finally, some authors are aware of the opportunities that everyday devices and platforms such as phones and social networks offer but find it difficult to take advantage of these tools due to lack of standardization, which reduces the target population for defining a study [92].

## Opportunities

Some studies introduced devices and platforms to improve the monitoring of patients. The application of everyday digital tools could reduce necessary resources and therefore facilitate data collection [99,127]. Promoting the use of frequently used devices combined with the application of AI models seems like a future trend that could improve the treatment of many conditions where the chance of treatment response decreases over time [126]. Further, it opens possibilities of internet-based treatments that could be conducted in real time with digital technology, easing the load on hospitals [99].

Data sharing and public databases should be encouraged to develop and implement more trustworthy AI models. AI models from clinical stage to clinical practice could be difficult but powerful tools to gain insights into predictor collection, human-based decisions, and AI biases while these techniques are being implemented in clinical world. Many studies report the high potential of AI in mental health for clinical support, computer-aided systems, and possibly preliminary screening [94,127].

Currently, many guidelines and initiatives exist to which researchers could adhere to in order to increase transparency and better use AI models. Currently, the EQUATOR (Enhancing the Quality and Transparency of Health Research) network initiative reports useful guidelines that could foster collaboration and implementation [147].

## Conclusion

AI algorithms are increasingly being incorporated into mental health research; however, it is still uneven between ICD-11 categories. Collaboration is merely anecdotal, and data and developed models mostly remain private. Significant methodological flaws exist involving the statistical process of AI applications and data preprocessing pipelines. Only 1 study

was found reporting second validation, and 13.7% (21/153) reported external validation. The evaluation of the risk of bias and transparent reporting was discouraging. Model hyperparameters or trained coefficients are rarely reported, nor are insights about the explainability of the AI models. The lack of transparency and methodological flaws are concerning, as

they delay the safe, practical implementation of AI. Furthermore, data engineering for AI models seems to be overlooked or misunderstood, and data are often not adequately managed. These significant shortcomings may indicate overly accelerated promotion of new AI models without pausing to assess their real-world viability.

---

## Acknowledgments

This study was funded by the Division of Country Health Policies and Systems, Regional Office for Europe of the World Health Organization.

---

## Disclaimer

DNO, NA-M and LL are staff members of the World Health Organization. The authors alone are responsible for the views expressed in this paper, and they do not necessarily represent the decisions, policies, or views of the World Health Organization.

---

## Conflicts of Interest

None declared.

---

## Multimedia Appendix 1

Search string queries; description of categories and indicators; evaluation of risk of bias; data collection for AI applications; and risk of bias stratified by disorders, study designs, and outcome. AI: artificial intelligence.

[[DOCX File , 611 KB-Multimedia Appendix 1](#)]

---

## Multimedia Appendix 2

Analysis results.

[[XLSX File \(Microsoft Excel File\), 81 KB-Multimedia Appendix 2](#)]

---

## References

1. The Pan-European Mental Health Coalition. WHO. URL: <https://www.euro.who.int/en/health-topics/health-policy/european-programme-of-work/flagship-initiatives/the-pan-european-mental-health-coalition> [accessed 2023-01-23]
2. WHO Regional Office for Europe. European Programme of Work, 2020–2025. WHO. 2020. URL: <https://www.who.int/europe/about-us/our-work/european-programme-of-work> [accessed 2023-01-23]
3. Reddy S, Fox J, Purohit MP. Artificial intelligence-enabled healthcare delivery. *J R Soc Med* 2019 Jan;112(1):22-28 [FREE Full text] [doi: [10.1177/0141076818815510](https://doi.org/10.1177/0141076818815510)] [Medline: [30507284](https://pubmed.ncbi.nlm.nih.gov/30507284/)]
4. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med* 2019 Jan;25(1):30-36 [FREE Full text] [doi: [10.1038/s41591-018-0307-0](https://doi.org/10.1038/s41591-018-0307-0)] [Medline: [30617336](https://pubmed.ncbi.nlm.nih.gov/30617336/)]
5. Borges do Nascimento IJ, Marcolino MS, Abdulazeem HM, Weerasekara I, Azzopardi-Muscat N, Gonçalves MA, et al. Impact of Big Data Analytics on People's Health: Overview of Systematic Reviews and Recommendations for Future Studies. *J Med Internet Res* 2021 Apr 13;23(4):e27275 [FREE Full text] [doi: [10.2196/27275](https://doi.org/10.2196/27275)] [Medline: [33847586](https://pubmed.ncbi.nlm.nih.gov/33847586/)]
6. Andaur Navarro CL, Damen JAA, Takada T, Nijman SWJ, Dhiman P, Ma J, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. *BMJ* 2021 Oct 20;375:n2281 [FREE Full text] [doi: [10.1136/bmj.n2281](https://doi.org/10.1136/bmj.n2281)] [Medline: [34670780](https://pubmed.ncbi.nlm.nih.gov/34670780/)]
7. Felzmann H, Fosch-Villaronga E, Lutz C, Tamò-Larrieux A. Towards Transparency by Design for Artificial Intelligence. *Sci Eng Ethics* 2020 Dec 16;26(6):3333-3361 [FREE Full text] [doi: [10.1007/s11948-020-00276-4](https://doi.org/10.1007/s11948-020-00276-4)] [Medline: [33196975](https://pubmed.ncbi.nlm.nih.gov/33196975/)]
8. Vollmer S, Mateen BA, Bohner G, Király FJ, Ghani R, Jonsson P, et al. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. *BMJ* 2020 Mar 20;368:l6927 [FREE Full text] [doi: [10.1136/bmj.l6927](https://doi.org/10.1136/bmj.l6927)] [Medline: [32198138](https://pubmed.ncbi.nlm.nih.gov/32198138/)]
9. 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. *Int J Surg* 2021 Apr;88:105906. [doi: [10.1016/j.ijssu.2021.105906](https://doi.org/10.1016/j.ijssu.2021.105906)] [Medline: [33789826](https://pubmed.ncbi.nlm.nih.gov/33789826/)]
10. Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* 2014 Oct 14;11(10):e1001744 [FREE Full text] [doi: [10.1371/journal.pmed.1001744](https://doi.org/10.1371/journal.pmed.1001744)] [Medline: [25314315](https://pubmed.ncbi.nlm.nih.gov/25314315/)]
11. Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.3 (updated February 2022). London, UK: Cochrane; 2022.

12. Wolff RF, Moons KG, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med* 2019 Jan 01;170(1):51. [doi: [10.7326/m18-1376](https://doi.org/10.7326/m18-1376)]
13. Higgins J. *Cochrane Handbook for Systematic Reviews of Interventions Version 5*. The Cochrane Collaboration. London, UK: Cochrane Collaboration; 2011. URL: <https://training.cochrane.org/handbook> [accessed 2023-01-24]
14. Caye A, Agnew-Blais J, Arseneault L, Gonçalves H, Kieling C, Langley K, et al. A risk calculator to predict adult attention-deficit/hyperactivity disorder: generation and external validation in three birth cohorts and one clinical sample. *Epidemiol Psychiatr Sci* 2019 May 15;29:e37. [doi: [10.1017/s2045796019000283](https://doi.org/10.1017/s2045796019000283)]
15. de Pierrefeu A, Löfstedt T, Laidi C, Hadj-Seleem F, Bourgin J, Hajek T, et al. Identifying a neuroanatomical signature of schizophrenia, reproducible across sites and stages, using machine learning with structured sparsity. *Acta Psychiatr Scand* 2018 Dec 21;138(6):571-580. [doi: [10.1111/acps.12964](https://doi.org/10.1111/acps.12964)] [Medline: [30242828](https://pubmed.ncbi.nlm.nih.gov/30242828/)]
16. Vieira S, Gong QY, Pinaya WHL, Scarpazza C, Tognin S, Crespo-Facorro B, et al. Using Machine Learning and Structural Neuroimaging to Detect First Episode Psychosis: Reconsidering the Evidence. *Schizophr Bull* 2020 Jan 04;46(1):17-26 [FREE Full text] [doi: [10.1093/schbul/sby189](https://doi.org/10.1093/schbul/sby189)] [Medline: [30809667](https://pubmed.ncbi.nlm.nih.gov/30809667/)]
17. Bae Y, Kumarasamy K, Ali IM, Korfiatis P, Akkus Z, Erickson BJ. Differences Between Schizophrenic and Normal Subjects Using Network Properties from fMRI. *J Digit Imaging* 2018 Apr 18;31(2):252-261 [FREE Full text] [doi: [10.1007/s10278-017-0020-4](https://doi.org/10.1007/s10278-017-0020-4)] [Medline: [28924878](https://pubmed.ncbi.nlm.nih.gov/28924878/)]
18. Chand GB, Dwyer DB, Erus G, Sotiras A, Varol E, Srinivasan D, et al. Two distinct neuroanatomical subtypes of schizophrenia revealed using machine learning. *Brain* 2020 Mar 01;143(3):1027-1038 [FREE Full text] [doi: [10.1093/brain/awaa025](https://doi.org/10.1093/brain/awaa025)] [Medline: [32103250](https://pubmed.ncbi.nlm.nih.gov/32103250/)]
19. Koutsouleris N, Wobrock T, Guse B, Langguth B, Landgrebe M, Eichhammer P, et al. Predicting Response to Repetitive Transcranial Magnetic Stimulation in Patients With Schizophrenia Using Structural Magnetic Resonance Imaging: A Multisite Machine Learning Analysis. *Schizophr Bull* 2018 Aug 20;44(5):1021-1034 [FREE Full text] [doi: [10.1093/schbul/sbx114](https://doi.org/10.1093/schbul/sbx114)] [Medline: [28981875](https://pubmed.ncbi.nlm.nih.gov/28981875/)]
20. Anderson JP, Icten Z, Alas V, Benson C, Joshi K. Comparison and predictors of treatment adherence and remission among patients with schizophrenia treated with paliperidone palmitate or atypical oral antipsychotics in community behavioral health organizations. *BMC Psychiatry* 2017 Oct 18;17(1):346 [FREE Full text] [doi: [10.1186/s12888-017-1507-8](https://doi.org/10.1186/s12888-017-1507-8)] [Medline: [29047368](https://pubmed.ncbi.nlm.nih.gov/29047368/)]
21. Chung Y, Addington J, Bearden CE, Cadenhead K, Cornblatt B, Mathalon DH, North American Prodrome Longitudinal Study (NAPLS) Consortium, the Pediatric Imaging, Neurocognition, Genetics (PING) Study Consortium. Use of Machine Learning to Determine Deviance in Neuroanatomical Maturity Associated With Future Psychosis in Youths at Clinically High Risk. *JAMA Psychiatry* 2018 Sep 01;75(9):960-968 [FREE Full text] [doi: [10.1001/jamapsychiatry.2018.1543](https://doi.org/10.1001/jamapsychiatry.2018.1543)] [Medline: [29971330](https://pubmed.ncbi.nlm.nih.gov/29971330/)]
22. Dillon K, Calhoun V, Wang Y. A robust sparse-modeling framework for estimating schizophrenia biomarkers from fMRI. *J Neurosci Methods* 2017 Jan 30;276:46-55 [FREE Full text] [doi: [10.1016/j.jneumeth.2016.11.005](https://doi.org/10.1016/j.jneumeth.2016.11.005)] [Medline: [27867012](https://pubmed.ncbi.nlm.nih.gov/27867012/)]
23. Schreiner M, Forsyth JK, Karlsgodt KH, Anderson AE, Hirsh N, Kushan L, et al. Intrinsic Connectivity Network-Based Classification and Detection of Psychotic Symptoms in Youth With 22q11.2 Deletions. *Cereb Cortex* 2017 Jun 01;27(6):3294-3306 [FREE Full text] [doi: [10.1093/cercor/bhx076](https://doi.org/10.1093/cercor/bhx076)] [Medline: [28383675](https://pubmed.ncbi.nlm.nih.gov/28383675/)]
24. Lefort-Besnard J, Varoquaux G, Derntl B, Gruber O, Aleman A, Jardri R, et al. Patterns of schizophrenia symptoms: hidden structure in the PANSS questionnaire. *Transl Psychiatry* 2018 Oct 30;8(1):237 [FREE Full text] [doi: [10.1038/s41398-018-0294-4](https://doi.org/10.1038/s41398-018-0294-4)] [Medline: [30375374](https://pubmed.ncbi.nlm.nih.gov/30375374/)]
25. Brodey B, Girgis R, Favorov O, Bearden C, Woods S, Addington J, et al. The Early Psychosis Screener for Internet (EPSI)-SR: Predicting 12 month psychotic conversion using machine learning. *Schizophr Res* 2019 Jun;208:390-396 [FREE Full text] [doi: [10.1016/j.schres.2019.01.015](https://doi.org/10.1016/j.schres.2019.01.015)] [Medline: [30777603](https://pubmed.ncbi.nlm.nih.gov/30777603/)]
26. Koutsouleris N, Kahn RS, Chekroud AM, Leucht S, Falkai P, Wobrock T, et al. Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. *The Lancet Psychiatry* 2016 Oct;3(10):935-946. [doi: [10.1016/s2215-0366\(16\)30171-7](https://doi.org/10.1016/s2215-0366(16)30171-7)]
27. Nieuwenhuis M, Schnack HG, van Haren NE, Lappin J, Morgan C, Reinders AA, et al. Multi-center MRI prediction models: Predicting sex and illness course in first episode psychosis patients. *Neuroimage* 2017 Jan 15;145(Pt B):246-253 [FREE Full text] [doi: [10.1016/j.neuroimage.2016.07.027](https://doi.org/10.1016/j.neuroimage.2016.07.027)] [Medline: [27421184](https://pubmed.ncbi.nlm.nih.gov/27421184/)]
28. Rozycki M, Satterthwaite TD, Koutsouleris N, Erus G, Doshi J, Wolf DH, et al. Multisite Machine Learning Analysis Provides a Robust Structural Imaging Signature of Schizophrenia Detectable Across Diverse Patient Populations and Within Individuals. *Schizophr Bull* 2018 Aug 20;44(5):1035-1044 [FREE Full text] [doi: [10.1093/schbul/sbx137](https://doi.org/10.1093/schbul/sbx137)] [Medline: [29186619](https://pubmed.ncbi.nlm.nih.gov/29186619/)]
29. Chen J, Zang Z, Braun U, Schwarz K, Harneit A, Kremer T, et al. Association of a Reproducible Epigenetic Risk Profile for Schizophrenia With Brain Methylation and Function. *JAMA Psychiatry* 2020 Jun 01;77(6):628-636 [FREE Full text] [doi: [10.1001/jamapsychiatry.2019.4792](https://doi.org/10.1001/jamapsychiatry.2019.4792)] [Medline: [32049268](https://pubmed.ncbi.nlm.nih.gov/32049268/)]
30. Fond G, Bulzacka E, Boucekine M, Schürhoff F, Berna F, Godin O, FACE-SZ (FondaMental Academic Centers of Expertise for Schizophrenia) group, et al. Machine learning for predicting psychotic relapse at 2 years in schizophrenia in the national

- FACE-SZ cohort. *Prog Neuropsychopharmacol Biol Psychiatry* 2019 Jun 08;92:8-18. [doi: [10.1016/j.pnpbp.2018.12.005](https://doi.org/10.1016/j.pnpbp.2018.12.005)] [Medline: [30552914](https://pubmed.ncbi.nlm.nih.gov/30552914/)]
31. Gibbons RD, Chattopadhyay I, Meltzer HY, Kane JM, Guinart D. Development of a computerized adaptive diagnostic screening tool for psychosis. *Schizophr Res* 2022 Jul;245:116-121. [doi: [10.1016/j.schres.2021.03.020](https://doi.org/10.1016/j.schres.2021.03.020)] [Medline: [33836922](https://pubmed.ncbi.nlm.nih.gov/33836922/)]
  32. Martinuzzi E, Barbosa S, Daoudlarian D, Bel Haj Ali W, Gilet C, Fillatre L, OPTiMiSE Study Group. Stratification and prediction of remission in first-episode psychosis patients: the OPTiMiSE cohort study. *Transl Psychiatry* 2019 Jan 17;9(1):20 [FREE Full text] [doi: [10.1038/s41398-018-0366-5](https://doi.org/10.1038/s41398-018-0366-5)] [Medline: [30655509](https://pubmed.ncbi.nlm.nih.gov/30655509/)]
  33. Skåtun KC, Kaufmann T, Doan NT, Alnæs D, Córdova-Palomera A, Jönsson EG, KaSP, et al. Consistent Functional Connectivity Alterations in Schizophrenia Spectrum Disorder: A Multisite Study. *Schizophr Bull* 2017 Jul 01;43(4):914-924 [FREE Full text] [doi: [10.1093/schbul/sbw145](https://doi.org/10.1093/schbul/sbw145)] [Medline: [27872268](https://pubmed.ncbi.nlm.nih.gov/27872268/)]
  34. Busk J, Faurholt-Jepsen M, Frost M, Bardram JE, Vedel Kessing L, Winther O. Forecasting Mood in Bipolar Disorder From Smartphone Self-assessments: Hierarchical Bayesian Approach. *JMIR Mhealth Uhealth* 2020 Apr 01;8(4):e15028 [FREE Full text] [doi: [10.2196/15028](https://doi.org/10.2196/15028)] [Medline: [32234702](https://pubmed.ncbi.nlm.nih.gov/32234702/)]
  35. Sun ZY, Houenou J, Duclap D, Sarrazin S, Linke J, Daban C, et al. Shape analysis of the cingulum, uncinata and arcuate fasciculi in patients with bipolar disorder. *J Psychiatry Neurosci* 2017 Jan;42(1):27-36 [FREE Full text] [doi: [10.1503/jpn.150291](https://doi.org/10.1503/jpn.150291)] [Medline: [28234596](https://pubmed.ncbi.nlm.nih.gov/28234596/)]
  36. Kim TT, Dufour S, Xu C, Cohen ZD, Sylvia L, Deckersbach T, et al. Predictive modeling for response to lithium and quetiapine in bipolar disorder. *Bipolar Disord* 2019 Aug;21(5):428-436. [doi: [10.1111/bdi.12752](https://doi.org/10.1111/bdi.12752)] [Medline: [30729637](https://pubmed.ncbi.nlm.nih.gov/30729637/)]
  37. Lei D, Li W, Tallman MJ, Patino LR, McNamara RK, Strawn JR, et al. Changes in the brain structural connectome after a prospective randomized clinical trial of lithium and quetiapine treatment in youth with bipolar disorder. *Neuropsychopharmacology* 2021 Jun;46(7):1315-1323 [FREE Full text] [doi: [10.1038/s41386-021-00989-5](https://doi.org/10.1038/s41386-021-00989-5)] [Medline: [33753882](https://pubmed.ncbi.nlm.nih.gov/33753882/)]
  38. Brown EC, Clark DL, Forkert ND, Molnar CP, Kiss ZHT, Ramasubbu R. Metabolic activity in subcallosal cingulate predicts response to deep brain stimulation for depression. *Neuropsychopharmacology* 2020 Sep;45(10):1681-1688 [FREE Full text] [doi: [10.1038/s41386-020-0745-5](https://doi.org/10.1038/s41386-020-0745-5)] [Medline: [32580207](https://pubmed.ncbi.nlm.nih.gov/32580207/)]
  39. Carrillo F, Sigman M, Fernández Slezak D, Ashton P, Fitzgerald L, Stroud J, et al. Natural speech algorithm applied to baseline interview data can predict which patients will respond to psilocybin for treatment-resistant depression. *J Affect Disord* 2018 Apr 01;230:84-86. [doi: [10.1016/j.jad.2018.01.006](https://doi.org/10.1016/j.jad.2018.01.006)] [Medline: [29407543](https://pubmed.ncbi.nlm.nih.gov/29407543/)]
  40. Sikora M, Heffernan J, Avery ET, Mickey BJ, Zubieta J, Peciña M. Saliency Network Functional Connectivity Predicts Placebo Effects in Major Depression. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2016 Jan;1(1):68-76 [FREE Full text] [doi: [10.1016/j.bpsc.2015.10.002](https://doi.org/10.1016/j.bpsc.2015.10.002)] [Medline: [26709390](https://pubmed.ncbi.nlm.nih.gov/26709390/)]
  41. Al-Kaysi AM, Al-Ani A, Loo CK, Powell TY, Martin DM, Breakspear M, et al. Predicting tDCS treatment outcomes of patients with major depressive disorder using automated EEG classification. *J Affect Disord* 2017 Jan 15;208:597-603. [doi: [10.1016/j.jad.2016.10.021](https://doi.org/10.1016/j.jad.2016.10.021)] [Medline: [28029427](https://pubmed.ncbi.nlm.nih.gov/28029427/)]
  42. Athreya A, Iyer R, Neavin D, Wang L, Weinshilboum R, Kaddurah-Daouk R, et al. Augmentation of Physician Assessments with Multi-Omics Enhances Predictability of Drug Response: A Case Study of Major Depressive Disorder. *IEEE Comput Intell Mag* 2018 Aug;13(3):20-31. [doi: [10.1109/mci.2018.2840660](https://doi.org/10.1109/mci.2018.2840660)]
  43. Bai R, Xiao L, Guo Y, Zhu X, Li N, Wang Y, et al. Tracking and Monitoring Mood Stability of Patients With Major Depressive Disorder by Machine Learning Models Using Passive Digital Data: Prospective Naturalistic Multicenter Study. *JMIR Mhealth Uhealth* 2021 Mar 08;9(3):e24365 [FREE Full text] [doi: [10.2196/24365](https://doi.org/10.2196/24365)] [Medline: [33683207](https://pubmed.ncbi.nlm.nih.gov/33683207/)]
  44. Bao Z, Zhao X, Li J, Zhang G, Wu H, Ning Y, et al. Prediction of repeated-dose intravenous ketamine response in major depressive disorder using the GWAS-based machine learning approach. *J Psychiatr Res* 2021 Jun;138:284-290. [doi: [10.1016/j.jpsychires.2021.04.014](https://doi.org/10.1016/j.jpsychires.2021.04.014)] [Medline: [33878621](https://pubmed.ncbi.nlm.nih.gov/33878621/)]
  45. Bartlett EA, DeLorenzo C, Sharma P, Yang J, Zhang M, Petkova E, et al. Pretreatment and early-treatment cortical thickness is associated with SSRI treatment response in major depressive disorder. *Neuropsychopharmacology* 2018 Oct 19;43(11):2221-2230 [FREE Full text] [doi: [10.1038/s41386-018-0122-9](https://doi.org/10.1038/s41386-018-0122-9)] [Medline: [29955151](https://pubmed.ncbi.nlm.nih.gov/29955151/)]
  46. Bremer V, Becker D, Kolovos S, Funk B, van Breda W, Hoogendoorn M, et al. Predicting Therapy Success and Costs for Personalized Treatment Recommendations Using Baseline Characteristics: Data-Driven Analysis. *J Med Internet Res* 2018 Aug 21;20(8):e10275 [FREE Full text] [doi: [10.2196/10275](https://doi.org/10.2196/10275)] [Medline: [30131318](https://pubmed.ncbi.nlm.nih.gov/30131318/)]
  47. Goerigk S, Hilbert S, Jobst A, Falkai P, Bühner M, Stachl C, et al. Predicting instructed simulation and dissimulation when screening for depressive symptoms. *Eur Arch Psychiatry Clin Neurosci* 2020 Mar 12;270(2):153-168. [doi: [10.1007/s00406-018-0967-2](https://doi.org/10.1007/s00406-018-0967-2)] [Medline: [30542818](https://pubmed.ncbi.nlm.nih.gov/30542818/)]
  48. Hopman H, Chan S, Chu W, Lu H, Tse C, Chau S, et al. Personalized prediction of transcranial magnetic stimulation clinical response in patients with treatment-refractory depression using neuroimaging biomarkers and machine learning. *J Affect Disord* 2021 Jul 01;290:261-271. [doi: [10.1016/j.jad.2021.04.081](https://doi.org/10.1016/j.jad.2021.04.081)] [Medline: [34010751](https://pubmed.ncbi.nlm.nih.gov/34010751/)]
  49. Monaro M, Toncini A, Ferracuti S, Tessari G, Vaccaro MG, De Fazio P, et al. The Detection of Malingering: A New Tool to Identify Made-Up Depression. *Front Psychiatry* 2018 Jun 8;9:249 [FREE Full text] [doi: [10.3389/fpsy.2018.00249](https://doi.org/10.3389/fpsy.2018.00249)] [Medline: [29937740](https://pubmed.ncbi.nlm.nih.gov/29937740/)]

50. Nie Z, Vairavan S, Narayan VA, Ye J, Li QS. Predictive modeling of treatment resistant depression using data from STAR\*D and an independent clinical study. *PLoS One* 2018 Jun 7;13(6):e0197268 [FREE Full text] [doi: [10.1371/journal.pone.0197268](https://doi.org/10.1371/journal.pone.0197268)] [Medline: [29879133](https://pubmed.ncbi.nlm.nih.gov/29879133/)]
51. Setoyama D, Kato TA, Hashimoto R, Kunugi H, Hattori K, Hayakawa K, et al. Plasma Metabolites Predict Severity of Depression and Suicidal Ideation in Psychiatric Patients-A Multicenter Pilot Analysis. *PLoS One* 2016 Dec 16;11(12):e0165267 [FREE Full text] [doi: [10.1371/journal.pone.0165267](https://doi.org/10.1371/journal.pone.0165267)] [Medline: [27984586](https://pubmed.ncbi.nlm.nih.gov/27984586/)]
52. Shepherd-Banigan M, Smith VA, Lindquist JH, Cary MP, Miller KEM, Chapman JG, et al. Identifying treatment effects of an informal caregiver education intervention to increase days in the community and decrease caregiver distress: a machine-learning secondary analysis of subgroup effects in the HI-FIVES randomized clinical trial. *Trials* 2020 Feb 14;21(1):189 [FREE Full text] [doi: [10.1186/s13063-020-4113-x](https://doi.org/10.1186/s13063-020-4113-x)] [Medline: [32059687](https://pubmed.ncbi.nlm.nih.gov/32059687/)]
53. Zhao M, Feng Z. Machine Learning Methods to Evaluate the Depression Status of Chinese Recruits: A Diagnostic Study. *NDT* 2020 Nov;16:2743-2752. [doi: [10.2147/ndt.s275620](https://doi.org/10.2147/ndt.s275620)]
54. Schultebrucks K, Choi KW, Galatzer-Levy IR, Bonanno GA. Discriminating Heterogeneous Trajectories of Resilience and Depression After Major Life Stressors Using Polygenic Scores. *JAMA Psychiatry* 2021 Jul 01;78(7):744-752 [FREE Full text] [doi: [10.1001/jamapsychiatry.2021.0228](https://doi.org/10.1001/jamapsychiatry.2021.0228)] [Medline: [33787853](https://pubmed.ncbi.nlm.nih.gov/33787853/)]
55. Castellani B, Griffiths F, Rajaram R, Gunn J. Exploring comorbid depression and physical health trajectories: A case-based computational modelling approach. *J Eval Clin Pract* 2018 Dec 02;24(6):1293-1309. [doi: [10.1111/jep.13042](https://doi.org/10.1111/jep.13042)] [Medline: [30277297](https://pubmed.ncbi.nlm.nih.gov/30277297/)]
56. Jacobson NC, Nemesure MD. Using Artificial Intelligence to Predict Change in Depression and Anxiety Symptoms in a Digital Intervention: Evidence from a Transdiagnostic Randomized Controlled Trial. *Psychiatry Res* 2021 Jan;295:113618 [FREE Full text] [doi: [10.1016/j.psychres.2020.113618](https://doi.org/10.1016/j.psychres.2020.113618)] [Medline: [33278743](https://pubmed.ncbi.nlm.nih.gov/33278743/)]
57. Kessler RC, Furukawa TA, Kato T, Luedtke A, Petukhova M, Sadikova E, et al. An individualized treatment rule to optimize probability of remission by continuation, switching, or combining antidepressant medications after failing a first-line antidepressant in a two-stage randomized trial. *Psychol Med* 2021 Mar 08;52(15):3371-3380. [doi: [10.1017/s0033291721000027](https://doi.org/10.1017/s0033291721000027)]
58. Bondar J, Caye A, Chekroud AM, Kieling C. Symptom clusters in adolescent depression and differential response to treatment: a secondary analysis of the Treatment for Adolescents with Depression Study randomised trial. *The Lancet Psychiatry* 2020 Apr;7(4):337-343. [doi: [10.1016/s2215-0366\(20\)30060-2](https://doi.org/10.1016/s2215-0366(20)30060-2)]
59. Kato M, Asami Y, Wajsbrot DB, Wang X, Boucher M, Prieto R, et al. Clustering patients by depression symptoms to predict venlafaxine ER antidepressant efficacy: Individual patient data analysis. *J Psychiatr Res* 2020 Oct;129:160-167 [FREE Full text] [doi: [10.1016/j.jpsychires.2020.06.011](https://doi.org/10.1016/j.jpsychires.2020.06.011)] [Medline: [32912597](https://pubmed.ncbi.nlm.nih.gov/32912597/)]
60. Bueno ML, Hommersom A, Lucas PJ, Janzing J. A probabilistic framework for predicting disease dynamics: A case study of psychotic depression. *J Biomed Inform* 2019 Jul;95:103232 [FREE Full text] [doi: [10.1016/j.jbi.2019.103232](https://doi.org/10.1016/j.jbi.2019.103232)] [Medline: [31201965](https://pubmed.ncbi.nlm.nih.gov/31201965/)]
61. Alexopoulos GS, Raue PJ, Banerjee S, Mauer E, Marino P, Soliman M, et al. Modifiable predictors of suicidal ideation during psychotherapy for late-life major depression. A machine learning approach. *Transl Psychiatry* 2021 Oct 18;11(1):536 [FREE Full text] [doi: [10.1038/s41398-021-01656-5](https://doi.org/10.1038/s41398-021-01656-5)] [Medline: [34663787](https://pubmed.ncbi.nlm.nih.gov/34663787/)]
62. Solomonov N, Lee J, Banerjee S, Flückiger C, Kanellopoulos D, Gunning FM, et al. Modifiable predictors of nonresponse to psychotherapies for late-life depression with executive dysfunction: a machine learning approach. *Mol Psychiatry* 2021 Sep 10;26(9):5190-5198 [FREE Full text] [doi: [10.1038/s41380-020-0836-z](https://doi.org/10.1038/s41380-020-0836-z)] [Medline: [32651477](https://pubmed.ncbi.nlm.nih.gov/32651477/)]
63. Bruijniks SJ, DeRubeis RJ, Lemmens LH, Peeters FP, Cuijpers P, Huibers MJ. The relation between therapy quality, therapy processes and outcomes and identifying for whom therapy quality matters in CBT and IPT for depression. *Behav Res Ther* 2021 Jan 28;139:103815. [doi: [10.1016/j.brat.2021.103815](https://doi.org/10.1016/j.brat.2021.103815)] [Medline: [33581481](https://pubmed.ncbi.nlm.nih.gov/33581481/)]
64. Di Y, Wang J, Li W, Zhu T. Using i-vectors from voice features to identify major depressive disorder. *J Affect Disord* 2021 Jun 01;288:161-166. [doi: [10.1016/j.jad.2021.04.004](https://doi.org/10.1016/j.jad.2021.04.004)] [Medline: [33895418](https://pubmed.ncbi.nlm.nih.gov/33895418/)]
65. Gunlicks-Stoessel M, Klimes-Dougan B, VanZomeran A, Ma S. Developing a data-driven algorithm for guiding selection between cognitive behavioral therapy, fluoxetine, and combination treatment for adolescent depression. *Transl Psychiatry* 2020 Sep 21;10(1):321 [FREE Full text] [doi: [10.1038/s41398-020-01005-y](https://doi.org/10.1038/s41398-020-01005-y)] [Medline: [32958758](https://pubmed.ncbi.nlm.nih.gov/32958758/)]
66. Iniesta R, Malki K, Maier W, Rietschel M, Mors O, Hauser J, et al. Combining clinical variables to optimize prediction of antidepressant treatment outcomes. *J Psychiatr Res* 2016 Jul;78:94-102 [FREE Full text] [doi: [10.1016/j.jpsychires.2016.03.016](https://doi.org/10.1016/j.jpsychires.2016.03.016)] [Medline: [27089522](https://pubmed.ncbi.nlm.nih.gov/27089522/)]
67. Liu Y, Admon R, Mellem MS, Belleau EL, Kaiser RH, Clegg R, et al. Machine Learning Identifies Large-Scale Reward-Related Activity Modulated by Dopaminergic Enhancement in Major Depression. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 2020 Feb;5(2):163-172. [doi: [10.1016/j.bpsc.2019.10.002](https://doi.org/10.1016/j.bpsc.2019.10.002)]
68. Maglanoc LA, Kaufmann T, Jonassen R, Hilland E, Beck D, Landrø NI, et al. Multimodal fusion of structural and functional brain imaging in depression using linked independent component analysis. *Hum Brain Mapp* 2020 Jan;41(1):241-255 [FREE Full text] [doi: [10.1002/hbm.24802](https://doi.org/10.1002/hbm.24802)] [Medline: [31571370](https://pubmed.ncbi.nlm.nih.gov/31571370/)]
69. Mennen AC, Turk-Browne NB, Wallace G, Seok D, Jaganjac A, Stock J, et al. Cloud-Based Functional Magnetic Resonance Imaging Neurofeedback to Reduce the Negative Attentional Bias in Depression: A Proof-of-Concept Study. *Biol Psychiatry*

- Cogn Neurosci Neuroimaging 2021 Apr;6(4):490-497 [FREE Full text] [doi: [10.1016/j.bpsc.2020.10.006](https://doi.org/10.1016/j.bpsc.2020.10.006)] [Medline: [33422469](https://pubmed.ncbi.nlm.nih.gov/33422469/)]
70. Salvetat N, Van der Laan S, Vire B, Chimienti F, Cleophax S, Bronowicki JP, et al. RNA editing blood biomarkers for predicting mood alterations in HCV patients. *J Neurovirol* 2019 Dec 22;25(6):825-836 [FREE Full text] [doi: [10.1007/s13365-019-00772-9](https://doi.org/10.1007/s13365-019-00772-9)] [Medline: [31332697](https://pubmed.ncbi.nlm.nih.gov/31332697/)]
  71. van Bronswijk SC, DeRubeis RJ, Lemmens LHJM, Peeters FPML, Keefe JR, Cohen ZD, et al. Precision medicine for long-term depression outcomes using the Personalized Advantage Index approach: cognitive therapy or interpersonal psychotherapy? *Psychol. Med* 2019 Nov 22;51(2):279-289. [doi: [10.1017/s0033291719003192](https://doi.org/10.1017/s0033291719003192)]
  72. van Bronswijk SC, Bruijniks SJE, Lorenzo-Luaces L, Derubeis RJ, Lemmens LHJM, Peeters FPML, et al. Cross-trial prediction in psychotherapy: External validation of the Personalized Advantage Index using machine learning in two Dutch randomized trials comparing CBT versus IPT for depression. *Psychother Res* 2021 Jan 23;31(1):78-91. [doi: [10.1080/10503307.2020.1823029](https://doi.org/10.1080/10503307.2020.1823029)] [Medline: [32964809](https://pubmed.ncbi.nlm.nih.gov/32964809/)]
  73. Webb CA, Trivedi MH, Cohen ZD, Dillon DG, Fournier JC, Goer F, et al. Personalized prediction of antidepressant v. placebo response: evidence from the EMBARC study. *Psychol Med* 2018 Jul 2;49(07):1118-1127. [doi: [10.1017/s0033291718001708](https://doi.org/10.1017/s0033291718001708)]
  74. Wu W, Zhang Y, Jiang J, Lucas MV, Fonzo GA, Rolle CE, et al. An electroencephalographic signature predicts antidepressant response in major depression. *Nat Biotechnol* 2020 Apr 10;38(4):439-447 [FREE Full text] [doi: [10.1038/s41587-019-0397-3](https://doi.org/10.1038/s41587-019-0397-3)] [Medline: [32042166](https://pubmed.ncbi.nlm.nih.gov/32042166/)]
  75. Fonzo GA, Etkin A, Zhang Y, Wu W, Cooper C, Chin-Fatt C, et al. Brain regulation of emotional conflict predicts antidepressant treatment response for depression. *Nat Hum Behav* 2019 Dec 23;3(12):1319-1331 [FREE Full text] [doi: [10.1038/s41562-019-0732-1](https://doi.org/10.1038/s41562-019-0732-1)] [Medline: [31548678](https://pubmed.ncbi.nlm.nih.gov/31548678/)]
  76. Hartmann A, von Wietersheim J, Weiss H, Zeeck A. Patterns of symptom change in major depression: Classification and clustering of long term courses. *Psychiatry Res* 2018 Sep;267:480-489. [doi: [10.1016/j.psychres.2018.03.086](https://doi.org/10.1016/j.psychres.2018.03.086)] [Medline: [29980128](https://pubmed.ncbi.nlm.nih.gov/29980128/)]
  77. Zilcha-Mano S, Brown PJ, Roose SP, Cappetta K, Rutherford BR. Optimizing patient expectancy in the pharmacologic treatment of major depressive disorder. *Psychol Med* 2018 Nov 13;49(14):2414-2420. [doi: [10.1017/s0033291718003343](https://doi.org/10.1017/s0033291718003343)]
  78. Kacem A, Hammal Z, Daoudi M, Cohn J. Detecting Depression Severity by Interpretable Representations of Motion Dynamics. *Proc Int Conf Autom Face Gesture Recognit* 2018 May;2018:739-745 [FREE Full text] [doi: [10.1109/FG.2018.00116](https://doi.org/10.1109/FG.2018.00116)] [Medline: [30271308](https://pubmed.ncbi.nlm.nih.gov/30271308/)]
  79. Bailey N, Hoy K, Rogasch N, Thomson R, McQueen S, Elliot D, et al. Responders to rTMS for depression show increased fronto-midline theta and theta connectivity compared to non-responders. *Brain Stimul* 2018 Jan;11(1):190-203. [doi: [10.1016/j.brs.2017.10.015](https://doi.org/10.1016/j.brs.2017.10.015)] [Medline: [29128490](https://pubmed.ncbi.nlm.nih.gov/29128490/)]
  80. Browning M, Kingslake J, Dourish CT, Goodwin GM, Harmer CJ, Dawson GR. Predicting treatment response to antidepressant medication using early changes in emotional processing. *Eur Neuropsychopharmacol* 2019 Jan;29(1):66-75 [FREE Full text] [doi: [10.1016/j.euroneuro.2018.11.1102](https://doi.org/10.1016/j.euroneuro.2018.11.1102)] [Medline: [30473402](https://pubmed.ncbi.nlm.nih.gov/30473402/)]
  81. Byun S, Kim AY, Jang EH, Kim S, Choi KW, Yu HY, et al. Detection of major depressive disorder from linear and nonlinear heart rate variability features during mental task protocol. *Comput Biol Med* 2019 Sep;112:103381 [FREE Full text] [doi: [10.1016/j.compbiomed.2019.103381](https://doi.org/10.1016/j.compbiomed.2019.103381)] [Medline: [31404718](https://pubmed.ncbi.nlm.nih.gov/31404718/)]
  82. Cash RFH, Cocchi L, Anderson R, Rogachov A, Kucyi A, Barnett AJ, et al. A multivariate neuroimaging biomarker of individual outcome to transcranial magnetic stimulation in depression. *Hum Brain Mapp* 2019 Nov 01;40(16):4618-4629 [FREE Full text] [doi: [10.1002/hbm.24725](https://doi.org/10.1002/hbm.24725)] [Medline: [31332903](https://pubmed.ncbi.nlm.nih.gov/31332903/)]
  83. Furukawa TA, Debray TPA, Akechi T, Yamada M, Kato T, Seo M, et al. Can personalized treatment prediction improve the outcomes, compared with the group average approach, in a randomized trial? Developing and validating a multivariable prediction model in a pragmatic megatrial of acute treatment for major depression. *Journal of Affective Disorders* 2020 Sep;274:690-697. [doi: [10.1016/j.jad.2020.05.141](https://doi.org/10.1016/j.jad.2020.05.141)]
  84. Horigome T, Sumali B, Kitazawa M, Yoshimura M, Liang K, Tazawa Y, et al. Evaluating the severity of depressive symptoms using upper body motion captured by RGB-depth sensors and machine learning in a clinical interview setting: A preliminary study. *Compr Psychiatry* 2020 Feb 20;98:152169 [FREE Full text] [doi: [10.1016/j.comppsy.2020.152169](https://doi.org/10.1016/j.comppsy.2020.152169)] [Medline: [32145559](https://pubmed.ncbi.nlm.nih.gov/32145559/)]
  85. Redlich R, Opel N, Grotegerd D, Dohm K, Zaremba D, Bürger C, et al. Prediction of Individual Response to Electroconvulsive Therapy via Machine Learning on Structural Magnetic Resonance Imaging Data. *JAMA Psychiatry* 2016 Jun 01;73(6):557-564. [doi: [10.1001/jamapsychiatry.2016.0316](https://doi.org/10.1001/jamapsychiatry.2016.0316)] [Medline: [27145449](https://pubmed.ncbi.nlm.nih.gov/27145449/)]
  86. Zhou X, Wang Y, Zeng D. Outcome-Weighted Learning for Personalized Medicine with Multiple Treatment Options. *Proc Int Conf Data Sci Adv Anal* 2018 Oct;2018:565-574 [FREE Full text] [doi: [10.1109/DSAA.2018.00072](https://doi.org/10.1109/DSAA.2018.00072)] [Medline: [30931437](https://pubmed.ncbi.nlm.nih.gov/30931437/)]
  87. Barrett BW, Abraham AG, Dean LT, Plankey MW, Friedman MR, Jacobson LP, et al. Social inequalities contribute to racial/ethnic disparities in depressive symptomology among men who have sex with men. *Soc Psychiatry Psychiatr Epidemiol* 2021 Feb 11;56(2):259-272 [FREE Full text] [doi: [10.1007/s00127-020-01940-7](https://doi.org/10.1007/s00127-020-01940-7)] [Medline: [32780176](https://pubmed.ncbi.nlm.nih.gov/32780176/)]

88. Chekroud AM, Zotti RJ, Shehzad Z, Gueorguieva R, Johnson MK, Trivedi MH, et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. *The Lancet Psychiatry* 2016 Mar;3(3):243-250. [doi: [10.1016/s2215-0366\(15\)00471-x](https://doi.org/10.1016/s2215-0366(15)00471-x)]
89. Foster S, Mohler-Kuo M, Tay L, Hothorn T, Seibold H. Estimating patient-specific treatment advantages in the 'Treatment for Adolescents with Depression Study'. *J Psychiatr Res* 2019 May;112:61-70. [doi: [10.1016/j.jpsychires.2019.02.021](https://doi.org/10.1016/j.jpsychires.2019.02.021)] [Medline: [30856378](https://pubmed.ncbi.nlm.nih.gov/30856378/)]
90. Kambeitz J, Goerigk S, Gattaz W, Falkai P, Benseñor IM, Lotufo PA, et al. Clinical patterns differentially predict response to transcranial direct current stimulation (tDCS) and escitalopram in major depression: A machine learning analysis of the ELECT-TDCS study. *J Affect Disord* 2020 Mar 15;265:460-467. [doi: [10.1016/j.jad.2020.01.118](https://doi.org/10.1016/j.jad.2020.01.118)] [Medline: [32090773](https://pubmed.ncbi.nlm.nih.gov/32090773/)]
91. Kessler RC, van Loo HM, Wardenaar KJ, Bossarte RM, Brenner LA, Cai T, et al. Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. *Mol Psychiatry* 2016 Oct;21(10):1366-1371 [FREE Full text] [doi: [10.1038/mp.2015.198](https://doi.org/10.1038/mp.2015.198)] [Medline: [26728563](https://pubmed.ncbi.nlm.nih.gov/26728563/)]
92. Pratap A, Atkins DC, Renn BN, Tanana MJ, Mooney SD, Anguera JA, et al. The accuracy of passive phone sensors in predicting daily mood. *Depress Anxiety* 2019 Jan 21;36(1):72-81 [FREE Full text] [doi: [10.1002/da.22822](https://doi.org/10.1002/da.22822)] [Medline: [30129691](https://pubmed.ncbi.nlm.nih.gov/30129691/)]
93. Rajpurkar P, Yang J, Dass N, Vale V, Keller AS, Irvin J, et al. Evaluation of a Machine Learning Model Based on Pretreatment Symptoms and Electroencephalographic Features to Predict Outcomes of Antidepressant Treatment in Adults With Depression: A Prespecified Secondary Analysis of a Randomized Clinical Trial. *JAMA Netw Open* 2020 Jun 01;3(6):e206653 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.6653](https://doi.org/10.1001/jamanetworkopen.2020.6653)] [Medline: [32568399](https://pubmed.ncbi.nlm.nih.gov/32568399/)]
94. Kautzky A, Möller HJ, Dold M, Bartova L, Seemüller F, Laux G, et al. Combining machine learning algorithms for prediction of antidepressant treatment response. *Acta Psychiatr Scand* 2021 Jan 27;143(1):36-49 [FREE Full text] [doi: [10.1111/acps.13250](https://doi.org/10.1111/acps.13250)] [Medline: [33141944](https://pubmed.ncbi.nlm.nih.gov/33141944/)]
95. Lee Y, Mansur RB, Brietzke E, Kapogiannis D, Delgado-Peraza F, Boutilier JJ, et al. Peripheral inflammatory biomarkers define biotypes of bipolar depression. *Mol Psychiatry* 2021 Jul 03;26(7):3395-3406 [FREE Full text] [doi: [10.1038/s41380-021-01051-y](https://doi.org/10.1038/s41380-021-01051-y)] [Medline: [33658605](https://pubmed.ncbi.nlm.nih.gov/33658605/)]
96. Ma Y, Ji J, Huang Y, Gao H, Li Z, Dong W, et al. Implementing machine learning in bipolar diagnosis in China. *Transl Psychiatry* 2019 Nov 18;9(1):305 [FREE Full text] [doi: [10.1038/s41398-019-0638-8](https://doi.org/10.1038/s41398-019-0638-8)] [Medline: [31740657](https://pubmed.ncbi.nlm.nih.gov/31740657/)]
97. Deng F, Wang Y, Huang H, Niu M, Zhong S, Zhao L, et al. Abnormal segments of right uncinate fasciculus and left anterior thalamic radiation in major and bipolar depression. *Prog Neuropsychopharmacol Biol Psychiatry* 2018 Feb 02;81:340-349. [doi: [10.1016/j.pnpbp.2017.09.006](https://doi.org/10.1016/j.pnpbp.2017.09.006)] [Medline: [28912043](https://pubmed.ncbi.nlm.nih.gov/28912043/)]
98. Nielsen SFV, Madsen KH, Vinberg M, Kessing LV, Siebner HR, Miskowiak KW. Whole-Brain Exploratory Analysis of Functional Task Response Following Erythropoietin Treatment in Mood Disorders: A Supervised Machine Learning Approach. *Front Neurosci* 2019 Nov 20;13:1246 [FREE Full text] [doi: [10.3389/fnins.2019.01246](https://doi.org/10.3389/fnins.2019.01246)] [Medline: [31824247](https://pubmed.ncbi.nlm.nih.gov/31824247/)]
99. Hoogendoorn M, Berger T, Schulz A, Stolz T, Szolovits P. Predicting Social Anxiety Treatment Outcome Based on Therapeutic Email Conversations. *IEEE J. Biomed. Health Inform* 2017 Sep;21(5):1449-1459. [doi: [10.1109/jbhi.2016.2601123](https://doi.org/10.1109/jbhi.2016.2601123)]
100. Ihmig FR, Neurohr-Parakenings F, Schäfer SK, Lass-Hennemann J, Michael T. On-line anxiety level detection from biosignals: Machine learning based on a randomized controlled trial with spider-fearful individuals. *PLoS One* 2020 Jun 23;15(6):e0231517 [FREE Full text] [doi: [10.1371/journal.pone.0231517](https://doi.org/10.1371/journal.pone.0231517)] [Medline: [32574167](https://pubmed.ncbi.nlm.nih.gov/32574167/)]
101. Sharma A, Verbeke WJMI. Understanding importance of clinical biomarkers for diagnosis of anxiety disorders using machine learning models. *PLoS One* 2021 May 10;16(5):e0251365 [FREE Full text] [doi: [10.1371/journal.pone.0251365](https://doi.org/10.1371/journal.pone.0251365)] [Medline: [33970950](https://pubmed.ncbi.nlm.nih.gov/33970950/)]
102. Demiris G, Corey Magan KL, Parker Oliver D, Washington KT, Chadwick C, Voigt JD, et al. Spoken words as biomarkers: using machine learning to gain insight into communication as a predictor of anxiety. *J Am Med Inform Assoc* 2020 Jun 01;27(6):929-933 [FREE Full text] [doi: [10.1093/jamia/ocaa049](https://doi.org/10.1093/jamia/ocaa049)] [Medline: [32374378](https://pubmed.ncbi.nlm.nih.gov/32374378/)]
103. Frick A, Engman J, Alaie I, Björkstrand J, Gingnell M, Larsson E, et al. Neuroimaging, genetic, clinical, and demographic predictors of treatment response in patients with social anxiety disorder. *J Affect Disord* 2020 Jan 15;261:230-237 [FREE Full text] [doi: [10.1016/j.jad.2019.10.027](https://doi.org/10.1016/j.jad.2019.10.027)] [Medline: [31655378](https://pubmed.ncbi.nlm.nih.gov/31655378/)]
104. Lebowitz ER, Zilcha-Mano S, Orbach M, Shimshoni Y, Silverman WK. Moderators of response to child-based and parent-based child anxiety treatment: a machine learning-based analysis. *J Child Psychol Psychiatry* 2021 Oct 24;62(10):1175-1182. [doi: [10.1111/jcpp.13386](https://doi.org/10.1111/jcpp.13386)] [Medline: [33624848](https://pubmed.ncbi.nlm.nih.gov/33624848/)]
105. Lenhard F, Sauer S, Andersson E, Månsson KN, Mataix-Cols D, Rück C, et al. Prediction of outcome in internet-delivered cognitive behaviour therapy for paediatric obsessive-compulsive disorder: A machine learning approach. *Int J Methods Psychiatr Res* 2018 Mar 28;27(1):e1576 [FREE Full text] [doi: [10.1002/mpr.1576](https://doi.org/10.1002/mpr.1576)] [Medline: [28752937](https://pubmed.ncbi.nlm.nih.gov/28752937/)]
106. Schultebrucks K, Qian M, Abu-Amara D, Dean K, Laska E, Siegel C, et al. Pre-deployment risk factors for PTSD in active-duty personnel deployed to Afghanistan: a machine-learning approach for analyzing multivariate predictors. *Mol Psychiatry* 2021 Sep 02;26(9):5011-5022 [FREE Full text] [doi: [10.1038/s41380-020-0789-2](https://doi.org/10.1038/s41380-020-0789-2)] [Medline: [32488126](https://pubmed.ncbi.nlm.nih.gov/32488126/)]



107. Rangaprakash D, Deshpande G, Daniel TA, Goodman AM, Robinson JL, Salibi N, et al. Compromised hippocampus-striatum pathway as a potential imaging biomarker of mild-traumatic brain injury and posttraumatic stress disorder. *Hum Brain Mapp* 2017 Jun 15;38(6):2843-2864 [FREE Full text] [doi: [10.1002/hbm.23551](https://doi.org/10.1002/hbm.23551)] [Medline: [28295837](https://pubmed.ncbi.nlm.nih.gov/28295837/)]
108. Schultebrucks K, Shalev AY, Michopoulos V, Grudzen CR, Shin S, Stevens JS, et al. A validated predictive algorithm of post-traumatic stress course following emergency department admission after a traumatic stressor. *Nat Med* 2020 Jul 06;26(7):1084-1088. [doi: [10.1038/s41591-020-0951-z](https://doi.org/10.1038/s41591-020-0951-z)] [Medline: [32632194](https://pubmed.ncbi.nlm.nih.gov/32632194/)]
109. Grisanzio KA, Goldstein-Piekarski AN, Wang MY, Rashed Ahmed AP, Samara Z, Williams LM. Transdiagnostic Symptom Clusters and Associations With Brain, Behavior, and Daily Function in Mood, Anxiety, and Trauma Disorders. *JAMA Psychiatry* 2018 Feb 01;75(2):201-209 [FREE Full text] [doi: [10.1001/jamapsychiatry.2017.3951](https://doi.org/10.1001/jamapsychiatry.2017.3951)] [Medline: [29197929](https://pubmed.ncbi.nlm.nih.gov/29197929/)]
110. Hinrichs R, van Rooij SJH, Michopoulos V, Schultebrucks K, Winters S, Maples-Keller J, et al. Increased Skin Conductance Response in the Immediate Aftermath of Trauma Predicts PTSD Risk. *Chronic Stress (Thousand Oaks)* 2019 Apr 24;3:247054701984444 [FREE Full text] [doi: [10.1177/2470547019844441](https://doi.org/10.1177/2470547019844441)] [Medline: [31179413](https://pubmed.ncbi.nlm.nih.gov/31179413/)]
111. Breen MS, Thomas KG, Baldwin DS, Lipinska G. Modelling PTSD diagnosis using sleep, memory, and adrenergic metabolites: An exploratory machine-learning study. *Hum Psychopharmacol* 2019 Mar 22;34(2):e2691. [doi: [10.1002/hup.2691](https://doi.org/10.1002/hup.2691)] [Medline: [30793802](https://pubmed.ncbi.nlm.nih.gov/30793802/)]
112. Jin C, Jia H, Lanka P, Rangaprakash D, Li L, Liu T, et al. Dynamic brain connectivity is a better predictor of PTSD than static connectivity. *Hum Brain Mapp* 2017 Sep 12;38(9):4479-4496 [FREE Full text] [doi: [10.1002/hbm.23676](https://doi.org/10.1002/hbm.23676)] [Medline: [28603919](https://pubmed.ncbi.nlm.nih.gov/28603919/)]
113. Salminen LE, Morey RA, Riedel BC, Jahanshad N, Dennis EL, Thompson PM. Adaptive Identification of Cortical and Subcortical Imaging Markers of Early Life Stress and Posttraumatic Stress Disorder. *J Neuroimaging* 2019 May;29(3):335-343 [FREE Full text] [doi: [10.1111/jon.12600](https://doi.org/10.1111/jon.12600)] [Medline: [30714246](https://pubmed.ncbi.nlm.nih.gov/30714246/)]
114. Yuan M, Qiu C, Meng Y, Ren Z, Yuan C, Li Y, et al. Pre-treatment Resting-State Functional MR Imaging Predicts the Long-Term Clinical Outcome After Short-Term Paroxetine Treatment in Post-traumatic Stress Disorder. *Front Psychiatry* 2018 Oct 30;9:532 [FREE Full text] [doi: [10.3389/fpsy.2018.00532](https://doi.org/10.3389/fpsy.2018.00532)] [Medline: [30425661](https://pubmed.ncbi.nlm.nih.gov/30425661/)]
115. Zandvakili A, Swearingen HR, Philip NS. Changes in functional connectivity after theta-burst transcranial magnetic stimulation for post-traumatic stress disorder: a machine-learning study. *Eur Arch Psychiatry Clin Neurosci* 2021 Feb 27;271(1):29-37. [doi: [10.1007/s00406-020-01172-5](https://doi.org/10.1007/s00406-020-01172-5)] [Medline: [32719969](https://pubmed.ncbi.nlm.nih.gov/32719969/)]
116. Marmar CR, Brown AD, Qian M, Laska E, Siegel C, Li M, et al. Speech-based markers for posttraumatic stress disorder in US veterans. *Depress Anxiety* 2019 Jul 22;36(7):607-616 [FREE Full text] [doi: [10.1002/da.22890](https://doi.org/10.1002/da.22890)] [Medline: [31006959](https://pubmed.ncbi.nlm.nih.gov/31006959/)]
117. Iceta S, Tardieu S, Nazare J, Dougkas A, Robert M, Disse E. An artificial intelligence-derived tool proposal to ease disordered eating screening in people with obesity. *Eat Weight Disord* 2021 Oct 02;26(7):2381-2385. [doi: [10.1007/s40519-020-01076-2](https://doi.org/10.1007/s40519-020-01076-2)] [Medline: [33387276](https://pubmed.ncbi.nlm.nih.gov/33387276/)]
118. Senger K, Schröder A, Kleinstäuber M, Rubel JA, Rief W, Heider J. Predicting optimal treatment outcomes using the Personalized Advantage Index for patients with persistent somatic symptoms. *Psychother Res* 2022 Feb 29;32(2):165-178. [doi: [10.1080/10503307.2021.1916120](https://doi.org/10.1080/10503307.2021.1916120)] [Medline: [33910487](https://pubmed.ncbi.nlm.nih.gov/33910487/)]
119. Bae S, Chung T, Ferreira D, Dey AK, Suffoletto B. Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults: Implications for just-in-time adaptive interventions. *Addict Behav* 2018 Aug;83:42-47 [FREE Full text] [doi: [10.1016/j.addbeh.2017.11.039](https://doi.org/10.1016/j.addbeh.2017.11.039)] [Medline: [29217132](https://pubmed.ncbi.nlm.nih.gov/29217132/)]
120. Joseph JE, Vaughan BK, Camp CC, Baker NL, Sherman BJ, Moran-Santa Maria M, et al. Oxytocin-Induced Changes in Intrinsic Network Connectivity in Cocaine Use Disorder: Modulation by Gender, Childhood Trauma, and Years of Use. *Front Psychiatry* 2019 Jul 19;10:502 [FREE Full text] [doi: [10.3389/fpsy.2019.00502](https://doi.org/10.3389/fpsy.2019.00502)] [Medline: [31379621](https://pubmed.ncbi.nlm.nih.gov/31379621/)]
121. Laska EM, Siegel CE, Lin Z, Bogenschutz M, Marmar CR. Gabapentin Enacarbil Extended-Release Versus Placebo: A Likely Responder Reanalysis of a Randomized Clinical Trial. *Alcohol Clin Exp Res* 2020 Sep 31;44(9):1875-1884 [FREE Full text] [doi: [10.1111/acer.14414](https://doi.org/10.1111/acer.14414)] [Medline: [33460198](https://pubmed.ncbi.nlm.nih.gov/33460198/)]
122. Pan Y, Liu H, Metsch LR, Feaster DJ. Factors Associated with HIV Testing Among Participants from Substance Use Disorder Treatment Programs in the US: A Machine Learning Approach. *AIDS Behav* 2017 Feb 8;21(2):534-546 [FREE Full text] [doi: [10.1007/s10461-016-1628-y](https://doi.org/10.1007/s10461-016-1628-y)] [Medline: [27933461](https://pubmed.ncbi.nlm.nih.gov/27933461/)]
123. Schmitgen MM, Niedtfeld I, Schmitt R, Mancke F, Winter D, Schmahl C, et al. Individualized treatment response prediction of dialectical behavior therapy for borderline personality disorder using multimodal magnetic resonance imaging. *Brain Behav* 2019 Sep 14;9(9):e01384 [FREE Full text] [doi: [10.1002/brb3.1384](https://doi.org/10.1002/brb3.1384)] [Medline: [31414575](https://pubmed.ncbi.nlm.nih.gov/31414575/)]
124. Bang S, Son S, Roh H, Lee J, Bae S, Lee K, et al. Quad-phased data mining modeling for dementia diagnosis. *BMC Med Inform Decis Mak* 2017 May 18;17(Suppl 1):60-10 [FREE Full text] [doi: [10.1186/s12911-017-0451-3](https://doi.org/10.1186/s12911-017-0451-3)] [Medline: [28539115](https://pubmed.ncbi.nlm.nih.gov/28539115/)]
125. Yao T, Sweeney E, Nagorski J, Shulman JM, Allen GI. Quantifying cognitive resilience in Alzheimer's Disease: The Alzheimer's Disease Cognitive Resilience Score. *PLoS One* 2020 Nov 5;15(11):e0241707 [FREE Full text] [doi: [10.1371/journal.pone.0241707](https://doi.org/10.1371/journal.pone.0241707)] [Medline: [33152028](https://pubmed.ncbi.nlm.nih.gov/33152028/)]
126. Grassi M, Rouleaux N, Caldirola D, Loewenstein D, Schruers K, Perna G, Alzheimer's Disease Neuroimaging Initiative. A Novel Ensemble-Based Machine Learning Algorithm to Predict the Conversion From Mild Cognitive Impairment to Alzheimer's Disease Using Socio-Demographic Characteristics, Clinical Information, and Neuropsychological Measures. *Front Neurol* 2019 Jul 16;10:756 [FREE Full text] [doi: [10.3389/fneur.2019.00756](https://doi.org/10.3389/fneur.2019.00756)] [Medline: [31379711](https://pubmed.ncbi.nlm.nih.gov/31379711/)]

127. Li Q, Zhao Y, Chen Y, Yue J, Xiong Y. Developing a machine learning model to identify delirium risk in geriatric internal medicine inpatients. *Eur Geriatr Med* 2022 Feb 23;13(1):173-183. [doi: [10.1007/s41999-021-00562-9](https://doi.org/10.1007/s41999-021-00562-9)] [Medline: [34553310](https://pubmed.ncbi.nlm.nih.gov/34553310/)]
128. Andersson S, Bathula DR, Iliadis SI, Walter M, Skalkidou A. Predicting women with depressive symptoms postpartum with machine learning methods. *Sci Rep* 2021 Apr 12;11(1):7877 [FREE Full text] [doi: [10.1038/s41598-021-86368-y](https://doi.org/10.1038/s41598-021-86368-y)] [Medline: [33846362](https://pubmed.ncbi.nlm.nih.gov/33846362/)]
129. Li Y, Rosenfeld B, Pessin H, Breitbart W. Bayesian Nonparametric Clustering of Patients with Advanced Cancer on Anxiety and Depression. 2017 Presented at: 16th IEEE International Conference on Machine Learning and Applications (ICMLA); December 18-21, 2017; Cancun, Mexico p. 674-678. [doi: [10.1109/icmla.2017.00-83](https://doi.org/10.1109/icmla.2017.00-83)]
130. Zhou S, Fei S, Han H, Li J, Yang S, Zhao C. A Prediction Model for Cognitive Impairment Risk in Colorectal Cancer after Chemotherapy Treatment. *Biomed Res Int* 2021 Feb 20;2021:6666453-6666413 [FREE Full text] [doi: [10.1155/2021/6666453](https://doi.org/10.1155/2021/6666453)] [Medline: [33688501](https://pubmed.ncbi.nlm.nih.gov/33688501/)]
131. Vitinius F, Escherich S, Deter H, Hellmich M, Jünger J, Petrowski K, et al. Somatic and sociodemographic predictors of depression outcome among depressed patients with coronary artery disease - a secondary analysis of the SPIRR-CAD study. *BMC Psychiatry* 2019 Feb 04;19(1):57 [FREE Full text] [doi: [10.1186/s12888-019-2026-6](https://doi.org/10.1186/s12888-019-2026-6)] [Medline: [30717711](https://pubmed.ncbi.nlm.nih.gov/30717711/)]
132. Wallert J, Gustafson E, Held C, Madison G, Norlund F, von Essen L, et al. Predicting Adherence to Internet-Delivered Psychotherapy for Symptoms of Depression and Anxiety After Myocardial Infarction: Machine Learning Insights From the U-CARE Heart Randomized Controlled Trial. *J Med Internet Res* 2018 Oct 10;20(10):e10754 [FREE Full text] [doi: [10.2196/10754](https://doi.org/10.2196/10754)] [Medline: [30305255](https://pubmed.ncbi.nlm.nih.gov/30305255/)]
133. Gu S, Zhou J, Yuan C, Ye Q. Personalized prediction of depression in patients with newly diagnosed Parkinson's disease: A prospective cohort study. *J Affect Disord* 2020 May 01;268:118-126. [doi: [10.1016/j.jad.2020.02.046](https://doi.org/10.1016/j.jad.2020.02.046)] [Medline: [32158001](https://pubmed.ncbi.nlm.nih.gov/32158001/)]
134. Montazeri F, de Bildt A, Dekker V, Anderson GM. Network Analysis of Behaviors in the Depression and Autism Realms: Inter-Relationships and Clinical Implications. *J Autism Dev Disord* 2020 May 18;50(5):1580-1595. [doi: [10.1007/s10803-019-03914-4](https://doi.org/10.1007/s10803-019-03914-4)] [Medline: [30778821](https://pubmed.ncbi.nlm.nih.gov/30778821/)]
135. Ekhtiari H, Kuplicki R, Yeh H, Paulus MP. Physical characteristics not psychological state or trait characteristics predict motion during resting state fMRI. *Sci Rep* 2019 Jan 23;9(1):419 [FREE Full text] [doi: [10.1038/s41598-018-36699-0](https://doi.org/10.1038/s41598-018-36699-0)] [Medline: [30674933](https://pubmed.ncbi.nlm.nih.gov/30674933/)]
136. Rosellini AJ, Stein MB, Benedek DM, Bliese PD, Chiu WT, Hwang I, et al. Predeployment predictors of psychiatric disorder-symptoms and interpersonal violence during combat deployment. *Depress Anxiety* 2018 Nov 13;35(11):1073-1080 [FREE Full text] [doi: [10.1002/da.22807](https://doi.org/10.1002/da.22807)] [Medline: [30102442](https://pubmed.ncbi.nlm.nih.gov/30102442/)]
137. Smith R, Feinstein JS, Kuplicki R, Forthman KL, Stewart JL, Paulus MP, Tulsa 1000 Investigators, et al. Perceptual insensitivity to the modulation of interoceptive signals in depression, anxiety, and substance use disorders. *Sci Rep* 2021 Jan 22;11(1):2108 [FREE Full text] [doi: [10.1038/s41598-021-81307-3](https://doi.org/10.1038/s41598-021-81307-3)] [Medline: [33483527](https://pubmed.ncbi.nlm.nih.gov/33483527/)]
138. Liang S, Vega R, Kong X, Deng W, Wang Q, Ma X, et al. Neurocognitive Graphs of First-Episode Schizophrenia and Major Depression Based on Cognitive Features. *Neurosci Bull* 2018 Apr 2;34(2):312-320 [FREE Full text] [doi: [10.1007/s12264-017-0190-6](https://doi.org/10.1007/s12264-017-0190-6)] [Medline: [29098645](https://pubmed.ncbi.nlm.nih.gov/29098645/)]
139. Popovic D, Ruef A, Dwyer DB, Antonucci LA, Eder J, Sanfelici R, PRONIA Consortium. Traces of Trauma: A Multivariate Pattern Analysis of Childhood Trauma, Brain Structure, and Clinical Phenotypes. *Biol Psychiatry* 2020 Dec 01;88(11):829-842. [doi: [10.1016/j.biopsych.2020.05.020](https://doi.org/10.1016/j.biopsych.2020.05.020)] [Medline: [32782139](https://pubmed.ncbi.nlm.nih.gov/32782139/)]
140. Zandvakili A, Philip NS, Jones SR, Tyrka AR, Greenberg BD, Carpenter LL. Use of machine learning in predicting clinical response to transcranial magnetic stimulation in comorbid posttraumatic stress disorder and major depression: A resting state electroencephalography study. *J Affect Disord* 2019 Jun 01;252:47-54 [FREE Full text] [doi: [10.1016/j.jad.2019.03.077](https://doi.org/10.1016/j.jad.2019.03.077)] [Medline: [30978624](https://pubmed.ncbi.nlm.nih.gov/30978624/)]
141. Perez Arribas I, Goodwin GM, Geddes JR, Lyons T, Saunders KEA. A signature-based machine learning model for distinguishing bipolar disorder and borderline personality disorder. *Transl Psychiatry* 2018 Dec 13;8(1):274 [FREE Full text] [doi: [10.1038/s41398-018-0334-0](https://doi.org/10.1038/s41398-018-0334-0)] [Medline: [30546013](https://pubmed.ncbi.nlm.nih.gov/30546013/)]
142. Tate AE, McCabe RC, Larsson H, Lundström S, Lichtenstein P, Kuja-Halkola R. Predicting mental health problems in adolescence using machine learning techniques. *PLoS One* 2020 Apr 6;15(4):e0230389 [FREE Full text] [doi: [10.1371/journal.pone.0230389](https://doi.org/10.1371/journal.pone.0230389)] [Medline: [32251439](https://pubmed.ncbi.nlm.nih.gov/32251439/)]
143. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK, SPIRIT-AICONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. *BMJ* 2020 Sep 09;370:m3164 [FREE Full text] [doi: [10.1136/bmj.m3164](https://doi.org/10.1136/bmj.m3164)] [Medline: [32909959](https://pubmed.ncbi.nlm.nih.gov/32909959/)]
144. Rivera SC, Liu X, Chan A, Denniston AK, Calvert MJ, SPIRIT-AICONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. *BMJ* 2020 Sep 09;370:m3210 [FREE Full text] [doi: [10.1136/bmj.m3210](https://doi.org/10.1136/bmj.m3210)] [Medline: [32907797](https://pubmed.ncbi.nlm.nih.gov/32907797/)]
145. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. *Nat Med* 2020 Sep 09;26(9):1320-1324 [FREE Full text] [doi: [10.1038/s41591-020-1041-y](https://doi.org/10.1038/s41591-020-1041-y)] [Medline: [32908275](https://pubmed.ncbi.nlm.nih.gov/32908275/)]
146. Collins GS, Dhiman P, Andaur Navarro CL, Ma J, Hooft L, Reitsma JB, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based

on artificial intelligence. *BMJ Open* 2021 Jul 09;11(7):e048008 [FREE Full text] [doi: [10.1136/bmjopen-2020-048008](https://doi.org/10.1136/bmjopen-2020-048008)] [Medline: [34244270](https://pubmed.ncbi.nlm.nih.gov/34244270/)]

147. Altman DG, Simera I. A history of the evolution of guidelines for reporting medical research: the long road to the EQUATOR Network. *J R Soc Med* 2016 Feb;109(2):67-77 [FREE Full text] [doi: [10.1177/0141076815625599](https://doi.org/10.1177/0141076815625599)] [Medline: [26880653](https://pubmed.ncbi.nlm.nih.gov/26880653/)]

## Abbreviations

**AI:** artificial intelligence

**CONSORT:** Consolidated Standards of Reporting Trials

**CHARMS:** Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies

**DOI:** digital object identifier

**EQUATOR:** Enhancing the Quality and Transparency of Health Research

**ICD-11:** International Classification of Diseases 11th Revision

**IEEE:** Institute of Electrical and Electronics Engineers

**MI-CLAIM:** Minimum Information About Clinical Artificial Intelligence Modeling

**PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses

**PROBAST:** Prediction Model Risk of Bias Assessment Tool

**SPIRIT:** Standard Protocol Items: Recommendations for Interventional Trials

**SVM:** support vector machine

**TRIPOD:** Transparent Reporting of a Multivariable Prediction Model of Individual Prognosis or Diagnosis

*Edited by J Torous; submitted 19.08.22; peer-reviewed by D Shung, Z Dai, C Price; comments to author 09.09.22; revised version received 02.11.22; accepted 20.11.22; published 02.02.23*

*Please cite as:*

*Tornero-Costa R, Martinez-Millana A, Azzopardi-Muscat N, Lazeri L, Traver V, Novillo-Ortiz D*

*Methodological and Quality Flaws in the Use of Artificial Intelligence in Mental Health Research: Systematic Review*

*JMIR Ment Health* 2023;10:e42045

URL: <https://mental.jmir.org/2023/1/e42045>

doi: [10.2196/42045](https://doi.org/10.2196/42045)

PMID: [36729567](https://pubmed.ncbi.nlm.nih.gov/36729567/)

©Roberto Tornero-Costa, Antonio Martinez-Millana, Natasha Azzopardi-Muscat, Ledia Lazeri, Vicente Traver, David Novillo-Ortiz. Originally published in *JMIR Mental Health* (<https://mental.jmir.org>), 02.02.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Mental Health*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.