Review Article

DIGITAL HEALTH

DIGITAL HEALTH
Volume 10: 1–23

DOI: 10.1177/20552076241260414
journals.sagepub.com/home/dhj

Sage

Remote sensing mental health: A systematic review of factors essential to clinical translation from validation research

Niranjan Bidargaddi[1], Richard Leibbrandt[2], Tamara L Paget[1], Johan Verjans[3,4,5], Jeffrey CL Looi[6] and Jessica Lipschitz[7,8]

Abstract

Background: Mental illness remains a major global health challenge largely due to the absence of definitive biomarkers applicable to diagnostics and care processes. Although remote sensing technologies, embedded in devices such as smartphones and wearables, offer a promising avenue for improved mental health assessments, their clinical integration has been slow.

Objective: This scoping review, following preferred reporting items for systematic reviews and meta-analyses guidelines, explores validation studies of remote sensing in clinical mental health populations, aiming to identify critical factors for clinical translation.

Methods: Comprehensive searches were conducted in six databases. The analysis, using narrative synthesis, examined clinical and socio-demographic characteristics of the populations studied, sensing purposes, temporal considerations and reference mental health assessments used for validation.

Results: The narrative synthesis of 50 included studies indicates that ten different sensor types have been studied for tracking and diagnosing mental illnesses, primarily focusing on physical activity and sleep patterns. There were many variations in the sensor methodologies used that may affect data quality and participant burden. Observation durations, and thus data resolution, varied by patient diagnosis. Currently, reference assessments predominantly rely on deficit focussed self-reports, and socio-demographic information is underreported, therefore representativeness of the general population is uncertain.

Conclusion: To fully harness the potential of remote sensing in mental health, issues such as reliance on self-reported assessments, and lack of socio-demographic context pertaining to generalizability need to be addressed. Striking a balance between resolution, data quality, and participant burden whilst clearly reporting limitations, will ensure effective technology use. The scant reporting on participants' socio-demographic data suggests a knowledge gap in understanding the effectiveness of passive sensing techniques in disadvantaged populations.

[1]Digital Health Research Lab, College of Medicine and Public Health, Flinders Health and Medical Research Institute, Flinders University, Adelaide, South Australia, Australia
[2]College of Science and Engineering, Flinders University, Adelaide, South Australia, Australia
[3]Australian Institute for Machine Learning, University of Adelaide, Adelaide, South Australia, Australia
[4]Lifelong Health, South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia
[5]Department of Cardiology, Royal Adelaide Hospital, Adelaide, South Australia, Australia
[6]Academic Unit of Psychiatry & Addiction Medicine, The Australian National University School of Medicine and Psychology, Garran, Australia
[7]Department of Psychiatry, Brigham and Women's Hospital, Boston, MA, USA
[8]Department of Psychiatry, Harvard Medical School, Boston, MA, USA

Corresponding author:
Niranjan Bidargaddi, Digital Health Research Lab, College of Medicine and Public Health, Flinders Health and Medical Research Institute, Flinders University, GPO Box 2100 Adelaide, South Australia 5001, Australia.
Email: niranjan.bidargaddi@flinders.edu.au

Keywords

Remote sensing, mental illness, mobile technology, improved mental health, wearables

Submission date: 1 January 2024; Acceptance date: 21 May 2024

Introduction

Mental illness remains a significant global health concern, ranking among the top ten leading burdens of disease since 1990.[1] For patients that seek treatment for mental illness, accurate diagnosis is necessary for provision of specific interventions that will improve mental illness. Unfortunately, unlike some other chronic conditions, e.g. diabetes, there are no objective biomarkers available for mental illness diagnosis, and the current paradigm relies on intermittent clinical interviews or subjective self-reporting, which may provide an incomplete and inaccurate picture of patients' conditions.[2]

To improve the quality of observations to inform reliable diagnoses and progress monitoring, researchers have turned to remote sensing digital technologies such as mobile phones, wearables, and actigraphs[3] to enable the collection of behavioral and physiological data without requiring direct contact or self-reporting. This process, known as remote sensing[4] allows for continuous observations of patients in their living environments, providing valuable observational data for mental health assessments and diagnosis.[5] For example, automatically captured sensor data has been linked to individuals' daily behaviors[6] or changes in mood symptoms, functioning levels and early detection of warning signs for suicidal ideation or onset and phase changes in bipolar disorder.[7]

The unprecedented, accelerated trend towards increase in smartphone and wearable device ownership worldwide suggests that remote sensing has the potential to transform clinical and research practices in mental health.[8,9] By combining clinical interview-like questionnaires with remote sensing, we can obtain frequent and near real-time measurements of core symptoms and behavioral mental health disorders from a patient's living environment, providing unprecedented proximity to their occurrence. This approach can mitigate the reliance on potentially unreliable retrospective accounts. However, to implement remote sensing in practice, we need evidence demonstrating concordance between passive sensor data and validated assessments. Existing experimental studies validating remote sensing with mental ill-health populations largely indicate the validity of remote sensor measurements when compared to validated psychiatrist assessments.[3,10,11]

While validity is important, it is not enough to integrate remote sensing into clinical and research practice. The accuracy and relevance of remote sensing data are also influenced by its temporal resolution, which can refer to both the duration of sampling and the time between samples. The temporal resolution needed for remote sensing varies depending on the symptoms or behaviors being monitored and the clinical context. For instance, mental health disorders with rapidly changing symptoms may require frequent and near real-time monitoring, while others with more stable or slowly changing symptoms can be monitored over longer intervals. The complexity of technical details, combined with differing reporting norms across various fields has led to duplicated and unreproducible research efforts.[10] As a result, the translation of remote sensing into mental health clinical practice has been slow. In response to this burgeoning gap, this study conducted a systematic review of remote sensing validation studies undertaken in populations with depression, schizophrenia, bipolar and addiction disorders to identify critical factors that are relevant to informing translation into clinical practice.

Specifically, we aimed to address the following questions in the studies where remote sensing was validated:

1. What were the distinctive characteristics of the population and the environment in which validation took place?
2. What was the context for the validation of remote sensing, including the sensing purpose, and the nature and method of reference assessments used for validation?
3. Which specific parameters were sensed, and what were the temporal factors influencing the sensing process?

Together this information aims to bridge the gap in knowledge related to translating remote sensing into practice in that it may assist in providing data that better informs the clinical diagnosis and treatment of patients with mental illness.

Methods

Search strategy

A systematic search of Medline OVID, Medline PubMed, IEEE, ACM, Scopus, and PsycINFO databases was performed on December 3, 2019 using search terms developed in collaboration with a research librarian at Flinders

Table 1. Search terms organized by PICO framework.

Category	Search Word Used
Population	Depression OR Anxiety OR Schizophrenia OR Bipolar OR Mood OR psychosis OR psychotic OR addict* OR "psychological stress")
Intervention	("cell phone" OR "cell phones" OR "mobile phone" OR "mobile phones" OR "mobile technology" OR smartphone OR smartphones OR "table device" OR "tablet computer" OR "data audio" OR "data analytic" OR "data movement" OR "data motion" OR "data GPS" OR "data global position" OR "data location" OR "data proximity" OR "data blue tooth" OR Physiological OR "heart rate" OR "respiratory rate" OR speech OR audio OR voice OR interactions OR microphone OR camera OR keyboard OR chat OR typing OR location OR activity OR sleep OR movement OR GPS
Comparator	[any]
Outcome	"monitor real?time" OR "monitor data" OR "sensing real?time" OR "sensing data" OR "sensor real-time" OR "sensor data" OR "multi sensor real-time" OR "multi sensor data" OR "passive sensor" OR "passive sensing" OR "passive collected" OR "passive sensed"

University, as shown in Table 1 (see Appendix 1 for detailed search strategy). Searches were repeated on July 15, 2022 to collect any further works that had been published since the initial search. The search results are presented in preferred reporting items for systematic reviews and meta-analyses (PRISMA) format in Figure 1.

Eligibility criteria

Only studies that (i) included populations with clinically diagnosed depression, anxiety, stress, schizophrenia, bipolar or substance use disorders as these conditions collectively represent the common constellation of patients encountered in clinical practice; (ii) gathered sensor data from participants using a wearable device or smartphone application reflecting the technology readily accessible to consumers, with potential for practical integration in real-world clinical settings; (iii) reported collection of validated assessment measurement of mental health outcomes from participants as a reference assessment; (iv) described a statistical relationship between the sensor data and reference mental health outcome assessment; (v) described the type and nature of collected sensor data; and (vi) involved collecting sensor data for at least 7 days. The stipulation of a minimum 7-day data collection period was established to align with real-world psychiatry assessments, capturing meaningful patterns and variations in mental health-related data, thus enhancing the clinical relatability and representativeness of the study findings. By specifically selecting studies that have employed validated assessment measurements of mental health outcomes and established a statistical relationship between the sensor data and the reference mental health outcome assessment, we ensure that only evidence-based real-world psychiatry assessments are considered in our analysis of insights into the translation and implementation issues. Non-English language publications, papers not published through peer review, studies providing no statistical outcomes, papers published before January 1, 2009, conference reports, protocols, in-vitro or lab-based studies, editorials or letters of opinion were excluded from the analysis. Search terms are included in Appendix 3.

Study selection processes

Screening of included papers was completed by three reviewers (AN, CL, YK), while data extraction was completed by another three reviewers (AN, MT, SI). Two authors were involved in screening and data extraction, and any disagreements at each of these stages were resolved to consensus between the authors through discussion. Unresolved discrepancies were resolved by a fourth author, NB. The process was carried out using Covidence software.

Data extraction and synthesis

Covidence 2.0 was used for screening and extracting information from included studies. The following variables were extracted from all eligible papers: (a) reported demographics, (b) clinical diagnosis of study population, (c) type and combination of reference assessments used for validation and (d) mode of administration (i.e. clinical interview, self-report, cognitive test) of reference assessments used, (e) types of modalities sensed (e.g. activity, sleep, speech, digital interactions, physiological signals), (f) intended purpose defined as types of decisions that were made using sensing data (e.g. detecting the potential for the disease, tracking the progression of disease that exists, forecasting future prognosis and intervention), (g) time resolution relevant to sensors—interval between consecutive samples, (h) duration of the study over which sensor data was gathered, (i) experimental design type—the chronological order of

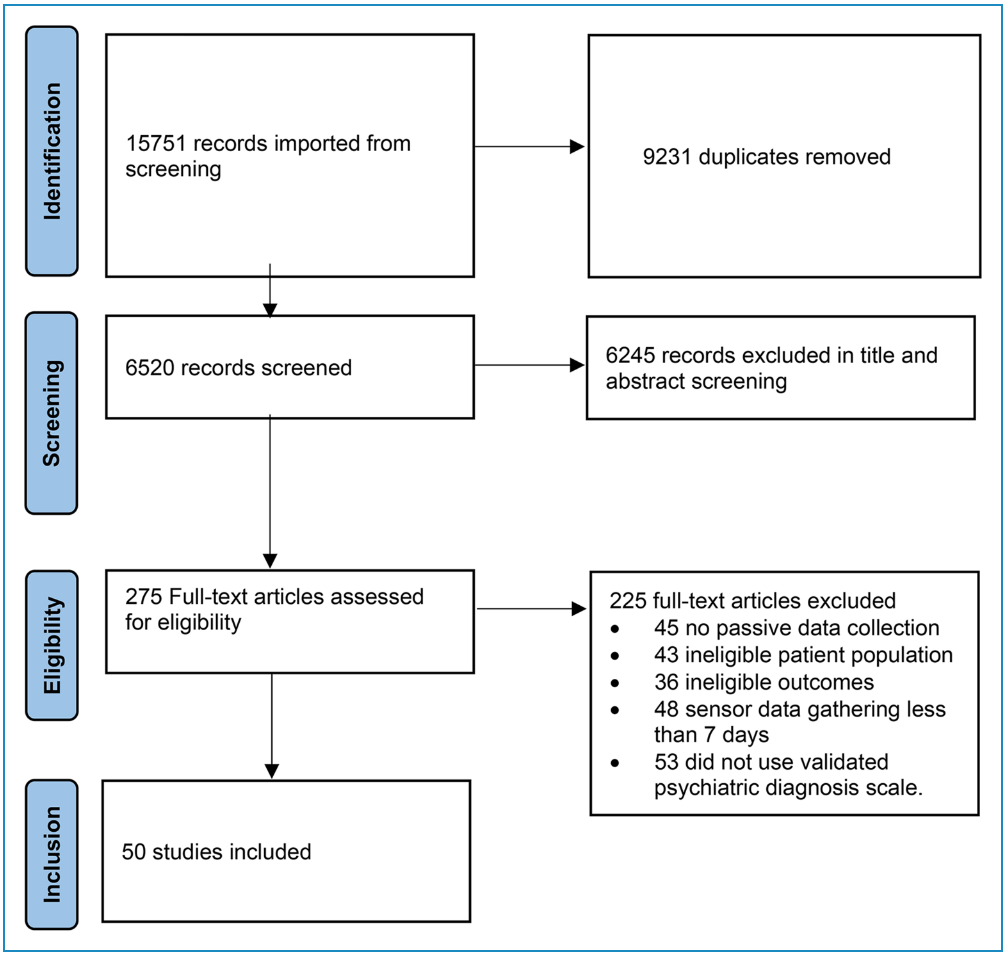

Figure 1. *PRISMA publication review chart. Adapted from:* Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. Doi: 10.1136/bmj.n71.

sensor data collection and reference assessments administration. Rationales used to include or exclude participants were derived qualitatively through thematic content analysis of recruitment criteria descriptions included in the paper.[12] Results are presented using descriptive statistics and narrative synthesis, and visualizations of the data were created using the Python software library seaborn (version 0.12.0).

Results

Overview of studies

The PRISMA chart shown in Figure 1 illustrates the entire study selection process. The search query returned 15,751 results, which reduced to 6520 studies after duplicates were removed. After reviewing the title and abstract further, another 6245 studies were excluded. Full text was retrieved for the remaining 275 studies, of which 225 studies were excluded for the reasons shown in the PRISMA chart. As a result, a total of 50 passive sensing experimental studies with mental illness were included[12–61] (Tables 1 and 2, Figure 2). Of these studies, 18 were in depression,[13–30] 11 in schizophrenia,[19,31–40] 17 in bipolar,[41–57] 3 in anxiety,[18,58,59] 2 in substance use disorder,[60,61] and 1 study was transdiagnostic[62] (Table 1). The median number of participants enrolled per study was 109 participants (range = 6–1784) (Table 1). Of the studies that reported gender of enrolled participants ($n = 48$), the average proportion of female participants were 53.23% per study (range = 0–92%) (Table 1). By geographical location, the United States ($n = 20$) was the most common country, and only three reported studies (two Brazil, one China) were from a location outside of high-income countries (Table 1).

Characteristics of population and study environment

The sociodemographic information of enrolled participants was not uniformly reported. Of the included studies, 94%

Table 2. Summary of reviewed studies.

Reference number	First Author	Year	Title	For Whom/For What
13	Alcantara	2016	Sleep disturbances and depression in the multi-ethnic study of atherosclerosis	To establish psycho/physiological features associated with different HYPOTHESIZED disease states
14	Averill	2018	Clinical response to treatment in inpatients with depression correlates with changes in activity levels and psychomotor speed	To establish possible passively sensed sleep markers for monitoring disease states
15	Ben-Zeev	2015	Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health	To establish possible passively sensed markers for detecting disease states
16	Bewernick	2017	Walking away from depression-motor activity increases ratings of mood and incentive drive in patients with major depression	To establish possible passively sensed activity markers for detecting disease states
17	Chikersal	2021	Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing: A machine learning approach with robust feature selection	To establish possible passively sensed markers for detecting disease states
18	Difrancesco	2019	Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disorders: A 2-week ambulatory assessment study	To establish psycho/physiological features associated with different diagnoses (/CONTROL)
19	Fasmer	2016	Distribution of active and resting periods in the motor activity of patients with depression and schizophrenia	To establish psycho/physiological features associated with different diagnoses
20	Finazzi	2009	Motor activity and depression severity in adolescent outpatients	To refine extant sensor markers for disease states by adjusting measurement for a paediatric population
21	Jacobson	2019	Using digital phenotyping to accurately detect depression severity	To establish possible passively sensed markers for detecting condition severity of a known disease state
22	Kumagai	2019	Predicting recurrence of depression using lifelog data: an explanatory feasibility study with a panel VAR approach	To establish possible passively sensed markers for detecting relapse into a known disease state
23	Merikanto	2017	Advanced phases and reduced amplitudes are suggested to characterize the daily rest-activity cycles in depressed adolescent boys	To establish psycho/physiological features associated with different hypothesized disease states using passive activity data
24	Mesquita	2016	Activity/rest rhythm of depressed adolescents undergoing therapy: case studies	To establish possible passively sensed sleep markers for monitoring disease states
25	Moukaddam	2019	Findings from a trial of the smartphone and online usage-based evaluation for depression (SOLVD) application: What do apps really tell us about patients with depression? concordance between app-generated data and standard psychiatric questionnaires for depression and anxiety	To refine extant short form questionairre markers for disease states by incorporating sensed states

(continued)

Table 2. Continued.

Reference number	First Author	Year	Title	For Whom/For What
26	O'Brien	2017	A study of wrist-worn activity measurement as a potential real-world biomarker for late-life depression	To establish psycho/physiological features associated with different hypothesized disease states using passive activity data
27	Saeb	2015	Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study	To establish possible passively sensed activity and phone usage markers for detecting disease states
28	Saeb	2016	The relationship between mobile phone location sensor data and depressive symptom severity	To refine extant sensor markers for disease states by adjusting measurement for different contexts
29	Stavrakakis	2015	Temporal dynamics of physical activity and affect in depressed and nondepressed individuals	To establish psycho/physiological features associated with different diagnoses
30	Wahle	2016	Mobile sensing and support for people with depression:A pilot trial in the wild	To establish possible passively sensed behavioral markers for detecting disease states and enabling appropriate interventions
31	Barnett	2018	Relapse prediction in schizophrenia through digital phenotyping: a pilot study	To establish possible passively sensed activity, phone engagement, and self-report mood markers for detecting disease states
32	Bromundt	2011	Sleep–wake cycles and cognitive functioning in schizophrenia	To establish possible passively sensed sleep markers for monitoring disease states
33	Bueno-Antequera	2018	Relationship between objectively measured sedentary behavior and health outcomes in schizophrenia patients: The PsychiActive project	To establish psycho/physiological features associated with different hypothesized disease states
34	Cella	2019	Blending active and passive digital technology methods to improve symptom monitoring in early psychosis	To establish possible passively sensed markers for detecting disease states
35	Chen	2016	Association between actigraphy-derived physical activity and cognitive performance in patients with schizophrenia	To establish psycho/physiological features associated with different hypothesized disease states using passive activity data
36	Chung	2018	Correlates of sleep irregularity in schizophrenia	To refine extant sensor markers for disease states by adjusting measurement by examining covariates
37	Depp	2019	GPS mobility as a digital biomarker of negative symptoms in schizophrenia: A case–control study	To establish psycho/physiological features associated with different diagnoses (/CONTROL)
38	Janney	2013	Sedentary behavior and psychiatric symptoms in overweight and obese adults with schizophrenia and schizoaffective disorders (WAIST Study)	To establish psycho/physiological features associated with different disease states

(continued)

Table 2. Continued.

Reference number	First Author	Year	Title	For Whom/For What
39	Kume	2015	A pilot study: Comparative research of social functioning, circadian rhythm parameters, and cognitive function among institutional inpatients, and outpatients with chronic schizophrenia and healthy elderly people	To establish psycho/physiological features associated with different diagnoses
40	Mulligan	2016	High resolution examination of the role of sleep disturbance in predicting functioning and psychotic symptoms in schizophrenia: A novel experience sampling study	To establish possible passively sensed sleep markers for detecting disease states
41	Abdullah	2016	Automatic detection of social rhythms in bipolar disorder	To establish psycho/physiological features associated with different hypothesized disease states using passive activity and phone usage data
42	Banihashemi	2016	Quantifying the effect of body mass index, age, and depression severity on 24-h activity patterns in persons with a life time history of affective disorders	To establish psycho/physiological features associated with different hypothesized disease states using passive activity data
43	Beiwinkel	2016	Using smartphones to monitor bipolar symptoms: A pilot study	To establish possible passively sensed location/geographic markers and phone usage markers for monitoring disease states
44	Faurholt-Jepsen	2014	Smartphone data as objective measures of bipolar disorder symptoms	To establish possible passively sensed markers for detecting condition severity of a known disease state
45	Faurholt-Jepsen	2015	Smartphone data as an electronic biomarker of illness activity in bipolar disorder	To establish possible passively sensed markers for detecting condition severity of a known disease state
46	Faurholt-Jepsen	2019	Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder	To establish possible passively sensed markers for detecting disease states
47	Faurholt-Jepsen	2016	Objective smartphone data as a potential diagnostic marker of bipolar disorder	To refine extant sensor markers for disease states by adjusting measurement for different contexts
48	Gonzalez	2014	The relationship between affective state and the rhythmicity of activity in bipolar disorder	To establish possible passively sensed markers for detecting condition severity of a known disease state
49	Grierson	2016	Circadian rhythmicity in emerging mood disorders: State or trait marker?	To establish possible passively sensed activity and phone usage markers for predicting disease states
50	Grunerbl	2015	Smartphone-based recognition of states and state changes in bipolar disorder patients	To establish possible passively sensed markers for detecting disease states

(continued)

Table 2. Continued.

Reference number	First Author	Year	Title	For Whom/For What
51	Janney	2014	Are adults with bipolar disorder active? Objectively measured physical activity and sedentary behavior using accelerometry	To establish possible passively sensed markers for detecting condition severity of a known disease state
52	Krane-Gartiser	2019	Which actigraphic variables optimally characterize the sleep-wake cycle of individuals with bipolar disorders?	To establish possible passively sensed markers for detecting disease states
53	McGlinchey	2014	Physical activity and sleep: Day-to-day associations among individuals with and without bipolar disorder	To establish psycho/physiological features associated with different diagnoses
54	Merikangas	2019	Real-time mobile monitoring of the dynamic associations among motor activity, energy, mood, and sleep in adults with bipolar disorder	To establish psycho/physiological features associated with different diagnoses (/CONTROL)
55	Ortiz	2016	Exponential state transition dynamics in the rest-activity architecture of patients with bipolar disorder	To establish psycho/physiological features associated with different hypothesized disease states using passive activity
56	Palmius	2017	Detecting bipolar depression from geographic location data	To establish possible passively sensed markers for detecting disease states
57	McGowan	2019	Circadian rest-activity patterns in bipolar disorder and borderline personality disorder	To establish psycho/physiological features associated with different diagnoses
58	Cohodes	2019	Novel insights from actigraphy: Anxiety is associated with sleep quantity but not quality during childhood	To refine extant sensor markers for disease states by adjusting measurement for a paediatric population
59	Jacobson	2021	Deep learning paired with wearable passive sensing data predicts deterioration in anxiety disorder symptoms across 17-18 years	To establish possible passively sensed markers for detecting condition severity of a known disease state
60	Epstein	2014	Real-time tracking of neighborhood surroundings and mood, in Urban drug misusers: Application of a new method to study behavior in its geographical context	To establish possible passively sensed location/geographic markers for detecting disease states
61	Kennedy	2015	Continuous in-the-field measurement of heart rate: Correlates of drug use, craving, stress, and mood in polydrug users	To establish possible passively sensed physiological markers for detecting disease states
62	Gloster	2021	The spatiotemporal movement of patients in and out of a psychiatric hospital: An observational GPS study	To establish possible passively sensed markers for detecting condition severity of a known disease state

reported gender (Table 3), 92% reported age (Table 3), 26% reported ethnicity (Table 3), 20% reported marital status (Table 3), 36% reported education (Table 3), and 28% reported employment status (Table 3).

Those studies that reported details on recruited participants included: 44 studies with healthcare patients, 2 of students, and 4 of community members (Table 3). Of those studies that reported their recruitment source, 38 were

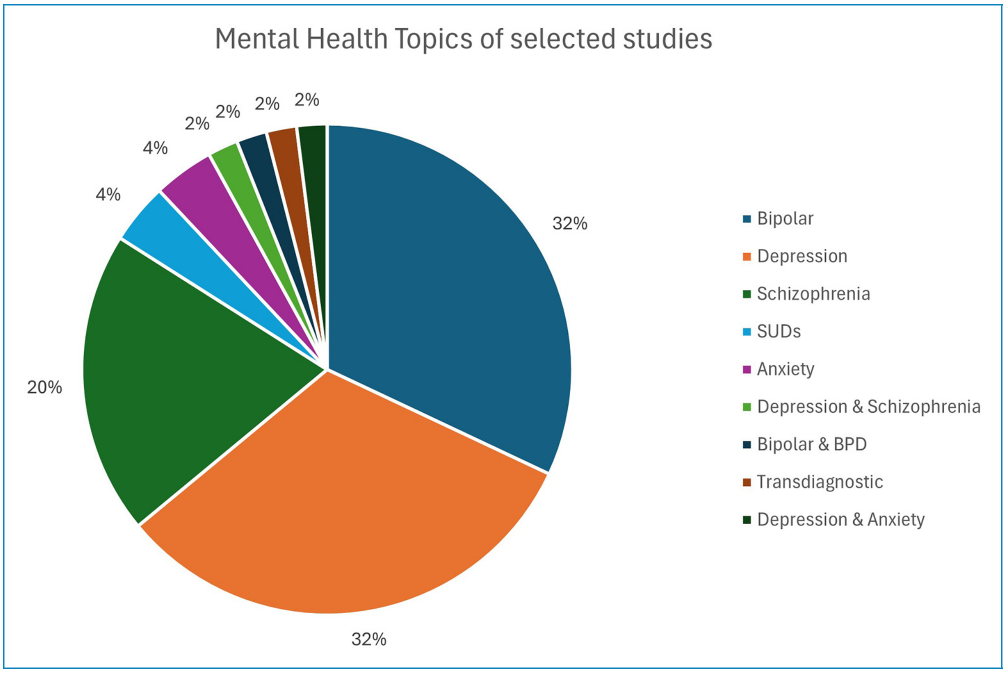

Figure 2. Mental health topics of selected studies.

enrolled through clinics,[14,16,19–26,29,31–50,52,54–57,60–62] 3 from the internet,[27,30,53] and 2 from University.[15,17]

Of the 50 studies, 39 reported a funding source[12–17,19,21–23,25–27,29–34,36,37,40–48,50–53,55–58,60–62] and 14 declared one or more authors had a conflict of interest.[21,36,37,40,41,43–48,50,53,56] Twelve of the 50 studies reported using incentives for participation. The incentives ranged from financial incentives[22,25,29,32,36,37,41,51,61] to gifting of the technology used in the study,[16] to gifting a t-shirt with an opportunity to win a smart phone[63] (Table 3).

A detailed overview of all variables that were used in deciding whether to include or exclude participants involved in the study are shown in Appendix 3. These variables were (a) diagnosis, (b) self-reported health behavior, (c) social, economic and family factors, and (d) severity of symptoms. The 253 identified variables were distributed across diagnosis (52.9%), health behaviors (19.8%), social and economic factors (14.2%) and other (3.9%).

Remote sensing validation context

Intended purpose of remote sensing. Studies of remote sensing used three methodologies: (i) to detect the potential for a disease to be present that is not known[15,17,25,39,49,50,52,56]; (ii) to track the progression of an established illness[14,16,18–24,26–28,32–38,40–46,48,51,53–55,57,58,61,62]; and (iii) to predict the progression of an established illness.[13,30,31,59]

Reference assessments type and mode of administration used for validation. The reference or comparator mental health assessments that informed diagnosis administered in the remote sensing studies varied substantially (see Table 4).

Clinician interviews,[18,35–37,40,43,45–48,52–54,56–59–61] cognitive assessment tasks,[21,26,35,39] and self-report scales[13–16,18–23,25–30,32–34,36–38,39,40,42–48,50–53,56–58,60–62] comprised the main modes of reference assessment.

For studies using self-report for reference mental health assessments ($n = 41/50$), the assessments scales used measured multiple mental health deficit (loss of function) diagnostic domains,[1,14,32,33,38,39,62] depression/mood,[12–15,17–22,24–27,29,35,36,39,41–43,45,49–52,55,56] anxiety and stress,[15,18,25,58] psychosis/mania[19,32,34,36–38,40,43–48,50–53,57] and other areas[13,15,29,32,33,58,60–62] (Table 3).

Studies that reported using interviews for reference or comparator mental health assessments ($n = 18$), included 10 for bipolar disorders,[43–48,52–54,56,57] 4 for schizophrenia,[35–37,40] 2 for substance use,[60,61] 1 for anxiety,[59] and 1 for depressive disorders.[18]

Of studies that used self-report scales for reference or comparator mental health assessments ($n = 41$), 16 were for depressive disorders,[13–16,18–25–30] 13 for bipolar disorders,[42–47,48,50–53,56,57] 9 for schizophrenia,[19,32–34,36–40] 2 for anxiety,[18,58] and 2 for substance use disorders.[60,61]

Experimental designs used. All experimental designs that gathered remote sensor data and reference or comparator mental health diagnostic assessments used varied methodologies regarding how data was collected.

Twenty-eight studies described the data collection timeline methodology with sufficient detail to discern that there were four overall designs for these studies:

Design 1: There were 12 studies, which involved a reference assessment, followed by sensor data collection, and then one or more reference assessments.[17,20,22–25,43,44–47,50,59]

Design 2: In 9 studies, a reference assessment was followed by sensor data collection, and then a single reference assessment was performed.[14,16,19,28,30,51–53,61]

Design 3: In 5 studies, reference assessment preceded sensor data collection.[23,26,34,38,39]

Design 4: In 2 studies, sensor data collection preceded the reference assessment.[33,56]

Sensor characteristics

Type of sensing parameters. Sleep and activity sensors used relatively short measurement intervals (in the range of 30 seconds to 2 minutes).

Location sensors were used in a small number of studies with relatively longer measurement intervals (in the order of 5–15 minutes).

Ten different sensor types were used across studies (Table 3).

In the order of the number of studies that used a particular sensor measurement:

(1) Activity[13–23,25,26,29,32,34–36,38,39,41–44,48,49,51–55,59,61]
(2) Sleep,[13,17–20,32,36,38,39,40,42,48,49,51,53,54,55,58,59] interaction with phone[24,25,27,41–47]
(3) Location[15,17,24,25,27,28,30,31,37,41,44,56,60,62]
(4) Interaction with phone[24,25,27,41,44–47]
(5) Other sensors used were: Bluetooth[22]; conversations[31,47,50]; skin temperature[22,34]; skin conductance[34,61]; heart rate[34,61]; ambient light.[21]

Time resolution for sensor sampling. The average pre-specified time period for gathering sensor data per study was 44 days (SD = 76.2).

The time interval between consecutively sensed parameters varied by the type of sensor used is shown in Table 4. The two most frequently sensed parameters—activity and sleep—had an average of 81 and 86 seconds between consecutive samples. The least two frequently sensed parameters—ambient light and Bluetooth—had a reported interval between consecutive samples of 900 and 600 seconds, respectively.

Time period of sensor data. There was substantial variation in the time period over which the sensor data was gathered. Studies that enrolled bipolar disorder patients had an average observation period of 61 days (SD = 90.8), 16.4 (SD = 24.2) for patients with schizophrenia and 44 (SD = 86.6) for patients with depressive disorders.

Figure 3 (Appendix 2) shows the relationship between the two aspects of sensor time resolution (measurement interval and duration of observation), as well as their interaction with sensor type, study purpose, study design and the types of condition assessment used. The data is further grouped by diagnosed condition. Only 29 of the 50 studies (58%) reported both measurement interval and duration. The other studies are not included in the figure.

Discussion

Reliance on self-reported assessments in remote sensing

The three objectives of remote sensing mental health—detecting unknown diseases, tracking the progression of known diseases, and forecasting future outcomes—are congruent with findings from other studies.[64] However, our study revealed shortcomings in the reference assessments employed. Self-reports were the most common form of reference assessments. Self-report scales can be prone to bias, and this finding suggests biases and limitations inherent in self-reports may also be perpetuated into sensing approach. Furthermore, the majority of selected assessments have focused on detecting deficits, such as symptoms and impairments, in contrast to measures of function and strengths, which are just as critical to gain a comprehensive understanding of an individual's mental health.[65]

Lack of socio-demographic details

Our study reveals a gap in understanding the applicability of remote sensing approaches to disadvantaged populations, partly due to insufficient reporting of sociodemographic information. We observed most research in this area is conducted in developed countries, leaving a dearth of studies on geographically diverse populations with limited mental health support options. Additionally, challenges in recruiting males have resulted in a gender imbalance in study participation, while the exclusion of low-functioning individuals raises concerns about the representativeness of findings and potential discrimination. Coordinated efforts toward shaping guidelines for ethical inclusive and representativeness in remote sensing mental health research are imperative. This effort should involve diverse perspectives from mental health researchers, ethicists, and representatives from different populations. Regulatory bodies can then monitor adherence to these guidelines and ensure benefits of remote mental health sensing research is accessible to all, regardless of their gender, socioeconomic status, or other sociodemographic factors.

Adoption details

Our study also revealed a gap in understanding into human factors that influence acceptability and adoption of remote sensing in existing research. Although some studies offered financial incentives, reporting on participant

Table 3. Demographic overview of included studies.

Ref	First Author	Year	Diagnosed Condition	Total # Subjects	Female	Average Age	Country	Incentives	Observational?	Ethnicity	Marital Status	Education	Employed	Participant Type	Study Setting	Funding Source	Conflict of interest declared
13	Alcantara	2016	Depression	1784	57%	66.47	U.S.	N	Y	Y	N	Y	Y	Established cohort	Clinic	NIH, University	N
14	Averill	2018	Depression	24	38%	38.08	U.S.	N	Y	Y	N	N	N	Patients	Clinic	Foundation, Uni	N
15	Ben-Zeev	2015	Depression	47	21%	22.5	U.S.	Y	Y	Y	N	N	N	Students	University	None stated	N
16	Bewernick	2017	Depression	12	67%	49	Germany	N	Y	N	N	N	N	Patients	Clinic	None stated	-
17	Chikersal	2021	Depression	138	N/A	N/A	U.S.	Y	Y	N	N	Y	Y	Students	Mail	None stated	-
18	Difrancesco	2019	Depression & Anxiety	359	62%	49.9	Netherlands	N	Y	N	N	N	N	Established cohort	N/A	European Cmsn	N
19	Fasmer	2016	Depression & Schizophrenia	76	41%	42.7	Norway	N	Y	N	N	N	N	Patients	N/A	None stated	-
20	Finazzi	2009	Depression	6	67%	16.5	Brazil	N	Y	N	N	N	N	Patients	Clinic	Foundation	-
21	Jacobson	2019	Depression	15	87%	47.6	U.S.	N	Y	N	N	N	N	Patients	Clinic	Industry	-
22	Kumagai	2019	Depression	89	45%	44.3	Japan	Y	Y	Y	Y	Y	Y	Patients	Clinic	Government, NGOs	N
23	Merikanto	2017	Depression	17	0%	16	Finland	N	Y	N	N	N	N	Patients	Clinic	Foundation	N
24	Mesquita	2016	Depression	6	67%	16.5	Brazil	N	Y	N	N	N	N	Patients	Clinic	None stated	N
25	Moukaddam	2019	Depression	22	76%	N/A	U.S.	Y	Y	Y	N	N	N	Patients	Clinic	None stated	N
26	O'Brien	2017	Depression	59	73%	74	UK	N	Y	N	N	Y	N	Patients	Clinic	Research Council	N
27	Saeb	2015	Depression	40	71%	28.9	U.S	N	Y	N	N	N	N	Community	Internet	Research Grant	N
28	Saeb	2016	Depression	48	21%	N/A	U.S.	N	Y	Y	Y	Y	N		N/A		N

(continued)

Table 3. Continued.

Ref	First Author	Year	Diagnosed Condition	Total # Subjects	Female	Average Age	Country	Incentives	Observational?	Ethnicity	Marital Status	Education	Employed	Participant Type	Study Setting	Funding Source	Conflict of interest declared
														Established cohort		Research Grant	
29	Stavrakakis	2015	Depression	20	70%	36.6	Netherlands	Y	Y	N	N	N	N	Patients	Clinic	NIH	-
30	Wahle	2016	Depression	12	64%	N/A	Switzerland	N	N	N	N	N	N	Community	Internet	Industry	N
31	Barnett	2018	Schizophrenia	15	N/A	N/A	U.S.	N	Y	N	N	N	N	Patients	Clinic	NIH	N
32	Bromundt	2011	Schizophrenia	14	7%	39.9	Switzerland	Y	Y	N	N	N	N	Patients	Clinic	Industry	N
33	Bueno-Antequera	2018	Schizophrenia	82	18%	41	Spain	N	Y	N	Y	Y	Y	Patients	Clinic	University	N
34	Cella	2019	Schizophrenia	14	20%	28.1	UK	N	Y	N	N	N	N	Patients	Clinic	Research Centre	-
35	Chen	2016	Schizophrenia	159	41%	42.6	Taiwan	N	Y	N	N	Y	N	Patients	Clinic	University	N
36	Chung	2018	Schizophrenia	66	55%	44.1	China	Y	Y	Y	N	N	N	Patients	Clinic	Nil	N
37	Depp	2019	Schizophrenia	142	32%	51.6	U.S.	Y	Y	Y	N	Y	Y	Patients	Clinic	Government	Y
38	Janney	2013	Schizophrenia	249	47%	45.6	U.S	N	Y	Y	Y	Y	Y	Patients	Clinic	Research Grant	Y
39	Kume	2015	Schizophrenia	20	51%	61.8	Japan	N	Y	N	N	Y	Y	Patients	Clinic	None stated	N
40	Mulligan	2016	Schizophrenia	22	41%	37.4	UK	N	Y	Y	N	N	N	Patients	Clinic	None stated	-
41	Abdullah	2016	Bipolar	7	56%	N/A	U.S.	Y	Y	N	N	N	N	Patients	Clinic	Grant, Govt	Y
42	Banihashemi	2016	Bipolar	272	62%	37.25	Australia	N	Y	N	N	N	N	Patients	Clinic	Research Council	Y
43	Beiwinkel	2016	Bipolar	13	39%	47.2	Germany	N	Y	Y	N	Y	Y	Patients	Clinic	European Cmsn	N
44	Faurholt-Jepsen	2014	Bipolar	17	71%	33.4	Denmark	N	Y	N	N	N	N	Patients	Clinic		Y

(continued)

Table 3. Continued.

Ref	First Author	Year	Diagnosed Condition	Total # Subjects	Female	Average Age	Country	Incentives	Observational?	Ethnicity	Marital Status	Education	Employed	Participant Type	Study Setting	Funding Source	Conflict of interest declared
																European Cmsn	
45	Faurholt-Jepsen	2015	Bipolar	61	78%	29.2	Denmark	N	Y	N	Y	Y	Y	Patients	Clinic	European Cmsn	Y
46	Faurholt-Jepsen	2019	Bipolar	66	57%	30.2	Denmark	N	Y	N	Y	Y	Y	Patients	Clinic	European Cmsn	Y
47	Faurholt-Jepsen	2016	Bipolar	182	62%	33.5	Denmark	N	Y	N	Y	Y	Y	Patients	Clinic	European Cmsn	Y
48	Gonzalez	2014	Bipolar	42	64%	41	U.S.	N	Y	N	N	N	N	Community	N/A	Award and Grant	Y
49	Grierson	2016	Bipolar	63	62%	19.3	Australia	N	Y	N	N	N	N	Patients	Clinic	Grant	Y
50	Grunerbl	2015	Bipolar	10	92%	N/A	Austria	N	Y	N	N	N	N	Patients	Clinic	None stated	-
51	Janney	2014	Bipolar	60	65%	45.3	U.S.	Y	Y	N	N	N	N	Established cohort	N/A	Industry	Y
52	Krane-Gartiser	2019	Bipolar	122	54%	N/A	France	N	Y	N	N	N	N	Patients	Clinic	Research Grant	N
53	McGlinchey	2014	Bipolar	68	57%	34	U.S.	N	Y	Y	Y	N	Y	Community	Internet	Grant	-
54	Merikangas	2019	Bipolar	242	62%	48	U.S.	Y	Y	N	N	N	N	Patients	Clinic	European Cmsn	Y
55	Ortiz	2016	Bipolar	20	55%	51.6	Canada	N	Y	N	N	N	N	Patients	Clinic	Research Council	N
56	Palmius	2017	Bipolar	36	75%	43	UK	N	Y	N	N	N	Y	Patients	Clinic	None stated	-
57	McGowan	2019	Bipolar & BPD	87	74%	37.6	UK	N	Y	N	N	N	Y	Patients	Clinic	Government, University	Y

(continued)

Table 3. Continued.

Ref	First Author	Year	Diagnosed Condition	Total # Subjects	Female	Average Age	Country	Incentives	Observational?	Ethnicity	Marital Status	Education	Employed	Participant Type	Study Setting	Funding Source	Conflict of interest declared
58	Cohodes	2019	Anxiety	92	56%	12.7	U.S.	N	Y	N	N	N	N	Established cohort	N/A	Award and Grant	-
59	Jacobson	2021	Anxiety	265	58%	44.3	U.S.	N	Y	Y	N	N	N	Established cohort	N/A	Government	N
60	Epstein	2014	SUDs	27	41%	41.2	U.S.	N	Y	Y	Y	Y	Y	Patients	Clinic	Government	N
61	Kennedy	2015	SUDs	57	30%	41.4	U.S.	Y	N	Y	Y	Y	Y	Patients	Clinic	Government	N
62	Gloster	2021	Transdiagnostic	84	46%	34.8	Switzerland	N	Y	N	Y	N	N	Patients	Clinic	Foundation	N

Table 4. Remote sensor parameters temporal resolution.

Remote Sensing Parameter	Samples	Mean Duration Between Consecutive Samples (in seconds)	Standard deviation	Minimum	Maximum	95.0% lower CL for mean	95.0% upper CL for mean
Activity	33	81	73	2	300	49	113
Sleep	19	86	133	30	600	18	155
Location	14	486	279	60	900	286	686
Interactions within phone	10	143	113	30	300	−38	323
Conversations	5	.	.	.	.	.	.
Heart rate	2	60	.	60	60	.	.
Skin conductance	2	.	.	.	.	.	.
Skin temperature	3	.	.	.	.	.	.
Ambient light	1	900	.	900	900	.	.
Bluetooth	1	600	.	600	600	.	.

incentives remained inadequate. It is important for the scientific community to understand financial incentives offered because this lends insight into the extent to which engagement in a given study is likely to reflect engagement in routine care. It will also be important to explore applications of remote sensing in routine care settings with a focus on documenting engagement level and identifying patient perspectives on the value of such tools.

While various sensor types were explored to infer mental health, activity and sleep emerged as the most commonly and frequently sensed parameters, and the average time period over which sensor data was gathered was 44 days. This duration of observation may not be long enough to harness the full potential of these methods, which typically rely on intraindividual comparisons over time. This limitation has been noted in viewpoints,[66] but our finding puts an objective number behind this concern within the field. Additionally, consideration of the clinical implications of varying sampling rates across sensor types is crucial. For instance, high sampling rates may be required for activity and sleep sensors to accurately capture changes in behavior, whereas ambient light and Bluetooth sensors may require lower sampling rates because of their slower changes. It is also essential to carefully consider the practicality and feasibility of using certain types of sensors in real-world settings. For instance, location and Bluetooth sensors which demand significant power, may inadvertently disrupt other phone functionalities, thereby escalating the participant burden. While the use of multiple sensors and higher resolution has the potential to enhance data quality,[67] they may also be perceived as invasive, raising concerns about privacy and affecting participant willingness to engage in real-world studies, and in turn reducing the adoption of remote sensing in real-world. The diversity in sensor types, sampling rate, and data collection durations observed in our study underscores the necessity for standardizing approaches to sensor data collection in validation studies to advance remote sensing as a viable alternative paradigm in mental health research.[68]

There remain a number of other considerations prompted by relevant recent literature. Based on this scoping review, it is unclear that the extant sensor methods have been sufficiently investigated and validated as proposed by other authors to constitute a "digital phenotype."[64,69] While the concept of a phenotype and corresponding genotype is relevant to mental illness, it is perhaps more achievable to derive endophenotypes[70] that describe some salient measurements of physiological characteristics of those with mental illness, such as activity levels in those who are depressed. In this context, our paper highlights that there remains much more to be investigated regarding what are suitable sensor measurements and whether these are valid characterizations of physiologic phenomena that are indicative of a particular mental illness. Similarly, other related research[71] points to another crucial area of investigation, that of user-

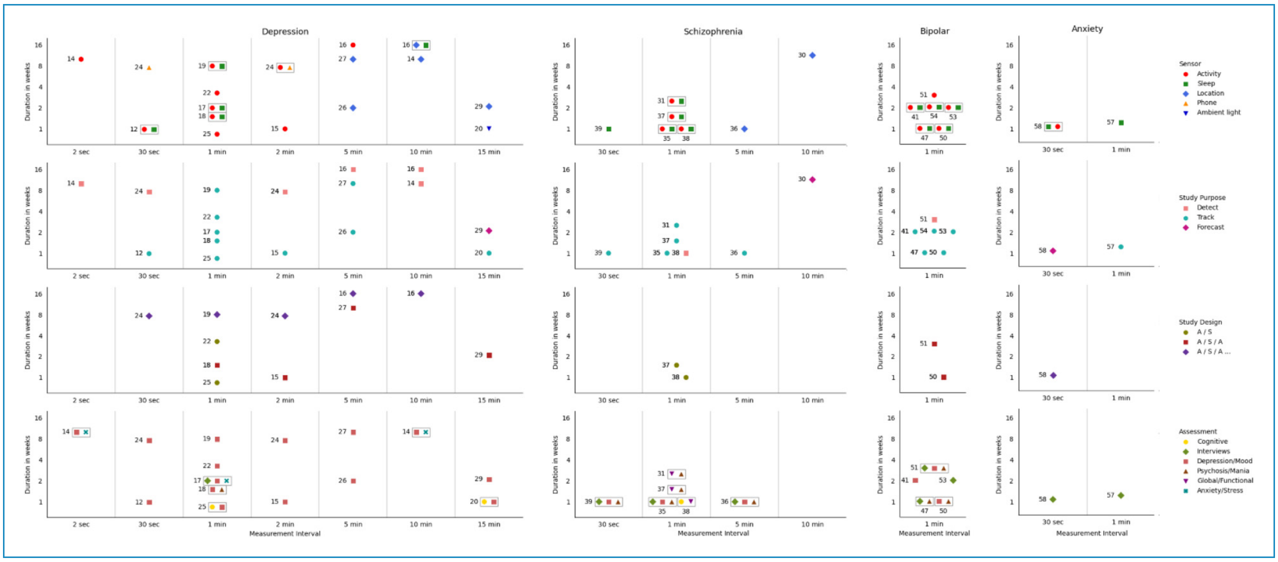

Figure 3. Graphs show the relationship between sensor measurement interval and duration of remote sensing (with duration on a logarithmic scale). Measurement interval is shown as one of seven discrete values. Numerical labels indicate the number of the publication in the References. Where multiple studies reported the same interval and duration, markers for the studies are displayed side by side with arbitrary ordering. Where multiple sensor types were used in the same study (top row) or multiple forms of assessment were used (bottom row), markers belonging to the same study are grouped together inside a rectangular box, with arbitrary ordering inside the box. Note that three studies[14,16,24] made use of multiple sensors at multiple measurement intervals, and hence are plotted in two separate locations each. Where a study shown in the top row has no corresponding marker in the same location in a lower row (e.g. study [14] has no marker for Study Design), this indicates that the information in question could not be obtained from the publication.

friendliness and acceptability of remote sensor measurements for patients and clinicians, and which will likely require customization for clinical use.

Future directions

Future research should address gaps identified in reliance on understanding the applicability of remote sensing to disadvantaged populations and comprehending human factors influencing acceptability and adoption. Potential research directions for maximizing the benefits of remote sensing technology and addressing ethical concerns could focus on: (i) designing best practice reporting standards to capture minimum sociodemographic data in remote mental health sensing research protocols; (ii) increasing research with longer monitoring windows to enhance the opportunities for intraindividual analyses; (iii) understanding feasibility of implementing these methods in routine care; and (iv) expanding application of these methods to other mental health diagnoses beyond mood disorders.

Limitations

The review has several limitations that need to be considered. First, owing to varied methodologies, objectives and data collection practices, we could not evaluate the accuracy of predictions or diagnoses based on remote sensing. Second, we focused on peer-reviewed, academic studies published and not studies or examples of applications from industry or other venues. Third, our review focused on mental health and does not reflect the state of the remote sensing literature in other clinical populations. Nevertheless, this scoping review provides a comprehensive overview of psychiatric passive sensing academic literature in a narrative form.

Conclusion

The study highlights the nascent potential of remote sensing in mental health research and clinical practice. However, this analysis also reveals gaps in the field, including limited reporting on sociodemographic information, a focus on resource-rich countries, gender imbalance in study populations, and exclusions based on social determinants. While remote sensing shows promise in mental health, there is a need for standardization in study designs, greater inclusivity and attention to sociodemographic factors, attention to underrepresented diagnoses, longer monitoring windows, and improve transparency in reporting all protocol components that may be relevant for future implementation. Finally, the promise of remote sensing also needs to consider the limitations of the current methodologies, evidence-based[64,69] and user (patient/clinician) acceptability,[71] well before the potential for characterization of mental illness using parameters can be realized.

Acknowledgments: The authors wish to acknowledge the contributions of Amy Nielson, Lidia Thorpe, Sarah Immanuel, Christine Huien Lee, Yixin Kwan, and Meseret Teferra in the data collection for this article and Dan Thorpe for analysis assistance.

Contributorship: NB conceived the study and wrote the first draft of the manuscript. RB and TP were involved in data analysis. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: Ethics approval was not required for this systematic review.

Funding: The authors received no financial support for the research, authorship, and/or publication of this article.

Informed consent: Informed consent was not required for this systematic review.

ORCID iDs: Niranjan Bidargaddi https://orcid.org/0000-0003-2868-9260
Tamara L Paget https://orcid.org/0000-0002-3664-8752

Supplemental material: Supplemental material for this article is available online.

References

1. “Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019,” *Lancet Psychiatry*, 9, , 137–150, 2022,.
2. Moore RC, Depp CA, Wetherell JL, et al. Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. *J Psychiatr Res* 2016; 75: 116–123.
3. Cornet VP and Holden RJ. Systematic review of smartphone-based passive sensing for health and wellbeing. *J Biomed Inform* 2018; 77: 120–132.
4. Dzedzickis A, Kaklauskas A and Bucinskas V. Human emotion recognition: review of sensors and methods. *Sensors* 2020; 20: 92.
5. Cho C-H, Lee T, Kim M-G, et al. Mood prediction of patients with mood disorders by machine learning using passive digital phenotypes based on the circadian rhythm: Prospective observational cohort study. *J Med Internet Res* 2019; 21. doi: 10.2196/11029
6. Asselbergs J, Ruwaard J, Ejdys M, et al. Mobile phone-based unobtrusive ecological momentary assessment of day-to-day mood: an explorative study. *J Med Internet Res* 2016; 18: 72.
7. Thompson WK, Gershon A, O’Hara R, et al. The prediction of study-emergent suicidal ideation in bipolar disorder: a pilot study using ecological momentary assessment data. *Bipolar Disord* 2014; 16: 669–677.
8. Lee SM and Lee D. Healthcare wearable devices: an analysis of key factors for continuous use intention. *Serv Bus* 2020; 14: 503–531.
9. Rosenberg S. Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center’s Global Attitudes Project. Accessed: Jan. 11, 2023. [Online]. Available: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/
10. De Angel V, Lewis S, White K, et al. Digital health tools for the passive monitoring of depression: a systematic review of methods. *Npj Digit Med* 2022; 5, Art. no. 1. doi: 10.1038/s41746-021-00548-8
11. Seppälä J, De Vita I, Jämsä T, et al. Mobile phone and wearable sensor-based mHealth approaches for psychiatric disorders and symptoms: systematic review. *JMIR Ment Health* 2019; 6: e9819.
12. Braun V and Clarke V. Thematic analysis. In: *APA Handbook of research methods in psychology, vol 2: research designs: quantitative, qualitative, neuropsychological, and biological, in APA handbooks in psychology®*., Washington, DC, US: American Psychological Association, 2012, pp.57–71. doi: 10.1037/13620-004.
13. Alcántara C, Biggs ML, Davidson KW, et al. Sleep disturbances and depression in the multi-ethnic study of atherosclerosis. *Sleep* 2016; 39: 915–925.
14. Averill IR, Crowe M, Frampton CM, et al. Clinical response to treatment in inpatients with depression correlates with changes in activity levels and psychomotor speed. *Aust N Z J Psychiatry* 2018; 52: 652–659.
15. Ben-Zeev D, Scherer EA, Wang R, et al. Next-generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health. *Psychiatr Rehabil J* 2015; 38: 218–226.
16. Bewernick BH, Urbach AS, Bröder A, et al. Walking away from depression-motor activity increases ratings of mood and incentive drive in patients with major depression. *Psychiatry Res* 2017; 247: 68–72.
17. Chikersal P, Doryab A, Tumminia M, et al. Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing: a machine learning approach with robust feature selection. *ACM Trans Comput-Hum Interact* 2021; 28: 1–41.
18. Difrancesco S, Lamers F, Riese H, et al. Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disorders: a 2-week ambulatory assessment study. *Depress Anxiety* 2019; 36: 975–986.
19. Fasmer OB, Hauge E, Berle JØ, et al. Distribution of active and resting periods in the motor activity of patients with depression and schizophrenia. *Psychiatry Investig* 2016; 13: 112–120.
20. Finazzi ME, Mesquita ME, Lopes JR, et al. Motor activity and depression severity in adolescent outpatients. *Neuropsychobiology* 2010; 61: 33–40.
21. Jacobson NC, Weingarden H and Wilhelm S. Using digital phenotyping to accurately detect depression severity. *J Nerv Ment Dis* 2019; 207: 893–896.

22. Kumagai N, Tajika A, Hasegawa A, et al. Predicting recurrence of depression using lifelog data: an explanatory feasibility study with a panel VAR approach. *BMC Psychiatry* 2019; 19: 91.
23. Merikanto I, Partonen T, Paunio T, et al. Advanced phases and reduced amplitudes are suggested to characterize the daily rest-activity cycles in depressed adolescent boys. *Chronobiol Int* 2017; 34: 967–976.
24. Mesquita ME, Finazzi ME, Gonçalves B, et al. Activity/rest rhythm of depressed adolescents undergoing therapy: case studies. *Trends Psychiatry Psychother* 2016; 38: 216–220.
25. Moukaddam N, Truong A, Cao J, et al. Findings from a trial of the smartphone and OnLine usage-based eValuation for depression (SOLVD) application: what do apps really tell us about patients with depression? Concordance between app-generated data and standard psychiatric questionnaires for depression and anxiety. *J Psychiatr Pract* 2019; 25: 365–373.
26. O'Brien JT, Gallagher P, Stow D, et al. A study of wrist-worn activity measurement as a potential real-world biomarker for late-life depression. *Psychol Med* 2017; 47: 93–102.
27. Saeb S, Zhang M, Karr CJ, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *J Med Internet Res* 2015; 17: e175.
28. Saeb S, Lattie EG, Schueller SM, et al. The relationship between mobile phone location sensor data and depressive symptom severity. *PeerJ* 2016; 4: e2537.
29. Stavrakakis N, Booij SH, Roest AM, et al. Temporal dynamics of physical activity and affect in depressed and nondepressed individuals. *Health Psychol* 2015; 34: 1268–1277.
30. Wahle F, Kowatsch T, Fleisch E, et al. Mobile sensing and support for people with depression: a pilot trial in the wild. *JMIR MHealth UHealth* 2016; 4: e111.
31. Barnett I, Torous J, Staples P, et al. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. *Neuropsychopharmacology* 2018; 43: 1660–1666.
32. Bromundt V, Köster M, Georgiev-Kill A, et al. Sleep-wake cycles and cognitive functioning in schizophrenia. *Br J Psychiatry J Ment Sci* 2011; 198: 269–276.
33. Bueno-Antequera J, Oviedo-Caro MÁ and Munguía-Izquierdo D. Relationship between objectively measured sedentary behavior and health outcomes in schizophrenia patients: the PsychiActive project. *Schizophr Res* 2018; 197: 87–92.
34. Cella M, He Z, Killikelly C, et al. Blending active and passive digital technology methods to improve symptom monitoring in early psychosis. *Early Interv Psychiatry* 2019; 13: 1271–1275.
35. Chen L-J, Steptoe A, Chung M-S, et al. Association between actigraphy-derived physical activity and cognitive performance in patients with schizophrenia. *Psychol Med* 2016; 46: 2375–2384.
36. Chung K-F, Poon YPY-P, Ng T-K, et al. Correlates of sleep irregularity in schizophrenia. *Psychiatry Res* 2018; 270: 705–714.
37. Depp CA, Bashem J, Moore RC, et al. GPS Mobility as a digital biomarker of negative symptoms in schizophrenia: a case control study. *NPJ Digit Med* 2019; 2: 108.
38. Janney CA, Ganguli R, Richardson CR, et al. Sedentary behavior and psychiatric symptoms in overweight and obese adults with schizophrenia and schizoaffective disorders (WAIST study). *Schizophr Res* 2013; 145: 63–68.
39. Kume Y, Sugita T, Oga K, et al. A pilot study: comparative research of social functioning, circadian rhythm parameters, and cognitive function among institutional inpatients, and outpatients with chronic schizophrenia and healthy elderly people. *Int Psychogeriatr* 2015; 27: 135–143.
40. Mulligan LD, Haddock G, Emsley R, et al. High resolution examination of the role of sleep disturbance in predicting functioning and psychotic symptoms in schizophrenia: a novel experience sampling study. *J Abnorm Psychol* 2016; 125: 788–797.
41. Abdullah S, Matthews M, Frank E, et al. Automatic detection of social rhythms in bipolar disorder. *J Am Med Inform Assoc JAMIA* 2016; 23: 538–543.
42. Banihashemi N, Robillard R, Yang J, et al. Quantifying the effect of body mass index, age, and depression severity on 24-h activity patterns in persons with a lifetime history of affective disorders. *BMC Psychiatry* 2016; 16: 17.
43. Beiwinkel T, Kindermann S, Maier A, et al. Using smartphones to monitor bipolar disorder symptoms: a pilot study. *JMIR Ment Health* 2016; 3: e2.
44. Faurholt-Jepsen M, Frost M, Vinberg M, et al. Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Res* 2014; 217: 124–127.
45. Faurholt-Jepsen M, Vinberg M, Frost M, et al. Smartphone data as an electronic biomarker of illness activity in bipolar disorder. *Bipolar Disord* 2015; 17: 715–728.
46. Faurholt-Jepsen M, Vinberg M, Frost M, et al. Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. *Int J Methods Psychiatr Res* 2016; 25: 309–323.
47. Faurholt-Jepsen M, Busk J, Þórarinsdóttir H, et al. Objective smartphone data as a potential diagnostic marker of bipolar disorder. *Aust N Z J Psychiatry* 2019; 53: 119–128.
48. Gonzalez R, Tamminga CA, Tohen M, et al. The relationship between affective state and the rhythmicity of activity in bipolar disorder. *J Clin Psychiatry* 2014; 75: e317–e322.
49. Grierson AB, Hickie IB, Naismith SL, et al. Circadian rhythmicity in emerging mood disorders: state or trait marker? *Int J Bipolar Disord* 2016; 4: 3.
50. Grünerbl A, Muaremi A, Osmani V, et al. Smartphone-based recognition of states and state changes in bipolar disorder patients. *IEEE J Biomed Health Inform* 2015; 19: 140–148.
51. Janney CA, Fagiolini A, Swartz HA, et al. Are adults with bipolar disorder active? Objectively measured physical activity and sedentary behavior using accelerometry. *J Affect Disord* 2014; 152–154: 498–504.
52. Krane-Gartiser K, Scott J, Nevoret C, et al. Which actigraphic variables optimally characterize the sleep-wake cycle of individuals with bipolar disorders? *Acta Psychiatr Scand* 2019; 139: 269–279.
53. McGlinchey EL, Gershon A, Eidelman P, et al. Physical activity and sleep: day-to-day associations among individuals with and without bipolar disorder. *Ment Health Phys Act* 2014; 7: 183–190.
54. Merikangas KR, Swendsen J, Hickie IB, et al. Real-time Mobile monitoring of the dynamic associations among motor activity, energy, mood, and sleep in adults with bipolar disorder. *JAMA Psychiatry* 2019; 76: 190–198.
55. Ortiz A, Bradler K, Radu L, et al. Exponential state transition dynamics in the rest-activity architecture of patients with bipolar disorder. *Bipolar Disord* 2016; 18: 116–123.
56. Palmius N, Tsanas A, Saunders KE, et al. Detecting bipolar depression from geographic location data. *IEEE Trans Biomed Eng* 2017; 64: 1761–1771.

57. McGowan NM, Goodwin GM, Bilderbeck AC, et al. Circadian rest-activity patterns in bipolar disorder and borderline personality disorder. *Transl Psychiatry* 2019; 9: 95.
58. Cohodes EM, Abusch A, Odriozola P, et al. Novel insights from actigraphy: anxiety is associated with sleep quantity but not quality during childhood. *Clin Child Psychol Psychiatry* 2020; 25: 189–199.
59. Jacobson NC, Lekkas D, Huang R, et al. Deep learning paired with wearable passive sensing data predicts deterioration in anxiety disorder symptoms across 17-18 years. *J Affect Disord* 2021; 282: 104–111.
60. Epstein DH, Tyburski M, Craig IM, et al. Real-time tracking of neighborhood surroundings and mood in urban drug misusers: application of a new method to study behavior in its geographical context. *Drug Alcohol Depend* 2014; 134: 22–29.
61. Kennedy AP, Epstein DH, Jobes ML, et al. Continuous in-the-field measurement of heart rate: correlates of drug use, craving, stress, and mood in polydrug users. *Drug Alcohol Depend* 2015; 151: 159–166.
62. Gloster AT, Meyer AH, Klotsche J, et al. The spatiotemporal movement of patients in and out of a psychiatric hospital: an observational GPS study. *BMC Psychiatry* 2021; 21: 65.
63. Ben-Zeev D, Young MA and Depp CA. Real-time predictors of suicidal ideation: mobile assessment of hospitalized depressed patients. *Psychiatry Res* 2012; 197: 55–59.
64. Mendes JPM, Moura IR, Van de Ven P, et al. Sensing apps and public data sets for digital phenotyping of mental health: systematic review. *J Med Internet Res* 2022; 24: e28735.
65. Huppert FA. Positive mental health in individuals and populations. In: *The science of well-being*, New York, NY. US: Oxford University Press, 2005, pp.306–340. doi: 10.1093/acprof:oso/9780198567523.003.0012
66. Faurholt-Jepsen M, Bauer M and Kessing LV. Smartphone-based objective monitoring in bipolar disorder: status and considerations. *Int J Bipolar Disord* 2018; 6: 6.
67. Teh HY, Kempa-Liehr AW and Wang KI-K. Sensor data quality: a systematic review. *J Big Data* 2020; 7: 11.
68. Schick A, Rauschenberg C, Ader L, et al. Novel digital methods for gathering intensive time series data in mental health research: scoping review of a rapidly evolving field. *Psychol Med* 2023; 53: 55–65.
69. Moura I, Teles A, Viana D, et al. Digital phenotyping of mental health using multimodal sensing of multiple situations of interest: a systematic literature review. *J Biomed Inform* 2023; 138: 104278.
70. Gottesman II and Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry* 2003; 160: 636–645.
71. Rogan J, Bucci S and Firth J. Health care Professionals' views on the use of passive sensing, AI, and machine learning in mental health care: systematic review with meta-synthesis. *JMIR Ment Health* 2024; 11: e49577.

Appendix 1

					Self-report						
ID	First Author	Year	Interview	Cognitive Assessments	*Global & Functional*	*Depression/ Mood*	*Anxiety & Stress*	*Psychosis/ Mania*	*Other*	Duration (days)	Experiment Design Type
12	Alcantara	2018				CES-D			WHIIRS, ESS	7	5
13	Averill	2018			CGIS, FAST	QIDS				3	2
14	Ben-Zeev	2015				PHQ-9	PSS		UCLA Loneliness Scale	70	5
15	Bewernick	2017				MADRS				7	2
16	Chikersal	2021								16	1
17	Difranceso	2019	CIDI			IDS	Beck A			14	5
18	Fasmer	2016				MADRS		BPRS		11	2
19	Finazzi	2009				K-SADS-PL, CDRS				63	1
20	Jacobson	2019		MINI		HAM-D					5
21	Kumagai	2019				PHQ-9				365	1
22	Merikanto	2017				K-SADS-PL, HAM-D				18	3
23	Mesquita	2016								91	1
24	Moukaddam	2019				PHQ-9, HAM-D	HAM-A			8	1
25	O'Brien	2017		MINI, NART, EHI		MADRS, GDS-15				7	3
26	Saeb	2015				PHQ-9				14	5
27	Saeb	2016				PHQ-9				10	2

(continued)

Appendix 1. Continued.

					Self-report						
ID	First Author	Year	Interview	Cognitive Assessments	*Global & Functional*	*Depression/ Mood*	*Anxiety & Stress*	*Psychosis/ Mania*	*Other*	Duration (days)	Experiment Design Type
28	Stavrakakis	2015							MCTQ	30	5
29	Wahle	2016				PHQ-9				14	2
30	Barnett	2018								84	5
31	Bromundt	2011			CGIS			BPRS, PANNS	SAS	21	5
32	Bueno-Antequera	2018			SF-36				BSI-18	7	4
33	Cella	2019						PANNS		10	3
34	Chen	2016	DSM-IV	VTS, GPT						7	5
35	Chung	2018	DSM-V			CDSS		PANNS		7	5
36	Depp	2019	CAINS 33			CDSS		BPRS, SANS		7	5
37	Janney	2013			GAF, CGIS			PANNS		7	3
38	Kume	2015		BACS	GAF					7	3
39	Mulligan	2016	DSM-IV			CDSS		PSYRATS, PANNS		7	5
40	Abdullah	2016								28	5
41	Banihashemi	2016				HAM-D				22	5
42	Beiwinkel	2016	DSM-IV-R			HAM-D		YMRS		365	1
43	Faurholt-Jepsen	2014				HAM-D		YMRS		91	1
44	Faurholt-Jepsen	2015	SCAN					YMRS		182	1

(continued)

Appendix 1. Continued.

					Self-report						
ID	First Author	Year	Interview	Cognitive Assessments	*Global & Functional*	*Depression/ Mood*	*Anxiety & Stress*	*Psychosis/ Mania*	*Other*	Duration (days)	Experiment Design Type
45	Faurholt-Jepsen	2019	SCAN					YMRS		84	1
46	Faurholt-Jepsen	2016	SCAN			HAM-D		YMRS		49	1
47	Gonzalez	2014	ISD-C-30					YMRS		7	5
48	Grierson	2016								14	5
49	Grunerbl	2015				HAM-D		YMRS		84	1
50	Janney	2014				HAM-D		YMRS		7	2
51	Krane-Gartiser	2019	DIGs, FIGs			MADRS		YMRS		21	2
52	McGlinchey	2014	DSM-IV			IDS		YMRS		60	2
53	Merikangas	2019	DSM-V							14	5
54	Ortiz	2016								14	5
55	Palmius	2017	DSM-IV			QIDS				7	4
56	McGowan	2019	IPDE			QIDS		ASRM		4	5
57	Cohodes	2019					MASC		PSQI	18	5
58	Jacobson	2021	CIDI							7	1
59	Epstein	2014	DIS-IV						ASI	112	5
60	Kennedy	2015	DIS-IV, SCID						ASI	49	2
61	Gloster	2021			MHC-SF				BSCL, Psyflex	7	5

Design (1: gold standard → sensor → gold standard etc.; 2: gold standard → sensor → gold standard (once off) 3: gold standard → sensor; 4: sensor → gold standard; 5: unknown

Appendix 2

Study Design

As might be expected, there was a relationship between duration of sensor measurement and study design, as studies that were of longer duration were also likely to follow a more elaborate design of multiple cycles of assessment and sensing ("A / S / A …" in Figure 1), while studies of shorter duration followed simpler study designs. Study design appeared to be unrelated to the measurement interval.