Protocol

Using Smartphone-Based Digital Phenotyping to Predict Relapse in Serious Mental Disorders Among Slum Residents in Dhaka, Bangladesh: Protocol for a Machine Learning Study

Nadia Alam[1], MSc; Chayon Kumar Das[2], MPhil; Neelabja Roy[3], MS; Domenico Giacco[2,4], MBBS, PhD; Swaran P Singh[1,4], MBBS, MD; Sagar Jilka[1], PhD

[1]Warwick Medical School, University of Warwick, Coventry, United Kingdom
[2]Department of Clinical Psychology, University of Dhaka, Dhaka, Bangladesh
[3]Department of Psychiatry, National Institute of Mental Health and Neurosciences, Bangalore, India
[4]Coventry and Warwickshire Partnership NHS Trust, Coventry, United Kingdom

Corresponding Author:
Nadia Alam, MSc
Warwick Medical School
University of Warwick
B 020, Gibbet Hill Road
Coventry CV4 7AL
United Kingdom
Phone: 07300311294
Email: nadia.alam@warwick.ac.uk

Abstract

Background: Serious mental illnesses (SMIs) are associated with high relapse rates and limited access to continuous care, particularly in low-resource settings such as urban slums. Traditional clinical monitoring is constrained by accessibility and scalability challenges. Digital phenotyping, through passive smartphone data, offers a novel approach to predict relapse by capturing real-world behavioral changes.

Objective: This study aims to evaluate the feasibility and predictive value of smartphone-based digital phenotyping for detecting relapse in individuals with SMIs living in the Korail slum of Dhaka, Bangladesh.

Methods: This prospective 6-month cohort study will recruit 430 participants diagnosed with SMIs who own Android (Google LLC) smartphones. Passive data (eg, screen time, mobility, and call or text frequency) will be continuously collected using a custom-built app (DataDoc). Monthly active data, including symptom and functioning assessments, will be collected via self-report and clinical engagement. Machine learning models will integrate these data to detect early warning signs and predict relapse trajectories.

Results: This study was funded by the NIHR (National Institute for Health and Care Research; award number NIHR200846) in October 2022. Data collection commenced in August 2025 and is ongoing. A total of 14 participants have been recruited, as of January 2026. Preliminary data analysis is ongoing, with expected results to be published in fall 2026.

Conclusions: This study is one of the first to apply smartphone-based digital phenotyping and machine learning for relapse prediction in low- and middle-income countries' slum settings. The findings will inform scalable, low-cost digital interventions to address the mental health treatment gap in underresourced communities.

International Registered Report Identifier (IRRID): PRR1-10.2196/79826

JMIR Res Protoc 2026;15:e79826; doi: 10.2196/79826

Keywords: mental health; digital phenotyping; passive sensing; low-income; urban settings; mobile phone

Introduction

Background

Poor outcomes for patients with serious mental illnesses (SMI) are associated with various factors involving patients, health care providers, and the health care system, in addition to the delivery of inadequate medical services [1-3]. The current clinical ways used to track the progression of SMI may be insufficient or inadequate. Traditionally, health care professionals rely on in-person consultations for assessment and diagnosis [4,5]. However, these clinic-based services come with a variety of logistical difficulties [4], especially for people from impoverished communities [6]. To monitor patients, they must frequently visit a clinical facility, often within restricted operating hours, which can be particularly challenging for individuals dealing with SMI [4]. Additionally, these approaches are demanding in terms of resources because they necessitate one-on-one interactions with a trained health care provider, making widespread adoption difficult. The inherent issues related to accessibility and scalability in these methods create significant obstacles to patient care [7,8].

Remotely collected digital markers may help overcome key challenges in monitoring SMI, such as the reliance on traveling for monitoring and in-person assessments [9, 10]. Continuous monitoring of behavioral and physiological data using personal devices has the potential to aid in the identification of personalized indicators for the initiation of illness [4], subsequently enabling the development of tailored treatment strategies [9,11,12]. Smartphones can provide a solution.

Digital Phenotyping (DP)

Smartphones can be used to collect data to understand mental health status. This technique, called digital phenotyping (DP) [13], is already in use in some behavioral problems and mental disorders, such as posttraumatic stress disorder (PTSD) [14], and SMI such as schizophrenia [15-17] and mood disorders [18-20] to predict relapse, symptom exacerbation, or mood fluctuations [21].

Digital phenotypes are collected either actively or passively [5,22]. Active data collection requires the individual's active participation through active engagement [23]. This would include self-reported thoughts and symptoms through surveys or scales, whereas passive data collection does not require an individual to perform any specific action outside their regular activities [24]. Passively collected data, such as screen behavior, including screen time, locking or unlocking events, and screen lighting has been shown to reflect mental health states in adults, offering insights into phases of agitation, depression, and variations in sleep quality and patterns [10,23,24]. Accelerometer data from smartphones has also been linked to mental health concerns such as depression and anxiety due to reduced activity, such as mobility and traveling, while audio data, such as read-aloud recordings, help indicate depressive states [23,25]. Within DP, sleep-related features, such as sleep

duration, variability, latency, and nighttime phone use, are increasingly recognized as important phenotypes that can reflect underlying mental states. Disturbed sleep, as captured through these DPs, has been linked to elevated stress and SMI [26]. In bipolar disorder, disruptions in sleep patterns can precipitate manic or depressive episodes and persist even during remission, affecting overall functioning [27-29]. Major depressive disorder (MDD) often features insomnia or hypersomnia, with sleep issues potentially preceding depressive episodes and exacerbating their severity [30]. In schizophrenia, sleep abnormalities frequently emerge before psychotic symptoms and are linked to more severe clinical outcomes [30]. Geolocation data offer insights into individuals' mobility patterns and daily routines [31]. In schizophrenia, reduced GPS-derived mobility, such as spending more time at home and traveling shorter distances, is associated with increased severity of negative symptoms such as social withdrawal and diminished motivation [32]. Similarly, in bipolar disorder, fluctuations in geolocation patterns can reflect mood episodes, with decreased movement often observed during depressive phases [33].

DP has gained significant traction in high-income countries (HICs), where studies have demonstrated its utility in detecting early warning signs (EWS), monitoring symptom fluctuations, and predicting relapse in individuals with SMIs [22,34-37]. For example, in a US-based study, passive sensing data, including GPS location, accelerometer activity, screen use, and call or text logs, was used to identify EWS of psychotic relapse. Machine learning (ML) techniques were applied to detect behavioral anomalies preceding relapse events, demonstrating the potential of DPs for relapse prediction in schizophrenia [22]. A narrative review demonstrated how various digital features, such as geolocation, communication patterns, and activity levels, have been effectively used in schizophrenia research to understand behavioral changes, and studies included were conducted in HICs [16]. In mood disorders, passive data from smartphones and wearable devices, such as GPS location, accelerometer activity, sleep pattern data, and heart rate, were used to estimate depressive and anxiety symptom severity [18,19].

Together, these passively collected DP, such as sleep, mobility, screen use, and activity patterns, offer a continuous, unobtrusive window into individuals' lived experiences. Their integration into mental health monitoring holds promise for early detection of symptom changes, relapse prediction, and timelier, personalized interventions in the care of individuals with SMI.

Relapse

Relapse is a significant concern in the management of SMIs. Relapse can lead to adverse outcomes, including poor psychosocial functioning, increased caregiver burden, and higher health care costs [38-40]. Therefore, continuous monitoring and adherence to treatment are crucial in mitigating the risk of relapse and improving long-term outcomes for individuals with SMIs. Most patients go through a period with changes in behavior, which precedes their psychotic relapse, commonly known as EWS [41]. EWS

includes changes in sleep patterns, hallucinations, delusions, hostile behavior, cognitive decline, depression, and paranoia [42]. Developing systems that can identify and monitor these EWS could help clinicians intervene early [41]. In a study, anomalies in passive data, such as geolocation, accelerometer readings, and screen state, were 2.12 times more frequent in the month preceding a relapse compared to nonrelapse periods [40,43]. In another study with participants with schizophrenia, the analysis revealed that anomalies in passive data (geolocation, accelerometer data, and screen state) were significantly more frequent before a relapse [43]. Importantly, models incorporating passive data outperformed those relying solely on active survey data in predicting relapse [43]. These findings show the potential of passive smartphone data to serve as early digital markers of relapse, offering a scalable and nonintrusive method for the timely intervention and continuous monitoring in individuals with SMI.

Low- and Middle-Income Countries (LMICs)

As described above, most of the DP research has been conducted in HICs, raising concerns about generalizability to low- and middle-income countries (LMICs) [44], where digital usage patterns, social determinants of health, and mental health care infrastructure differ substantially [42]. This motivates the need to evaluate whether DP approaches can be adapted and applied effectively in LMIC settings such as Bangladesh, where the burden of untreated SMIs is high and innovative, low-cost monitoring solutions are urgently needed [45,46].

Mental and behavioral disorders account for 12% of the global disease burden, with over 70% of this impact affecting LMICs [47]. In Bangladesh, the treatment gap (ie, the gap between people who need care and the people who obtain care) is over 92% [47], which means that less than 1 in 10 people get the mental health care they need, and the gap is more evident in slums [6,45,47]. People living in slum communities have high rates of SMIs, limited access to mental health services, and conditions of chronic hardship [48]. Therefore, a low-cost, easy-to-use solution to support the identification and monitoring of mental disorders is needed, and DP may provide a solution. DP has paved its way in mental health research [9-11], but most of the research conducted has been in high-income or developed countries; these do not account for findings in LMIC settings. Yet, 80% of people with mental disorders live in LMICs [42]. Social factors such as sudden changes in life, urbanization, and poverty often lead to a high burden of mental illness in several LMICs [42,49-51]. Additionally, there is a significant lack of awareness regarding mental health conditions among LMICs [49]. Slum dwellers here have a high burden of mental health conditions, which remains unnoticed due to a lack of help-seeking or lack of resources [48,49]. The addition of DP will enhance the detection of mental disorders among these communities with little to no cost, as the use of smart devices such as smartphones is common among slum dwellers in Dhaka.

Aims and Objectives

Overview

The proposed study aims to determine whether DPs can effectively predict relapse among residents from the Korail Slum, Dhaka. One of the key outputs of this study will be to provide evidence of the feasibility and reliability of using DP in LMIC settings, particularly slum settings, for more effective mental health monitoring and intervention [47].

Primary Objective

We aimed to develop and validate predictive models using smartphone-based DP data to detect and forecast relapse in individuals with serious mental disorders living in the Korail Slum.

Secondary Objectives

For secondary objectives, our aims were:

- To explore associations between DPs and clinical facets such as mood states, social withdrawal or functioning, sleep disturbances, and activity levels.
- To assess the feasibility, acceptability, and reliability of smartphone-based DP among slum residents in a low-resource setting.
- To identify contextual and user-level factors (eg, socioeconomic, environmental, and technological) that influence data completeness and engagement with DP tools.

Methods

Outcome: Relapse

The primary outcome for prediction modeling will be the occurrence of a relapse event within a 6-month follow-up period defined using prespecified, quantitative, diagnosis-specific criteria based on validated clinical scales, supplemented by structured clinician adjudication. For participants with psychotic disorders, relapse will be defined as either (1) a ≥25% increase from baseline on the Positive and Negative Syndrome Scale (PANSS), sustained for at least 7 days, or (2) deterioration in global functioning as indicated by a decline of ≥10 points on the Global Assessment of Functioning (GAF), confirmed by clinician rating. These thresholds align with relapse definitions used in prior studies in psychosis, where relapse was operationalized using scale-based symptom worsening and functional decline [17,22,43]. For participants with nonpsychotic disorders, including MDD and PTSD, relapse will be defined as clinically significant symptom worsening on the Brief Symptom Inventory (BSI), operationalized as a ≥30% increase from baseline on the relevant symptom subscale and/or meeting caseness on the BSI Global Severity Index (GSI), defined as a total GSI T-score ≥63 together with an increase of at least 7 points from the mean of the 2 preceding assessment scores, indicating a move into or further into the clinically significant distress range [52]. In repeated-measurement and longitudinal studies involving patients with mood, anxiety, and trauma-related disorders,

changes in BSI symptom subscales and the GSI have been widely used to capture clinically meaningful worsening over time [52-55]. In these populations, sustained increases from an individual's own baseline, rather than isolated score fluctuations, are commonly interpreted as reflecting true symptom deterioration. To reduce the risk of misclassifying temporary or situational symptom changes as relapse, clinically meaningful worsening is typically operationalized using relative change thresholds or standardized deviations from baseline sustained across multiple assessments.

Across all diagnostic groups, any psychiatric hospital admission or emergency psychiatric presentation during follow-up will be classified as a relapse event. All algorithm-identified relapse events will undergo blind clinician adjudication by an independent clinician using only symptom scale trajectories and clinical information; the clinician will be blinded to all DP features. Clinician judgment may confirm or refute events meeting quantitative thresholds but will not introduce relapse events in the absence of predefined criteria, thereby preventing post hoc outcome redefinition. Each relapse will be timestamped at the earliest point at which the criteria are met, enabling prediction within predefined risk horizons (eg, relapse within 30 d).

Study Design

This study is a 6-month prospective cohort study designed to explore the feasibility and reliability of identifying DP to predict relapse of SMIs among slum residents in Dhaka, Bangladesh. This study is part of the broader NIHR (National Institute for Health and Care Research) initiative, Transforming Access to Care for Serious Mental Disorders in Slums (TRANSFORM project) [47], which provides the overarching framework. The TRANSFORM project is dedicated to improving access to mental health care for individuals with SMIs residing in urban slum environments [47]. The TRANSFORM study aims to bridge the treatment gap by developing, implementing, and assessing innovative community-based interventions [47].

Setting

Participants will be recruited from the Korail slum, one of the largest and most densely populated slums in Dhaka [47]. It is situated between 2 affluent areas and is home to an estimated 200,000 people living in densely populated conditions with limited access to basic services [47,51,56, 57]. The community consists of a mix of long-term residents and recent migrants from rural areas, and many inhabitants face precarious employment, housing instability, and health inequities [56,57]. Mental health needs are high, yet access to formal mental health services remains scarce [47].

Participants and Recruitment

Participants will be selected using purposive sampling. Participants will be identified through the TRANSFORM project [47], where community engagement representatives are well known and respected in the area. A community-based approach will be used to identify and enroll participants, leveraging existing partnerships with local stakeholders. We collaborated with community health workers, nongovernmental organizations, and TRANSFORM [47] community representatives, who had well-established relationships with the residents. These trusted individuals facilitated introductions and provided culturally appropriate explanations of this study to potential participants. Importantly, recruitment will not be limited to individuals already enrolled in TRANSFORM. Community representatives will help identify potential participants, introduce this study, provide culturally appropriate explanations, and facilitate trust with potential participants. Participants will progress through clearly defined stages: approached and screened for eligibility, enrolled if inclusion criteria are met, followed up monthly with both active and passive data collection, and retained in the analytic sample provided sufficient data completeness is achieved. Numbers at each stage will be reported to ensure transparency.

This study will recruit participants diagnosed with SMIs, including MDD, psychotic disorders (such as schizophrenia and bipolar disorder), and PTSD, as per the *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*) [58]. These conditions were selected due to their high burden in low-resource settings, their recurrent nature, and the critical need for early detection of symptom exacerbation [42,59-61]. Inclusion criteria require participants to be at least 18 years of age, residents of the Korail slum, diagnosed with SMI, and in possession of an Android smartphone. Exclusion criteria include those younger than 18 years of age, those without a smartphone, and individuals diagnosed with substance-induced psychosis or mental illness.

Sample Size

Total Enrollment

Based on findings from the National Mental Health Survey (2019), ≈17% of adults in Bangladesh are estimated to experience a mental health disorder [62]. While specific prevalence rates of SMIs in Korail, Dhaka, are not available, this national estimate provides a useful contextual baseline. Population estimates for Korail vary widely, from 50,000 to 2,00,000 individuals, but for planning purposes, a conservative midpoint estimate of 100,000 residents was used [47,51, 57].

To determine an appropriate sample size for estimating a proportion with maximum variability, the standard formula with finite population correction was applied [63-65]:

$$\text{Sample size} = \frac{\frac{z^2 \times p(1-p)}{e^2}}{1 + \left(\frac{z^2 \times p(1-p)}{e^2 N}\right)}$$

Where Z=1.96 (for a 95% confidence level), *P*=0.5 (maximum variability), e=0.05 (margin of error), and n=100,000 (estimated Korail population). Substituting these parameters yields n of ≈384.16. To account for potential nonresponse or incomplete data, a 90% response rate was assumed:

$$n_{\text{adjusted}} = 384.16/0.90 - 426.84$$
$$n_{\text{final}} = 426.84 \approx 430$$

Prediction Model Adequacy (Relapse Events)

To power the prediction model, we planned recruitment of 430 participants and assumed a 15% attrition rate (N is ≈366) over 6 months. The primary justification for the sample size is based on the widely accepted events per predictor (EPP) effective degree of freedom (EDF) principle for prediction model development, which mandates a minimum of 10 events per effective predictor.

Based on previous literature, relapse incidences range between 25%-40% for psychosis [66,67], 15%-30% for MDD [68], and 10%-25% for PTSD [69]. Using these rates, we estimated the expected number of relapse events for our planned cohort. To account for diagnostic heterogeneity within the sample, we used existing diagnostic compositions from the TRANSFORM study [70]: (1) 45% psychosis/42% MDD/20% PTSD) – expected ≈ 66-121 relapse events across the incidence ranges, (2) psychosis: 0.40×366 × 0.25-0.40=37-59 events, (3) MDD: 0.40×366 × 0.15-0.30=22-44 events, and (4) PTSD: 0.20×366 × 0.10-0.25=7-18 events.

Concerning anticipated events (E), after accounting for a 15% attrition rate and the calculated relapse incidence ranges (66-121 events), the minimum number of events (E) is 66 relapses.

Of the required EDF, given a minimum of E=66, the maximum number of effective predictors (EDF) that the model can reliably support while maintaining the EPP>10 [71] ratio is 6.6 predictors.

About model strategy alignment, our study involves high-dimensional data (numerous digital phenotypes). To ensure compliance with the estimated EDF (≈6-7 predictors), the modeling pipeline will reduce the large number of raw features into a smaller set of effective predictors through penalized modeling (eg, LASSO [least absolute shrinkage and selection operator], ridge regression, or regularized tree-based methods) and dimensionality reduction or feature aggregation techniques.

Procedure

Participants will be informed about this study through participant information sheets (PIS) and will be given 48 hours to provide written informed consent. The PIS will be provided in Bengali, the local language, and will clearly explain the purpose of this study, the type of data to be collected, what participation involves, risks and benefits, data security, and the right to withdraw at any time without affecting access to services. Participants will be contacted through community engagement representatives and ongoing recruitment activities under the TRANSFORM project [47]. These representatives, who have established relationships within the community, will distribute the PIS and introduce this study to potential participants. Trained research assistants will obtain written informed consent at the local field office with a witness present from participants willing to take part in this study. This will be done in a private setting to ensure confidentiality, and the information sheet will be explained in Bengali. After obtaining consent, the DataDoc app will be installed on participants' smartphones to passively collect data. The research team will assist them in downloading the app, granting required permissions (eg, access to device usage), setting up the app to ensure proper functionality, and troubleshooting any technical issues. Participants will be trained on app installation and basic use. Monthly clinical assessments will be conducted either in-person or via telephone, while passive data will be captured continuously in the background. Technical support will be available through community fieldworkers to address issues such as app functioning, battery management, or data sync.

Participants will install DataDoc on their smartphones for data collection as outlined below. This data will be collected via the app automatically and passively once the user gives the app permissions. A link to the app will be sent to participants via email or text (based on their preference). This is an .apk file which they can install on their Android phone. We will support participants to do this (eg, in person in the Korail field office) if needed. Upon downloading the app, they will be asked to give the app permission to access their device usage. Please note, the app also has the potential to connect to music and health data, for example, music data (ie, permission to access their Spotify [Spotify AB] data) and health data (ie, permission to access their Google Health app data), but we will only be using app usage for this study.

Relapse events will be identified using a combined approach. First, thresholds on validated scales (active data) will trigger a potential relapse flag. These events will then be reviewed by 2 independent clinicians and a researcher with expertise in psychiatry and psychology. The researcher will be blinded to the sensor-based DP data and will adjudicate relapse status using only clinical and scale-based information. In case of disagreement, a third senior clinician will review the case to reach a consensus.

DataDoc App

We have developed an app, DataDoc (Figure 1), to collect DP data, which we will use for data collection in this study. Some active data will be collected directly through the DataDoc app, while assessments requiring clinician input will be obtained separately. After giving permissions, the app will sync this data, and then the participant will be presented with a screen asking them to complete the questionnaires outlined below in the active data section. Participants will be asked to complete the questionnaire via the app only. The questionnaire must be completed within 2 weeks of syncing the device, music, and health data. Participants will be called with reminders if they have not completed the questionnaire after 1 week and again after 2 weeks. At this point, we will ask them to resync their data and complete the questionnaire, so all their data is synced to the same time point.

Figure 1. The DataDoc app.
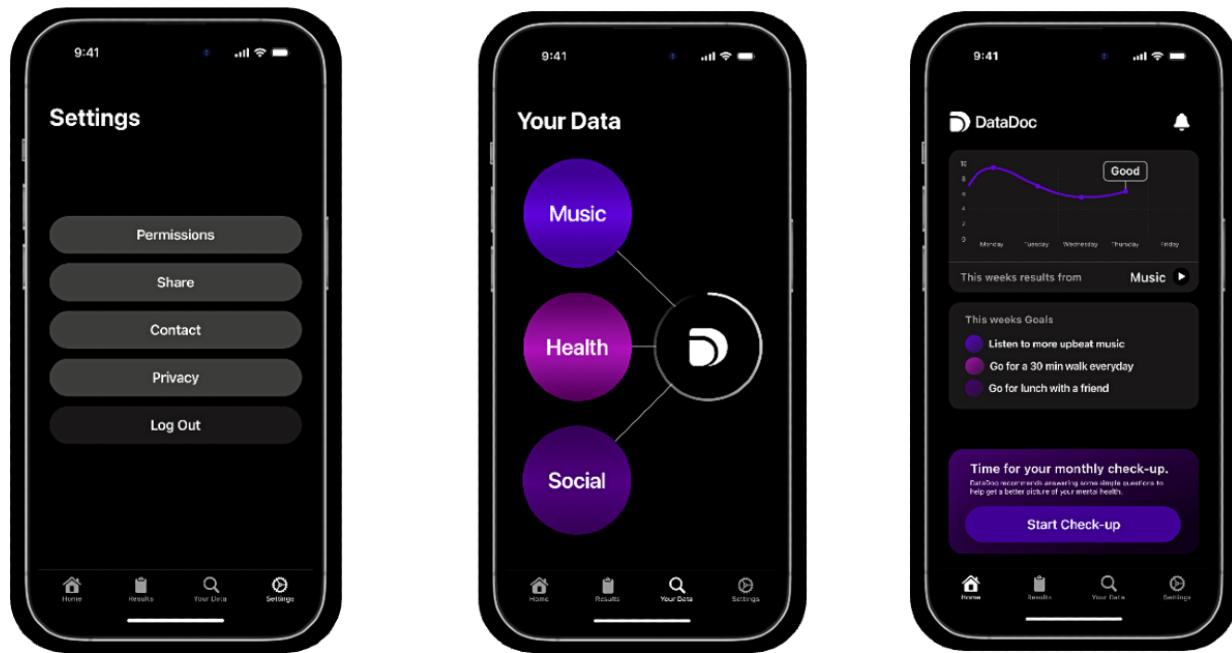
Metrics

Participants will complete a set of baseline assessments and assessments at intervals as shown in Table 1.

Table 1. Timeline of data collection.

Domain, subdomains, and tools	Assessment time points							
	Baseline	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Continuously
Active data								
Sociodemographic profile	✓							
Clinical profile and history								
Diagnosis	✓							
Timeline of diagnosis	✓							
Clinical assessments								
Global Assessment of Functioning (GAF)	✓	✓	✓	✓	✓	✓	✓	
WHOQOL (World Health Organization Quality of Life)	✓	✓	✓	✓	✓	✓	✓	
Brief Symptom Inventory (BSI) or the Positive and Negative Syndrome Scale (PANSS)	✓	✓	✓	✓	✓	✓	✓	
Perceived Stress Scale	✓	✓	✓	✓	✓	✓	✓	
Sleep questionnaire	✓	✓	✓	✓	✓	✓	✓	
Thought patterns								
Journaling (optional)	✓	✓	✓	✓	✓	✓	✓	
Passive data								
Device interactions								
Phone usage								✓
Application usage								✓
Screen time								✓

Domain, subdomains, and tools	Assessment time points							
	Baseline	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Continuously
Digital interaction								
Calls (frequency)								✓
Duration of calls								✓
Messages (frequency)								✓
Activities (if applicable)								
Activities app (sleep and health)								✓
Sleep								✓
Movement and location								
Geolocation (longitude and latitude)								✓
Duration at location								✓
Movement time								✓

Passive Data

Passive data will be collected continuously, on an event-by-event basis, and will include:

- Device interactions: this study will collect data on how a participant interacts with their digital devices. This includes phone usage, application usage (application history such as application on the foreground, number of application notifications, when an application is installed and removed), and screen time.
- Digital interaction: calls and messaging habits (number of incoming and outgoing messages and calls), frequencies of responses (duration of call; the content of communications will not be accessed or stored).
- Activities (if applicable): sleep and health application installed or removed, and total sleep duration (integrated within DataDoc app).
- Movement and location: geolocation (longitude and latitude), duration at one location, and movement time.

Active Data

Demographic Questionnaire

Each participant will be provided with a demographic questionnaire containing questions regarding age, gender, level of education, occupation, diagnosis profile, and home address.

Active data will consist of the following and will be collected monthly.

Clinical Data

The clinical data consisted of the following:

- GAF is a clinician-rated scale to assess global functioning. GAF represents a latent continuous measure score from 1 (lowest score representing worst symptomology) to 100 (high score representing optimal level of functioning) [72]. This data will be collected from the ongoing TRANSFORM Project [47].

- WHOQOL-BREF (World Health Organization Quality of Life-Brief) is a 26-item self-reported scale to assess quality of life, where scores range from 16 to 112, and a higher score represents a higher quality of life [73].
- BSI is a 53-item self-reported instrument on a 5-point scale used to evaluate psychological symptoms over the past week in patients, nonpatients, and participants involved in experimental research [74]. The BSI consists of 9 main symptom categories: somatization, obsessive-compulsive tendencies, sensitivity in relationships, feelings of depression, anxiety, hostility, phobic anxiety, paranoid thoughts, and tendencies toward psychosis, and higher scores represent more severe symptoms [74].
- The PANSS is a scale used for measuring symptom severity with 3 components of positive, negative, and general psychopathology scales [75]. The total PANSS score is the sum of the scores of the 3 scales, ranging from a minimum of 30 to a maximum of 210, with higher scores indicating more severe symptoms [75].
- Perceived Stress Scale [76] to assess perception of stress with scores ranging from 0 to 40, where higher scores represent higher perceived stress.
- The Nottingham Onset Schedule is a structured interview tool designed to assess the onset of psychosis, and it provides a systematic method of gathering detailed information about the initial presentation and progression of symptoms [77].

Activities

The activities consisted of the following:

- Self-reported sleep patterns through a dedicated general questionnaire for this study derived from the sleep questionnaire for adults (part two) by the National Health Service, which will explore the time of the participant going to bed, hours of sleep, number of times the participant woke up, perceived quality of sleep, etc [78].

- For journaling, participants will be asked to "journal," that is, write down their thoughts and feelings if they wish [79].

Data Privacy and Security

As illustrated in Figure 2, data is first collected on the participant's phone using the DataDoc app. Data are then securely uploaded to a General Data Protection Regulation–compliant cloud server, where they are stored in encrypted form at rest. From there, encrypted anonymized data are transferred to the University of Warwick's secure server for long-term storage and analysis, in line with Warwick's 10-year research data policy. No personally identifiable information (eg, phone numbers and message content) is collected at any stage. Access to both the cloud server and Warwick's secure server is restricted to authorized study personnel using password-protected accounts.

Figure 2. Data flow.
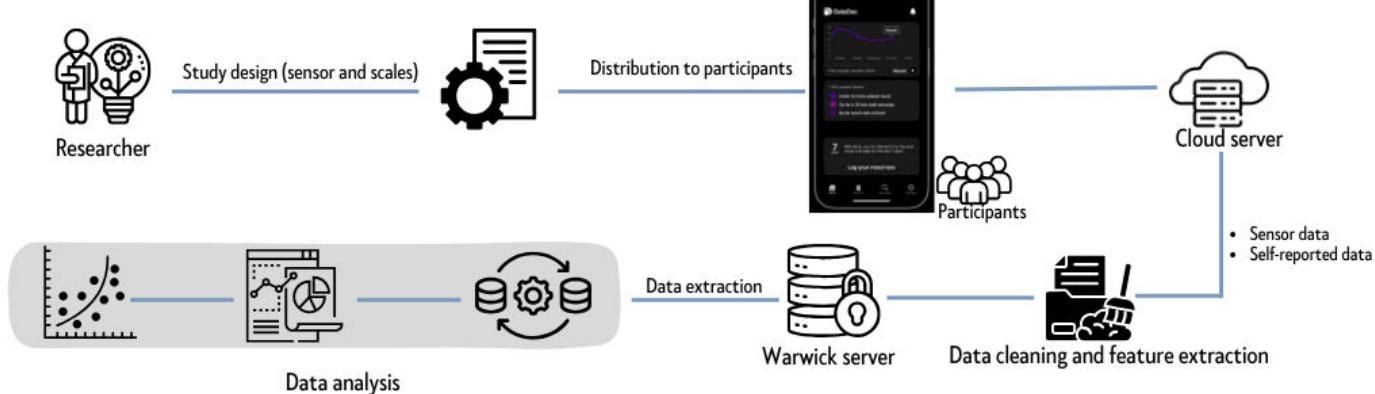
Analysis

Overview

This study adopts a multimodal ML-driven analytical framework to integrate and analyze both active and passive data streams collected via the DataDoc app. The goal is to develop predictive models for relapse detection and to explore associations between specific digital phenotypes and clinical dimensions. The analysis pipeline will involve preprocessing of the collected active and passive data, feature engineering, labeling (when needed), ML model selection, model fitting, and evaluating the predictive validity of our multimodal data, followed by further interpretation and deployment when needed.

Analysis of Active Data

Data Preprocessing

Cleaning and Standardization

We will remove any inconsistencies and standardize the format of the active data to ensure uniformity. For demographic questionnaire variables such as gender, occupation, or diagnosis, any typographical errors will be corrected, and if any absurd numerical value is recorded for age or education, it will be flagged for replacement. For clinical assessment instruments, it must be ensured that item responses are within the appropriate range of the corresponding scales. If not, the out-of-range and/or strange entries will be flagged for further systematic treatments.

Handling Missing Values

Missing responses will be addressed through multiple imputation techniques or median replacement where feasible, and guided by manuals as necessary. For demographic variables, if there are randomly missing variables, then mean or median imputation for numerical variables and mode for categorical variables may be considered. In cases of substantial missingness, a separate category such as "unknown" may be assigned. A threshold for what constitutes "substantial" will be prespecified during analysis planning. For the clinical assessment scales, if there is item-level missingness, then the person's mean imputation will be used. If there is any block-level missingness that affects entire subscales, then the participant would be flagged and dropped if needed.

Encoding Categorical Variables

Categorical variables, such as demographic data like gender, will be converted into numerical representations suitable for computational steps and ML models. Variables such as education level can be mapped to ascending integers, and nominal ones such as occupation can be encoded with 1-hot or label encoding.

Feature Engineering

Extraction of Relevant Features

We will identify and extract features from active data that are indicative of behavioral and psychological factors associated with relapse. Demographic variables such as occupation or education level can be grouped into a factor such as socioeconomic status if deemed necessary for future analysis. From the responses on the clinical assessment instruments such as

GAF and others, scoring will be done as per their respective manuals after ensuring all responses are in their appropriate ranges. Additionally, for scales such as WHOQOL-BREF, where specific items such as items 3 and 4 are also reverse-worded, they will be reverse-scored before further processing. Its domains, such as physical health, psychological health, social relationships, and environment, will also have their domain scores from averaging across items of each domain. Similarly, for BSI, subscale scores will be computed for all the symptom dimensions as well as the global indices. For PANSS, subscores of domains (positive, negative, and general psychopathology) will be computed from summing items of each domain, while total PANSS will be the sum of all items. For a scale such as Nottingham Onset Schedule, where key outputs are dates such as that of the onset of prodrome or psychosis, the date-field variables will be converted to the computational datetime format, following which durations such as that of untreated psychosis will be computed comparing as intervals between symptom onset and treatment start dates or likewise.

Symptom Severity Trends

Longitudinal changes in clinical assessments (eg, PANSS and BSI scores) will be modeled to track deterioration or improvement.

Sentiment and Thematic Analysis

Journaling entries will undergo natural language processing (NLP)–based sentiment and thematic analysis to extract emotional patterns. Given the linguistic context of Dhaka slum residents, entries are often written in a mixture of Bengali and romanized Bengali ("Banglish"), and off-the-shelf pretrained models are limited for such code-mixed and colloquial data. We therefore will adopt a staged approach: (1) start with strong multilingual transformers (mBERT and XLM-R); (2) fine-tune Bengali models (BanglaBERT [80] or IndicBERT [81]) on a small, study-specific, annotated code-mixed corpus; and (3) add light normalization (transliteration handling; domain lexicon for common Banglish expressions). Multilingual transformers (especially XLM-R) generally outperform vanilla mBERT on low-resource languages, and Bengali-specific models (BanglaBERT and IndicBERT) provide monolingual priors we can adapt via task-specific fine-tuning. To ensure validity, we will manually label a stratified subsample (≈10%-20% of entries) for sentiment or topic, compute interannotator agreement, and use this "gold" to calibrate and evaluate the NLP models; NLP features will be treated as exploratory covariates rather than primary end points. Recent work has released Bengali-English code-mixed sentiment datasets (eg, BnSentMix) [82] and shows that fine-tuning or further pretraining on code-mixed data substantially improves over vanilla multilingual models, which supports our feasibility plan.

Analysis of Passive Data

Overview

High-dimensional smartphone sensor features (eg, mobility, sleep, social communication, and screen-use variability) will be condensed into a limited number of aggregated, clinically interpretable composite variables before model training. Analyses will use penalized regression and regularized tree-based algorithms (eg, gradient boosting) to mitigate overfitting.

Data Preprocessing

Noise Reduction

Continuous, high-frequency passive raw data, as captured from the digital sensors, are often noisy, containing irrelevant signals that need to be smoothed out for more reliable analysis and modeling. Therefore, artifacts such as random or closely spaced idle screen event sessions can be merged, and spatial smoothing and clustering can be applied over GPS jitters. Zero-duration, irrelevant calls or text communications will be filtered out. Kalman filters or moving averages will be used for smoothing out sensor spikes for motion data.

Outlier Detection

Tailored to the behavioral context in question, anomaly detection techniques for variables from such passive data sources will be used. To start, the fast and easily interpretable methods, such as IQR-based filtering, will be applied to handle and assess the quality of individual behavioral variables such as screen time, call counts, and their durations, duration of app use or of movement, and of movement distance as well. For more multivariate, time-varying, and longitudinal contexts, techniques of isolation forests and rolling z scores will identify and address deviations and outliers from the standard range and/or baseline.

Normalization

We will normalize passive data to ensure consistent scales across different features. To maintain comparability across different devices and participants, minimum-maximum scaling or z score normalization will be applied to passive data features.

Feature Engineering

Feature Extraction

We will be extracting relevant information from continuous passive data streams.

ML Steps

In this study, predictive modeling using ML will focus on leveraging the DPs collected through participants' smartphones as the primary features (predictors), with relapse as the primary target outcome. Clinical symptom scores (eg, PANSS, BSI, and WHOQOL-BREF) will serve as secondary outcomes for exploratory analyses and for validating relapse definitions. By analyzing patterns in the DPs, such as device interactions, sleep patterns, movement, and app usage, the model will learn to identify subtle behavioral changes that may be indicative of mental health shifts, such as the onset of relapse. By training the model on historical data from participants, where mental health conditions have been

assessed via standardized clinical questionnaires, the system will aim to forecast future mental health status or relapse risk with high accuracy. This approach will enable the integration of both passive smartphone data and active clinical assessments to provide a comprehensive and early-warning prediction system for serious mental disorders in low-income settings.

Primary Analysis

Overview

For the ML analysis, we define the prediction horizon as the risk of relapse within the next 30 days. This choice is informed by prior DP research demonstrating elevated behavioral anomalies in the 2 weeks preceding relapse and increased anomaly frequency within a 30-day window surrounding relapse [17,83]. Features from passive and active smartphone data will be derived using a 14-day look-back window before the relapse index date. To reduce bias from the clustering of events, a 30-day washout period after each relapse will be applied, during which no new relapse event will be coded.

The primary supervised modeling approach follows a structured analytical pipeline, beginning with temporal feature aggregation, model training, and validation. Passive DP features, including screen usage, mobility, and sleep-related indicators, will be aggregated across predefined sliding windows (eg, last 1, 7, and 30 d) to condense high-frequency time series data into summary statistics. This aggregation facilitates noise reduction, captures short-term behavioral trends, and converts raw sensor streams into fixed-length feature vectors suitable for supervised learning.

As aggregation may introduce feature redundancy and multicollinearity, dimensionality reduction and feature selection techniques, including principal component analysis, singular value decomposition, regularized LASSO regression, mutual information filtering, and SHAP (Shapley Additive Explanations)-based feature importance assessment, will be applied to retain the most informative behavioral features while preserving interpretability. Interaction features will also be derived to examine nonlinear relationships between digital phenotypes (eg, movement variability and screen usage), enabling representation of higher-level behavioral constructs such as mobility, rest, activity, social behavior, and device use. These derived components will be mapped to labeled relapse and nonrelapse windows and aligned with static demographic or clinical covariates, when applicable.

All primary models will be trained with relapse as the target outcome. To ensure robust evaluation and prevent temporal and subject-level data leakage, the dataset will be divided into time-aware training (70%-80%) and testing (20%-30%) partitions, ensuring that future behavioral data are not used to predict earlier outcomes. Model selection and hyperparameter tuning will be conducted using nested cross-validation within training folds. Temporal validation procedures, including rolling window and expanding window validation, will be used to assess model stability over time. Nested cross-validation will also be used to minimize overfitting and optimistic bias.

Given the relatively low frequency of relapse events, class imbalance will be addressed using case-control sampling by time (matching nonrelapse windows from the same observation period) and cost-sensitive learning approaches, with higher penalties assigned to misclassified relapse events. Synthetic oversampling techniques, such as the synthetic minority oversampling technique, will not be applied due to the risk of temporal leakage in longitudinal data.

Primary Model Specific and Evaluation

The prespecified primary supervised model will be gradient boosting (XGBoost or LightGBM), selected a priori for its robust performance with multimodal behavioral data and demonstrated effectiveness in prior DP studies of relapse prediction [7,17]. The gradient boosting model, together with the feature construction and validation strategy described above, constitutes the single prespecified primary modeling pipeline for this study. Model performance will be internally validated using temporal cross-validation and bootstrap optimism correction, with reporting of discrimination (area under the receiver operating characteristic curve) and calibration metrics (calibration slope and Brier score). At a prespecified probability threshold, sensitivity, specificity, positive predictive value, and negative predictive value will also be reported for relapse prediction within the 30-day horizon. Calibration quality will be further assessed using expected calibration error, and decision curve analysis will be conducted to evaluate potential clinical utility across a range of risk thresholds. Explainability methods, including SHAP values and permutation importance, will be applied to identify the most influential digital phenotypes contributing to relapse risk.

Secondary and Exploratory Analyses

Alternative Supervised Models

As sensitivity analyses, the performance of the primary gradient boosting model will be benchmarked against regularized logistic regression and random forest models. These models provide interpretable and clinically acceptable baselines and will be trained and evaluated using the same feature set, outcome definition, and validation framework as the primary analysis.

Unsupervised and Deep Learning Approaches

As exploratory analyses, unsupervised anomaly detection methods (eg, autoencoders and clustering approaches) and time-series deep learning models (eg, long short-term memory and transformer-based architectures) will be examined to explore temporal behavioral patterns and relapse trajectories. These analyses will be contingent on data volume and relapse event frequency and are intended for hypothesis generation and pattern discovery rather than primary inference.

Exploratory Use of Active Data and Data Fusion

Analyses using clinical symptom scores (eg, PANSS, BSI, and WHOQOL-BREF) will be conducted as secondary exploratory models to contextualize relapse prediction and support interpretation of behavioral features. To maximize predictive insight, multimodal fusion approaches will be explored at the feature and decision levels. Feature-level fusion will integrate passive DP data with active clinical assessments, while decision-level fusion will involve combining outputs from multiple models using ensemble techniques such as stacking and boosting. Multimodal fusion and ensemble methods will be used strictly as secondary analyses to complement, but not replace, the prespecified primary supervised model [7,17,83].

Ethical Considerations

This study was approved by the Biomedical and Scientific Research Ethics Committee at the University of Warwick (reference number Biomedical and Scientific Research Ethics Committee 98/23-24). Data collected from participants are securely stored and will not be available for sharing due to privacy considerations. The smartphone app (DataDoc) used for data collection and the feature processing code are available upon request for research purposes.

Participants were informed about the nature, procedures, potential risks, benefits, and voluntary nature of the research before enrollment. Informed consent was obtained from all individual participants before data collection. Participants received detailed information sheets and provided written consent confirming their voluntary participation and understanding of their rights, including the right to withdraw at any time without penalty. The consent process ensured comprehension of this study's aims, data collection procedures, privacy protections, and how data would be used and stored.

To recognize participants' time and engagement in this study, compensation was provided at a rate of 1200 BDT per month for the duration of their involvement. This compensation was disclosed to participants during the consent process and was approved as part of the ethics review.

All data collected from participants are securely stored in accordance with local and international data protection standards and are not publicly available due to privacy and confidentiality considerations. Identifiable personal information has been protected and anonymized where possible to prevent unintended disclosure.

Results

Pilot data collection began in August 2025 in the Korail slum, Dhaka, and as of October 2025, recruitment is ongoing. Data analysis, including feasibility and predictive modeling, commenced in December 2025, with preliminary findings anticipated in mid-2026 and expected results to be published in fall 2026. A total of 14 participants have been recruited, as of January 2026.

Discussion

Principal Findings

The results of this study will offer critical insights into the application of DP for predicting the relapse of SMIs in low-resource settings, specifically among slum residents in Dhaka, Bangladesh. By using passive data collected from participants' smartphones, such as device interactions, sleep patterns, movement, and social engagement, this study aims to assess the ability of ML models to detect early signs of relapse. Our findings will contribute to the broader understanding of how DP can be applied in real-world settings to improve mental health care, especially in LMICs where access to traditional mental health services is difficult or limited [42,49].

This study's innovative approach of combining passive data with clinical assessments allows for the development of personalized, dynamic interventions tailored to individuals' behavioral patterns. This could represent a significant advance over static mental health monitoring methods, offering more timely and effective interventions. However, several challenges and limitations must be acknowledged. First, while DP offers a novel way to monitor behavioral changes, the accuracy and validity of these markers in predicting clinical outcomes require further exploration. The reliance on self-reported clinical assessments may introduce bias [4], as participants' perceptions of their mental health may not always align with objective symptoms. Moreover, such assessments often lack ecological validity, as they are typically conducted in clinical settings and may not capture the nuances of daily life. In contrast, DP offers the potential to monitor individuals in real-world contexts, providing richer insights into behavior and symptom fluctuations outside the clinic [12,13]. Furthermore, privacy concerns surrounding the continuous collection of personal data must be addressed [12], even though strict data protection protocols are in place.

The results will also have implications beyond the scope of this study. The potential for DP biomarkers to serve as dynamic tailoring variables in adaptive interventions could reshape how mental health treatments are delivered, particularly in resource-poor settings. As smartphones become increasingly prevalent in LMICs, leveraging these devices for mental health care could provide a scalable, cost-effective solution that reaches underserved populations. The findings from this study will help guide future research and intervention strategies, paving the way for larger, more comprehensive studies in similar settings.

Limitations

A key limitation of this study is the requirement for participants to own an Android smartphone, which may systematically exclude individuals with SMIs who are older, poorer, or have lower educational attainment. These subgroups may be among those most vulnerable to relapse, meaning our

sample could underrepresent the most at-risk individuals [84,85]. To mitigate bias, we will document sociodemographic and clinical characteristics of enrolled participants and compare these with available data on the broader Korail slum population. In addition, subgroup analyses stratified by socioeconomic and demographic variables will be conducted to evaluate whether predictive performance varies across groups. These steps will help us interpret findings in light of potential exclusion and inform future efforts to reach harder-to-engage individuals with SMIs.

Finally, our sample size estimation was based on the national prevalence of mental disorders in Bangladesh (17%). We acknowledge that this may not accurately capture the epidemiology of SMI in slum settings, where prevalence may be higher or differently distributed due to socioeconomic disadvantage, overcrowding, and limited access to care. As such, our study may be under- or overpowered relative to the true prevalence in Korail. This limitation will be transparently acknowledged in reporting and highlights the need for more robust epidemiological studies of SMI within urban slum populations.

Several practical limitations may affect feasibility. First, requiring participants to install the app as an .apk file may present challenges for individuals with low technical literacy. To address this, in-person installation support and illustrated step-by-step instructions in Bengali will be provided. Second, continuous passive data collection, particularly from GPS sensors, has the potential to increase battery consumption. Given that there is potentially unstable access to electricity in slum environments, participants may disable or uninstall the app to conserve power. While the DataDoc app has been optimized for low-frequency sampling to reduce power demand, battery drain remains a possible barrier to adherence. Furthermore, to minimize battery drainage in a setting with intermittent electricity, the app will upload data only when connected to Wi-Fi or charging [86]. Third, concerns about data privacy and security may limit participant engagement. Although strict protocols are in place, including anonymization and secure server storage, participants may still perceive privacy risks, which could affect long-term acceptability. To help mitigate this, a prior qualitative study was conducted with the same Korail community to explore participants' understanding, perspectives, and willingness to share digital data [87]. Insights from this study informed the design of our consent process and data collection strategy. Nevertheless, residual concerns about privacy remain a possible barrier. We will therefore monitor such issues throughout this study and systematically document dropout or data loss attributable to these barriers.

Conclusions

This study represents a step forward in understanding how DP and ML can be used to predict relapse in individuals with serious mental disorders living in low-resource settings. By combining passive smartphone data with clinical assessments, we aim to build a predictive model that can identify the EWS of relapse, allowing for timely interventions.

While there are challenges associated with the use of DP, such as data privacy and the interpretation of behavioral markers, the potential benefits for early detection and intervention in mental health are substantial. The findings from this study will not only contribute to the growing body of research on DP but also offer practical solutions for addressing the mental health treatment gap in underserved populations. Ultimately, this research could help improve the quality of mental health care in LMIC settings and impoverished communities and provide a model for scalable, technology-driven interventions in other low-resource environments. This research has the potential to inform both local and global strategies for improving access to mental health care in low-resource settings.

Data Availability

The data generated and analyzed during this study are not publicly available due to privacy and ethical considerations involving sensitive participant information. However, deidentified data or a subset of the data may be made available from the principal investigator upon reasonable request and are subject to appropriate ethical approvals. The smartphone app (DataDoc) and associated data processing code used in this study are available for research purposes on request.

Authors' Contributions

NA and SJ conceived this study, formulated the research questions, and led this study's design, protocol development, and manuscript writing. CKD contributed to the clinical components of the study design and manuscript editing. NR contributed

to the statistical and machine learning analysis plan, data pipeline design, and provided critical revisions of this paper. DG provided supervision on study methodology, ethical considerations, and contributed to paper refinement. SPS provided conceptual input during this study's formulation and offered senior academic oversight and critical revisions. All authors reviewed and approved this final paper.

Conflicts of Interest

None declared.

References

1. Viron MJ, Stern TA. The impact of serious mental illness on health and healthcare. Psychosomatics. 2010;51(6):458-465. [doi: 10.1176/appi.psy.51.6.458] [Medline: 21051676]

2. Agenagnew L, kassaw C. The lifetime prevalence and factors associated with relapse among mentally ill patients at Jimma University Medical Center, Ethiopia: cross sectional study. J Psychosoc Rehabil Ment Health. Dec 2020;7(3):211-220. [doi: 10.1007/s40737-020-00176-7]

3. Ustün TB. The global burden of mental disorders. Am J Public Health. Sep 1999;89(9):1315-1318. [doi: 10.2105/ajph.89.9.1315] [Medline: 10474545]

4. Abdullah S, Choudhury T. Sensing technologies for monitoring serious mental illnesses. IEEE MultiMedia. Jan 2018;25(1):61-75. [doi: 10.1109/MMUL.2018.011921236]

5. Onnela JP, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. Neuropsychopharmacology. Jun 2016;41(7):1691-1696. [doi: 10.1038/npp.2016.7] [Medline: 26818126]

6. Gruebner O, Khan MMH, Lautenbach S, et al. A spatial epidemiological analysis of self-rated mental health in the slums of Dhaka. Int J Health Geogr. May 20, 2011;10:36. [doi: 10.1186/1476-072X-10-36] [Medline: 21599932]

7. Henson P, D'Mello R, Vaidyam A, Keshavan M, Torous J. Anomaly detection to predict relapse risk in schizophrenia. Transl Psychiatry. Jan 11, 2021;11(1):28. [doi: 10.1038/s41398-020-01123-7] [Medline: 33431818]

8. Torous J, Kiang MV, Lorme J, Onnela JP. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Ment Health. Jan 2016;3(2):e16. [doi: 10.2196/mental.5165]

9. Coghlan S, D'Alfonso S. Digital phenotyping: an epistemic and methodological analysis. Philos Technol. Dec 2021;34(4):1905-1928. [doi: 10.1007/s13347-021-00492-1]

10. Huckvale K, Venkatesh S, Christensen H. Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. NPJ Digit Med. Sep 6, 2019;2(1):88. [doi: 10.1038/s41746-019-0166-1]

11. Oudin A, Maatoug R, Bourla A, et al. Digital phenotyping: data-driven psychiatry to redefine mental health. J Med Internet Res. Oct 4, 2023;25:e44502. [doi: 10.2196/44502] [Medline: 37792430]

12. Henson P, Barnett I, Keshavan M, Torous J. Towards clinically actionable digital phenotyping targets in schizophrenia. npj Schizophr. May 5, 2020;6(1):13. [doi: 10.1038/s41537-020-0100-1] [Medline: 32372059]

13. Jilka S, Giacco D. Digital phenotyping: how it could change mental health care and why we should all keep up. J Ment Health. Aug 2024;33(4):439-442. [doi: 10.1080/09638237.2024.2395537] [Medline: 39301756]

14. Bourla A, Ferreri F, Ogorzelec L, Guinchard C, Mouchabac S. Évaluation des troubles thymiques par l'étude des données passives : le concept de phénotype digital à l'épreuve de la culture de métier de psychiatre. L'Encéphale. Apr 2018;44(2):168-175. [doi: 10.1016/j.encep.2017.07.007]

15. Difrancesco S, Fraccaro P, van der Veer SN, et al. Out-of-home activity recognition from GPS data in schizophrenic patients. Presented at: 2016 IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS); 324-328; Dublin. 2016.[doi: 10.1109/CBMS.2016.54]

16. Lane E, D'Arcey J, Kidd S, et al. Digital phenotyping in adults with schizophrenia: a narrative review. Curr Psychiatry Rep. Nov 2023;25(11):699-706. [doi: 10.1007/s11920-023-01467-z] [Medline: 37861979]

17. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela JP. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacol. Jul 2018;43(8):1660-1666. [doi: 10.1038/s41386-018-0030-z]

18. Pedrelli P, Fedor S, Ghandeharioun A, et al. Monitoring changes in depression severity using wearable and mobile sensors. Front Psychiatry. 2020;11:584711. [doi: 10.3389/fpsyt.2020.584711] [Medline: 33391050]

19. Jacobson NC, Lekkas D, Huang R, Thomas N. Deep learning paired with wearable passive sensing data predicts deterioration in anxiety disorder symptoms across 17-18 years. J Affect Disord. Mar 1, 2021;282:104-111. [doi: 10.1016/j.jad.2020.12.086] [Medline: 33401123]

20. Lu J, Shang C, Yue C, et al. Joint modeling of heterogeneous sensing data for depression assessment via multi-task learning. Proc ACM Interact Mob Wearable Ubiquitous Technol. Mar 26, 2018;2(1):1-21. [doi: 10.1145/3191753]

21. Bufano P, Laurino M, Said S, Tognetti A, Menicucci D. Digital phenotyping for monitoring mental disorders: systematic review. J Med Internet Res. Dec 13, 2023;25:e46778. [doi: 10.2196/46778] [Medline: 38090800]

22. Adler DA, Ben-Zeev D, Tseng VWS, et al. Predicting early warning signs of psychotic relapse from passive sensing data: an approach using encoder-decoder neural networks. JMIR mHealth uHealth. Aug 31, 2020;8(8):e19962. [doi: 10.2196/19962] [Medline: 32865506]

23. Melcher J, Lavoie J, Hays R, et al. Digital phenotyping of student mental health during COVID-19: an observational study of 100 college students. J Am Coll Health. Mar 24, 2023;71(3):736-748. [doi: 10.1080/07448481.2021.1905650]

24. Liang Y, Zheng X, Zeng DD. A survey on big data-driven digital phenotyping of mental health. Inf Fusion. Dec 2019;52:290-307. [doi: 10.1016/j.inffus.2019.04.001]

25. Stasak B, Epps J, Schatten HT, Miller IW, Provost EM, Armey MF. Read speech voice quality and disfluency in individuals with recent suicidal ideation or suicide attempt. Speech Commun. Sep 2021;132:10-20. [doi: 10.1016/j.specom.2021.05.004]

26. Ered A, Cooper S, Ellman LM. Sleep quality, psychological symptoms, and psychotic-like experiences. J Psychiatr Res. Mar 2018;98:95-98. [doi: 10.1016/j.jpsychires.2017.12.016] [Medline: 29331930]

27. Harvey AG, Kaplan KA, Soehner AM. Interventions for sleep disturbance in bipolar disorder. Sleep Med Clin. Mar 2015;10(1):101-105. [doi: 10.1016/j.jsmc.2014.11.005] [Medline: 25750600]

28. Robillard R, Hermens DF, Lee RSC, et al. Sleep-wake profiles predict longitudinal changes in manic symptoms and memory in young people with mood disorders. J Sleep Res. Oct 2016;25(5):549-555. [doi: 10.1111/jsr.12413] [Medline: 27242171]

29. Harvey AG, Talbot LS, Gershon A. Sleep disturbance in bipolar disorder across the lifespan. Clin Psychol: Sci Pract. Jun 2009;16(2):256-277. [doi: 10.1111/j.1468-2850.2009.01164.x]

30. Mendlewicz J. Sleep disturbances: core symptoms of major depressive disorder rather than associated or comorbid disorders. World J Biol Psychiatry. 2009;10(4):269-275. [doi: 10.3109/15622970802503086] [Medline: 19921968]

31. Palmius N, Tsanas A, Saunders KEA, et al. Detecting bipolar depression from geographic location data. IEEE Trans Biomed Eng. Aug 1, 2017;64(8):1761-1771. [doi: 10.1109/TBME.2016.2611862]

32. Raugh IM, James SH, Gonzalez CM, et al. Digital phenotyping adherence, feasibility, and tolerability in outpatients with schizophrenia. J Psychiatr Res. Jun 2021;138:436-443. [doi: 10.1016/j.jpsychires.2021.04.022] [Medline: 33964681]

33. Fraccaro P, Beukenhorst A, Sperrin M, et al. Digital biomarkers from geolocation data in bipolar disorder and schizophrenia: a systematic review. J Am Med Inform Assoc. Nov 1, 2019;26(11):1412-1420. [doi: 10.1093/jamia/ocz043] [Medline: 31260049]

34. Currey D, Torous J. Digital phenotyping data to predict symptom improvement and mental health app personalization in college students: prospective validation of a predictive model. J Med Internet Res. 2023;25:e39258. [doi: 10.2196/39258]

35. Andrea A, Agulia A, Serafini G, Amore M. Digital biomarkers and digital phenotyping in mental health care and prevention. Eur J Public Health. Sep 1, 2020;30(Supplement_5):ckaa165.1080. [doi: 10.1093/eurpub/ckaa165.1080]

36. Hays R, Keshavan M, Wisniewski H, Torous J. Deriving symptom networks from digital phenotyping data in serious mental illness. BJPsych Open. Nov 3, 2020;6(6):e135. [doi: 10.1192/bjo.2020.94] [Medline: 33138889]

37. Henson P, Rodriguez-Villa E, Torous J. Investigating associations between screen time and symptomatology in individuals with serious mental illness: longitudinal observational study. J Med Internet Res. Mar 1, 2021;23(3):e23144. [doi: 10.2196/23144]

38. Amha H, Getnet A, Munie BM, et al. Relapse rate and predictors among people with severe mental illnesses at Debre Markos Comprehensive Specialized Hospital, Northwest Ethiopia: a prospective follow up study. Eur Arch Psychiatry Clin Neurosci. Sep 18, 2024. [doi: 10.1007/s00406-024-01900-1] [Medline: 39292261]

39. Leucht S, Tardy M, Komossa K, et al. Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. Lancet. Jun 2, 2012;379(9831):2063-2071. [doi: 10.1016/S0140-6736(12)60239-6] [Medline: 22560607]

40. Zipursky RB, Menezes NM, Streiner DL. Risk of symptom recurrence with medication discontinuation in first-episode psychosis: a systematic review. Schizophr Res. Feb 2014;152(2-3):408-414. [doi: 10.1016/j.schres.2013.08.001] [Medline: 23972821]

41. Gumley AI, Bradstreet S, Ainsworth J, et al. Digital smartphone intervention to recognise and manage early warning signs in schizophrenia to prevent relapse: the EMPOWER feasibility cluster RCT. Health Technol Assess. 2022;26(27):1-174. [doi: 10.3310/HLZE0479]

42. Rathod S, Pinninti N, Irfan M, et al. Mental health service provision in low- and middle-income countries. Health Serv Insights. Jan 1, 2017;10. [doi: 10.1177/1178632917694350] [Medline: 28469456]

43. Cohen A, Naslund JA, Chang S, et al. Relapse prediction in schizophrenia with smartphone digital phenotyping during COVID-19: a prospective, three-site, two-country, longitudinal study. Schizophrenia (Heidelb). Jan 27, 2023;9(1):6. [doi: 10.1038/s41537-023-00332-5] [Medline: 36707524]

44. Naslund JA, Deng D. Addressing mental health stigma in low-income and middle-income countries: a new frontier for digital mental health. Ethics Med Public Health. Dec 2021;19:100719. [doi: 10.1016/j.jemep.2021.100719] [Medline: 35083375]

45. Hossain MD, Ahmed HU, Chowdhury WA, Niessen LW, Alam DS. Mental disorders in Bangladesh: a systematic review. BMC Psychiatry. Jul 30, 2014;14(1):216. [doi: 10.1186/s12888-014-0216-9] [Medline: 25073970]

46. Ahmed T, Rizvi SJR, Rasheed S, et al. Digital health and inequalities in access to health services in Bangladesh: mixed methods study. JMIR mHealth uHealth. Jul 21, 2020;8(7):e16473. [doi: 10.2196/16473] [Medline: 32706736]

47. Singh SP, Jilka S, Abdulmalik J, et al. Transforming access to care for serious mental disorders in slums (the TRANSFORM Project): rationale, design and protocol. BJPsych Open. Oct 13, 2022;8(6):e185. [doi: 10.1192/bjo.2022.584] [Medline: 36226591]

48. Gruebner O, Khan MMH, Lautenbach S, et al. Mental health in the slums of Dhaka - a geoepidemiological study. BMC Public Health. Mar 9, 2012;12(1):177. [doi: 10.1186/1471-2458-12-177] [Medline: 22404959]

49. Alloh FT, Regmi P, Onche I, Teijlingen EV, Trenoweth S. Mental health in low-and middle income countries (LMICs): going beyond the need for funding. Health Prospect. Jun 19, 2018;17(1):12-17. [doi: 10.3126/hprospect.v17i1.20351]

50. Gamieldien F, Galvaan R, Myers B, Syed Z, Sorsdahl K. Exploration of recovery of people living with severe mental illness (SMI) in low/middle-income countries (LMICs): a scoping review. BMJ Open. Mar 24, 2021;11(3):e045005. [doi: 10.1136/bmjopen-2020-045005] [Medline: 33762242]

51. Sinthia SA. Analysis of urban slum: case study of Korail Slum, Dhaka. Int J Urban Civil Eng. 2020;14(11):416-430. URL: https://www.researchgate.net/publication/352785100 [Accessed 2026-01-24]

52. Derogatis LR. BSI, Brief Symptom Inventory: Administration, Scoring, & Procedures Manual. National Computer Systems, 1993; 1993. ISBN: 9780749154356

53. Franke GH, Jagla-Franke M, Küch D, Petrowski K. A new routine for analyzing brief symptom inventory profiles in chronic pain patients to evaluate psychological comorbidity. Front Psychol. 2021;12:692545. [doi: 10.3389/fpsyg.2021.692545] [Medline: 34659009]

54. Michel G, Baenziger J, Brodbeck J, Mader L, Kuehni CE, Roser K. The Brief Symptom Inventory in the Swiss general population: presentation of norm scores and predictors of psychological distress. PLoS ONE. Jul 1, 2024;19(7):e0305192. [doi: 10.1371/journal.pone.0305192]

55. Mohammadkhani P, Dobson KS, Amiri M, Ghafari FH. Psychometric properties of the Brief Symptom Inventory in a sample of recovered Iranian depressed patients. Int J Clin Health Psychol. 2010;10(3):541-551. URL: https://www.redalyc.org/pdf/337/33714079009.pdf [Accessed 2026-01-28]

56. Biplob P, Sarker DC, Sarker RC. Assessment of water supply and sanitation facilities for Korail Slum in Dhaka City. IJCEE-IJENS. 2013. URL: https://www.researchgate.net/publication/282062799_Assessment_of_Water_Supply_and_Sanitation_Facilities_for_Korail_Slum_in_Dhaka_City [Accessed 2025-05-28]

57. Bashar R, Tommy SS, Ria AF, Khan NA. Assessing the real-life socio-economic scenario of established slums in Dhaka: the cases of Korail and Sattola. Eur Online J Nat Soc Sci. May 21, 2020;9(2):455-466. URL: https://european-science.com/eojnss/article/view/6029/pdf [Accessed 2026-01-24]

58. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. American Psychiatric Association; 2013. [doi: 10.1176/appi.books.9780890425596] ISBN: 0-89042-555-8

59. Patel V. Mental health in low- and middle-income countries. Br Med Bull. 2007;81-82:81-96. [doi: 10.1093/bmb/ldm010] [Medline: 17470476]

60. Merchant R, Torous J, Rodriguez-Villa E, Naslund JA. Digital technology for management of severe mental disorders in low-income and middle-income countries. Curr Opin Psychiatry. Sep 2020;33(5):501-507. [doi: 10.1097/YCO.0000000000000626] [Medline: 32520747]

61. Javed A, Lee C, Zakaria H, et al. Reducing the stigma of mental health disorders with a focus on low- and middle-income countries. Asian J Psychiatr. Apr 2021;58:102601. [doi: 10.1016/j.ajp.2021.102601] [Medline: 33611083]

62. Alam MF. National mental health survey 2019. National Institute of Mental Health; 2021. URL: https://nimh.gov.bd/wp-content/uploads/2021/11/Mental-Health-Survey-Report.pdf [Accessed 2026-01-24]

63. Shen Y, Zhao W. Knowledge, attitude and practice of urticaria patients towards urticaria medication treatment: a cross-sectional study at the first affiliated hospital of Kunming Medical University. BMJ Open. Nov 27, 2024;14(11):e091425. [doi: 10.1136/bmjopen-2024-091425] [Medline: 39609011]

64. Kim HY. Statistical notes for clinical researchers: sample size calculation 1. comparison of two independent sample means. Restor Dent Endod. 2016;41(1):74. [doi: 10.5395/rde.2016.41.1.74]

65. Gogtay NJ. Principles of sample size calculation. Indian J Ophthalmol. 2010;58(6):517-518. [doi: 10.4103/0301-4738.71692] [Medline: 20952836]

66. Leucht S, Barnes TRE, Kissling W, Engel RR, Correll C, Kane JM. Relapse prevention in schizophrenia with new-generation antipsychotics: a systematic review and exploratory meta-analysis of randomized, controlled trials. Am J Psychiatry. Jul 2003;160(7):1209-1222. [doi: 10.1176/appi.ajp.160.7.1209] [Medline: 12832232]

67. Bergen J, Hunt G, Armitage P, Bashir M. Six-month outcome following a relapse of schizophrenia. Aust N Z J Psychiatry. Dec 1998;32(6):815-822. [doi: 10.3109/00048679809073871]

68. Kishi T, Sakuma K, Hatano M, et al. Relapse and its modifiers in major depressive disorder after antidepressant discontinuation: meta-analysis and meta-regression. Mol Psychiatry. Mar 2023;28(3):974-976. [doi: 10.1038/s41380-022-01920-0]

69. Purnell L, Graham A, Chiu K, Trickey D, Meiser-Stedman R. A systematic review and meta-analysis of PTSD symptoms at mid-treatment during trauma-focused treatment for PTSD. J Anxiety Disord. Oct 2024;107:102925. [doi: 10.1016/j.janxdis.2024.102925] [Medline: 39270371]

70. Jilka S, Siddiqi B, Winsper C, et al. Influences on help-seeking for serious mental illness in Dhaka, Bangladesh: a mixed-methods study. Soc Psychiatry Psychiatr Epidemiol. Oct 31, 2025. [doi: 10.1007/s00127-025-03012-0] [Medline: 41174073]

71. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ. Mar 18, 2020;368:m441. [doi: 10.1136/bmj.m441] [Medline: 32188600]

72. Aas IM. Global Assessment of Functioning (GAF): properties and frontier of current knowledge. Ann Gen Psychiatry. Dec 2010;9(1):20. [doi: 10.1186/1744-859X-9-20]

73. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med. Nov 1995;41(10):1403-1409. [doi: 10.1016/0277-9536(95)00112-K]

74. Wideman TH, Sullivan MJL, Inada S. Brief Symptom Inventory. In: Encyclopedia of Behavioral Medicine. Springer; 2013:269-270. [doi: 10.1007/978-1-4419-1005-9_3]

75. Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr Bull. Jan 1, 1987;13(2):261-276. [doi: 10.1093/schbul/13.2.261]

76. Chan SF, La Greca AM. Perceived Stress Scale (PSS). In: Encyclopedia of Behavioral Medicine. Springer; 2020:1646-1648. [doi: 10.1007/978-3-030-39903-0_773]

77. Singh SP, Cooper JE, Fisher HL, et al. Determining the chronology and components of psychosis onset: the Nottingham Onset Schedule (NOS). Schizophr Res. Dec 2005;80(1):117-130. [doi: 10.1016/j.schres.2005.04.018]

78. Sleep questionnaire for adults. Oxford University Hospitals NHS Trust; 2025. URL: https://www.ouh.nhs.uk/media/vv1jq1qj/sleep-questionnaire-over-11.pdf [Accessed 2026-01-24]

79. MacIsaac A, Mushquash AR, Wekerle C. Writing yourself well: dispositional self-reflection moderates the effect of a smartphone app-based journaling intervention on psychological wellbeing across time. Behav Change. Dec 2023;40(4):297-313. [doi: 10.1017/bec.2022.24]

80. Rahman MH, Uddin MA, Ria ZF, Rahman RM. Optimizing BERT for Bengali emotion classification: evaluating knowledge distillation, pruning, and quantization. CMES. 2025;142(2):1637-1666. [doi: 10.32604/cmes.2024.058329]

81. Subhash PM, Gupta D, kanjirangat V, CR K. Indo-Aryan dialect identification using deep learning ensemble model. Procedia Comput Sci. 2024;235:2886-2896. [doi: 10.1016/j.procs.2024.04.273]

82. Alam S, Ishmam F, Alvee H, et al. BnSentMix: a diverse bengali-english code-mixed dataset for sentiment analysis. arXiv. Preprint posted online on Dec 10, 2024. [doi: 10.48550/arXiv.2408.08964]

83. Cohen A, Naslund J, Lane E, et al. Digital phenotyping data and anomaly detection methods to assess changes in mood and anxiety symptoms across a transdiagnostic clinical sample. Acta Psychiatr Scand. Mar 2025;151(3):388-400. [doi: 10.1111/acps.13712] [Medline: 38807465]

84. Kirkbride JB, Anglin DM, Colman I, et al. The social determinants of mental health and disorder: evidence, prevention and recommendations. World Psychiatry. Feb 2024;23(1):58-90. [doi: 10.1002/wps.21160] [Medline: 38214615]

85. Clausen W, Watanabe-Galloway S, Bill Baerentzen M, Britigan DH. Health literacy among people with serious mental illness. Community Ment Health J. May 2016;52(4):399-405. [doi: 10.1007/s10597-015-9951-8] [Medline: 26443671]

86. Alam NB, Surani M, Das CK, Giacco D, Singh SP, Jilka S. Challenges and standardisation strategies for sensor-based data collection for digital phenotyping. Commun Med (Lond). Aug 19, 2025;5(1):360. [doi: 10.1038/s43856-025-01013-3] [Medline: 40830260]

87. Alam N, Giacco D, Siddiqi B, Singh SP, Jilka S. Investigating awareness and acceptance of digital phenotyping in Dhaka's Korail slum: qualitative study. JMIR Form Res. Jun 23, 2025;9:e65530. [doi: 10.2196/65530] [Medline: 40549359]

Abbreviations

BSI: Brief Symptom Inventory

DP: digital phenotyping
DSM-5: *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*
EDF: effective degrees of freedom
EPP: events per predictor
EWS: early warning sign
GAF: Global Assessment of Functioning
GSI: Global Severity Index
HIC: high-income country
LASSO : least absolute shrinkage and selection operator
LMIC: low- and middle-income country
MDD: major depressive disorder
ML: machine learning
NIHR: National Institute for Health and Care Research
NLP: natural language processing
PANSS: Positive and Negative Syndrome Scale
PIS: participant information sheet
PTSD: posttraumatic stress disorder
SHAP: Shapley Additive Explanations
SMI: serious mental illness
TRANSFORM: Transforming Access to Care for Serious Mental Disorders in Slums
WHOQOL-BREF: World Health Organization Quality of Life-Brief