

REVIEW ARTICLE

Digital Phenotyping for Relapse Prediction in Depression and Schizophrenia: A Narrative Review of Signals, Models, and Clinical Translation

Qian Wu,^{#,†} Xixi Fan,^{#,‡} and Xin Qi^{††}

[†]The First Clinical College of Gansu University of Chinese Medicine, Lanzhou, Gansu, China

[‡]James Cook University Singapore, Singapore

[#] The authors contributed equally.

^{††}Corresponding author. Email: qixin13895932514@163.com

Abstract

Digital phenotyping brings measurement of daily behaviour and physiology into the study of psychiatric relapse. Smartphones capture mobility, communication-related activity, and device interactions, while wearable sensors trace sleep, rest-activity rhythms, and selected autonomic signals. This narrative review synthesises evidence on these approaches in depression and schizophrenia and examines their development into clinically useful warning systems. Smartphone studies demonstrate the value of longitudinal behavioural observation, pilot work in schizophrenia has directly examined relapse-related changes, and prospective actigraphy research in depression links sleep and activity patterns with subsequent relapse. Multimodal models provide a framework for integrating these complementary signals. The central challenge is to turn measurement into dependable decisions: a relapse-warning system needs a defined clinical outcome, a meaningful forecast horizon, reproducible sensor features, independent validation, and an agreed response to alerts. Patient-reported information strengthens the interpretation of passive data, while privacy, consent, equitable access, and usability shape implementation. We propose an evidence-to-care framework connecting digital signals, prospective risk estimation, and clinical action. This approach positions digital phenotyping as a practical component of future longitudinal mental-health care, with clinical effectiveness to be established through prospective implementation studies.

Keywords: Digital phenotyping; Depression; Schizophrenia; Relapse prediction; Smartphone sensing; Wearable devices; Digital mental health

1. Introduction

Relapse prevention is a central concern in the continuing care of depression and schizophrenia. Recurrent depressive episodes can disrupt recovery and functioning, while relapse after an initial episode of schizophrenia can alter longer-term clinical trajectories (Monroe and Harkness 2011; Robinson et al. 1999). Effective follow-up therefore depends on recognising changes in everyday functioning as they develop. Clinical consultations, symptom scales, and patient or caregiver reports provide indispensable assessment, but they sample a small part of life between visits. Smartphones and



wearable technologies expand this observation window by recording behavioural and physiological variation in natural settings (Torous *et al.* 2016; Bufano *et al.* 2023).

Digital phenotyping translates everyday technology use into repeated measurements of behaviour, physiology, and their temporal patterns. Geolocation and movement sensors describe routines and mobility; phone interaction data reflect patterns of digital activity; wearables add information about sleep, physical activity, and selected autonomic processes. Ecological momentary assessments and brief self-reports can anchor these signals in patients' own experience (Kamath *et al.* 2022; Shen *et al.* 2025). Together, these sources can document changes that precede, accompany, or follow clinical deterioration. Their clinical value rests on identifying which changes offer usable advance warning of a specified relapse event.

This review follows the progression from measurement to clinical action. It first considers the behavioural and physiological meaning of digital features, then compares evidence from smartphones, wearables, and multimodal systems. It next examines model design, longitudinal validation, and integration into psychiatric services. A central distinction guides the analysis: symptom monitoring describes the current state, whereas relapse prediction estimates a future event over a defined interval. The literature contains important contributions at each stage, including direct pilot work in schizophrenia and prospective observational evidence in depression (Barnett *et al.* 2018; Tonon *et al.* 2026). Bringing these strands together clarifies what a clinically credible warning pathway requires.

2. Review Scope and Evidence Framework

We conducted a question-led narrative synthesis of primary studies and reviews included in the source literature collection. We selected and organised publications for their relevance to five linked questions: which signals are measured, what clinical outcomes they represent, how temporal relationships are assessed, how prediction models are evaluated, and how digital monitoring can inform care. The review discusses depression and schizophrenia separately where their evidence bases or clinical endpoints differ. This approach supports critical comparison of concepts, study designs, and translation pathways across a developing interdisciplinary field.

The synthesis distinguishes four levels of evidence: reliable acquisition of digital signals; association with concurrent symptoms or functioning; longitudinal association with later clinical deterioration or relapse; and validated prediction linked to a usable clinical response. Each level answers a different question. For example, a mobility-symptom correlation establishes the relevance of a behavioural measure, while a prospectively tested relapse model addresses its forecasting value. We use these categories to interpret individual studies and to identify the next research step for each digital modality.

3. Conceptual and Measurement Foundations

3.1 *The biopsychosocial perspective and behavioural context*

The biopsychosocial perspective provides a coherent basis for selecting and interpreting digital features. Mental-health trajectories involve interactions among biological regulation, psychological experience, and social conditions (Engel 1977). Wearable heart-rate and sleep measurements characterise aspects of physiological regulation; mobility and activity capture changes in daily behaviour; and communication-related features describe parts of social participation. These measures acquire clinical meaning through their relation to an individual's context. Reduced movement, for example, can accompany depressive withdrawal, medication effects, illness, work routines, or restricted travel. Integrating patient history and contemporaneous assessment gives sensor patterns a more precise interpretation.

A dynamical-systems perspective extends this framework from isolated measurements to trajectories. Sleep disruption, reduced activity, and altered interaction may co-occur and evolve together as a person's condition changes (Bosl, Enlow, and Nelson 2025). Longitudinal digital data make it

possible to examine these sequences at a finer temporal resolution than periodic clinic visits. The central analytical task is to distinguish meaningful deviations from a person's baseline from ordinary variation introduced by context, devices, and measurement conditions.

3.2 Behavioural activation and digital biomarkers

Behavioural activation theory helps explain why patterns of engagement are relevant to depression. Reductions in activity, reward-seeking, and social participation interact with mood and functioning. Related changes in movement and engagement also occur in some presentations of schizophrenia, particularly where negative symptoms are prominent. Smartphone mobility, activity timing, and interaction records operationalise selected aspects of these processes (Mow et al. 2022; Fisher et al. 2025). Their temporal measurement creates an opportunity to examine how behavioural changes develop in relation to symptom trajectories and subsequent relapse.

The path from sensor output to digital biomarker has several defined steps. Investigators first specify the raw signal and recording conditions, then extract reproducible features, establish measurement reliability, and connect the feature to a clinically meaningful outcome. They must also document gaps in recording and the influence of device hardware, operating systems, and proprietary processing. These choices determine whether results can be compared across cohorts and reproduced on other devices (Alam et al. 2025; Hassan et al. 2025). A clinically useful biomarker combines technical stability with an interpretable link to prognosis or patient management.

4. Evidence from Smartphones and Wearable Devices

4.1 Smartphone mobility and digital behaviour

Smartphones reveal patterns of mobility and digital activity at the scale of everyday life. Geolocation data describe time spent at home, location diversity, and changes in routine, while interaction logs capture aspects of device engagement. In a multicentre observational study, Laiou et al. (2022) examined smartphone-derived home-stay behaviour in relation to symptoms of major depressive disorder. Their findings establish a connection between daily mobility and clinical state. Research on phone interaction and passive behavioural features adds detail about individual differences and the range of signals associated with depression (Kamath et al. 2022; Lampe et al. 2026). Together, these studies build the measurement basis for future longitudinal warning models.

Research in schizophrenia connects smartphone-derived behaviour with both functioning and relapse-related change. Mow et al. (2022) examined mobility metrics in relation to daily social motivation and behaviour. Barnett et al. (2018) took the next step by investigating relapse prediction from longitudinal smartphone features in a pilot study. This distinction is clinically important: the former clarifies behavioural relevance, while the latter directly examines temporal forecasting. A systematic review of schizophrenia-spectrum digital phenotyping maps the broader evidence across diagnosis, symptoms, and clinical trajectories, highlighting the diversity of measures and endpoints in use (Vecchio et al. 2025).

Context determines how mobility and phone-use features should enter clinical models. Time at home can reflect withdrawal, remote work, physical illness, or household responsibilities. Communication patterns also depend on application preferences and changes in technology use. Stronger models therefore combine personal baselines with transparent feature definitions, device-quality information, and relevant contextual variables. Patient accounts and clinical assessment help interpret an alert by identifying the circumstances behind an observed behavioural change.

4.2 Wearable sleep, activity, and autonomic measurements

Wearable sensors extend digital phenotyping from behavioural observation to repeated physiological measurement. Rest-activity timing, sleep-related features, heart rate, and heart-rate variability have been examined in psychiatric and sleep research (Boettger et al. 2006; Chouchou and Desseilles 2014).

These measures can reveal how biological rhythms vary over days and weeks. Their interpretation requires attention to measurement conditions: physical exertion, medication, illness, sensor placement, and device algorithms all influence recorded values (Hassan et al. 2025). Longitudinal designs can evaluate whether patterns in these signals contribute prognostic information beyond clinical history and concurrent symptoms.

Prospective actigraphy research offers a direct example of relapse-focused measurement. Tonon et al. (2026) followed a multicentre observational cohort of 93 adults with remitted major depressive disorder and investigated sleep and rest-activity rhythms in relation to subsequent relapse. The study links longitudinal rhythmic features with a clinically specified future outcome, strengthening the evidence for the temporal relevance of wearable signals. Related multimodal follow-up research demonstrates how repeated data streams can be collected alongside psychiatric assessment (Aledavood et al. 2025). These studies provide a foundation for testing calibrated risk estimates, independent replication, and interventions triggered by changes in rhythms.

The value of wearables is greatest when their measurements add useful information to existing follow-up. Consumer-device research in severe mental illness identifies physiological and activity signals that can be obtained outside clinical facilities (Hassan et al. 2025). For relapse-warning services, the next evaluation compares these signals with practical alternatives such as symptom ratings, known relapse history, and routine clinician contact. This comparative design determines the incremental value of continuous sensing for patients and care teams.

4.3 Multimodal monitoring and limits of aggregate performance

Multimodal phenotyping combines information that single sensors capture only partially. Mobility records behavioural opportunity, wearables register sleep and physiological rhythms, and symptom reports describe subjective experience. Integrated systems have examined combinations of these channels in psychiatric monitoring (Zlatintsi et al. 2022; Shen et al. 2025). The most productive integration strategy tests whether each modality contributes new information after accounting for basic clinical predictors. Clear temporal alignment, missing-data handling, and evaluation in an independent cohort are central to determining the added value of multimodal fusion.

Evidence syntheses show both the momentum of passive-sensing prediction research and the variation in reported model performance. Fang and Chen (2026) reviewed studies covering multiple psychiatric diagnoses, including depression, bipolar disorder, and schizophrenia-spectrum conditions, and reported discrimination results across heterogeneous settings, including AUC values of approximately 0.70–0.88. Because those estimates arise from different samples and endpoints, they describe a range of model findings across the reviewed field. The schizophrenia-spectrum review by Vecchio et al. (2025) likewise documents substantial methodological diversity. These findings direct attention to diagnosis-specific outcomes, validation design, and the operating characteristics that will determine practical clinical performance.

4.4 Speech, self-report, and the distinction between active and passive sensing

Speech offers another channel for describing changes in psychiatric state. Acoustic features such as pace, pauses, intonation, and voice quality have been investigated as potential markers of psychosis-related deterioration (Zaher et al. 2024). Their clinical interpretation depends on language, task design, microphone characteristics, conversational context, and medication. Research protocols should specify whether recordings are obtained during active speech tasks or through ambient collection, since these approaches produce different data and privacy obligations. Clearly defined acoustic features can be integrated with behavioural and clinical measurements when their incremental value has been demonstrated.

Active reports give meaning to passive measurements. Brief symptom ratings and ecological momentary assessments record mood, distress, functioning, and recent experiences in the patient's

own terms. They also help explain changes in mobility, sleep, or phone use that have several plausible causes (Torous et al. 2016; Bufano et al. 2023). Combining active and passive channels requires attention to the circumstances of missing observations, because survey nonresponse and temporary sensor inactivity are themselves shaped by participant experience and technical conditions. Comparative analyses can show how much each channel contributes to a given prediction task (Shen et al. 2025).

5. Prediction Models: From Discrimination to Clinical Use

5.1 Defining the clinical event and prediction horizon

Clinical relapse must be defined before a prediction model is built. Studies variously use symptom thresholds, clinician-adjudicated episodes, treatment escalation, emergency presentations, or hospital admission as outcomes. Each definition identifies a different clinical event and carries different implications for early warning. The prediction horizon is equally important: a seven-day warning supports immediate follow-up, whereas a one-year risk estimate informs longer-term planning. Prospective protocols should specify the outcome, monitoring interval, forecast window, and handling of repeated events. These decisions connect statistical performance to an actual care decision.

5.2 Model development and validation

Time-series prediction requires evaluation that follows the sequence of clinical observation. Flexible algorithms can model nonlinear relationships and individual patterns, but repeated measurements from the same participant also create opportunities for leakage. Participant-level separation, chronological testing, and independent-site validation prevent models from gaining artificial advantages through familiar individuals or future observations. All preprocessing, feature selection, and imputation procedures should be fitted within the training data for each validation split. Reviews of psychiatric prediction models identify transportability and replication as central criteria for judging progress (Meehan et al. 2022; Schmaal 2023; Chekroud et al. 2024).

Prediction performance is best described through a linked set of measures. AUC summarises discrimination across thresholds. Calibration shows whether estimated risks correspond to observed event frequencies; sensitivity and specificity describe operating trade-offs; and positive predictive value indicates how frequently an alert identifies an actual event. Short warning windows can make alert frequency and false-positive burden particularly important to clinical teams. Reports should state confidence intervals, participant and event counts, missing-data coverage, clinical thresholds, and the unit of prediction. Benchmarking against standard clinical follow-up reveals whether digital features improve decisions in a meaningful way.

5.3 Personalisation, interpretability, and comparative utility

Personalised models offer a way to accommodate differences in baseline mobility, sleep, and communication. Within-person trajectories can identify departures from established patterns, while clinically interpretable features allow care teams to investigate the reason for a warning. Model updates should account for medication changes, altered routines, sensor replacements, and periods with sparse data. Comparative evaluations can then establish whether complex personalised methods outperform transparent baselines based on recent symptoms, clinical history, and simple longitudinal rules. Interpretability matters most when it supports a clear and proportionate clinical response.

5.4 Frequent methodological failure modes

Several methodological choices strongly influence apparent prediction performance. Follow-up length determines the opportunity to observe relapse, while overlapping feature and outcome windows can transform a detection task into an apparent forecast. Threshold selection after inspecting

test outcomes exaggerates estimated performance. Non-random data loss can coincide with clinical deterioration or device disengagement. Finally, splitting individual device-days at random allows data from the same participant to enter both training and test sets. Transparent reporting of event definitions, sampling, splitting, and missingness makes these risks visible and supports credible replication.

A useful warning explains what the service should do next. Reduced daily movement, disrupted sleep, or unusual phone activity can prompt a supportive assessment that places the signal in context. Clinical teams may respond with a check-in, review of physical health or medication, or a fuller symptom assessment. The algorithm's role is to prioritise attention using longitudinal information. Clear escalation responsibilities and an accessible record of why an alert occurred make that information actionable for patients and clinicians.

6. Feasibility, Feedback, and Clinical Workflow Integration

A clinically integrated digital phenotyping service links four activities: sensing, risk estimation, communication, and response. Sensor data are converted into interpretable features, evaluated against a predefined alert threshold, and routed to an identified recipient. That recipient needs a practical plan for follow-up, including timelines, escalation arrangements, and documentation. Protocols for digital psychiatry illustrate how monitoring and needs-based interventions can be connected within a planned service pathway (Fowler *et al.* 2019; Oakey-Neate *et al.* 2020). Evaluating the entire pathway shows whether information generated outside the clinic improves the timeliness and quality of care.

Sustained participation depends on fitting technology into daily life. Charging, connectivity, permissions, interface design, and symptom fluctuations all affect adherence. Feasibility research in severe mental illness has therefore focused on monitoring burden and usability in real settings (Machleid *et al.* 2025). Participatory work on mental-health data visualisation shows how accessible explanations can support engagement with recorded information (Lim *et al.* 2026). Successful systems give users understandable feedback, meaningful control over participation, and a direct route to clinical support when a warning requires discussion.

The outcome of a digital warning service depends on the decisions made after an alert. Evaluation should measure clinical response time, false-alert burden, patient experience, intervention delivery, service workload, and subsequent relapse-related outcomes. Studies comparing an alert-enabled pathway with appropriate usual care can establish its incremental effect. This focus shifts evaluation from an isolated model score to the patient and service consequences of using that model. It also clarifies where prediction improves care and where workflow design remains the decisive factor.

7. Ethics, Governance, and Equity

7.1 Privacy and meaningful consent

Digital phenotyping produces sensitive information about location, routines, relationships, and health. Longitudinal recordings can reveal substantial detail even when obvious personal identifiers are removed. Ethical scholarship therefore calls for clear governance of access, secondary use, and continuous observation (Martinez-Martin *et al.* 2018; Tomičić, Malešević, and Čartolovni 2021). A sound design begins with data minimisation, separates research and care purposes, specifies retention and deletion periods, and gives participants understandable control over collection and sharing. Security provisions should match the sensitivity and intended use of each data stream.

Consent should remain meaningful throughout a long monitoring period. Clinical circumstances and an individual's ability to make a particular decision can change, making reassessment and supported decision-making relevant in some settings (Palmer *et al.* 2018). Patients need accessible explanations of which signals are measured, which parties can see them, how an alert is interpreted,

and how consent can be changed. Transparent control of raw audio, location histories, and model updates provides a practical basis for trust.

7.2 Bias, cultural context, and digital exclusion

Digital behaviour varies across social and cultural contexts. Work schedules, transport access, household structure, disability, and device ownership influence mobility and communication features. A model trained in one service setting may consequently perform differently elsewhere. The broader psychiatric prediction literature demonstrates the importance of testing generalisation across independent populations (Chekroud et al. 2024). Evaluation should examine subgroup discrimination, calibration, completeness of sensing, and participation patterns. Such analyses help distinguish clinical signals from inequalities in technology access and routine behaviour.

Equitable implementation also concerns who can participate. Monitoring depends on a usable device, connectivity, regular charging, understandable instructions, and accessible support. Cognitive difficulties and resource constraints can interrupt any of these requirements. Literature on serious mental illness and digital-health implementation identifies these practical needs as central to adoption (Hatch et al. 2018; Alam et al. 2026). Services should offer assistance with digital tools alongside alternative routes to follow-up so that continuity of care does not depend on constant sensor participation.

7.3 Accountability across the data and care pathway

Accountability connects the technology to the institution using it. Healthcare organisations and technology providers need clear responsibilities for model updates, data access, system outages, unexpected alerts, and response delays. Clinical teams should know who reviews a warning and what actions are expected. Patients should receive equally clear information about the monitoring schedule and the availability of a human response. Protocols, training, audit processes, and incident reporting turn these responsibilities into a dependable care pathway.

8. Priority Research Agenda

The first research priority is prospective, diagnosis-specific evaluation of relapse. Studies should define clinical events and monitoring windows before data analysis and follow remitted patients over periods appropriate to the intended warning. Depression and schizophrenia warrant separate outcome reporting even when the same sensor platform is used. Harmonised definitions within each condition will permit meaningful comparisons of sensor signatures and model performance across sites.

The second priority is reproducible measurement. Investigators should report sensor specifications, device and operating-system versions, sampling frequency, uptime, missingness, and methods for constructing each feature. Medication changes, employment routines, seasons, and other relevant contextual variables belong in the data dictionary. Shared measurement definitions and interoperable datasets will make multicentre comparison more informative and improve the prospects for replication (Alam et al. 2025; Shen et al. 2025).

The third priority is clinically oriented prediction validation. Models should undergo temporal and external testing, probability calibration, evaluation at actionable thresholds, and comparison with simple clinical baselines. Full reporting of cohort selection, prediction cutoffs, data partitioning, uncertainty intervals, and missing-data procedures will make performance interpretable. Generalisation across devices, services, and patient populations deserves particular attention given the challenges documented in psychiatric prediction research (Meehan et al. 2022; Schmaal 2023).

The fourth priority is evaluation of the warning-to-care pathway. Prospective implementation studies should specify who receives alerts, how follow-up is initiated, and which responses are appropriate for different levels of risk. Trials and pragmatic evaluations can assess patient outcomes,

workload, alert fatigue, and the acceptability of communication. Patient and caregiver involvement helps align monitoring practices with everyday needs and preferences (Lim et al. 2026; Machleid et al. 2025).

The fifth priority is privacy- and equity-centred design. Projects should minimise sensitive data collection, document storage and deletion practices, establish oversight of secondary data use, and test whether models perform comparably across relevant groups. Researchers should report who could not participate and why, including device access and digital-literacy barriers. These practices strengthen the interpretability and fairness of the evidence produced by digital phenotyping (Martinez-Martin et al. 2018; Tomičić, Malešević, and Čartolovni 2021).

A further scientific opportunity lies in direct comparison between diagnoses. Sleep disruption and reduced activity occur across several psychiatric conditions, while their relationship to relapse may differ according to symptom profile, medication, and baseline functioning. Comparative cohorts can test diagnosis-specific features, shared predictors, and individual baseline approaches under harmonised follow-up procedures. This design will identify which signals support common warning mechanisms and which require condition-specific interpretation.

9. Scope and Limitations

This review uses a question-led narrative selection of published studies and reviews. Its synthesis therefore emphasises conceptual and methodological comparison rather than comprehensive study enumeration or pooled effect estimation. Variations in source reporting, possible overlap between datasets covered by reviews, and selective publication of model results affect the available evidence. A future systematic review with prespecified eligibility criteria and structured risk-of-bias assessment could quantify these patterns and compare performance within clinically comparable outcomes.

10. Conclusion

Digital phenotyping creates a bridge between the everyday course of psychiatric illness and structured clinical follow-up. Smartphone studies link mobility and interaction patterns with behaviour and symptoms; pilot work has examined relapse-related smartphone signals in schizophrenia; and prospective actigraphy has connected sleep and rest-activity rhythms with later relapse in depression (Barnett et al. 2018; Laiou et al. 2022; Tonon et al. 2026). The next advance is an integrated system that defines the clinical event, measures reliable longitudinal features, validates risk estimates independently, and turns alerts into proportionate care. Multimodal measurement, patient-reported context, and transparent governance form complementary elements of that system. By connecting evidence quality with implementation design, this review outlines a practical research agenda for making digital phenotyping useful in longitudinal psychiatric care.

Declarations

Author Contributions

The authors' contributions are described according to the Contributor Roles Taxonomy (CRediT): **Conceptualization:** Qian Wu and Xixi Fan; **Investigation and Literature Search:** Qian Wu, Xixi Fan, and Xin Qi; **Writing – Original Draft:** Qian Wu and Xixi Fan; **Writing – Review & Editing:** Qian Wu, Xixi Fan, and Xin Qi; **Supervision:** Xin Qi. All authors reviewed and approved the final version of the manuscript and agreed to be accountable for their respective contributions.

Funding

This study received sponsorship from the Panorama Research Institute (Sponsorship No. PSG-PRI-SPN-2026-000016).

Conflicts of Interest

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

Ethics Approval and Informed Consent

This study is a narrative review based exclusively on previously published literature. It did not involve the recruitment of human participants, human experimentation, or the collection of new individual-level data. Therefore, ethics committee approval and informed consent were not required.

Data Availability

No new participant-level data were collected or generated during this study. All materials examined in this review were obtained from previously published literature, and the relevant sources are cited in the reference list.

Declaration of Generative AI Use

During the preparation of this manuscript, the authors used OpenAI ChatGPT solely for language editing and improving clarity of expression. All AI-assisted revisions were reviewed and verified by the authors, who take full responsibility for the accuracy, integrity, and final content of the manuscript.

Clinical Trial Registration

Not applicable. This study did not involve a clinical trial.

Preprint and Previous Dissemination

This manuscript has not previously been published or disseminated as a preprint or in any other formal publication.

References

1. Alam, Nadia Binte, Mohsin Surani, Chayon Kumar Das, Domenico Giacco, Swaran P. Singh, and Sagar Jilka. 2025. "Challenges and standardisation strategies for sensor-based data collection for digital phenotyping." *Communications Medicine* 5, no.1: Article 360. <https://doi.org/10.1038/s43856-025-01013-3>
2. Alam, Nadia Binte, Tahsinul Haque, Sanjana Subedar, Domenico Giacco, Swaran P. Singh, and Sagar Jilka. 2026. "Digital phenotyping for mental health conditions:a systematic review of implementation and application." *Frontiers in Digital Health* 8: Article 1772744. <https://doi.org/10.3389/fdgth.2026.1772744>
3. Aledavood, Talayah, Nguyen Luong, Ilya Baryshnikov, Richard Darst, Roope Heikkilä, Joel Holmén, Arsi Ikäheimonen, Annasofia Martikkala, Kirsi Riihimäki, Outi Saleva, Ana Maria Triana, and Erkki Isometsä. 2025. "Multimodal Digital Phenotyping Study in Patients With Major Depressive Episodes and Healthy Controls (Mobile Monitoring of Mood):Observational Longitudinal Study." *JMIR Mental Health* 12: Article e63622. <https://doi.org/10.2196/63622>
4. Barnett, Ian, John Torous, Patrick Staples, Luis Sandoval, Matcheri Keshavan, and Jukka-Pekka Onnela. 2018. "Relapse prediction in schizophrenia through digital phenotyping:a pilot study." *Neuropsychopharmacology* 43, no.8: 1660-1666. <https://doi.org/10.1038/s41386-018-0030-z>

5. Boettger, Silke, Dirk Hoyer, Katrin Falkenhahn, Martin Kaatz, Vikram K. Yeragani, and Karl-Jürgen Bär. 2006. "Altered diurnal autonomic variation and reduced vagal information flow in acute schizophrenia." *Clinical Neurophysiology* 117, no.12: 2715–2722. <https://doi.org/10.1016/j.clinph.2006.08.009>
6. Bosl, William J., Michelle Bosquet Enlow, and Charles A. Nelson. 2025. "A dynamical systems framework for precision psychiatry." *npj Digital Medicine* 8, no.1: Article 586. <https://doi.org/10.1038/s41746-025-01984-6>
7. Bufano, Pasquale, Marco Laurino, Sara Said, Alessandro Tognetti, and Danilo Menicucci. 2023. "Digital Phenotyping for Monitoring Mental Disorders: Systematic Review." *Journal of Medical Internet Research* 25: Article e46778. <https://doi.org/10.2196/46778>
8. Chekroud, Adam M., Matt Hawrilenko, Hieronimus Loho, Julia Bondar, Ralitza Gueorguieva, Alkomiet Hasan, Joseph Kambeitz, Philip R. Corlett, Nikolaos Koutsouleris, Harlan M. Krumholz, John H. Krystal, and Martin Paulus. 2024. "Illusory generalizability of clinical prediction models." *Science* 383, no.6679: 164–167. <https://doi.org/10.1126/science.adg8538>
9. Chouchou, Florian, and Martin Desseilles. 2014. "Heart rate variability: a tool to explore the sleeping brain?" *Frontiers in Neuroscience* 8: Article 402. <https://doi.org/10.3389/fnins.2014.00402>
10. Engel, George L. 1977. "The Need for a New Medical Model: A Challenge for Biomedicine." *Science* 196, no.4286: 129–136. <https://doi.org/10.1126/science.847460>
11. Fang, Shih-Shuan, and Sheng-Han Chen. 2026. "Digital phenotyping for predicting relapse in psychiatric disorders: a systematic review of passive sensing approaches." *BMC Psychiatry* 26, no.1: Article 507. <https://doi.org/10.1186/s12888-026-08157-z>
12. Fisher, Hadar, Nigel M. Jaffe, Habiballah Rahimi-Eichi, Erika E. Forbes, Diego A. Pizzagalli, Justin T. Baker, and Christian A. Webb. 2025. "Measuring activation during behavioral activation therapy: a proof-of-concept study using smartphone sensors and LLM-derived ratings in adolescents with anhedonia." *NPP—Digital Psychiatry and Neuroscience* 3: Article 24. <https://doi.org/10.1038/s44277-025-00045-w>
13. Fowler, J Corey, Nathan Cope, Jonathan Knights, Peter Phiri, Andrew Makin, Tim Peters-Strickland, and Shanaya Rathod. 2019. "Hummingbird Study: a study protocol for a multicentre exploratory trial to assess the acceptance and performance of a digital medicine system in adults with schizophrenia, schizoaffective disorder or first-episode psychosis." *BMJ Open* 9, no.6: Article e025952. <https://doi.org/10.1136/bmjopen-2018-025952>
14. Hassan, Lamiece, Alyssa Milton, Chelsea Sawyer, Alexander J Casson, John Torous, Alan Davies, Bernalyn Ruiz-Yu, and Joseph Firth. 2025. "Utility of Consumer-Grade Wearable Devices for Inferring Physical and Mental Health Outcomes in Severe Mental Illness: Systematic Review." *JMIR Mental Health* 12: Article e65143. <https://doi.org/10.2196/65143>
15. Hatch, Ainslie, Julia E Hoffman, Ruth Ross, and John P Docherty. 2018. "Expert Consensus Survey on Digital Health Tools for Patients With Serious Mental Illness: Optimizing for User Characteristics and User Support." *JMIR Mental Health* 5, no.2: Article e46. <https://doi.org/10.2196/mental.9777>
16. Kamath, Jayesh, Roberto Leon Barriera, Neha Jain, Efraim Keisari, and Bing Wang. 2022. "Digital phenotyping in depression diagnostics: Integrating psychiatric and engineering perspectives." *World Journal of Psychiatry* 12, no.3: 393–409. <https://doi.org/10.5498/wjp.v12.i3.393>
17. Laiou, Petroula, Dzmityr A Kaliukhovich, Amos A Folarin, Yatharth Ranjan, Zulqarnain Rashid, Pauline Conde, Callum Stewart, Shaoxiong Sun, Yuezhou Zhang, Faith Matcham, Alina Ivan, Grace Lavelle, Sara Siddi, Femke Lamers, Brenda WJH Penninx, Josep Maria Haro, Peter Annas, Nicholas Cummins, Srinivasan Vairavan, Nikolay V Manyakov, Vaibhav A Narayan, Richard JB Dobson, Matthew Hotopf, and RADAR-CNS. 2022. "The Association Between Home Stay and Symptom Severity in Major Depressive Disorder: Preliminary Findings From a Multicenter Observational Study Using Geolocation Data From Smartphones." *JMIR m Health and u Health* 1

- 0, no.1: Article e28095. <https://doi.org/10.2196/28095>
18. Lampe, Elizabeth W., Amanda C. Collins, Ahhyun Lee, Nicholas Enbar-Salo, Tess Z. Griffin, Arvind Pillai, Michael V. Heinz, Damien Lekkas, Matthew D. Nemesure, Daniel M. Mackin, Subigya Nepal, Andrew T. Campbell, and Nicholas C. Jacobson. 2026. "Interindividual differences in digital phenotypes of major depressive disorder:A passive sensing study using smartphone and wearable sensor data." *Behaviour Research and Therapy* 199: Article 104986. <https://doi.org/10.1016/j.brat.2026.104986>
19. Lim, Katherine, Rebekah Bodner, Kathryn Taylor Ledley, Shiwei Wang, Luke S Scheuer, and John Torous. 2026. "Integrating Data Visualizations Into Digital Mental Health Care for Adults With Anxiety and Depression:Participatory Design and Case Study." *Journal of Participatory Medicine* 18: Article e90255. <https://doi.org/10.2196/90255>
20. Machleid, Felix, Anette Schönewald, Esther Quinlivan, Linda Kokwaro, Louisa Schröder-Frekes, Toni Muffel, Caspar Wiegmann, and Jakob Kaminski. 2025. "Highly adaptable smartphone-based monitoring for patients with severe mental illness:Feasibility and usability study." *Digital Health* 11: Article 20552076251393283. <https://doi.org/10.1177/20552076251393283>
21. Martinez-Martin, Nicole, Thomas R. Insel, Paul Dagum, Henry T. Greely, and Mildred K. Cho. 2018. "Data mining for health:staking out the ethical territory of digital phenotyping." *npj Digital Medicine* 1: Article 68. <https://doi.org/10.1038/s41746-018-0075-8>
22. Meehan, Alan J., Stephanie J. Lewis, Seena Fazel, Paolo Fusar-Poli, Ewout W. Steyerberg, Daniel Stahl, and Andrea Danese. 2022. "Clinical prediction models in psychiatry: a systematic review of two decades of progress and challenges." *Molecular Psychiatry* 27, no.6: 2700–2708. <https://doi.org/10.1038/s41380-022-01528-4>
23. Monroe, Scott M., and Kate L. Harkness. 2011. "Recurrence in major depression:A conceptual analysis." *Psychological Review* 118, no.4: 655–674. <https://doi.org/10.1037/a0025190>
24. Mow, Jessica L., David E. Gard, Kim T. Mueser, Jasmine Mote, Kathryn Gill, Lawrence Leung, Tairmae Kangarloo, and Daniel Fulford. 2022. "Smartphone-based mobility metrics capture daily social motivation and behavior in schizophrenia." *Schizophrenia Research* 250: 13–21. <https://doi.org/10.1016/j.schres.2022.09.025>
25. Oakey-Neate, Lydia, Geoff Schrader, Jörg Strobel, Tarun Bastiampillai, Yasmin van Kasteren, and Niranjan Bidargaddi. 2020. "Using algorithms to initiate needs-based interventions for people on antipsychotic medication: implementation protocol." *BMJ Health Care Inform* 27, no.1: Article e100084. <https://doi.org/10.1136/bmjhci-2019-100084>
26. Palmer, Barton W., Gauri N. Savla, Scott C. Roesch, and Dilip V. Jeste. 2018. "Changes in capacity to consent over time in patients involved in psychiatric research." *British Journal of Psychiatry* 202, no.6: 454–458. <https://doi.org/10.1192/bjp.bp.112.121160>
27. Robinson, Delbert, Margaret G. Woerner, Jose Ma. J. Alvir, Robert Bilder, Robert Goldman, Stephen Geisler, Amy Koreen, Brian Sheitman, Miranda Chakos, David Mayerhoff, and Jeffrey A. Lieberman. 1999. "Predictors of Relapse Following Response From a First Episode of Schizophrenia or Schizoaffective Disorder." *Archives of General Psychiatry* 56, no.3: 241–247. <https://doi.org/10.1001/archpsyc.56.3.241>
28. Schmaal, Lianne. 2023. "The curse and opportunity of heterogeneity in the pursuit of psychiatric biomarkers." *World Psychiatry* 22, no.2: 271–272. <https://doi.org/10.1002/wps.21085>
29. Shen, ShiYing, Wenhao Qi, Jianwen Zeng, Sixie Li, Xin Liu, Xiaohong Zhu, Chaoqun Dong, Bin Wang, Yankai Shi, Jiani Yao, Bingsheng Wang, Xiajing Lou, Simin Gu, Pan Li, Jinghua Wang, Guowei Jiang, and Shihua Cao. 2025. "Passive Sensing for Mental Health Monitoring Using Machine Learning With Wearables and Smartphones:Scoping Review." *Journal of Medical Internet Research* 27: Article e77066. <https://doi.org/10.2196/77066>
30. Tomićić, Ana, Anamaria Malešević, and Anto Čartolovni. 2021. "Ethical,Legal and Social Issues of Digital Phenotyping as a Future Solution for Present-Day Challenges:A Scoping Review."

- Science and Engineering Ethics* 28, no.1: Article 1. <https://doi.org/10.1007/s11948-021-00354-1>
31. Tonon, Andre C., Adile Nexha, Jasmyn E. A. Cunningham, Jason d'Eon, Trisha Chakrabarty, Faranak Farzan, Jane A. Foster, Kate L. Harkness, Stefanie Hassel, Keith Ho, Raymond W. Lam, Roumen Milev, Luciano Minuzzi, Daniel J. Müller, Abraham Nunes, Sagar V. Parikh, Lena C. Quilty, Susan Rotzinger, Claudio N. Soares, Valerie H. Taylor, Gustavo Turecki, Rudolf Uher, Sidney H. Kennedy, and Benicio N. Frey. 2026. "One-Year Actigraphy Study of Sleep and Rest-Activity Rhythms as Markers of Relapse in Depression." *JAMA Psychiatry* 83, no.4: 379-388. <https://doi.org/10.1001/jamapsychiatry.2025.4453>
32. Torous, John, Mathew V Kiang, Jeanette Lorme, and Jukka-Pekka Onnela. 2016. "New Tools for New Research in Psychiatry:A Scalable and Customizable Platform to Empower Data Driven Smartphone Research." *JMIR Mental Health* 3, no.2: Article e16. <https://doi.org/10.2196/mental.5165>
33. Vecchio, Ivan, Lucas Mifsud, Sofia Castro E Almeida, and Johannes Passecker. 2025. "Diagnostic digital phenotyping in schizophrenia-spectrum disorders:a systematic review." *npj Digital Medicine* 9, no.1: Article 16. <https://doi.org/10.1038/s41746-025-02194-w>
34. Zaher, Farida, Mariama Diallo, Amélie M. Achim, Ridha Joober, Marc-André Roy, Marie-France Demers, Priya Subramanian, Katie M. Lavigne, Martin Lepage, Daniela Gonzalez, Irnes Zeljkovic, Kristin Davis, Michael Mackinley, Priyadharshini Sabesan, Shalini Lal, Alban Voppel, and Lena Palaniyappan. 2024. "Speech markers to predict and prevent recurrent episodes of psychosis:A narrative overview and emerging opportunities." *Schizophrenia Research* 266: 205-215. <https://doi.org/10.1016/j.schres.2024.02.036>
35. Zlatintsi, Athanasia, Panagiotis P. Filntisis, Christos Garoufis, Niki Efthymiou, Petros Maragos, Andreas Menychtas, Ilias Maglogiannis, Panayiotis Tsanakas, Thomas Sounapoglou, Emmanouil Kalisperakis, Thomas Karantinos, Marina Lazaridi, Vasiliki Garyfalli, Asimakis Mantas, Leonidas Mantonakis, and Nikolaos Smyrnis. 2022. "E-Prevention: Advanced Support System for Monitoring and Relapse Prevention in Patients with Psychotic Disorders Analyzing Long-Term Multimodal Data from Wearables and Video Captures." *Sensors* 22, no.19: Article 7544. <https://doi.org/10.3390/s22197544>

Tables

Table 1. Representative studies and the clinical questions addressed

Study	Population	Study design	Clinical question and contribution	Evidence level
Barnett et al. (2018)	Schizophrenia	Pilot longitudinal smart-phone study	Can longitudinal phone-derived behaviour help identify impending relapse?	Exploratory relapse prediction
Laiou et al. (2022)	Major depressive disorder	Multicentre observational geolocation study	How does time at home relate to contemporaneous depressive symptoms?	Symptom association
Mow et al. (2022)	Schizophrenia	Smartphone mobility and daily behaviour	How do mobility patterns relate to social motivation and behaviour?	Behavioural association
Tonon et al. (2026)	Remitted major depressive disorder	Prospective observational actigraphy cohort	How do sleep and rest-activity patterns relate to subsequent relapse?	Longitudinal prognosis
Vecchio et al. (2025)	Schizophrenia spectrum	Systematic review	What has digital phenotyping measured across the diagnostic spectrum?	Evidence synthesis
Fang and Chen (2026)	Mixed psychiatric disorders	Systematic review of passive sensing	What does the literature report about passive-sensing relapse models?	Cross-diagnostic synthesis

Note: The examples map distinct evidence types; model discrimination across diagnostic groups is not an estimate for either disorder individually.