



OPEN ACCESS

EDITED BY

David M. A. Mehler,
University Hospital RWTH Aachen, Germany

REVIEWED BY

Karl Andersson,
Skillsta Teknik, Sweden
Victor Okoye,
Saint Louis University, United States

*CORRESPONDENCE

Ghufran Alsalloum
✉ g00100830@aus.edu

RECEIVED 26 June 2025

ACCEPTED 03 November 2025

PUBLISHED 18 November 2025

CITATION

Alsalloum G, Dalibalta S and Hadijat Y (2025)
Harnessing digital health interventions to
address the heterogeneity of depression: a
systematic review.
Front. Digit. Health 7:1654745.
doi: 10.3389/fdgth.2025.1654745

COPYRIGHT

© 2025 Alsalloum, Dalibalta and Hadijat. This
is an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

Harnessing digital health interventions to address the heterogeneity of depression: a systematic review

Ghufran Alsalloum^{1*}, Sarah Dalibalta² and Yacine Hadijat³

¹Department of Bioscience and Bioengineering, College of Engineering, American University of Sharjah, Sharjah, United Arab Emirates, ²Department of Biology, Chemistry and Environmental Sciences, College of Arts and Sciences, American University of Sharjah, Sharjah, United Arab Emirates, ³Innovation in Health Science and Digital Health, College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai, United Arab Emirates

Background: Depression affects over 229 million people worldwide and ranks among the leading causes of disability and death, particularly in young adults, where suicide is a top contributor to mortality. Standard diagnostic and treatment approaches often overlook the marked clinical and biological heterogeneity of depression, resulting in low first-line remission rates and prolonged trial-and-error care, underscoring an urgent need for precision strategies in mental health practice.

Objective: This review explores the recent literature (January 2020–September 2025) on personalized digital health interventions for depression, with an emphasis on how these technologies address heterogeneity in symptomatology, biological underpinnings, and treatment response across diverse patient populations.

Methods: The study followed PRISMA guidelines, searching Scopus, IEEE Xplore, and [ClinicalTrials.gov](#) for English-language peer-reviewed articles and trials published and registered between January 2020 and September 2025. Only studies relevant to depression heterogeneity and digital health were included, and studies focusing solely on generic digital health tools without a personalized or adaptive component were excluded. Findings were synthesized narratively.

Findings: 29 publications were reviewed: 20 studies and 9 clinical trial reports, representing over 5,000 participants. Personalized machine-learning models using mobile sensing and ecological momentary assessments improved mood-forecasting accuracy by up to 25%. Randomized trials of just-in-time adaptive interventions (e.g., the Mello app) demonstrated moderate to large effect sizes for reductions in depression ($d = 0.50$), anxiety ($d = 0.61$), and repetitive negative thinking (RNT) ($d = 0.87$). Smart-messaging post-Cognitive Behavioral Therapy yielded sustained well-being improvements over 12 months, while neuromodulation-based digital therapeutics targeting apathy networks in late-life depression showed significant gains in executive function and motivation. Most studies featured small, convenience samples, variable outcome measures, and limited external validation; risk-of-bias concerns included lack of blinding and incomplete handling of missing data. Equity analyses across demographic and clinical subgroups were seldom reported.

Conclusions: and Relevance: Digital mental health technologies exhibit substantial promise for delivering personalized interventions that accommodate inter-individual variability in depression. High-quality evidence supports their capacity to enhance prediction, engagement, and clinical outcomes. However, broader implementation requires standardized multidimensional outcome measures, equity-focused algorithm validation, and integration of established clinical phenotypes.

KEYWORDS

depression, digital mental health, personalization, heterogeneity, precision psychiatry, systematic review

1 Introduction

Depression affects approximately 229 million people globally, imposing a heavy toll on individuals and society (1). In 2021, suicide, often linked to severe depression, claimed over 700,000 lives and ranked as the third leading cause of death among individuals aged 15–29 (2, 3). The disorder significantly diminishes productivity in work, education, and relationships, with depression and anxiety together accounting for 12 billion lost workdays annually and nearly \$1 trillion in global economic costs (4). The Coronavirus disease (COVID-19) pandemic further intensified this burden, with depressive disorders rising by 18% and anxiety disorders by 15% from 2019 to 2020 (5). By 2030, depressive disorders are projected to become the 12th leading cause of death, with a 35% increase in Disability-Adjusted Life Years (DALYs) (6).

1.1 Heterogeneity of depression

Depression manifests with remarkable variability across individuals, presenting diverse symptom profiles, disease trajectories, and treatment responses that complicate standardized approaches (7). This heterogeneity extends to biological foundations, as research demonstrates that genetic polymorphisms, hormonal factors, inflammatory markers, and neural connectivity patterns all contribute to individual differences in depression presentation (8). Key symptoms of depressive disorders include reduced motivation and pleasure (anhedonia), difficulties in managing anxiety and worry, inflexible thought patterns leading to self-reproach and guilt, impaired processing of sensory and social information, cognitive deficits in attention and memory, and various physical disturbances such as changes in weight, appetite, and sleep patterns (9). Figure 1 maps the different neural dysfunctions associated with depressive episodes and their manifested symptoms (9, 10).

Genetic evidence strongly supports the notion that depression comprises multiple biologically distinct subtypes. For instance, early-onset, recurrent, and postpartum depression exhibit higher Single Nucleotide Polymorphisms (SNP)-based heritability, indicating stronger genetic loading.

Conversely, late-onset and milder subtypes show lower genetic contributions (11). At the symptomatic level, clinical evidence suggests that the internal structure of depression is not uniform. For example, melancholic depression typically includes features such as anhedonia, early morning awakening, and psychomotor disturbance, whereas atypical depression includes mood reactivity, hypersomnia, and increased appetite. Such symptom-specific differences imply that some subtypes, particularly severe and suicidal depression, may be more treatment-resistant and require intensive, tailored interventions, while milder or situational forms might respond well to brief or low-intensity treatments (12). The biological correlates and symptom profiles of different depression subtypes are summarized in Table 1. Recent research using both Diagnostic Statistical Manual (DSM)-5 (25, 26) and data-driven methods consistently supports the existence of biologically distinct depression subtypes. Data-driven clustering reveals unique biological signatures, particularly in inflammatory markers and brain connectivity, for symptom groups such as neurovegetative, anhedonic, and anxious profiles (27, 28). However, progress is hindered by methodological variability, limited replication, and inconsistent reporting, underscoring the need for more standardized, multimodal approaches to fully delineate and validate these subtypes. This variability calls for more personalized interventions that address biological correlates in addition to symptom profiles (14, 24).

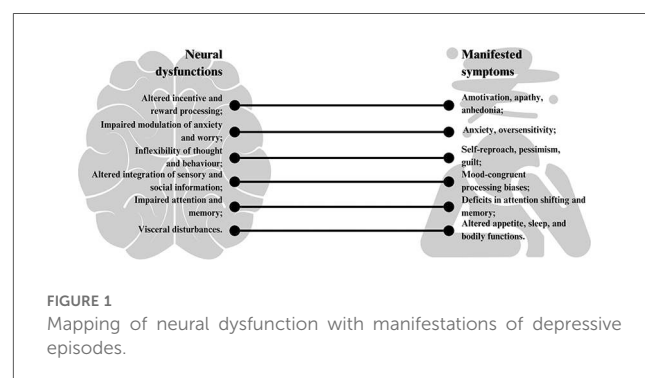


TABLE 1 Symptom profiles and biological underpinnings of depression subtypes.

Biological correlate	Associated symptoms	Depression subtypes
HPA-Axis Dysfunction Elevated basal cortisol; altered cortisol awakening response (hyper- or hypo-cortisolemia) (13, 14)	Insomnia or early-morning awakening; hyperarousal; anxiety sensitivity; weight/appetite loss or gain.	Melancholic/typical depression (HPA hyperactivity); Atypical/neurovegetative depression (blunted or normal cortisol response)
Inflammatory Markers ↑ CRP, IL-6, TNF- α , IL-8 (12, 13, 15)	Persistent fatigue; hypersomnia; somatic pain; appetite changes; “sickness” malaise	Atypical/neurovegetative depression (elevated inflammation); no elevation in melancholic/typical depression
Metabolic Indicators ↑ BMI, waist circumference, triglycerides, metabolic-syndrome markers (13, 16)	Weight gain; increased appetite; fatigue	Atypical/neurovegetative depression (metabolic dysregulation); no abnormalities in melancholic/typical depression
Neurotransmitter Dysregulation CRH-LC-NE hyperactivity; dopamine-agonist responsiveness (14, 17)	Anhedonia; low motivation; mood instability; irritability; anxiety; psychomotor changes	Melancholic depression (CRH-LC-NE hyperactivity); Atypical depression (dopamine-agonist responsive with reduced CRH-LC-NE tone)
Neural Circuitry Alterations Aberrant connectivity in DMN, salience, and cognitive-control networks (18–20)	Rumination and self-referential thought; psychomotor retardation; apathy; impaired executive function	Anxious-ruminative subtype (DMN hyperconnectivity); Melancholic/severe MDD (DMN hypoconnectivity); plus four data-driven “biotypes” predictive of treatment response
Genetic Risk Profiles Polygenic scores for MDD, schizophrenia, BMI- and metabolic-trait loci (10, 19)	Early-onset/recurrent course; cognitive impairment; diurnal mood variation; metabolic dysregulation	Familial/early-onset depression (high MDD-PRS); MDD with psychotic features (schizophrenia-PRS); Atypical/metabolic subtype (BMI- and triglyceride-PRS)
Kynurenine Pathway Dysregulation ↑ QUIN/KYNA ratio; ↓ kynurenic acid (16, 17)	Cognitive deficits (impaired concentration, memory); increased suicidality risk; anhedonia via serotonin depletion	Neuroinflammatory/cognitive-impairment subtype; TRD with neurotoxic kynurenine signature
BDNF Alterations ↓ Serum BDNF levels (21)	Anhedonia; reduced motivation; impaired synaptic plasticity; memory problems	Neurotrophic-deficit depression; TRD characterized by diminished neuroplasticity
Immune-Cell Changes ↑ WBC, lymphocytes, platelets; altered lymphocyte proliferation (15, 21)	Somatic aches and pains; heightened stress reactivity; poor antidepressant response	Immune-driven/inflammatory depression; TRD with pronounced cellular-immune activation
Hippocampal Alterations Reduced volume, impaired neurogenesis, altered regulation of HPA axis (22)	Contextual memory deficits and impaired pattern separation leading to overgeneralization of negative contexts	Chronic/recurrent MDD (volume loss after ≥ 2 years or multiple episodes); Stress-related depression (volume loss secondary to acute/chronic stress); Cognitive-impairment subtype characterized by poor memory and learning
Catatonic Syndrome Motor abnormalities (stupor, mutism, posturing, negativism, echophenomena); GABAergic dysregulation in basal ganglia–cortical circuits; right orbitofrontal hypoactivity (23, 24)	Stupor; mutism; negativism; posturing; echolalia/echopraxia; extreme anxiety; contextual overgeneralization	MDD with catatonic features (catatonic subtype)
Postpartum Endocrine & Immune Dysregulation Rapid withdrawal of estradiol/progesterone; reduced oxytocin; HPA-axis hypoactivation; ↑IL-6 and proinflammatory cytokines; serotonergic and BDNF gene polymorphisms (4, 25, 26)	Depressed mood; anhedonia; sleep/appetite disturbances; fatigue; anxiety; cognitive impairment	MDD with peripartum onset (Postpartum depression)

HPA, hypothalamic–pituitary–adrenal axis; CRP, C-reactive protein; IL-6, IL-8, interleukin-6, interleukin-8 (pro-inflammatory cytokines); TNF- α , tumor necrosis factor alpha; BMI, body mass index; CRH, corticotropin-releasing hormone; LC, locus coeruleus; NE, norepinephrine; DMN, default mode network; MDD, major depressive disorder; PRS, polygenic risk score; QUIN, quinolinic acid; KYNA, kynurenic acid; TRD, treatment-resistant depression; BDNF, brain-derived neurotrophic factor; WBC, white blood cells; GABA, gamma-aminobutyric acid.

1.2 Depression management in clinical practice

Clinical guidelines such as the DSM-5 and National Institute for Health and Care Excellence (NICE) acknowledge clinical heterogeneity in depression by incorporating specifiers and stratified care approaches rather than redefining novel biological subtypes. The DSM-5 introduces specifiers “with melancholic features”, “with atypical features”, “with anxious distress”, and others to capture distinct symptom patterns that may guide treatment selection—for example, considering electroconvulsive therapy in severe melancholic depression or MAO inhibitors in atypical presentations (25). Similarly, NICE stratifies patients by episode severity (“less severe” vs. “more severe”) and recognizes subgroups such as chronic depression, psychotic depression, and treatment-resistant depression, offering stepped-care algorithms and augmentation strategies tailored to each subgroup (29). By refining diagnostic descriptors and matching intervention

intensity and modality to individual clinical profiles, these guidelines operationalize depression’s heterogeneity within routine practice.

The diverse manifestations of depression directly impact clinical practice, particularly regarding diagnosis, treatment selection, and long-term management. Despite the well-documented heterogeneity of depression, primary care settings, where most depression cases are initially encountered, often rely on broad diagnostic criteria and generalized treatment guidelines, resulting in inconsistent treatment implementation (30). Only 30% of patients achieve remission with their first prescribed medication, raising the need for multiple treatment trials before finding effective interventions, extending suffering, and increasing healthcare costs. Furthermore, approximately 55% experience side effects, highlighting the urgent need for personalized treatment strategies (8, 16, 20).

Precision psychiatry aims to tailor mental health treatments to groups and ultimately to individual patients (personalized

psychiatry) by integrating biological, clinical, and digital data to predict treatment response and optimize therapeutic outcomes. Despite its potential, precision psychiatry remains difficult to implement. A key challenge is the lack of biomarker-based diagnostics. While psychiatric diagnoses are often reproducible, their biological validity remains weak, making precise treatment selection difficult (20). Measurement-Based Care (MBC) can improve treatment outcomes by using validated scales to systematically track symptoms, guide clinical decisions, and enhance patient adherence. Research demonstrates that MBC significantly increases remission rates (74% vs. 29% in standard care) and doubles treatment response odds in primary care settings (31).

1.3 Personalizing depression care

Recognition of depression's heterogeneous nature has driven momentum toward personalized medicine approaches that tailor treatment strategies to individual patient characteristics. These personalization strategies incorporate biological, psychological, and digital tools to refine diagnosis, predict treatment response, and improve outcomes as detailed below.

1.3.1 Pharmacogenetics and biomarkers

Pharmacogenetic testing examines how genetic variations influence individual responses to antidepressants (32). Evidence suggests that genetic markers affecting serotonin metabolism (e.g., SLC6A4 polymorphisms) and liver enzyme activity (e.g., CYP450 variants) can predict treatment efficacy and side effects, enabling more informed medication selection. Despite the promise, pharmacogenetics has not yet achieved widespread implementation, as questions regarding cost-effectiveness and clinical utility persist (33).

Biomarkers play a crucial role in advancing psychiatry by aiding in the diagnosis, treatment, and potential prevention of major psychiatric disorders such as depression, schizophrenia, and anxiety. Central (brain imaging) and peripheral (blood proteins, immune markers) biomarkers provide biological signatures that help distinguish between disorders and predict treatment responses (21). On the other hand, digital biomarkers, derived from mobile apps, wearables, and other digital health technologies, are transforming mental healthcare by enabling real-time, objective monitoring of mental states. These biomarkers include speech patterns, sleep metrics, heart rate variability (HRV), activity levels, and human-device interactions, which can provide continuous, passive assessment of mental health conditions (34, 35).

1.3.2 Personality subtypes and psychological stratification

Recent investigations suggest personality traits play a key role in depression subtyping and treatment optimization (36, 37). Personality traits like high neuroticism, low extraversion, and low conscientiousness are associated with greater responsiveness to intensive interventions such as CBT (38). Furthermore, high

levels of self-criticism, social avoidance, and personal reserve predict poorer responses to psychological therapies, underscoring the importance of tailored care not only to biological profile but also to psychological characteristics of patients (39, 40). Stratified care models—where treatments are assigned based on initial psychological assessments—have demonstrated greater clinical and cost-effectiveness compared to traditional stepped care approaches (23).

1.3.3 Digital phenotyping and machine learning (ML)

Digital phenotyping, which refers to the real-time and passive collection of behavioral data through digital devices, offers an unprecedented opportunity to understand the nuanced, moment-to-moment patterns of mental illness. Smartphones and wearables can unobtrusively track digital biomarkers, creating a digital fingerprint of each individual's mental state. These data, when paired with ML algorithms, enable the identification of personalized behavioral signatures linked to psychiatric conditions and treatment responses (41, 42). Algorithms trained on vast multimodal datasets—including active inputs like ecological momentary assessments (EMAs) and passive data from sensors—can predict depressive episodes, monitor treatment efficacy, and match patients to the most effective therapeutic modalities (42). Wearable devices gather continuous physiological and activity signals [e.g., heart rate (HR), movement, sleep patterns], which can be processed into metrics like resting heart rate or sleep efficiency. EMAs involve time-stamped, real-time self-reports via smartphones on mood, stress, context, and behaviors. Together, wearables offer objective sensor data, while EMA provides subjective, contextual insights (43).

1.4 Objectives of the review

This is perhaps the first review to discuss the heterogeneity of depression in relation to digital health. This systematic review aims to explore and analyze recent digital technology targeting personalized depression care, thereby making the case for the digital transformation of mental healthcare. The objectives are:

- To summarize and critique the recent literature on personalized digital interventions.
- To critically assess how personalized digital health interventions address the heterogeneity of depression
- To identify key challenges and future directions in the integration of digital health into personalized mental health care.

2 Methods

This paper adopts a systematic review approach to critically synthesize and evaluate the current literature on the role of digital health interventions in addressing the clinical heterogeneity of depression.

2.1 Search strategy

This review focuses on a specific niche in depression care: the personalization of digital health solutions to target the biological correlates of depression subtypes. While much research has explored either depression heterogeneity or digital tools independently, few have bridged both areas. The search strategy targeting this specific niche is shown in [Table 2](#). This review intentionally focused on personalized digital interventions for depression; studies without an explicit personalization component were excluded by design. Searches covered peer-reviewed literature in Scopus and IEEE Xplore, complemented by targeted grey literature from [ClinicalTrials.gov](#) and hand-searching of reference lists and forward citations. Searches were last run on 15 September 2025. Records published after this date were not considered. Following the PRISMA guidelines, 29 studies were identified that address heterogeneity of depression to varying levels: 20 papers and 9 clinical trial reports. The PRISMA scheme is shown in [Figure 2](#).

2.2 Data extraction and synthesis

Data was extracted using Microsoft Excel by a single author, G.A. The extracted information included research objectives, methods, personalization strategy, key findings, and other study details. Due to the heterogeneous nature of the studies, the data extraction and synthesis follow a narrative, descriptive approach organized along two orthogonal frameworks: (i) modality of intervention (passive sensing/forecasting; just-in-time adaptive interventions; conversational/coach-supported tools; decision-support/recommenders; physiology-coupled therapeutics) and (ii) personalization axis (timing & context, treatment format & dose, content & target). Data items include:

- Bibliographic and Design Details: author(s), year, country, study design (e.g., randomized controlled trials, cohort, case series).

- Participants: sample size, clinical population (e.g., MDD diagnosis, subclinical), age range, and demographic composition.
- Intervention Characteristics: digital modality (e.g., smartphone app, wearable), personalization strategy (e.g., speaker-specific layers, transfer learning, co-design), duration, and comparator.
- Outcomes and Measures: primary and secondary outcomes (e.g., PHQ-2/9 scores, mood correlation coefficients, Cohen's d, prediction accuracy), assessment instruments, and follow-up intervals.
- Implementation and Contextual Variables: technology platform, engagement metrics (e.g., retention rates), and any human-centred design or equity considerations reported.

2.3 SWOT analysis

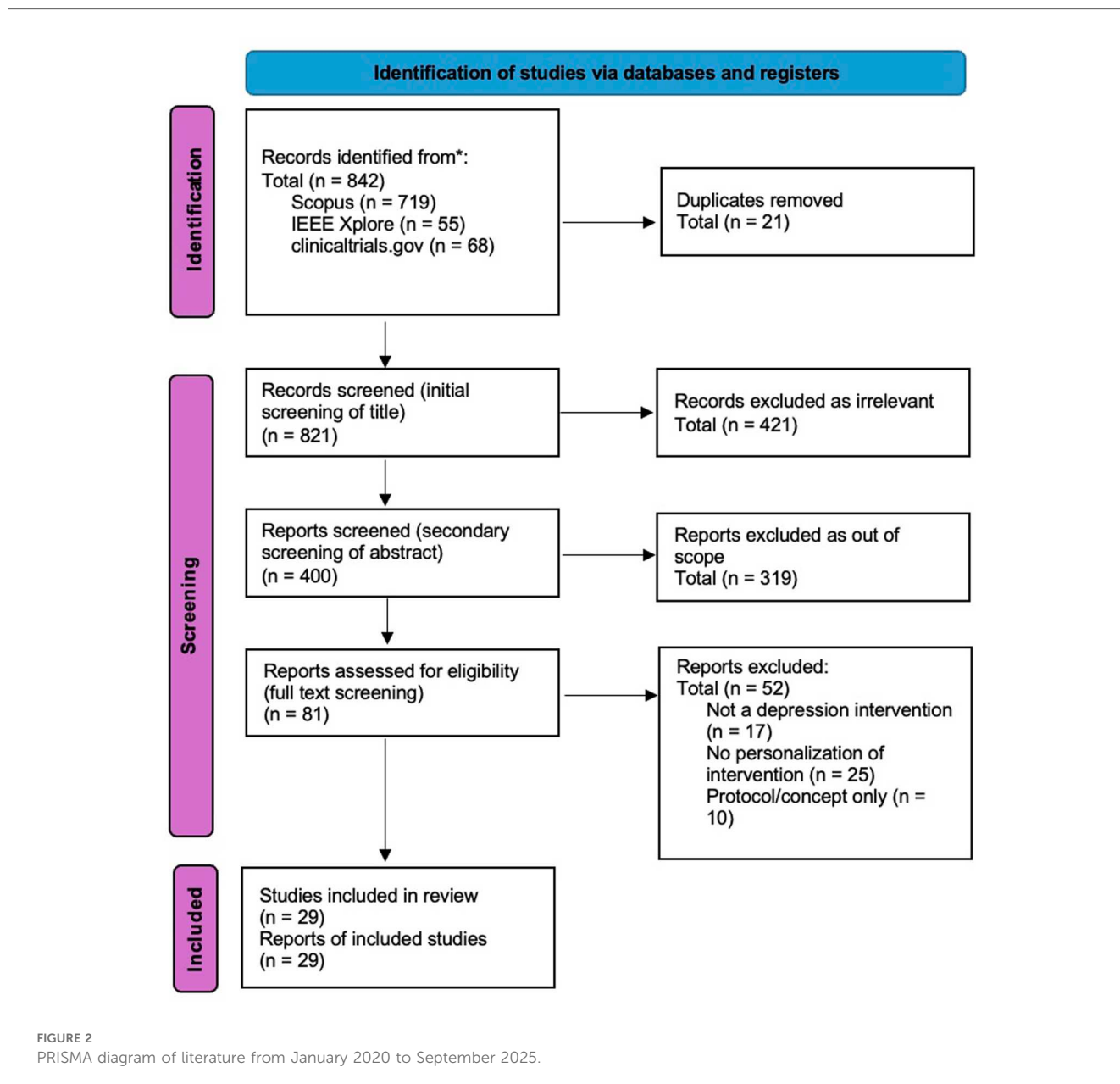
To systematically evaluate the empirical literature on personalized digital interventions for depression, SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis was employed. This strategic framework is commonly used in healthcare research to critically assess innovation potential, implementation challenges, and contextual fit. In this review, SWOT was applied to 20 peer-reviewed studies, each offering a unique technological or methodological approach to personalized mental health care.

2.4 Risk of bias assessment

To evaluate the methodological rigor and reliability of 14 prospective studies, a qualitative Risk of Bias (RoB) (44) assessment was conducted. This process considers five core domains adapted from the Cochrane Risk of Bias framework and digital intervention research best practices. Each study was categorized as Low, Moderate, or High risk in each domain, leading to an overall risk of bias judgment. This qualitative appraisal supported a nuanced interpretation of findings across heterogeneous methodologies.

TABLE 2 Search strategy.

Databases searched	Scopus, IEEE Xplore, ClinicalTrials.gov
Search keywords	TITLE-ABS-KEY [("Mental Health" OR "Depression" OR "Mood Disorders") AND ("Digital Health" OR "mHealth" OR "eHealth" OR "Mobile Mental Health Apps" OR "Digital intervention" OR "internet-delivered") AND ("personalized" OR "individualized" OR "patient-specific" OR "Adaptive"))] AND PUBYEAR>2019 AND PUBYEAR<2026 AND [LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "cp")] In ClinicalTrials.gov , the search was for Major Depressive Disorder, other terms are Depression and Digital.
Search period	From January 2020 and September 2025, focusing on recent advancements in digital health and personalized psychiatry.
Language	Only articles published in English were included.
Inclusion criteria	1. Peer-reviewed research articles, randomized controlled trials, observational studies, and conference papers relevant to depression and digital health. 2. Studies focusing specifically on depression and depressive symptoms. 3. Studies presenting a digital intervention or mHealth tool with explicit discussion of personalization technique.
Exclusion criteria	1. Review papers, concept papers, proposals, and theoretical frameworks, design studies or studies assessing the design of intervention rather than the outcome. 2. Studies focusing on other disorders or general mental health and wellbeing without a focus on depression. 3. Studies unrelated to a digital interventions. 4. Studies focusing solely on generic digital health tools without a personalized or adaptive component.



The Prediction Model Risk Of Bias Assessment Tool (PROBAST) (45) was used to assess the quality of the models used in the 3 modeling studies. The models were evaluated for signal problems in the four domains of PROBAST, and the results were categorized as “yes”, “probably yes”, “probably not”, “no”, or “no information”. This leaves 3 articles (13, 46, 47) that were not fit for these tools and were narratively appraised instead. Only one researcher, G.A., performed the risk of bias assessments.

3 Results

The development of personalized digital interventions is transforming depression care by incorporating user-specific data

and adaptive technologies into treatment delivery. Recent empirical studies demonstrate that personalization can significantly enhance mental health outcomes.

One promising avenue is real-time monitoring and forecasting using passive mobile sensing and EMAs to improve symptom prediction. Individualized models consistently outperform pooled approaches. For example, ML models trained on individual-level mobile data improved depression forecasting accuracy by up to 25% over non-personalized models (48). Similarly, personalized deep learning models using speech features outperformed general models in predicting mood states, highlighting the value of speaker-specific adaptations (17). In another study leveraging wearable and EMA data, Chatterjee et al. (49) developed explainable deep learning models that predicted mood scores with as little as 6% error for some

participants. Their approach not only enabled high accuracy but also facilitated insight into which biophysical indicators contributed to mood changes, underscoring the potential of explainable, individualized systems to support tailored interventions. However, both studies are limited by small sample sizes and missing data. Digital monitoring of depressive symptoms in older adults was studied using a smartwatch, a motion-sensing camera, and a chatbot. This living-lab platform gives them daily, baseline-anchored updates on their mental and physical health to support self-care, and strengthens social support by sharing daily status and emergency alerts with community caregivers (50).

Personalized support delivered “in the moment” demonstrates additional benefits. An automated, personalized smartphone program targeting repetitive negative was tested in a pilot randomized trial. The Mello app, a just-in-time adaptive intervention (JITAI), outperformed controls, reducing repetitive negative thinking, anxiety, and depressive symptoms (51). Another rumination-focused intervention, by Wang and Miller (52), pilot-tested a fully automated JITAI using CBT to identify and block depressive rumination. They used EMAs to drive a just-in-time system where every few hours participants reported recent stressors and, if a rumination trigger was present, the app identified the trigger type and checked whether the person was currently receptive (e.g., not driving or walking), then delivered tailored support. The study participants reported reduced episode counts and duration when support was triggered just after rumination was detected. Future validation with a larger sample size is essential (52). Cue, a precision smartphone program for outpatient care improved symptoms by timing small interventions to each person’s daily routines or “social rhythms” such as sleep timing, daily mood, and energy self-ratings. Cue is a smartphone-based platform that continuously tracks behavior and collects self-reports. It pairs psychoeducation modules with personalized “micro-interventions”; very short, targeted therapeutic actions, usually seconds to a few minutes, delivered in the flow of daily life, for example, reminders to keep a consistent wake time if weekend sleep-ins are detected. The aim is to regularize daily routines and, in turn, reduce depressive symptoms and sustain wellness with minimal patient effort (53). Another JITAI, delivered via m-Path smartphone app, was designed to provide psychological support to individuals experiencing subclinical and clinical levels of depressive symptoms while awaiting psychotherapy. Once triggered, the intervention asks participants to name the kind of social support they require, or surfaces a list of past contacts to identify who could help right now, or delivers one of six evidence-based support-seeking strategies. Participants then choose how to reach out. Microrandomization is performed at each eligible decision point, where the participant was found in need of support, they are randomized to either intervention or control. Microrandomized feasibility work showed high adherence, small immediate distress reductions, and the highest perceived appropriateness when individuals themselves signaled need in real time (19).

Communication-centered supports, ranging from smart messaging to agent-guided use, provide lower-intensity

personalization pathways that can scaffold care. A novel study by Malins et al. (54) applied personalized “smart-messaging” to support CBT follow-up care. Participants prewrote advice tailored to their future emotional states, which was delivered via text after therapy ended. Over 12 months, users of this personalized system showed greater improvements and stability in well-being than non-users, suggesting that low-intensity digital personalization can meaningfully support relapse prevention in clinical populations (54). While there is a lot of research on chatbots targeting mental health support, few offer personalized therapeutic support to depressed populations. Woebot® (WB001) is a prescription, 8-week mobile digital therapeutic used for postpartum depression, among other mental health conditions, that uses a conversational agent to deliver brief, personalized CBT, daily mood tracking, and psychoeducation under clinician oversight. It tailors chats to user-selected problem areas, maintains rapport via Natural Language Processing (NLP), and includes crisis detection with SOS escalation. Users show symptom reduction and high satisfaction, with evidence of therapeutic alliance and preliminary efficacy in young adults and postpartum samples (13). Another chatbot, MindBot, powered by AI, delivers personalized mental-health support by combining classic NLP like tokenization and sentiment scoring with large language models (LLMs) for context-aware replies. Bench testing showed it can track shifting emotions, handle typos, and maintain stable performance. However, evaluation focused on accuracy, “emotional accuracy”, and reliability, not clinical outcomes (46).

While this review is not focused on clinician-guided personalization of care, as that has been addressed in previous literature and is out of scope, certain studies have incorporated human-guided personalization with algorithmic tailoring of therapeutic content. For example, HAYT (“How Are You Today?”) is a mobile app for anxiety and depression that allows patients to directly chat with clinicians instead of chatbots. It combines a suite of services and features: a digital diary analyzed with NLP to detect sentiment and symptoms; DSM-5-derived symptom questionnaires; a predictive model for near-term risk; personalized CBT prompts; secure messaging/video with clinicians; scheduling/reminders; and clinician-facing reports. The preliminary results, though based on synthetic data, indicate a strong correlation between sentiment analysis and self-reported depressive symptoms, suggesting its utility in monitoring mental health in a clinical setting (47). Another app, mindLAMP was used to test the effect of having personalized recommendations given by a Digital Navigator “Guide” vs. general untailored content ‘Support’. A Digital Navigator is a care-team specialist in digital health who supports patients with technical troubleshooting, tailors and optimizes app use to boost engagement, and helps integrate digital tools into clinical care. Both groups had full access to the mindLAMP app, which included modules like Thought Patterns, Mindfulness, Journaling, Distraction Games, Gratitude Journaling, Behavioral Activation, and Strengths. Guide users completed significantly more activities overall ($p < .001$), while Support users “binged” early then tapered. Notably, even though the two coaching

groups didn't differ statistically, the people who used the app more were more likely to get big improvements; participants with $\geq 25\%$ drops in PHQ-9/GAD-7 had higher Digital Working Alliance Inventory scores than non-improvers (55).

Decision rules and recommender logic translate personalization into actionable choices about “what to do” or “what level of care to provide”. A data-driven, personalized activity recommender for mood disorders learns which specific activities boost an individual's mood. Data was gathered over several weeks from two mobile apps spanning clinical and non-clinical populations. In the MORIBUS clinical sample, 7 patients with unipolar or bipolar disorder logged 1,684 entries, selecting or typing specific activities and rating each activity's positivity. Text and labels from activity logs are preprocessed and modeled with Naive Bayes and SVM, comparing a pooled (“general”) vs. person-specific model. After roughly 59 activities per user, personalized models significantly outperform general ones with error rates as low as 10% for some participants (56). Another recommender, the Personalised and Optimised Therapy algorithm, trained on the 4,469 participants in the RESILIENT trial using regularized prediction models and early Patient Health Questionnaire-9 response. The aim is to estimate individual treatment effects across cognitive behavioral therapy skills and select the option with the highest probability of benefit. It recommends single skills for very low baseline severity and two-skill combinations for higher severity. In a simulated trial, it significantly increased overall treatment effect with approximately 35% greater benefit than the non-personalized group average (15).

Measurement-based, stratified systems of care that adjust level and type of support to ongoing assessments offer personalization in the clinical practice. STAND is a stepped-care model for university students that screens with the Computer Adaptive Test for Mental Health, assigns a care tier (T0-monitoring only, T1-digital therapy with coaches, T2-digital therapy assisted by clinicians in training, and T3-clinical care), and adapts level based on continuous symptom and suicide-risk monitoring. In deployment, hundreds of suicide-risk alerts were detected and managed in real time. Acceptability was high across tiers. Depression and anxiety improved significantly ($P < .001$) in all tiers and engaged participants showed $\geq 30\%$ symptom reductions (57).

Finally, physiology-coupled and lifestyle-tailored therapeutics personalize timing and content using signals beyond self-report. An e-health program, NEVERMIND, combined a sensorized smart shirt (electrocardiogram, respiration, movement) with a mobile app that administers questionnaires, forecasts depressive symptoms, and delivers personalized feedback and lifestyle guidance, and online cognitive behavioral therapy via Deprexis. The program significantly reduced depressive symptoms and suicidal ideation compared to control ($p < 0.001$), with a clinically relevant effect size (Cohen's $d = 0.39$) (58). In active-duty personnel, a randomized study in military personnel tested CBT alone vs. CBT plus a smartwatch-linked mHealth app that detects physiological stress, delivers real-time alerts with guided coping, and shares data with therapists via a provider portal.

The app group attended more sessions and showed significant reductions in depression, anxiety, stress, and anger, approaching asymptomatic levels by approximately 45 days (59). Another randomized waitlist-controlled trial tested FeelDTx. By integrating a mobile CBT-based program, a wearable emotion sensor (EDA, HRV, skin temperature) that triggers personalized, in-the-moment prompts, and weekly 15-minute Digital Navigator check-ins, the study reported high engagement and larger symptom reductions than the control. In the experimental symptomatic group, depressive and anxiety symptoms fell by 45% and 50% from baseline, respectively, with 32 and 39 subjects exceeding the Minimal Clinically Important Difference, respectively (60). In a unique approach to managing depressive symptoms, Campisi et al. (61) conducted an 8-week single-arm pilot of a personalized nutrition program for adolescents with MDD. They employed bi-weekly counseling with a Mediterranean-style co-created menu, as well as weekly groceries and eHealth messages (61). For a small sample of 10 parent-teenager pairs, feasibility was moderate (40% recruitment; 77% completion) with moderate–high acceptability. Preliminary effects showed small improvements in depressive symptoms ($d \approx 0.36$), parent food modeling ($d \approx 0.24$), and adolescent nutrition attitudes ($d \approx 0.36$) (61).

Taken together, the evidence supports a thematic map in which real-time monitoring and forecasting quantify within-person dynamics; just-in-time support and messaging deliver timely, symptom-oriented help; agents and human coaches personalize communication and adherence; decision rules and recommenders choose content or level of care; and physiology- or lifestyle-coupled therapeutics align delivery to physiological state and daily context. Across these modalities, personalization functions as the unifying mechanism most closely associated with improved engagement and clinical improvement. The details of these studies are summarized in Table 3, which organizes results by modality and adaptivity to clarify similarities and differences across approaches.

Across registered clinical trials and device studies (summarized in Table 4), personalization is being operationalized along three complementary axes: (i) timing and context; detecting when an individual most needs support and delivering it in the moment, (ii) treatment format and dose; adapting the level of human support or modality based on early response, and (iii) content and target; matching therapeutic ingredients to individual symptom mechanisms, cognitive profiles, or neurobiological signals. Together, these trials test whether precision in *when*, *how*, and *what* is delivered can improve outcomes, adherence, and scalability across routine-care and home settings. While peer-reviewed outcomes for many of these trials are pending, their methodologies and design indicate a significant move toward addressing the complexity of depression at the individual level.

Using timing and context for personalization, the Motor Activity–Subjective Energy (MASE) Project (NCT07059234) learns each person's within-day association between incidental, non-exercise activity and felt energy using accelerometry and ecological diaries, then uses neurobiological profiling (from

TABLE 3 Summary of reviewed papers grouped by modality.

Study; sample size; design	Personalization axis	Data sources used	Target mechanism or phenotype	Adaptivity	Primary outcome and effect metric (as reported)
A. Passive sensing and forecasting: Learns within-person patterns from ambient data; output is prediction/monitoring, not necessarily an intervention.					
(48) — Personalized depression forecasting using mobile sensor data and ecological momentary assessment; $N = 65$; observational modeling within a digital intervention cohort	Model-based (subject-dependent standardization; transfer learning; subgroup models)	Ecological momentary assessments, smartphone sensors	Depressive symptom severity prediction and next-day symptom forecasting	Static tailoring (modeling; not a real-time intervention)	End-of-day Patient Health Questionnaire-2 mean absolute error 0.801 (approximately 25% better than baseline 1.062); next-day mean absolute error 1.349 (approximately 12% better than baseline 1.539)
(49) — Towards personalized mood prediction and explanation for depression from biophysical data; $N = 14$; observational modeling	Model-based deep learning with model explainability	Ecological momentary assessments, wearable lifestyle data, neurocognitive assessments	Current mood state and depressive symptom severity	Static tailoring (offline personalized models)	Per-person prediction error reported as low as approximately six percent; deep learning models exceeded classical machine-learning baselines
(17) — Personalized deep learning for monitoring depressed mood from speech; $N = 41$ (30 major depressive disorder and 11 subclinical); observational modeling	Model-based	Speech captured via smartphone during ecological momentary assessment sessions	Depressed mood monitoring on a visual analogue mood scale	Static tailoring (offline personalized models)	Personalized models outperformed population models; higher correlation between predicted and self-reported mood (details in Supplement)
(50) — Socially vulnerable older adults; six-week single-arm living-lab pilot; $N = 25$	Rule-based individualized daily feedback dashboards	Daily Patient Health Questionnaire assessments via chatbot; wearable sensor data including heart rate variability, sleep, and physical activity	Day-to-day depressive symptom fluctuations in geriatric populations	Static daily feedback without just-in-time logic	Within-person modeling showed sleep fragmentation and sleep efficiency associated with higher same-day depressive symptoms; pre-post improvement in depressive symptoms and sleep quality; usability unchanged
B. Just-in-time digital support: Automated micro-interventions delivered at high-value moments, often driven by ecological momentary assessment or sensors.					
(51) — A personalized, transdiagnostic smartphone intervention (Mello) targeting repetitive negative thinking; $N = 55$; pilot randomized controlled trial	Model-based adaptive intervention (fully automated and personalized)	Ecological momentary assessments via smartphone	Repetitive negative thinking as a transdiagnostic mechanism; depression and anxiety symptoms	Adaptive in real time (just-in-time delivery)	Depression standardized effect size approximately 0.50; anxiety standardized effect size approximately 0.61; repetitive negative thinking standardized effect size approximately 0.87 over twelve weeks
(53) — Social rhythm-focused precision digital intervention (Cue) augmenting outpatient care; intent-to-treat $N = 133$; depressed-at-entry subgroup $N = 28$	Model-based personalization using smartphone behavior to time micro-interventions	Continuous smartphone behavioral patterns; symptom self-reports	Repetitive negative thinking and depressive symptoms with anxiety symptoms as secondary outcomes	Adaptive in real time with just-in-time smartphone delivery	Greater improvement from baseline to sixteen weeks in the full sample; larger reduction in Patient Health Questionnaire-8 scores in depressed-at-entry subgroup versus monitoring only
(52) — Pilot randomized controlled trial of rumination-focused mobile cognitive behavioral therapy just-in-time adaptive intervention; $N = 18$	Personalized timing using each participant's rumination pattern	Intensive self-reports via smartphone text messages; mobile intervention content	Depressive rumination episodes, duration, and carryover	Just-in-time adaptive delivery after detected rumination	Greater reductions in rumination episodes and minutes ruminating; evidence of reduced rumination carryover versus control
(19) — Social support just-in-time adaptive intervention while awaiting psychotherapy; microrandomized feasibility study; $N = 25$	Rule-based and personalized decision rules including fixed cutoffs, Shewhart control charts, or self-reported need	High-frequency ecological momentary assessments of negative affect, stress, loneliness, and rumination; in-	Distress reduction by mobilizing social support during vulnerable moments	Adaptive in real time; intervention triggered by real-time assessments according to microrandomized decision rules	High feasibility and compliance; interventions triggered by self-reported need were rated most appropriate and helpful; exploratory distress

(Continued)

TABLE 3 Continued

Study; sample size; design	Personalization axis	Data sources used	Target mechanism or phenotype	Adaptivity	Primary outcome and effect metric (as reported)
		application prompts and support-seeking logs			reductions with small effect sizes
C. Smart messaging and conversational agents: Scheduled or light-tailoring text/app messages that maintain gains or nudge behavior, Agent or therapist-guided online CBT or supportive chat.					
(54) — Smart-messaging as relapse prevention following psychological therapy; observational cohort in routine care; [Study 1: 53 out of 79 completed CBT, divided into smart-messaging (15) and no-messaging (38) groups. Study 2: 14 participants used smart-messaging.]	Rule-based tailoring	Short message service check-ins and follow-up symptom measures	Symptom improvement and relapse prevention after therapy	Static delivery (scheduled tailored messages; not real-time adaptive)	Greater twelve-month symptom improvement for smart-messaging users compared with non-users; stability at six months in routine practice
(13) — Woebot WB001 for postpartum depression; device profile and synthesis of efficacy signals; $N = 36,070$	Model-guided and agent-guided tailoring	In-application conversational exchanges; measurement-based care elements	Postpartum depression symptoms and interpersonal stressors	Adaptive conversational guidance	Device profile summarizing design and supportive efficacy data; no single definitive randomized outcome reported in this article
(46) — MindBot conversational agent (engineering and evaluation report, not tested with users)	Model-based personalization using real-time sentiment monitoring and large language models	In-application conversations; sentiment analysis; predefined templates and dynamic responses	Depressive sentiments and supportive engagement	Adaptive conversational responses	Usability and engagement-oriented results; no clinical depression outcomes reported in this paper
(47) — “How Are You Today?” mobile application using natural language processing to support diagnosis and treatment of anxiety and depression; feasibility engineering report ($N = 63$ synthetic diary entries of a single depressed subject over a nine-week period)	Model-based natural language processing with clinician-tailored follow up	Free-text diary entries; diagnostic screening based on the Diagnostic and Statistical Manual of Mental Disorders Fifth Edition; in-application interactions and notifications	Depressive and anxiety symptom monitoring and prediction of anxiety or panic episodes	Adaptive prompts and feedback based on language analysis with clinician escalation available	Preliminary feasibility with synthetic data; correlation between diary sentiment analysis and self-reported depressive symptom scores; no randomized clinical outcomes reported
(55) — Digital Navigator coaching: guided versus supportive models with mindLAMP; $N = 156$; six-week comparative study	Clinician-guided personalization of application recommendations (Digital Navigator “Guide” versus “Support”)	Smartphone application use logs, activities completed; survey outcomes	Depression and anxiety symptoms; engagement with therapeutic activities	Human-guided adaptation over time with scheduled contacts and tailored suggestions	Guide group completed more activities; thirty-four percent showed at least twenty-five percent decrease in Patient Health Questionnaire-9; thirty-eight percent showed at least twenty-five percent decrease in Generalized Anxiety Disorder-7 overall
D. Decision-support and recommenders: Personalized suggestions or level-of-care decisions learned from responses and preferences, Algorithms choose the best skills/modules for a person.					
(56) — Recommending activities for mental health and well-being: insights from two user studies; a clinical sample ($N = 318$ activities/user) and a non-clinical sample ($N = 59$ activities/user); observational modeling	Model-based	Ecological momentary assessment activity logs and ratings	Behavioral activation target through positive-affect activities	Static tailoring (offline recommendations; not just-in-time)	Personalized models outperformed pooled models; approximately fifty-nine activities per user required before personalized models surpassed general models
(15) — Personalized and optimized therapy algorithm for subthreshold depression (RESILIENT trial); randomized smartphone cognitive behavioral therapy	Model-based prescriptive algorithm recommending best skill or combination	In-application Patient Health Questionnaire-9 and usage data	Depressive symptom reduction in subthreshold depression	Adaptive selection at assignment stage (not continuous just-in-time)	Simulated randomized comparison: personalized and optimized therapy outperformed health information control with standardized mean

(Continued)

TABLE 3 Continued

Study; sample size; design	Personalization axis	Data sources used	Target mechanism or phenotype	Adaptivity	Primary outcome and effect metric (as reported)
with prescriptive modeling; N = 4,469					difference approximately –0.37 and approximately thirty-five percent greater benefit than group-average best
(57) — Screening and Treatment for Anxiety and Depression; N = 516 treated from 5,000 screened; open trial in a university system	Rule-based triage and adaptation by symptom severity and suicide risk	Computerized adaptive testing delivered remotely; ongoing symptom monitoring	Depression and anxiety symptom burden; suicide risk	Dynamic adaptation of level of care over forty weeks	Significant symptom improvements across tiers; feasibility and acceptability reported (no randomized comparison)
E. Wearable-integration and lifestyle support digital interventions: Apps tightly coupled with physiology to personalize timing/content, Tailoring lifestyle inputs as part of a digital care plan.					
(61) — Personalized nutrition for adolescent major depressive disorder; N = 10; single-arm mixed-methods feasibility	Clinician-tailored menus, stepped dietary goals, family context tailoring	Virtual counseling sessions, menu plans, grocery delivery, educational electronic health messages	Depressive symptoms via dietary-mechanism change; family food environment	Scheduled step-up across four bi-weekly sessions	Feasibility achieved; depressive symptoms improved with small-to-moderate effect (Cohen's d approximately 0.36; wide confidence interval)
(60) — Randomized controlled study of a digital data-driven therapeutic for depressive and generalized anxiety symptoms; N = 200 randomized; sixteen weeks	Data-driven personalization of intervention timing and content	Wearable physiology such as skin conductance, activity, and sleep; mobile application interactions	Depressive and generalized anxiety symptoms	Data-triggered adaptive delivery throughout treatment	Intervention achieved larger reductions than waitlist control with high engagement reported
(58) — NEVERMIND pragmatic randomized controlled trial; N = 425	Personalized behavioral content within application modules	Wearable physiological data via smart shirt; mobile application interactions; questionnaires	Depressive symptoms among patients with severe somatic conditions	Scheduled content; not just-in-time	Lower depressive symptoms at twelve weeks versus standard care; effect maintained in per-protocol analysis
(59) — Randomized controlled trial in military personnel (N = 30, divided over three arms)	Data-triggered personalized guidance and provider communication	Wearable physiology; mobile application stress alerts; symptom measures	Symptoms of depression, anxiety, stress, and anger in active-duty populations	Real-time alerts prompting immediate coping techniques	Application group completed therapy and showed significant symptom reductions; control cognitive behavioral therapy group had high dropout

brain images) to identify who benefits from which micro-activities. The goal is a smartphone system that prescribes “what/when/where” brief activities to raise energy, reduce depressive symptoms, and prevent relapse (62). In a similar paradigm, the CRM mobile application (NCT05400785) combines a wearable activity tracker with daily symptom check-ins and provides personalized mood predictions and prevention prompts to the active arm. The aim is to prevent recurrence by turning forecasts into targeted self-management guidance (63). The mHELP study (NCT07017569) extends real-time adaptation to a campus setting. Students use a watch and phone app for 10 weeks with machine-learning stress detection, on-demand skills (breathing, journaling, media), and two telehealth sessions. The trial tests whether state-contingent prompts and light clinician touchpoints improve anxiety, depression, stress, engagement, and service uptake relative to a monitoring-only control (64).

Following a response-guided format, an adaptive, transdiagnostic internet trial (NCT07051148) begins all participants on a 12-module self-applied program then classifies early vs. late responders after three modules. The program personalizes the format of support by randomizing late

responders to add brief therapist sessions or continue self-guided care. Outcomes include symptom change, emotion regulation, and alliance, directly testing whether early-trajectory signals can right-size human support (65).

Multiple programs tailor the training material to individual cognitive or neural profiles. Yale's feasibility and efficacy study of digital neurotherapy (NCT04961047) delivers eight weeks of personalized cognitive exercises to cancer survivors and patients on dialysis, using the REJUVENATE™ system (66). REJUVENATE™ is an at-home, adaptive digital neurotherapy delivering seven game-based exercises that train attention, inhibition, working memory, cognitive flexibility, processing speed, pattern recognition, categorization, and multitasking (67). Earlier studies of the system showed promising results, neuroplasticity-based computerized cognitive remediation (nCCR) produced greater improvements than control in depression severity and cognition, most notably executive function and verbal fluency (68). Concurrently, brain networks showed restoration toward a more efficient, hub-centric architecture: increased rich-club connectivity (69). Aalto University's MEL-T01 game-based therapeutic (NCT05426265)

TABLE 4 Summary of personalized depression digital health clinical trials.

Title	Population	Intervention Type	Addressing Heterogeneity	Personalization Approach
Motor Activity–Subjective Energy (MASE) Project (NCT07059234)	Adults with major depressive disorder (N = 180)	Just-in-time, state-contingent micro-activity prescription informed by within-person activity–energy associations and neurobiological profiling	High (integrates timing and context with neurobiological phenotyping)	Learns individual activity–energy coupling from accelerometry and ecological diaries; applies brain network analyses to stratify who benefits from which micro-activity; delivers when/where/what prompts in daily life
Digital neurotherapy with REJUVENATE (NCT04961047)	Cancer survivors and adults with end-stage kidney disease on dialysis with depressive symptoms (N = 36)	Mechanism-targeted digital neurotherapy that adapts cognitive task difficulty across attention, inhibition, working memory, flexibility, processing speed, pattern recognition, categorization, and multitasking	High (content and target tailored to cognitive phenotype; neuroplasticity rationale)	In-app telemetry drives individualized progression rules; session-by-session calibration of task parameters to the participant's evolving cognitive profile
MEL-T01 “Meliora” game-based digital therapeutics (NCT05426265)	Adults with major depressive disorder (N = 1,001)	Personalized cognitive training embedded in gameplay with continuous adaptation to executive function performance	High (content and target tailored to neurocognitive profile)	Real-time performance monitoring adjusts task difficulty, stimulus characteristics, and progression schedules to maintain individualized challenge and target executive-control deficits
Targeting network dysfunction in apathy of late-life depression (NCT05877885)	Older adults with late-life depression and clinically significant apathy (N = 84)	Customized cognitive-control training targeting attention, salience detection, and cognitive control networks	High (content and target aligned to a neurobiological subtype: apathy)	Training tasks and schedules are tuned to engage hypothesized neural circuits; clinical monitoring supports adherence and safety; seeks circuit-level remediation of apathy-related dysfunction
MIRAI trial of CT-152 (Rejoyn) mobile prescription digital therapeutics (NCT04770285)	Adults with major depressive disorder on stable antidepressant therapy (N = 386)	Mechanism-based digital therapeutics that train networks integrating emotion recognition/processing with cognition; measurement-based progression	High (content and target grounded in neuroplasticity and circuit integration)	Structured neurobehavioral exercises progress according to performance and symptom feedback to promote adaptive re-weighting of emotion–cognition circuitry
Circadian Rhythm for Mood (CRM) mobile application (NCT05400785)	Adults with a history of mood episodes (N = 93)	Just-in-time relapse prevention via daily mood prediction and personalized alerts	Moderate (timing and context; behavioral signals only)	Personalized next-day risk estimation from wearable activity and daily symptom entries; triggers user-specific prevention guidance when predicted risk exceeds threshold
Personalized, response-based transdiagnostic internet intervention (NCT07051148)	Adults with clinically significant anxiety and/or depressive symptoms (N = 366)	Response-adaptive format and dose (self-applied program versus hybrid program with brief therapist sessions for late responders)	Moderate (treatment format and dose adaptation)	Uses early symptom trajectory to classify early versus late responders; late responders randomized to add synchronous therapist sessions; early responders continue or discontinue per protocol
mHELP: Interactive mobile health for high anxiety and depression in college students (NCT07017569)	University students with elevated anxiety, stress, or depressive symptoms (N = 125)	Just-in-time support linked to machine-learning stress detection; adjunct telehealth encounters	Moderate (timing and context with minimal clinician input)	Watch and phone signals detect physiologic or behavioral stress; triggers real-time coping tasks and on-demand skills; two scheduled telehealth sessions used as light-touch dose adaptation
MENTINA: Effect of digital markers in self-management of depressive symptoms (NCT06919133)	Adults with current or prior depressive episodes or elevated depressive symptoms (N = 660)	Rule-based self-management with escalation: questionnaire and sensor monitoring drive feedback and safety suggestions	Low (content and target via rule-based personalization; safety escalation)	Predefined feedback rules map self-reports and sensor patterns to tailored psychoeducational content and recommendations, including prompts to contact emergency care when indicated

embeds neurocognitive training and therapeutic content into game mechanics. The main therapeutic component consisted of continuous in-game performance measurement coupled with adjustment of the neurocognitive training content to individual executive function levels, which also dynamically change over time. Adults with major depressive disorder are randomized to the active device, a comparator game, or treatment-as-usual, with symptom and cognition outcomes at 4, 8, 12, and 24 weeks (70, 71). Another trial by AdventHealth in late-life depression with apathy (NCT05877885) targets network dysfunction using a customized cognitive-training protocol on the Posit Science platform, with weekly care-manager support. Primary aims

include changes in brain connectivity, apathy severity, and cognitive control (72).

Additionally, the MIRAI trial (NCT04770285) evaluates the mobile digital therapeutic CT-152, known as Rejoyn, in adults with major depressive disorder on antidepressant monotherapy. Rejoyn is a prescription app-based digital therapeutic designed to leverage neuroplasticity by training networks that integrate emotion recognition and processing with cognition (73). Data showed consistent symptom improvements across clinician- and patient-reported scales with continued gains one month post-treatment (74). Finally, The MENTINA trial is an international, multicenter randomized controlled trial (Denmark, Germany,

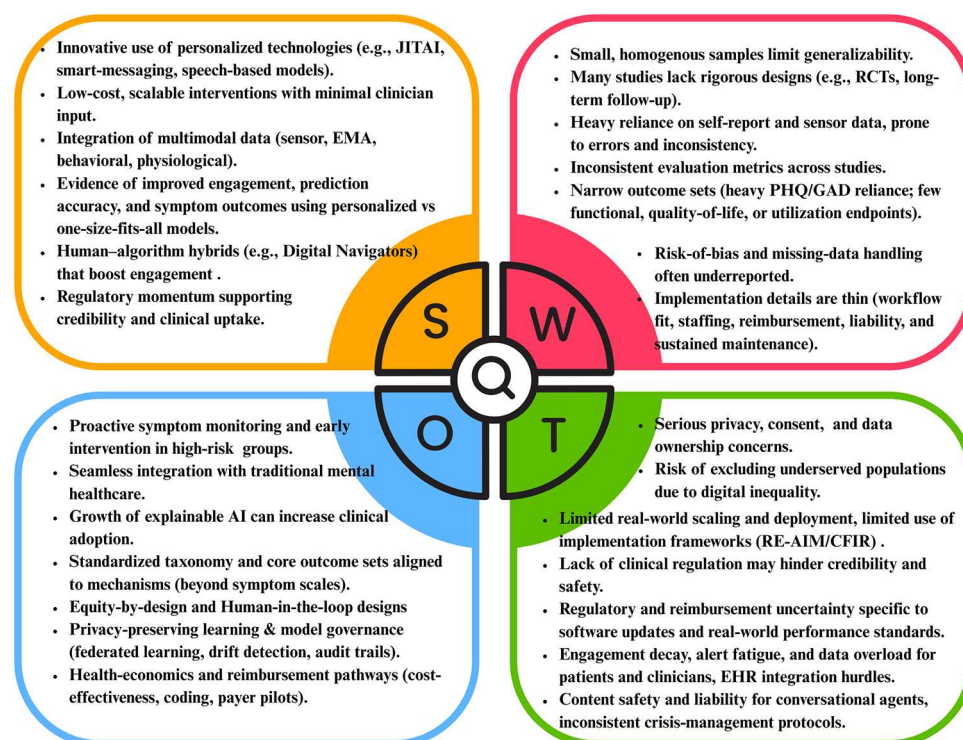


FIGURE 3
SWOT analysis of reviewed literature.

Spain) testing a smartphone self-management app for depression. Participants are randomized to active rule-based feedback vs. monitoring-only control. The rule-based feedback is generated based on self-monitored data and sensor data collected from the smartphones. The purpose of this rule-based feedback is to suggest supportive actions to participants, such as reading items from a content library within the app or contacting emergency healthcare facilities (75, 76).

3.1 SWOT analysis

Personalized digital interventions for depression show strong potential due to their scalability, multimodal integration (e.g., EMA, sensors), and promising engagement outcomes. Opportunities include early intervention, Explainable Artificial Intelligence (XAI), and integration with traditional care. However, weaknesses such as small samples, inconsistent evaluation, and reliance on self-report data persist. Major threats include data privacy issues and digital exclusion of underserved populations. The SWOT results are shown in Figure 3.

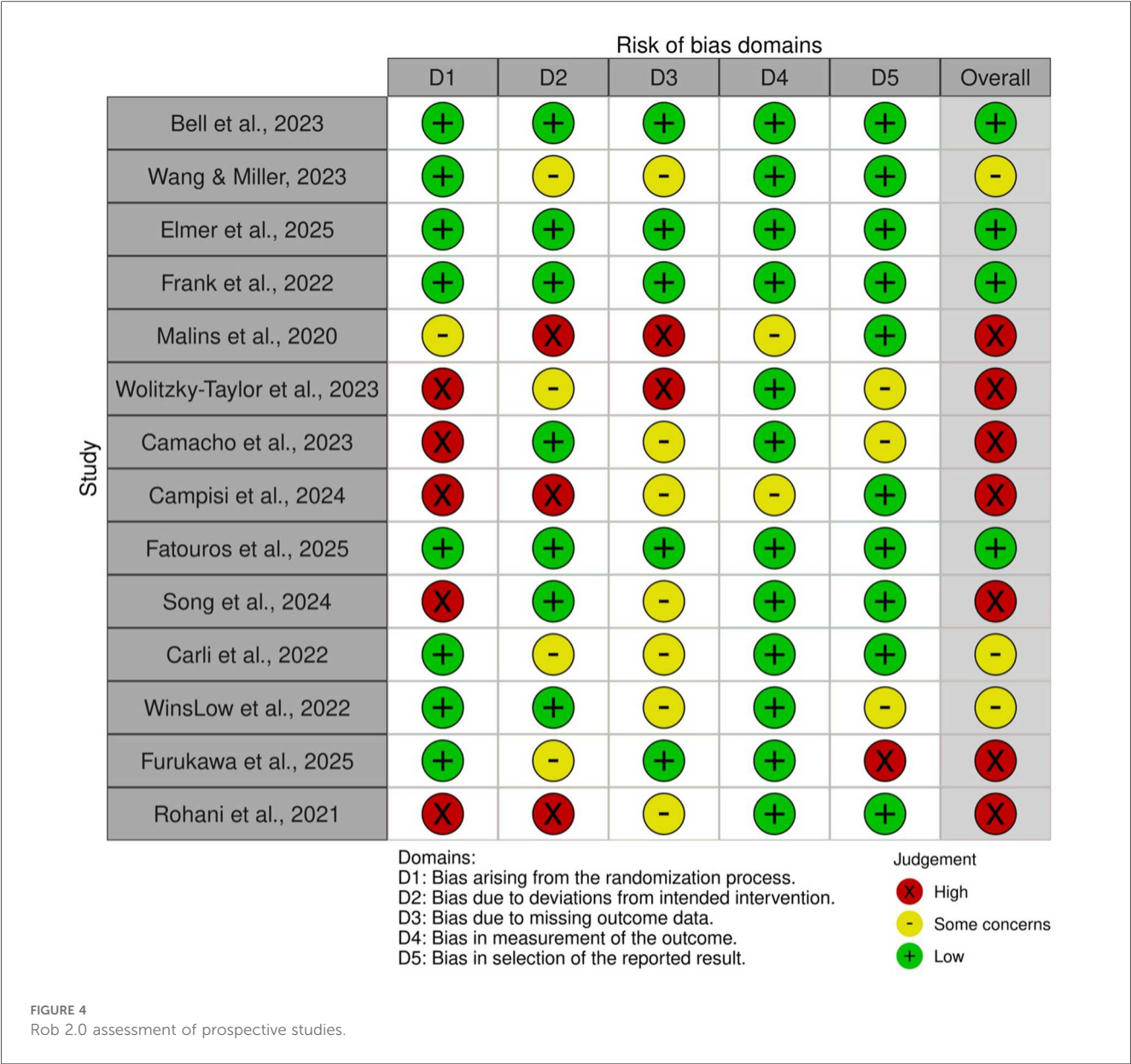
3.2 Risk of bias assessment

4 studies (19, 51, 53, 60) showed a low overall risk of bias, as shown in Figure 4. Most others raised concerns, especially

regarding randomization, missing data, and selective reporting. The predictive models assessed in the 3 retrospective studies were found to exhibit low to intermediate risk of bias, as shown in Table 5. These limitations highlight the need for more rigorous designs and standardized outcome reporting in future digital mental health research.

Across three non-scored narrative appraisals, Woebot (WB001), MindBot, and HAYT are clearly described but differ in clinical maturity and evidentiary strength. WB001 provides the most complete clinical framing, an FDA Breakthrough, prescription 8-week Agent-delivered CBT and interpersonal therapy program with daily mood tracking and NLP-based crisis detection, yet the brief lacks detailed data-governance disclosures (13). MindBot offers a rigorous engineering overview (preprocessing, sentiment thresholds, LLM-augmented replies), but it provides limited bias-monitoring and escalation specifics (46). HAYT delineates a clinician-integrated workflow (NLP of diary entries, DSM-5 questionnaires, CBT prompts, secure messaging) with transparent data flow in principle, though results are based on synthetic data and do not validate real-world safety or clinical impact (47). Across all three, safety provisions are conceptually present but unspecified in performance terms. Credibility would be strengthened by explicit reporting on privacy, human-in-the-loop escalation, and algorithm update policies.

In terms of transparent research practices and open scientific publication, we observed heterogeneous adoption of open-science



practices across the corpus. Trial registration was common in RCTs, but public protocols, pre-specified analysis plans, de-identified data, and analysis code were infrequently shared in most studies. We therefore flag transparent preregistration (e.g., [ClinicalTrials.gov/OSF](#)), protocol publication, and routine data- and code-sharing with privacy safeguards as concrete steps to reduce selective-reporting risk and improve reproducibility in this rapidly evolving field.

4 Discussion

Depression represents a multifaceted neuropsychiatric condition characterized by pronounced phenotypic heterogeneity across symptomatology, etiology, pathophysiology, and treatment response trajectories. This intrinsic variability manifests through diverse clinical presentations, ranging from

predominantly somatic manifestations to primarily cognitive dysfunctions, thereby challenging the efficacy of standardized therapeutic approaches. Contemporary nosological frameworks such as the DSM-5 provide categorical diagnostic parameters; however, these fail to capture the dimensional complexity of depressive phenomenology as revealed through advanced digital phenotyping methodologies and precision psychiatry initiatives.

The suboptimal efficacy of conventional interventions may be attributed to their inability to accommodate inter-individual variability in symptom constellations, neurobiological substrates, and psychosocial determinants. This problematic homogenization of heterogeneous depressive states necessitates a paradigmatic shift toward personalized intervention strategies informed by multimodal assessment protocols and computational modeling techniques.

The examined literature corpus demonstrates variable engagement with depression heterogeneity, reflecting a spectrum of

TABLE 5 PROBAST results.

Study	D1 Participants		D2 Predictors			D3 Outcome						Overall
	1.1	1.2	2.1	2.2	2.3	3.1	3.2	3.3	3.4	3.5	3.6	
Gerczuk et al. (17)	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟡
Chatterjee et al. (49)	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Kathan et al. (48)	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🔴	🟢	🟢	🟢	🔴
Study	D4 Analysis										Overall	
	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9			
Gerczuk et al. (17)	🟢	🟢	🟡	🟡	🟢	🟢	🟢	🟡	🔵?	🟡		
Chatterjee et al. (49)	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢		
Kathan et al. (48)	🟢	🟢	🔴	🟢	🟢	🟢	🟢	🟢	🟢	🔴		
D1: 1.1 Appropriate data sources used?; 1.2 Inclusions/exclusions appropriate? D2: 2.1 Predictors defined & assessed uniformly?; 2.2 Predictor assessment blinded to outcome?; 2.3 Predictors available at intended use? D3: 3.1 Outcome determined appropriately?; 3.2 Pre-specified/standard outcome definition?; 3.3 Predictors excluded from outcome definition?; 3.4 Outcome measured uniformly?; 3.5 Outcome assessment blinded to predictors?; 3.6 Time interval appropriate? D4: 4.1 Reasonable number of participants with outcome?; 4.2 Continuous/categorical predictors handled appropriately?; 4.3 All enrolled participants included in analysis?; 4.4 Missing data handled appropriately?; 4.5 Univariable predictor screening avoided?; 4.6 Complexities (e.g., censoring, competing risks) appropriately handled?; 4.7 Relevant performance measures evaluated appropriately?; 4.8 Overfitting & optimism accounted for?; 4.9 Predictor weights correspond to reported analysis?										Judgement 🟢 Yes 🟡 Probably not 🔴 No 🔵? No information		

methodological sophistication in addressing inter-individual and intra-individual symptom variability. Here, a two-dimensional evaluation framework (Figure 5) is proposed to assess the studies according to: (1) their level of heterogeneity engagement (low vs. high) and (2) their implementation stage (model building vs. clinical deployment). This taxonomic approach reveals significant disparities in how depression heterogeneity is operationalized across the research spectrum.

4.1 Models with limited heterogeneity integration

Gerczuk et al. (17) personalize mood estimation to speaker identity, capturing inter-speaker variance in acoustic features, yet remain uncustomized to symptom mechanisms (17). MindBot and HAYT demonstrate NLP and LLM-assisted pipelines and, in HAYT’s case, a clinician-messaging architecture using synthetic data (46, 47). These works illustrate that personalization can be embedded at the interface layer; nonetheless, they largely bypass clinical variance, no discussion of clinical endpoints, no subtype stratification, and limited attention to safety, governance, or model drift. Proxy metrics (e.g., “emotional accuracy”) and synthetic diaries risk overstating clinical readiness while under-representing failure modes in non-standard language, high-distress states, or low-literacy populations.

4.2 Clinical applications with limited heterogeneity integration

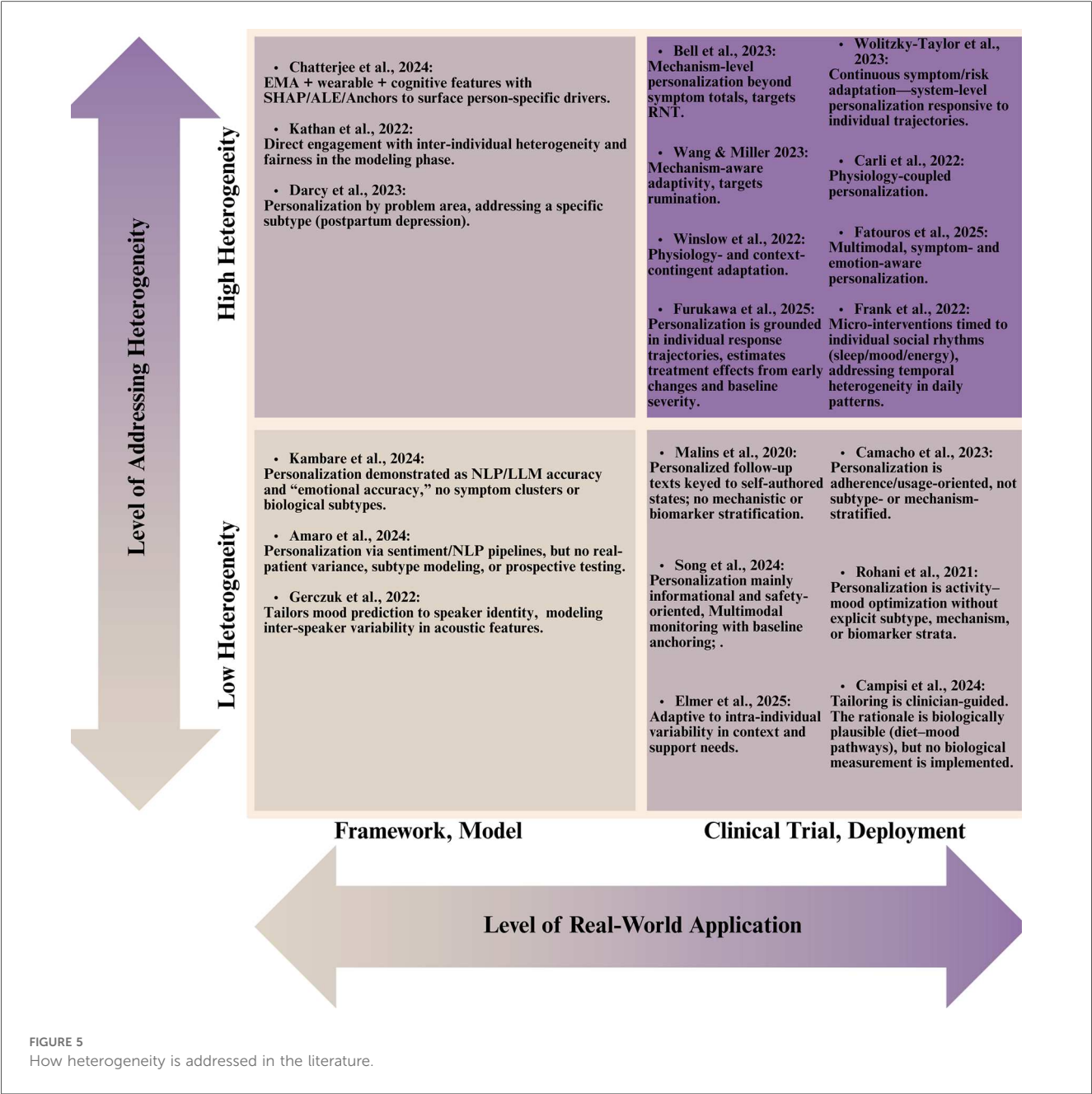
These studies personalize primarily by schedule, adherence support, or single-channel signals. Rohani et al. (56) evaluated recommendation algorithms for pleasant event scheduling across clinical and non-clinical populations. While their approach

accommodates intra-individual variability in activity-mood associations, it does not explicitly model symptom heterogeneity or depression subtypes, instead focusing on behavioral intervention optimization through reinforcement learning principles (56).

Malins et al. (54) investigated personalized relapse prevention messaging CBT, with messages tailored to individual recovery trajectories across distinct clinical states (wellness maintenance, early warning signs, full relapse). This strategy acknowledges heterogeneity in relapse vulnerability and symptom recognition patterns, though it relies on clinician-guided personalization without mechanistic or biomarker stratification (54). The m-Path social-support JITAI tailors timing and content via self-signaled need and momentary affective context (19). Adaptation is person-specific but not mechanism- or biotype-aware. This creates person-level adaptation on two axes: detection thresholds based to individual baselines and action menus refined by user preference. Yet the mechanism model remains indifferent to subtype. There is no personalization by symptoms, no biophysical state inputs, and no contextual signals (19).

Personalization in the living-lab for socially vulnerable older adults is primarily achieved through individualized baselining and longitudinal feedback rather than just-in-time control (50). Participants wear multimodal sensors (e.g., heart-rate variability, sleep efficiency/fragmentation, activity) and complete high-frequency chatbot PHQ check-ins. These streams are harmonized to compute each person’s baseline and day-over-day deltas for mood and physiology. Personalization therefore resides in the personal reference model and tailored visualizations that support self-care and caregiver awareness. However, there is no policy that optimizes timing or content based on estimated treatment effects, and no stratification by symptom dimension or biotype (50).

Digital-navigator coaching personalizes engagement rather than content; coaches review mindLAMP activity and symptom self-



reports and then tailor next steps like what module to try next and which homework to emphasize, via human-in-the-loop heuristics (55). Adaptation is driven by recent completion patterns, working alliance scores, and user-stated goals, with cadence modulated to sustain adherence. This raises dose and breadth of app use at the person level but personalization is pragmatic, not mechanistic (55). The personalized nutrition program for adolescent MDD individualizes targets at three layers: diet goals, eating-behavior, and a weekly menu plan tuned to family preferences and shopping logistics (61). Although the intervention is biologically plausible, leveraging Mediterranean-style patterns to influence inflammation, gut–brain signaling, and metabolic rhythms, the personalization engine does not measure or stratify by biological markers like CRP or microbiota, circadian rhythm, or cognitive symptoms. As a result,

“biology-aware” matching of diet to patient is not implemented and personalization remains preference-based rather than biology-directed (61). Most of the reviewed clinical trials implement personalization strategies centered primarily on behavioral parameters (e.g., mood fluctuations, patient-reported outcomes, application engagement metrics) or symptom severity indices, without establishing explicit linkages to biological subtypes or neurobiological mechanisms. CRM forecasts personalized next-day mood from wearables and check-ins to trigger risk-contingent prompts (NCT05400785) (63). The response-based transdiagnostic internet trial adapts treatment format and dose by using early symptom trajectories to add brief therapist sessions for late responders (NCT07051148) (65). mHELP combines ML stress detection on phone and smartwatch streams

with on-demand skills and two light telehealth touchpoints (NCT07017569) (64). MENTINA applies personalized rule-based feedback from questionnaire and sensor data, escalating to safety guidance when indicated (NCT06919133) (75). These approaches, while clinically valuable, conceptualize depression predominantly as a psychological or behavioral syndrome, neglecting to differentiate or specifically target distinct biological subtypes that may respond differentially to intervention strategies.

4.3 Models with advanced heterogeneity integration

Chatterjee et al. (49) implemented explainable deep learning architectures for mood prediction using multimodal data streams (EMA, physiological parameters, cognitive assessments). Their methodological innovation lies in the application of interpretability techniques (SHAP, ALE, Anchors) to elucidate person-specific predictors of mood disturbance. This approach enables computational phenotyping aligned with natural symptom expression, though it does not explicitly classify clinical subtypes or stratify based on symptom dimensions (49). On the other hand, the Woebot WB001 device profile engages a specific clinical subtype (postpartum depression), illustrating how tailoring by life context can anchor personalization even at the caregiving stage (13).

Kathan et al. (48) conducted a comparative evaluation of multiple personalization strategies for symptom prediction, including transfer learning with shared and individualized components, subject-specific data preprocessing, and gender-based stratification. Their explicit assessment of model equity across demographic subgroups represents one of the few studies directly addressing fairness in computational personalization. This is a critical consideration given the documented disparities in depression presentation across demographic strata. Their multifaceted approach to heterogeneity encompasses both methodological innovation and equity considerations (48).

Together, these studies show that individualized models outperform pooled baselines and that equity assessment is feasible, marking a substantive step beyond undifferentiated prediction. However, they stop short of prospective decision rules: neither specifies thresholds that trigger action, nor do they test whether model outputs change behavior, adherence, or outcomes in real life. Small, intensively monitored cohorts also risk selection bias and limit generalizability. Moreover, depressive mechanisms are largely statistical rather than biological, in other words, predictive features are not validated against circuitry, inflammatory markers, or subtype taxonomies.

4.4 Clinical applications with advanced heterogeneity integration

Several interventions tailor when and what to deliver using mechanisms that vary across people. Bell et al. (51) evaluated the Mello application, which delivers personalized cognitive-behavioral interventions based on real-time assessment of mood, rumination, location, and activity. By targeting a transdiagnostic

mechanism, RNT, that presents with substantial inter-individual variability, their approach shows sophisticated engagement with heterogeneity beyond symptomatic expression. The intervention's significant efficacy ($d = 0.50$ for depression, $d = 0.61$ for anxiety, $d = 0.87$ for RNT) and mediation findings support the clinical utility of mechanism-focused personalization strategies (51). Similarly, the JITAI by Wang & Miller (52) measures receptivity and targets rumination, showing medium-to-large effects and demonstrating that targeting heterogeneous cognitive processes can outperform symptom-total heuristics (52). Still, mechanisms are inferred from self-reports, passive indicators that could address cognitive states (speech, mobility, physiology) would add more mechanistic information.

Across other programs, personalization is implemented through distinct sensing and decision layers. Cue derives an individualized "social rhythm" baseline from passive smartphone traces (sleep-wake cycle, mobility, communication patterns) augmented by brief mood/energy check-ins. Deviations from that baseline trigger rule-mapped micro-interventions at a cadence designed to minimize alert fatigue by suppressing and rotating content and timing based on recent engagement. While this addresses symptom level and temporal heterogeneity, a fuller account should discuss biological correlates, particularly circadian regulation, as mechanistic levers for personalization (53).

Physiology-coupled systems (NEVERMIND, FeelDTx, smartwatch-CBT) learn per-user baselines of HR and HRV, electrodermal activity, sleep efficiency and fragmentation, temperature and activity; state detectors (rule-based or lightweight ML) fire context-appropriate prompts (paced breathing/HRV biofeedback for sympathetic arousal, mindfulness or reframing during negative-affect windows, sleep-hygiene guidance nocturnally) with refractory periods and weekly navigator/therapist touchpoints where applicable (58–60). While the addition of physiological markers introduces symptomatic and temporal heterogeneity to the intervention, future iterations should incorporate biotype-aware targeting, subgroup calibration, and fairness audits to raise the overall level.

STAND and RESILIENT operationalize personalization at the system level, addressing heterogeneity in treatment effects even if not biomarker-defined. Through repeated symptom and suicide-risk assessments, the STAND algorithm assigns the user to one of stepped tiers of care with real-time alerts and scripted escalation pathways (57). The RESILIENT models use baseline PHQ-9 patterns and early symptoms to estimate individual treatment effects, then assign the "best next" skill block (15).

Multiple trials showed sophisticated engagement with neurobiological heterogeneity, representing a paradigmatic advancement in precision psychiatry. The Digital Therapeutics for Apathy in Late-life Depression trial (NCT05877885) (72) specifically targets network dysfunction underlying apathy, a clinically distinct dimension frequently observed in geriatric depression presentations. The intervention architecture is meticulously designed to modulate specific neural circuitry, particularly networks subserving attention allocation, salience detection, and cognitive control functions; all systems that are consistently implicated in motivational deficits across neuroimaging and neurobiological investigations.

Similarly, the FDA-regulated clinical trial Effectiveness of Digital Therapeutics in Major Depressive Disorder (74) evaluates *Rejoyn*, a digital therapeutic explicitly engineered to engage neural networks implicated in major depressive psychopathology, with particular emphasis on neuroplasticity mechanisms. Unlike conventional approaches that conceptualize depression as a unitary construct, this trial acknowledges fundamental neural circuitry heterogeneity in MDD, implementing structured neurobehavioral exercises specifically designed to recalibrate dysfunctional neural circuits. Notably, the trial documentation explicitly addresses limitations of traditional pharmacological interventions that primarily target neurochemical dysregulation, highlighting the critical importance of circuit-based therapeutic approaches.

The MASE project personalizes timing and content by learning each person's Activity-Subjective Energy Association from accelerometry and ecological diaries, then combining this with neurobiological profiling to prescribe micro-activities matched to an individual's brain phenotype (NCT07059234) (62). REJUVENATE adapts task parameters and progression rules across multiple cognitive domains using in-app telemetry, aiming to remediate person-specific cognitive control deficits via neuroplasticity-based training (NCT04961047) (66). MEL-T01 embeds continuous performance sensing into gameplay to dynamically adjust difficulty, stimuli, and reinforcement schedules to an individual's executive-function profile (NCT05426265) (70).

Collectively, these studies operationalize heterogeneity at multiple levels, behavioral dynamics, cognitive phenotype, and neural circuitry, thus moving beyond severity-only tailoring. Methodological challenges remain (e.g., ensuring reliability of neurobehavioral markers, external validity across biotypes, and fairness audits for model-guided decisions), but the personalization mechanisms are technically rich and mechanism-aligned.

4.5 Methodological limitations and implementation challenges

Despite promising advances, several critical limitations persist across the reviewed literature:

- Limited Integration of Established Subtypes

Few studies explicitly incorporate validated depression subtypes (e.g., melancholic, atypical, anxious) or empirically derived symptom dimensions into their personalization frameworks. This disconnect between clinical phenotyping research and digital intervention development represents a significant translational gap that restricts mechanistic specificity and may obscure heterogeneous treatment effects across clinically meaningful subgroups.

- Predominance of Unidimensional Outcome Measures

Many studies rely on composite mood scores or generalized depression severity metrics (e.g., PHQ-9, PHQ-8) rather than multidimensional symptom assessments. This approach potentially obscures differential effects on specific symptom

clusters, such as sleep dysregulation, anergia, or cognitive slowing, that may vary across depression subtypes. Symptom-network analyses and domain-specific endpoints would better capture mechanism-aligned change and enable more nuanced estimation of heterogeneous treatment effects.

- Sample Limitations and Generalizability of Machine-Learning Findings

Across studies, samples are typically small (often $N < 100$), convenience-based, and demographically narrow, which amplifies overfitting risk and limits transportability of findings. Reported performance gains typically reflect within-sample or cross-validated performance against naïve baselines rather than out-of-distribution performance across settings, devices, or clinical strata.

- Insufficient Attention to Algorithmic Equity

With the notable exception of Kathan et al. (48), most studies inadequately address potential algorithmic biases across demographic and clinical subgroups. Few models report disaggregated performance metrics by gender, age, race/ethnicity, or socioeconomic status, nor do they examine differential false-alert rates or error patterns that could exacerbate existing health disparities. Given well-documented differences in depression presentation, help-seeking behavior, and digital access across demographic contexts, this represents a critical oversight. Routine fairness diagnostics, including performance parity, error symmetry, and calibration equity analyses, should be standard practice, accompanied by corrective strategies (e.g., stratified sampling, group-aware decision thresholds, re-weighting) and transparent reporting of any performance-fairness trade-offs.

- Limited External Validation and Deployment Science

Even prototype systems demonstrating efficacy within controlled research settings lack replication across independent health systems, diverse payer contexts, and real-world clinical workflows. Implementation outcomes such as reach, adoption fidelity, maintenance, and cost-effectiveness are consistently underreported. Without integration of established implementation science frameworks (for example, the Reach, Efficacy, Adoption, Implementation, and Maintenance framework (RE-AIM) (22) and the Consolidated Framework for Implementation Research (CFIR) (18)) and inclusion of payer-relevant endpoints, questions of scalability and sustainability remain unanswered. Prospective evaluations should ideally embed models within actual clinical workflows through silent deployment or randomized alert configurations to estimate real-world effectiveness and detect performance drift over time.

- Inadequate Temporal Resolution

Despite theoretical emphasis on dynamic symptom fluctuation, many studies employ relatively sparse assessment protocols that fail to capture rapid symptom oscillations potentially indicative of specific depression subtypes or vulnerability patterns. Current sampling cadences often miss context switches, diurnal mood variations, and event-triggered state changes that are critical for just-in-time adaptive interventions. Future designs should favor denser, temporally aligned multimodal data collection with

principled handling of missingness through pattern diagnostics and multiple imputation methods tailored to time-series contexts.

4.6 Future directions and methodological recommendations

Future work must begin by grounding personalization algorithms in empirically derived depressive subtypes. Rather than treating depression as a unidimensional construct, researchers could apply unsupervised learning techniques to multimodal datasets (77). These data-driven subgroups can then inform model architectures that tailor predictions to the unique symptom constellations of melancholic, atypical, anxious, or other clinically meaningful clusters. At the same time, interventions should move beyond composite mood scores and incorporate multidimensional symptom assessments, using network-analysis approaches to map the dynamic interrelationships among symptoms (78). By capturing how fatigue, anhedonia, sleep disturbance, and cognitive dysfunction co-activate and cascade over time, digital tools can offer more precise, subtype-specific feedback and treatment recommendations. Following the promising approaches discussed earlier, interventions should prioritize targeting transdiagnostic mechanisms with established heterogeneity in addition to general behavioral patterns and preferences.

Progress will hinge on moving beyond small, convenience samples toward adequately powered, diverse, and prospectively enrolled cohorts. Multi-site recruitment with stratified targets (age, sex/gender, race/ethnicity, socioeconomic status, language, device/OS) should be pre-specified and monitored. To curb overfitting and analytic bias, modeling work should adopt participant-level, temporally blocked, nested cross-validation and report both discrimination (e.g., MAE/RMSE, AUROC) and calibration (slope, intercept, expected calibration error).

Equity must become a core design principle rather than an afterthought. Future personalization frameworks should adopt fairness-aware optimization strategies—such as demographic parity or equalized odds constraints—to ensure consistent performance across gender, age, and cultural groups (79). Algorithmic bias can systematically skew who benefits and who is harmed. Disparate error rates (e.g., higher false positives for crisis alerts in one group, higher false negatives in another) distort triage, amplify clinician workload unevenly, and may delay care for those already underserved. Interventions would benefit from routine bias audits that quantify disparities in predictive accuracy and treatment suggestions, followed by algorithmic recalibration where needed (80). In parallel, there is a need for consensus on reporting standards, research protocols ought to include a “heterogeneity specification” checklist that details how subtypes were defined, which symptom dimensions were assessed, and what bias-mitigation techniques were implemented. Such standardized reporting will facilitate cross-study comparison, meta-analysis, and the cumulative advancement of the field. Moreover, to strengthen transparency and reproducibility, depression-focused digital trials should adopt open-science practices such as preregistration and open sharing of code and de-identified data. Practical steps like

registering hypotheses and analysis plans before data collection, archiving code and data in trusted repositories, and publishing preprints, are feasible now and would materially improve credibility in this fast-moving field (81, 82).

Additionally, truly user-centered innovation requires embedding human-centered design (HCD) throughout the development lifecycle. Researchers should engage patients, clinicians, and caregivers in participatory co-design workshops, iteratively refining wireframes, feature sets, and interaction flows based on real-world feedback (83, 84). Usability testing, employing think-aloud protocols and standardized measures (e.g., the System Usability Scale), can uncover interaction bottlenecks before large-scale deployment. Accessibility and inclusivity audits, evaluating readability, language support, digital literacy, and disability accommodations, will further ensure that personalized digital mental health tools are equitable and resonate with diverse user populations. By integrating robust analytics, fairness safeguards, and rigorous HCD practices, the next generation of interventions can fulfill the promise of precision psychiatry in a way that is both scientifically sound and deeply humane.

Finally, clinical integration should move beyond app availability to workflow-embedded, measurement-based care. Health systems can operationalize digital tools through (i) EHR-integrated screening and triage that route patients to matched interventions; (ii) a defined digital navigation role for onboarding, troubleshooting, and engagement support; and (iii) scheduled reassessment checkpoints that trigger escalation, switching, or augmentation within stepped-care pathways (85, 86). Interoperability and clear reimbursement pathways are critical to sustain routine use. Implementation should be guided by RE-AIM/CFIR with pragmatic and adaptive trials embedded in care to learn which components drive benefit (87). Systems must add equity and safety guardrails, like stratified performance rates, multilingual/low-literacy designs, transparent data-governance, and ethical frameworks. Finally, routine workload and cost accounting should inform scalable resourcing decisions (85, 86). Together, these steps shift digital interventions from promising pilots to reliable, equitable infrastructure for depression care.

5 Conclusion

The heterogeneous nature of depression necessitates personalized intervention approaches that accommodate inter-individual variability in symptom presentation, etiological factors, and treatment response patterns. Digital mental health technologies offer unprecedented opportunities for implementing such precision approaches at scale, yet current methodologies demonstrate variable engagement with heterogeneity, from superficial customization to sophisticated computational phenotyping.

The most promising approaches, exemplified by Bell et al. (51), Wang and Miller (52), and Frank et al. (53), integrate advanced computational methods with clinically informed conceptualizations of depression heterogeneity. By targeting

transdiagnostic mechanisms, employing interpretable ML, and evaluating algorithmic equity, these studies point toward a future where digital interventions can truly accommodate the multidimensional nature of depressive psychopathology.

However, significant methodological challenges remain, particularly regarding the integration of established clinical phenotypes, multidimensional outcome assessment, and cross-population validation. Addressing these limitations will require interdisciplinary collaboration between clinical researchers, computational scientists, and implementation specialists to ensure that technological innovations translate into meaningful clinical outcomes across the heterogeneous spectrum of depressive disorders.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

Author contributions

GA: Formal analysis, Writing – original draft, Conceptualization, Investigation, Data curation, Writing – review & editing, Methodology. SD: Writing – review & editing, Conceptualization, Supervision. YH: Writing – review & editing.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This study was

supported in part by the American University of Sharjah and Mohammed Bin Rashid University of Medicine and Health Sciences.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Major depressive disorder - Level 4 cause | Institute for Health Metrics and Evaluation [WWW Document]. (n.d.). Available online at: <https://www.healthdata.org/research-analysis/diseases-injuries-risks/factsheets/2021-major-depressive-disorder-level-4-disease> (Accessed September 12, 2025)
2. Mental health of adolescents [WWW Document]. (n.d.). Available online at: <https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health> (Accessed March 10, 2025)
3. Weaver ND, Bertolacci GJ, Rosenblad E, Ghoba S, Cunningham M, Ikuta KS, et al. Global, regional, and national burden of suicide, 1990–2021: a systematic analysis for the global burden of disease study 2021. *Lancet Public Health*. (2025) 10:e189–202. doi: 10.1016/S2468-2667(25)00006-4
4. Depressive disorder (depression) [WWW Document]. (n.d.). Available online at: <https://www.who.int/news-room/fact-sheets/detail/depression> (Accessed September 9, 2024)
5. IHME. Mental health [WWW Document]. (n.d.). Available online at: <https://www.healthdata.org/research-analysis/health-risks-issues/mental-health> (Accessed March 10, 2025)
6. Santomauro DF, Mantilla Herrera AM, Shadid J, Zheng P, Ashbaugh C, Pigott DM, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet*. (2021) 398:1700–12. doi: 10.1016/S0140-6736(21)02143-7
7. Monroe SM, Anderson SE. Depression: the shroud of heterogeneity. *Curr Dir Psychol Sci*. (2015) 24:227–31. doi: 10.1177/0963721414568342
8. Nierenberg AA. Advancing the treatment of depression with personalized medicine. *J Clin Psychiatry*. (2012) 73:e(17). doi: 10.4088/jcp.11031tx3c
9. Rolls ET. Cortical neurodynamics, schizophrenia, depression, and obsessive-compulsive disorder. In: Mishara AL, Moskalewicz M, Schwartz MA, Kranjec A, editors. *Phenomenological Neuropsychiatry: How Patient Experience Bridges the Clinic with Clinical Neuroscience*. Cham: Springer International Publishing (2024). p. 119–43. doi: 10.1007/978-3-031-38391-5_11
10. Davidson RJ, Lewis DA, Alloy LB, Amaral DG, Bush G, Cohen JD, et al. Neural and behavioral substrates of mood and mood regulation. *Biol Psychiatry*. (2002) 52:478–502. doi: 10.1016/S0006-3223(02)01458-0
11. Nguyen T-D, Harder A, Xiong Y, Kowalec K, Hägg S, Cai N, et al. Genetic heterogeneity and subtypes of major depression. *Mol Psychiatry*. (2022) 27:1667–75. doi: 10.1038/s41380-021-01413-6
12. Fried EI, Proppert RKK, Rieble CL. Building an early warning system for depression: rationale, objectives, and methods of the WARN-D study. *Clin Psychol Eur*. (2023) 5:e10075. doi: 10.32872/cpe.10075
13. Darcy A, Beaudette A, Chiauzzi E, Daniels J, Goodwin K, Mariano TY, et al. Anatomy of a woebot® (WB001): agent guided CBT for women with postpartum depression. *Expert Rev Med Devices*. (2023) 20:1035–49. doi: 10.1080/17434440.2023.2280686
14. Fried E. Moving forward: how depression heterogeneity hinders progress in treatment and research. *Expert Rev Neurother*. (2017) 17:423–5. doi: 10.1080/14737175.2017.1307737
15. Furukawa TA, Noma H, Tajika A, Toyomoto R, Sakata M, Luo Y, et al. Personalised & optimised therapy (POT) algorithm using five cognitive and behavioural skills for subthreshold depression. *NPJ Digit Med*. (2025) 8:531. doi: 10.1038/s41746-025-01906-6

16. Deif R, Salama M. Depression from a precision mental health perspective: utilizing personalized conceptualizations to guide personalized treatments. *Front Psychiatry*. (2021) 12:650318. doi: 10.3389/fpsy.2021.650318
17. Gerczuk M, Triantafyllopoulos A, Amiriparian S, Kathan A, Bauer J, Berking M, et al. Personalised deep learning for monitoring depressed mood from speech. *2022 E-Health and Bioengineering Conference (EHB) Presented at the 2022 E-Health and Bioengineering Conference (EHB)* (2022). p. 1–5. doi: 10.1109/EHB55594.2022.9991737
18. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. *Implement Sci*. (2022) 17:75. doi: 10.1186/s13012-022-01245-0
19. Elmer T, Wolf M, Snippe E, Scholz U. A social support just-in-time adaptive intervention for individuals with depressive symptoms: feasibility study with a microrandomized trial design. *JMIR Ment Health*. (2025) 12:e74103. doi: 10.2196/74103
20. Falkai P, Koutsouleris N. Why is it so difficult to implement precision psychiatry into clinical care? *Lancet Reg Health Eur*. (2024) 43:100952. doi: 10.1016/j.lanepe.2024.100952
21. Glannon W. Biomarkers in psychiatric disorders. *Camb Q Health Ethics*. (2022) 31:444–52. doi: 10.1017/S0963180122000056
22. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health*. (1999) 89:1322–7. doi: 10.2105/ajph.89.9.1322
23. Delgadillo J, Ali S, Fleck K, Agnew C, Southgate A, Parkhouse L, et al. Stratified care vs stepped care for depression: a cluster randomized clinical trial. *JAMA Psychiatry*. (2022) 79:101–8. doi: 10.1001/jamapsychiatry.2021.3539
24. Goldberg D. The heterogeneity of “major depression”. *World Psychiatry*. (2011) 10:226–8. doi: 10.1002/j.2051-5545.2011.tb00061.x
25. Diagnostic and statistical manual of mental disorders: DSM-5TM. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5TM*. 5th ed. Arlington, VA, US: American Psychiatric Publishing, Inc. (2013). doi: 10.1176/appi.books.9780890425596
26. DSM-5 Criteria for Major Depressive Disorder [WWW Document]. (n.d.). MDCalc. Available online at: <https://www.mdcalc.com/calc/10195/dsm-5-criteria-major-depressive-disorder> (Accessed May 9, 2025)
27. Beijers L, Wardenaar KJ, van Loo HM, Schoevers RA. Data-driven biological subtypes of depression: systematic review of biological approaches to depression subtyping. *Mol Psychiatry*. (2019) 24:888–900. doi: 10.1038/s41380-019-0385-5
28. Hannon K, Easley T, Zhang W, Lew D, Sotiras A, Sheline Y, et al. 15. Heterogeneity in depression: evidence for distinct clinical and neurobiological profiles. *Biol Psychiatry*. (2023) 93:S100. doi: 10.1016/j.biopsych.2023.02.255
29. Overview | Depression in adults: treatment and management | Guidance | NICE [WWW Document]. (2022). Available online at: <https://www.nice.org.uk/guidance/ng222> (Accessed May 15, 2025)
30. Park LT, Zarate CA. Depression in the primary care setting. *N Engl J Med*. (2019) 380:559–68. doi: 10.1056/NEJMc1712493
31. Hong RH, Murphy JK, Michalak EE, Chakrabarty T, Wang Z, Parikh SV, et al. Implementing measurement-based care for depression: practical solutions for psychiatrists and primary care physicians. *Neuropsychiatr Dis Treat*. (2021) 17:79–90. doi: 10.2147/NDT.S283731
32. Perlis RH, Mehta R, Edwards AM, Tiwari A, Imbens GW. Pharmacogenetic testing among patients with mood and anxiety disorders is associated with decreased utilization and cost: a propensity-score matched study. *Depress Anxiety*. (2018) 35:946–52. doi: 10.1002/da.22742
33. Maj M, Stein DJ, Parker G, Zimmerman M, Fava GA, De Hert M, et al. The clinical characterization of the adult patient with depression aimed at personalization of management. *World Psychiatry*. (2020) 19:269–93. doi: 10.1002/wps.20771
34. Chen I-M, Chen Y-Y, Liao S-C, Lin Y-H. Development of digital biomarkers of mental illness via Mobile apps for personalized treatment and diagnosis. *J Pers Med*. (2022) 12:936. doi: 10.3390/jpm12060936
35. McGinnis EW, Cherian J, McGinnis RS. The state of digital biomarkers in mental health. *Digit Biomark*. (2024) 8:210–7. doi: 10.1159/000542320
36. Koorevaar Aml, Hegeman Jm, Lamers F, Dhondt Adf, van der Mast Rc, Stek ML, et al. Big five personality characteristics are associated with depression subtypes and symptom dimensions of depression in older adults. *Int J Geriatr Psychiatry*. (2017) 32:e132–40. doi: 10.1002/gps.4670
37. Li Y, Wei D, Qiu J. Personality subtypes of depressive disorders and their functional connectivity basis. *Acta Psychol Sin*. (2023) 55:740–51. doi: 10.3724/SP.J.1041.2023.00740
38. Wardenaar KJ, Conradi HJ, Bos EH, De Jonge P. Personality modulates the efficacy of treatment in patients with major depressive disorder. *J Clin Psychiatry*. (2014) 75:e916–23. doi: 10.4088/JCP.13m08855
39. Altaewel N, Upthegrove R, Surtees A, Durdurak B, Marwaha S. Personality traits as risk factors for relapse or recurrence in major depression: a systematic review. *Front Psychiatry*. (2023) 14:1176355. doi: 10.3389/fpsy.2023.1176355
40. Kudo Y, Nakagawa A, Wake T, Ishikawa N, Kurata C, Nakahara M, et al. Temperament, personality, and treatment outcome in major depression: a 6-month preliminary prospective study. *Neuropsychiatr Dis Treat*. (2017) 13:17–24. doi: 10.2147/NDT.S123788
41. Bernardos AM, Pires M, Ollé D, Casar JR. Digital phenotyping as a tool for personalized mental healthcare. *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth'19*. New York, NY, USA: Association for Computing Machinery (2019). p. 403–8. doi: 10.1145/3329189.3329240
42. Stein DJ, Shoptaw SJ, Vigo DV, Lund C, Cuijpers P, Bantjes J, et al. Psychiatric diagnosis and treatment in the 21st century: paradigm shifts versus incremental integration. *World Psychiatry*. (2022) 21:393–414. doi: 10.1002/wps.20998
43. Adler DA, Wang F, Mohr DC, Choudhury T. Machine learning for passive mental health symptom prediction: generalization across different longitudinal mobile sensing studies. *PLoS One*. (2022) 17:e0266516. doi: 10.1371/journal.pone.0266516
44. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. Rob 2: a revised tool for assessing risk of bias in randomised trials. *Br Med J*. (2019) 366:14898. doi: 10.1136/bmj.14898
45. Wolff RF, Moons KGM, 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) 170:51–8. doi: 10.7326/M18-1376
46. Kambare SM, Jain K, Kale I, Kumbhare V, Lohote S, Lonare S. Design and evaluation of an AI-powered conversational agent for personalized mental health support and intervention (MindBot). *2024 International Conference on Sustainable Communication Networks and Application (ICSCNA) Presented at the 2024 International Conference on Sustainable Communication Networks and Application (ICSCNA)* (2024). p. 1394–402. doi: 10.1109/ICSCNA63714.2024.10863855
47. Amaro I, Greca AD, Tortora G. HAYT Application: the use of NLP to improve the diagnosis and treatment of anxiety and depression, in: 2024 IEEE international conference on bioinformatics and biomedicine (BIBM). *Presented at the 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (2024). p. 6774–81. doi: 10.1109/BIBM62325.2024.10822039
48. Kathan A, Harrer M, Küster L, Triantafyllopoulos A, He X, Milling M, et al. Personalised depression forecasting using mobile sensor data and ecological momentary assessment. *Front Digit Health*. (2022) 4:964582. doi: 10.3389/fdgth.2022.964582
49. Chatterjee S, Mishra J, Sundram F, Roop P. Towards personalised mood prediction and explanation for depression from biophysical data. *Sensors*. (2024) 24:164. doi: 10.3390/s24010164
50. Song S, Seo Y, Hwang S, Kim H-Y, Kim J. Digital phenotyping of geriatric depression using a community-based digital mental health monitoring platform for socially vulnerable older adults and their community caregivers: 6-week living lab single-arm pilot study. *JMIR MHealth UHealth*. (2024) 12:e55842. doi: 10.2196/55842
51. Bell I, Arnold C, Gilbertson T, D'Alfonso S, Castagnini E, Chen N, et al. A personalized, transdiagnostic smartphone intervention (mello) targeting repetitive negative thinking in young people with depression and anxiety: pilot randomized controlled trial. *J Med Internet Res*. (2023) 25:e47860. doi: 10.2196/47860
52. Wang L, Miller L. Assessment and disruption of ruminative episodes to enhance Mobile cognitive behavioral therapy just-in-time adaptive interventions in clinical depression: pilot randomized controlled trial. *JMIR Form Res*. (2023) 7:e37270. doi: 10.2196/37270
53. Frank E, Wallace ML, Matthews MJ, Kendrick J, Leach J, Moore T, et al. Personalized digital intervention for depression based on social rhythm principles adds significantly to outpatient treatment. *Front Digit Health*. (2022) 4:870522. doi: 10.3389/fdgth.2022.870522
54. Malins S, Biswas S, Patel S, Levene J, Moghaddam N, Morris R. Preventing relapse with personalized smart-messaging after cognitive behavioural therapy: a proof-of-concept evaluation. *Br J Clin Psychol*. (2020) 59:241–59. doi: 10.1111/bjc.12244
55. Camacho E, Chang SM, Currey D, Torous J. The impact of guided versus supportive coaching on mental health app engagement and clinical outcomes. *Health Informatics J*. (2023) 29:14604582231215872. doi: 10.1177/14604582231215872
56. Rohani DA, Springer A, Hollis V, Bardram JE, Whittaker S. Recommending activities for mental health and well-being: insights from two user studies. *IEEE Trans Emerg Top Comput*. (2021) 9:1183–93. doi: 10.1109/TETC.2020.2972007
57. Wolitzky-Taylor K, LeBeau R, Arnaudova I, Barnes-Horowitz N, Gong-Guy E, Fears S, et al. A novel and integrated digitally supported system of care for depression and anxiety: findings from an open trial. *JMIR Ment Health*. (2023) 10:e46200. doi: 10.2196/46200
58. Carli V, Petros NG, Hadlaczky G, Vitcheva T, Berchialla P, Bianchi S, et al. The NEVERMIND e-health system in the treatment of depressive symptoms among patients with severe somatic conditions: a multicentre, pragmatic randomised controlled trial. *eClinicalMedicine*. (2022) 48:101423. doi: 10.1016/j.eclinm.2022.101423
59. Winslow BD, Kwasinski R, Hullfish J, Ruble M, Lynch A, Rogers T, et al. Automated stress detection using mobile application and wearable sensors improves symptoms of mental health disorders in military personnel. *Front Digit Health*. (2022) 4:919626. doi: 10.3389/fdgth.2022.919626
60. Fatouros P, Tsirmpas C, Andrikopoulos D, Kaplow S, Kontoangelos K, Papageorgiou C. Randomized controlled study of a digital data driven intervention

for depressive and generalized anxiety symptoms. *NPJ Digit Med.* (2025) 8:1–16. doi: 10.1038/s41746-025-01511-7

61. Campisi SC, Liang M, Anthony SJ, Dettmer E, Korczak DJ. A personalised nutrition intervention for adolescent depression: a mixed-methods feasibility pilot study. *Br J Nutr.* (2024) 132:1–13. doi: 10.1017/S0007114524001338

62. University of Bern. The Motor Activity - Subjective Energy (MASE) Project: Neurobiological and Digital Phenotyping Towards Digital Mental Health Interventions in Depression (Clinical trial registration No. NCT07059234). (2025). Available online at: <https://www.clinicaltrials.gov> (Accessed September 17, 2025).

63. Hucircadian. Efficacy of the Mobile Phone Application (Circadian Rhythms for Mood) for Prediction and Prevention of Mood Episode Recurrence in Mood Disorders Based on Machine Learning of Daily Digital Phenotype Variables: A Sham-controlled Randomized Clinical Trial (Clinical trial registration No. NCT05400785). (2025). Available online at: <https://www.clinicaltrials.gov> (Accessed September 17, 2025).

64. Texas A&M University. Design and Development of an Interactive Mobile Health (mHealth) App to Provide Momentary and Periodic Intervention for High Anxiety and Depression (Clinical trial registration No. NCT07017569). (2025). Available online at: <https://www.clinicaltrials.gov> (Accessed September 17, 2025).

65. Universitat Jaume I. Efficacy of a Personalized, Response-based Transdiagnostic Intervention for Emotional Disorders Delivered Via the Internet: A Protocol for an Adaptive Trial (Clinical trial registration No. NCT07051148). (2025). Available online at: <https://www.clinicaltrials.gov> (Accessed September 17, 2025).

66. Yale University. Feasibility and Efficacy of a Digital Neurotherapy Protocol for Management of Depression, Cognition, and Quality of Life in Patients With Major Medical Illnesses (Clinical trial registration No. NCT04961047). (2025). Available online at: <https://www.clinicaltrials.gov> (Accessed September 17, 2025).

67. RejuvenateProduct – C8Sciences. (n.d.). Available online at: <https://c8sciences.com/rejuvenateproduct/> (Accessed October 1, 2025)

68. Morimoto SS, Altizer RA, Gunning FM, Hu W, Liu J, Cote SE, et al. Targeting cognitive control deficits with neuroplasticity-based computerized cognitive remediation in patients with geriatric Major depression: a randomized, double-blind, controlled trial. *Am J Geriatr Psychiatry.* (2020) 28:971–80. doi: 10.1016/j.jagp.2020.05.023

69. Shu M, Feng S, Liu J. Targeting disrupted rich-club network organization with neuroplasticity-based computerized cognitive remediation in major depressive disorder patients. *Psychiatry Res.* (2022) 316:114742. doi: 10.1016/j.psychres.2022.114742

70. Aalto University. (2025). The Effects of Videogames on Depression Symptoms and Brain Dynamics (Clinical trial registration No. NCT05426265). Available online at: <https://www.clinicaltrials.gov> (Accessed September 17, 2025).

71. Palva JM, Juvonen J, Lukka L, Vesterinen M, Salonen A, Bergman V-R, et al. (2025). Game-Based Digital Intervention for Neurocognitive Training in Major Depressive Disorder: A Randomized Double-Blinded Comparator-Controlled Clinical Trial. doi: 10.2139/ssrn.5409098

72. AdventHealth. (2025). Targeting Network Dysfunction in Apathy of Late-life Depression Using Digital Therapeutics (Clinical trial registration No.

NCT05877885). Available online at: <https://www.clinicaltrials.gov> (Accessed September 17, 2025).

73. Rejoyn: An add-on prescription app for depression symptoms [WWW Document]. (n.d.). Available online at: <https://www.rejoyn.com/> (Accessed October 1, 2025)

74. Otsuka Pharmaceutical Development & Commercialization, Inc. A Multi-center, Randomized, Controlled Trial to Evaluate the Effectiveness of a Digital Therapeutic (CT-152) as Adjunctive Therapy in Adult Subjects Diagnosed With Major Depressive Disorder (Clinical trial registration No. NCT04770285). (2024). Available online at: <https://www.clinicaltrials.gov> (Accessed September 17, 2025).

75. Kessing LV. The Effect of Digital Markers in the Self-management of Depressive Symptoms - The MENTINA Randomized Controlled Parallel Group Trial (Clinical trial registration No. NCT06919133). (2025). Available online at: <https://www.clinicaltrials.gov> (Accessed September 17, 2025).

76. Personalized Digital Mental Health Support: Meeting People Where They Are - MENTBEST [WWW Document]. (n.d.). Available online at: <https://mentbest.com/blog/personalized-digital-mental-health-support-meeting-people-where-they-are/> (Accessed October 1, 2025)

77. van Loo HM, de Jonge P, Romeijn J-W, Kessler RC, Schoevers RA. Data-driven subtypes of major depressive disorder: a systematic review. *BMC Med.* (2012) 10:156. doi: 10.1186/1741-7015-10-156

78. Borsboom D, Cramer AOJ. Network analysis: an integrative approach to the structure of psychopathology. *Annu Rev Clin Psychol.* (2013) 9:91–121. doi: 10.1146/annurev-clinpsy-050212-185608

79. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. *ACM Comput Surv.* (2021) 54:115:1–115:35. doi: 10.1145/3457607

80. Joseph J. Algorithmic bias in public health AI: a silent threat to equity in low-resource settings. *Front Public Health.* (2025) 13:1643180. doi: 10.3389/fpubh.2025.1643180

81. Allen C, Mehler DMA. Open science challenges, benefits and tips in early career and beyond. *PLoS Biol.* (2019) 17:e3000246. doi: 10.1371/journal.pbio.3000246

82. Bell V. Open science in mental health research. *Lancet Psychiatry.* (2017) 4:525–6. doi: 10.1016/S2215-0366(17)30244-4

83. ISO 9241-210:2019 [WWW Document]. (n.d.). ISO. Available online at: <https://www.iso.org/standard/77520.html> (Accessed June 21, 2025)

84. Sanders EB-N, Stappers PJ. Co-creation and the new landscapes of design. *CoDesign.* (2008) 4:5–18. doi: 10.1080/15710880701875068

85. Nordberg SS, Jaso-Yim BA, Sah P, Schuler K, Eyllon M, Pennine M, et al. Evaluating the implementation and clinical effectiveness of an innovative digital first care model for behavioral health using the RE-AIM framework: quantitative evaluation. *J Med Internet Res.* (2024) 26:e54528. doi: 10.2196/54528

86. Youn SJ, Jaso B, Eyllon M, Sah P, Hoyler G, Barnes JB, et al. Leveraging implementation science to integrate digital mental health interventions as part of routine care in a practice research network. *Adm Policy Ment Health Ment Health Serv Res.* (2024) 51:348–57. doi: 10.1007/s10488-023-01292-9

87. He Y, Murphy JK, Wang X, Michalak EE, Yang T, Yang X, et al. The state of implementation science in major depressive disorder: a narrative review. *Psychiatry Res.* (2025) 344:116330. doi: 10.1016/j.psychres.2024.116330

Glossary

ALE	accumulated local effects	LC	locus coeruleus
BDNF	brain-derived neurotrophic factor	LLM	large language model
BMI	body mass index	MAO	monoamine oxidase
CBT	cognitive behavioral therapy	MASE	motor activity–subjective energy
CFIR	consolidated framework for implementation research	MBC	measurement-based care
COVID-19	coronavirus disease	MDD	major depressive disorder
CRH	corticotropin-releasing hormone	ML	machine learning
CRP	C-reactive protein	NE	norepinephrine
DALYs	disability-adjusted life years	NICE	National Institute for Health and Care Excellence
DL	deep learning	NLP	natural language processing
DMN	default mode network	PHQ-9	patient health questionnaire-9
DSM-5	diagnostic and statistical manual of mental disorders, fifth edition	PRISMA	preferred reporting items for systematic reviews and meta-analyses
EDA	electrodermal activity	PROBAST	prediction model risk of bias assessment tool
EEG	electroencephalography	PRS	polygenic risk score
EHR	electronic health record	QUIN	quinolinic acid
EMA	ecological momentary assessment	RCT	randomized controlled trial
FDA	U.S. Food and Drug Administration	RE-AIM	reach, efficacy, adoption, implementation, and maintenance
GABA	gamma-aminobutyric acid	RNT	repetitive negative thinking
GAD-7	generalized anxiety disorder-7	RoB	risk-of-bias tool for randomized trials
HCD	human-centered design	SHAP	SHapley Additive exPlanations
HPA	hypothalamic–pituitary–adrenal axis	SMS	short message service
HR	heart rate	SNP	single nucleotide polymorphisms
HRV	heart rate variability	SVM	support vector machine
IL-6	interleukin-6	SWOT	strengths, weaknesses, opportunities, threats
IL-8	interleukin-8	TNF- α	tumor necrosis factor alpha
JITAI	just-in-time adaptive intervention	TRD	treatment-resistant depression
KYNA	kynurenic acid	WBC	white blood cells
		XAI	explainable artificial intelligence.