Evaluating a smartphone-based symptom self-monitoring app for psychosis in China (YouXin): A non-randomised validity and feasibility study with a mixed-methods design

Xiaolong Zhang[1,2] (iD), Shôn Lewis[1,3], Lesley-Anne Carter[4], Xu Chen[2,5], Jiaojiao Zhou[2,5], Xingyu Wang[1] and Sandra Bucci[1,3] (iD)

Abstract

Background: Psychosis causes a significant burden globally, including in China, where limited mental health resources hinder access to care. Smartphone-based remote monitoring offers a promising solution. This study aimed to assess the validity, feasibility, acceptability, and safety of a symptom self-monitoring smartphone app, YouXin, for people with psychosis in China.

Methods: A pre-registered non-randomised validity and feasibility study with a mixed-methods design. Participants with psychosis were recruited from a major tertiary psychiatric hospital in Beijing, China. Participants utilised the YouXin app to self-monitor psychosis and mood symptoms for four weeks. Feasibility outcomes were recruitment, retention and outcome measures completeness. Active symptom monitoring (ASM) validity was tested against corresponding clinical assessments (PANSS and CDS) using Spearman correlation. Ten participants completed qualitative interviews at study end to explore acceptability of the app and trial procedures.

Results: Feasibility parameters were met. The target recruitment sample of 40 participants was met, with 82.5% completing outcome measures, 60% achieving acceptable ASM engagement (completing >33% of all prompts), and 33% recording sufficient passive monitoring data to extract mobility indicators. Five ASM domains (hallucinations, suspiciousness, guilt feelings, delusions, grandiosity) achieved moderate correlation with clinical assessment. Both quantitative and qualitative evaluation showed high acceptability of YouXin. Clinical measurements indicated no symptom and functional deterioration. No adverse events were reported, suggesting YouXin is safe to use in this clinical population.

Conclusions: The trial feasibility, acceptability and safety parameters were met and a powered efficacy study is indicated. However, refinements are needed to improve ASM validity and increase passive monitoring data completeness.

Keywords

smartphone app, digital mental health, psychosis, self-management, China

Submission date: 6 September 2023; Acceptance date: 5 December 2023

[1]Division of Psychology and Mental Health, School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, The University of Manchester, Manchester, United Kingdom
[2]The National Clinical Research Center for Mental Disorders & Beijing Key Laboratory of Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China
[3]Greater Manchester Mental Health NHS Foundation Trust, Manchester, United Kingdom
[4]Division of Population Health, Health Services Research & Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health,

Manchester Academic Health Science Centre, The University of Manchester, Manchester, United Kingdom
[5]Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing, China

Corresponding author:
Sandra Bucci. 2nd Floor, Zochonis Building, Brunswick Street, The University of Manchester, Manchester, M13 9PL, United Kingdom.
Email: Sandra.bucci@manchester.c.uk

Background

Psychosis is a severe mental health condition that causes significant personal, social and economic burden.[1] China has the highest prevalence of schizophrenia in the world.[2] A recent Chinese national mental health survey showed that the lifetime prevalence rate of schizophrenia and other psychotic disorders was 0.7%.[3] Moreover, schizophrenia accounted for 13.7% of all mental and substance disorder disability-adjusted life years (DALYs) in 2019 in China.[4] However, staffing and resources to support people with severe mental health problems is scarce, with an estimated shortage of 40,000 mental health professionals relative to the population need.[5]

Digital mental health is a promising solution for the shortfall of mental healthcare resources in China.[6,7] The widespread availability of digital technologies offers an unparalleled opportunity to expand service delivery on an unprecedented scale. Previous studies have shown that the internet access rate in people with a diagnosis of schizophrenia in Hong Kong reached 90.9%,[8] and people with mental health problems in China in general are willing to accept digital health interventions (DHIs) for their mental health and considered them helpful.[9] Integrating digital health technologies (DHTs) into clinical workflows holds promise in enhancing evidence-based care, fostering shared decision making, and facilitating continuity of care.[10]

Smartphone-based symptom monitoring has the potential to generate near real-time data through active and passive monitoring methods to capture symptom fluctuation,[11] which is not achievable by conventional face-to-face methods of assessments in current service settings. Active monitoring is typically defined as the collection of user self-reported data (e.g., psychological, behavioural, physiological) by ecological momentary assessment.[12] Passive monitoring refers to using sensors embedded in a smartphone or a wearable device to record the user's behaviour pattern (e.g., app usage, activity, mobility) or contextual information (e.g., location, social interaction).[13] By combining active and passive monitoring, smartphone-based monitoring has the potential to provide a dynamic, personal, and valid representation of a person's emotional and behavioural state.[14,15] Additionally, a recent randomised controlled trial (RCT) study has shown that active self-monitoring of symptoms can reduce positive psychotic symptoms in people with early psychosis in the UK.[16]

Smartphone-based symptom monitoring in people with psychosis has proved to be feasible in the Western context[16–18]; however, only one published study has been conducted in non-western, low- and middle-income countries (LMICs) that tested a digital platform with a symptom monitoring element.[19] To date, no smartphone-based symptom monitoring tool has been developed and evaluated for people with psychosis in China.[6] Therefore, given the potential of smartphone-based symptom monitoring to improve the outcome of psychosis and reduce the burden of mental health system in China, we developed the YouXin smartphone app, a digital remote monitoring tool combining active and passive monitoring, and tested its validity, feasibility, acceptability and safety for people with psychosis in China. Further details of the development of the app are published elsewhere.[20]

Methods

The study was approved by the University of Manchester and Beijing Anding Hospital Research Ethics Committee. A pre-published trial protocol is published elsewhere.[20]

Participants

Participants were recruited from one of the major tertiary psychiatric hospitals located in the Beijing metropolitan area from August to November 2022. Staff identified and approached prospective participants and evaluated their willingness and eligibility to take part. With consent, those considered eligible and interested were referred to author XZ who confirmed eligibility. Eligible prospective participants were given an information sheet and at least 48 h to consider taking part. After this time, author XZ contacted the potential participant to explain the details of the study and answer any questions. Consent and assessments took place face-to-face at the hospital. Written informed consent was obtained from all participants prior to study initiation. Eligibility criteria were deliberately inclusive to fully assess feasibility of the app and to reach generalisable conclusions to a wide range of service users with psychosis. Inclusion criteria were: (1) diagnosed with a DSM-5 (F20-F29) schizophrenia spectrum disorder or met criteria for being at clinical high risk for psychosis according to the Structured Interview for Prodromal Syndromes (SIPS); (2) aged 16 to 65; (3) receiving care from Beijing Anding Hospital, China and continued to be actively supported by the hospital for the trial duration (5 weeks); (4) owned and able to use a smartphone; and (5) clinically stable as judged by the responsible clinician. Exclusion criteria were: (1) diagnosed with organic or substance induced psychosis; (2) lack of capacity to provide informed consent as judged by the responsible clinician; and (3) judged to be at risk for self-harm or harm others.

The intervention: YouXin app

YouXin is a smartphone app-based symptom self-monitoring tool designed specifically for people with psychosis, available for both Android and iOS platforms. The app is adapted from the ClinTouch app.[21] Palmier-Claus et al.[21] found that the ClinTouch items were significantly correlated with the Positive and Negative Syndrome Scale (PANSS) positive and general psychopathology subscales and the Calgary

Depression Scale (CDS) but not PANSS negative subscale. Therefore, to develop the YouXin app, we selected the validated ClinTouch psychosis and mood items equivalent to PANSS positive and general psychopathology subscales and the CDS, and translated the items into Chinese. To assess negative symptoms, we added a passive monitoring function to the app to extract mobility and activity indicators corresponding to the PANSS negative subscale. The YouXin app was developed by a multidisciplinary team, including academics, clinicians, software engineers and experts by experience. We utilised a systematic co-production approach to development to maximise user usability and acceptability from both patient and clinician perspectives. Author XZ carried out four online co-production consultation meetings with patients and clinicians (three patient meetings and one clinician meeting) at Beijing Anding Hospital, China. The co-production procedure is shown in Supplementary Text 1.

The core functions of YouXin consist of active monitoring of current symptoms (termed active symptom monitoring (ASM); including psychosis and mood symptoms; contextual information) in real-time and passive monitoring of behavioural activity (i.e., global positioning system (GPS) and step counting). Examples of items are shown in Supplemental Table 1. The ASM component uses a time-contingent design. That is, prompts were set to alert at two pseudo-random timepoints per day in a 12-h interval from 10:00 a.m. to 10:00 p.m. seven days a week for four weeks. This prompt frequency was set to minimise participant assessment burden. To prevent clustering of responses, each prompt on a particular day had a minimum 1-h interval. Participants could either snooze the prompt for 5 or 30 min or decline to answer. GPS data was passively collected every 5 min from smartphone sensors to measure users' mobility and activity levels. All collected data were wirelessly uploaded in real-time to a secure server managed by Beijing Anding Hospital.

Procedure

Participants were given a training session to help them navigate the app. Support was given to help the participant download the app onto their smartphone and a demonstration of the app was given. Participants had the opportunity to practice using the app and ask questions during the process; they had access to the app from the training session until the four-week assessment timepoint. An opt-out approach was applied to give participants autonomy to switch on/off the passive sensing section during the study. Participants were informed at the training sessions that the app was not suitable for seeking urgent medical care.

The timeline of outcome assessments is summarised in Table 1. Participants were assessed in person at baseline, one and four weeks. To assess the validity of the ASM items, the PANSS and CDS were administered at the one-week timepoint to calculate the correlation between the relevant ASM and clinical assessment items. The assessor was a clinician with experience conducting assessments in clinical trials and was trained in the assessment measures used here. The assessor was blind to the ASM results. Semistructured qualitative interviews were conducted after the four-week timepoint to explore the acceptability of the app and trial procedures. Author XZ provided technical support with the app as needed; participants were referred to their treating clinician for appropriate mental healthcare if any clinical concerns arose.

Measures

Sociodemographic and clinical characteristics. We gathered descriptive demographic (age, gender, education level, marriage status, average income, accommodation and living situation, employment status, ethnicity), clinical, service use and digital and mental health literacy information. Service use and clinical information were collected from the participant's medical record, with consent.

Feasibility parameters. To assess feasibility, we collected detailed recruitment and retention data according to the CONSORT statement for feasibility studies,[22] including (i) number of eligible participants consenting; (ii) completeness of outcome measures; (iii) number lost to follow-up; and (iv) reason for withdrawal. We defined a threshold of > 80% retention rate as feasible and no changes needed. The app usage and engagement parameters were explored using the Analysing and Measuring Usage and Engagement Data (AMUsED) framework.[23] The checklists of AMUsED framework are shown in Supplemental Tables 2–4. We defined acceptable engagement as completing at least 33% of the prompts, a criterion widely accepted within the digital mental health field,[16,18,24] despite its arbitrary nature.

Acceptability. The Quantitative Feedback Questionnaire (QFQ) was used to collect quantitative information regarding acceptability at the four-week timepoint. This is a 20-item questionnaire designed to assess the acceptability of utilising smartphone-based DHIs and has been used in previous DHI trials in psychosis.[25,26] Qualitative exit interviews were conducted to explore participants' experience of using Youxin, detailed findings of which are reported elsewhere (summary-level findings are reported here).

Clinical measures. The Chinese version of the Mini-International Neuropsychiatric Interview[27] was used to assess DSM-5 diagnosis. The Chinese version of the SIPS[28] was used to determine the presence of an At Risk Mental State. Psychotic symptoms were rated using the Chinese version of the PANSS,[29] a 30-item scale that assesses positive and negative symptoms and general psychopathology. Mood symptoms were assessed using the Chinese version of the CDS,[30] a nine-item scale designed

Table 1. Timeline of data collection.

Timepoint	Enrolment	Baseline	1-week follow up	4-week follow up
Enrolment				
Eligibility screen	X			
Informed consent		X		
Training session		X		
Intervention				
Smartphone monitoring		←——————→		
Assessments				
MINI	X			
SIPS	X			
PANSS		X	X	X
CDS		X	X	X
PHE-s		X		X
PSP		X		X
EQ-5D-5L		X		X
Socio-demographic characteristics		X		
Clinical characteristics		X		
QFQ				X
Qualitative interview				X
Electronic health record				
Number of hospitalisations		←——————→		
Time of first onset		X		
Number of episodes		←——————→		
Medication(s) or therapy used		←——————→		

Abbreviations: MINI: mini-international neuropsychiatric interview; SIPS: structured interview for prodromal syndromes; PANSS: Positive and Negative Syndrome Scale; CDS: Calgary Depression Scale for Schizophrenia; PHE-s: Patient Health Engagement Scale; PSP: Personal and Social Performance; EQ5D: EuroQoL 5D 5 level version; QFQ, Quantitative Feedback Questionnaire.

specifically for assessing depression in schizophrenia. Higher scores of PANSS and CDS correspond to more severe symptoms. Participant engagement was measured using the Patient Health Engagement (PHE) Scale,[31] a self-report scale assessing patient attitudes of engagement in healthcare. Social functioning was rated using the Chinese version of the Personal and Social Performance (PSP) scale[32] a 100-point single-item rating scale assessing four areas of functioning (socially useful activities, personal and social relationships, self-care, and disturbing and aggressive behaviours). Quality of life was measured by the EuroQoL 5D 5 level version (EQ-5D-5L),[33] a self-report questionnaire designed

Table 2. Demographic and clinical characteristics: numbers (%) of participants.

Characteristics	Values
Age range	
16–25	14 (35.0)
26–35	16 (40.0)
36–45	7 (17.5)
46–55	2 (5.0)
56–65	1 (2.5)
Gender	
Male	24 (60.0)
Female	16 (40.0)
Ethnicity	
Han	38 (95.0)
Hui	1 (2.5)
Mongolian	1 (2.5)
Education level	
Primary school	1 (2.5)
Middle school	1 (2.5)
High school	16 (40.0)
College	10 (25.0)
University	12 (30.0)
Marriage status	
Single	25 (62.5)
Divorced	2 (5.0)
Married	13 (32.5)
Average monthly income	
1000–5000	20 (50.0)
5000–10,000	12 (30.0)
10,000–15,000	6 (15.0)

(continued)

Table 2. Continued.

Characteristics	Values
>15,000	2 (5.0)
Accommodation and living situation	
Lived alone	2 (5.0)
Lived with others	38 (95.0)
Employment status	
Full time	11 (27.5)
unemployed	18 (45.0)
student	9 (22.5)
Parttime	2 (5.0)
Diagnosis	
Brief psychotic disorder	2 (5.0)
Clinical high risk	1 (2.5)
Delusional disorder	1 (2.5)
Schizophrenia	35 (87.5)
Schizophreniform disorder	1 (2.5)
Number of hospitalisations	
0	13 (32.5)
1	18 (45.0)
2	6 (15.0)
3	2 (5.0)
4	1 (2.5)
Number of episodes	
1	31 (77.5)
2	5 (12.5)
3	3 (7.5)
4	1 (2.5)
Medication(s) or therapy used	
Antipsychotics	40 (100.0)

(continued)

Table 2. Continued.

Characteristics	Values
Antidepressants	11 (27.5)
Antianxiety medications	4 (10.0)
Mood stabiliser	3 (7.5)
Other	7 (17.5)

for assessing health status and health-related quality of life. Higher scores of PSP and EQ-5D-5L correspond to better outcomes. Assessors received training and ongoing supervision in all study procedures.

Safety. Adverse events (AEs) were monitored throughout the study and reported using standardised procedures in line with the National Institute for Health and Care Research and Health Research Authority safety reporting procedures.

Data analysis

All quantitative analysis was conducted using R (Version 4.0.5).[34] Descriptive statistics were performed to analyse feasibility, acceptability and safety data. Logistic regression was used to explore socio-demographic and clinical characteristics and device attributes as predictors of feasibility. We also reported descriptive statistics of the clinical outcomes at baseline, one-week, and four-week assessments. Paired t-test and Wilcoxon signed-rank test were conducted to explore the changes in clinical outcomes during the four-week study period. To measure the validity of ASM domains, subject-level active monitoring data mean score of the first week was used to calculate the association between the active monitoring items score and the relevant symptom domain score of the PANSS and CDS at the week one follow-up assessment. Cronbach's alpha was used to calculate the internal consistency of the ASM item sets. The mean squared successive difference (MSSD) and intraclass correlation coefficient (ICC) were used to assess the instability of active monitoring domain scores over the first week. MSSD shows the day-to-day variability within participants. ICC indicates the proportion of variance that can be attributed between participants. GPS data was analysed following Müller et al.'s[35] approach. The R code was modified to fit the data set of this study. We used at least 12 h GPS data collected in a day as the criteria of valid GPS recording, as it is in line with the active monitoring interval. In order to determine "home" location, GPS data must include night-time hours (i.e., 9 p.m. to 6 a.m.). Locations were clustered as "unique" if a person spent at least 30 min within a 200-m radius. We did not perform imputation for missing data;

instead, we used available data to extract GPS mobility indicators, including unique locations, places visited, and time spent at home. Spearman correlation was used for exploratory analysis of the association between GPS mobility indicators and PANSS negative symptom score.

Qualitative data was transcribed verbatim in Chinese and then translated to English by author XZ. The English transcripts were then independently back-translated to Chinese by coauthor (XW) to check the validity of the translation. A framework analysis approach was employed following Braun and Clarke's[36] protocol of six phases of thematic analysis, including familiarisation, transcription, generating initial codes, searching for themes, reviewing themes, and then defining and deciding on meaningful themes. Nvivo software (Version 12)[37] was used to support qualitative analysis.

Results

Sample characteristics

Participant sociodemographic and clinical characteristics are shown in Table 2. In total, 40 people participated in the study. The median age of participants was 29 years (IQR = 22–35.25). Participants were predominantly male (n = 24, 60%), Han Chinese (n = 38, 95%), single (n = 25, 63%), lived with others (n = 38, 95%), and unemployed (n = 18, 45%). Approximately half the sample had college-level or higher education (n = 22, 55%) and reported an average monthly income between 1000–5000 RMB (n = 20, 50%). Schizophrenia was the most common diagnosis received (n = 35, 87.5%). The median age at first onset was 22 years (IQR = 18–27.25). Nearly half of the participants had one prior hospitalisation (n = 18, 45%), and more than three-quarters of participants had one episode of psychosis (n = 31, 77.5%). All participants were prescribed antipsychotic medication (n = 40/40, 100%). Regarding ownership, use and awareness of DHTs, most participants (n = 38, 95%) used smartphones with the Android operating system. Smartphones, social media, and smartphone apps were the most frequently used digital technologies, with almost all participants reporting using these technologies daily, suggesting relatively high digital literacy levels in the sample. Participants showed a good level of mental health literacy as indicated by an average of 14.8 (SD = 3.0) correct answers out of the 20-item true or false mental health literacy questionnaire.[38]

Recruitment and retention

From August to October, 2022, 47 people were approached to take part, of which 41 agreed to participate (see Figure 1). Reasons for not participating included concerns that answering app questions would be a burden and family member concerns that answering app questions might be triggering for the individual. One person did not meet the inclusion criteria, leaving 40 consenting study participants
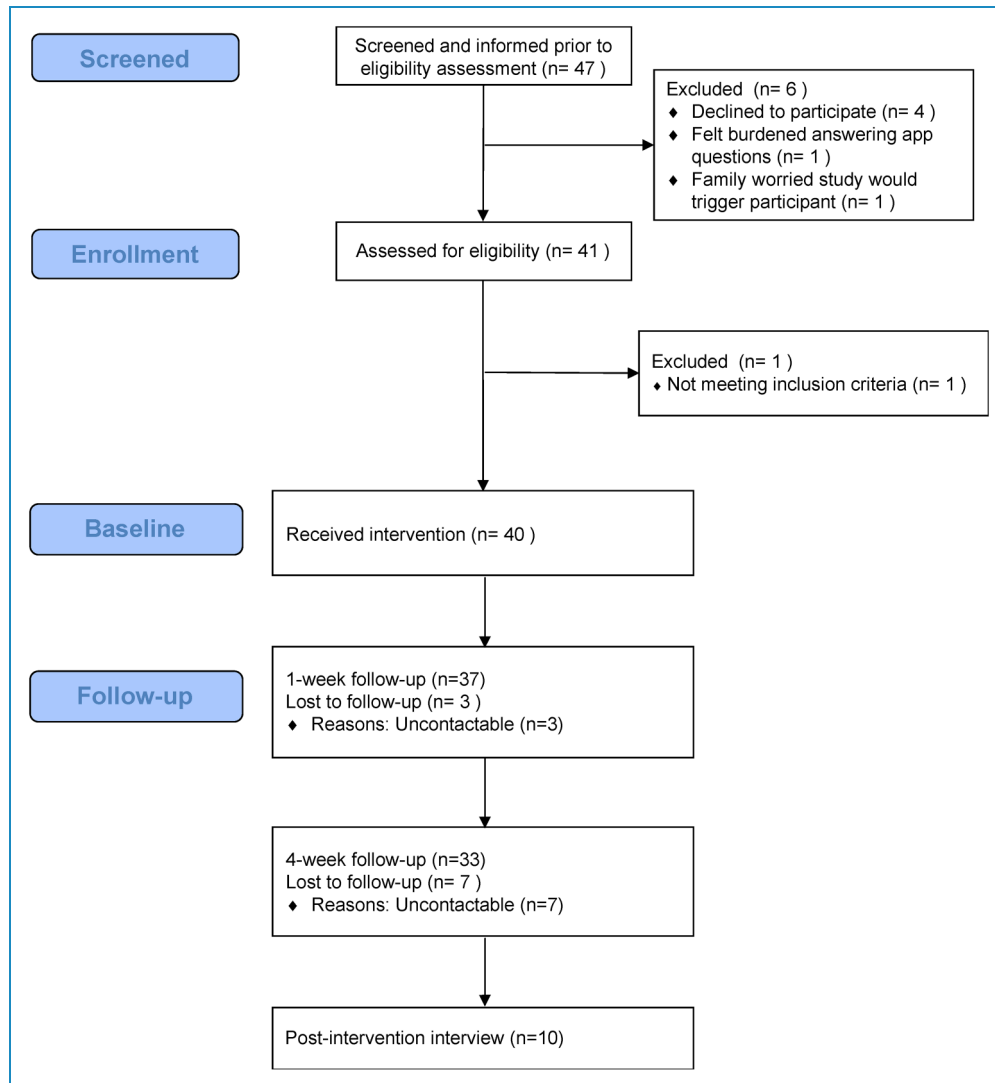
Figure 1. Flow diagram of recruitment and retention.

who completed the baseline assessment and installed the YouXin app on their smartphones. Of the 40 participants, 37 (92.5%) completed the one-week follow-up assessment, and 33 (82.5%) completed the four-week follow-up assessment. Ten participants took part in the post-intervention qualitative acceptability interview. Logistic regression indicated that age, education level, diagnosis, mental health literacy level, operating systems of the smartphone, baseline symptom severity and functioning scores were not associated with study drop-out (i.e., discontinued from the study).

ASM feasibility and completion rates

On average, participants completed 45.9% (SD = 33.7, range 0–94.6) of scheduled prompts and engaged with the app for 20.5 days (SD = 10.6, range 0–28) during the four-week exposure period (see Supplemental Figure 1 for a line graph of completion rate over four weeks). Twenty-four (60%) participants met the acceptable engagement criterion as defined by completing at least 33% of all prompts. The median time spent to complete each ASM question set was 79.5 s (IQR = 59–133.75); two prompts took >1 h to be completed, probably due to the participant using other apps (e.g., taking a phone call) while the question set was live. The average time interacting with the app, defined by the number of log-ins (i.e., opening the app on the smartphone), was 185.7 s (SD = 173.2, range 13–794).

Passive monitoring feasibility and completion rate

As we permitted participants to turn on and off the passive monitoring throughout the study, we examined the number of participants who remained using passive monitoring at the end of the study. At the four-week time point, 78% (n = 31) and 80% (n = 32) participants continued to have the GPS recording and step count function respectively enabled. There

is no consensus on the threshold of acceptable data completeness of GPS recording.[35] To explore the GPS data completeness for this study, we used the criterion of at least one data point recorded in a day (the minimum data availability threshold). Over the four-week exposure period, GPS data of 33 (82.5%) participants was recorded. The average number of hours of GPS data recorded per day was 3.41 (SD = 3.89). Seven participants recorded no GPS data, which might have been caused by turning off GPS recording at study start. For the 33 participants where GPS data was recorded, the average number of days recorded was 18.7 (SD = 9.37), with a mean GPS recording completion rate (% days recorded GPS data) of 66.78%. The mean number of days with useable GPS data, defined by recording at least 12 h GPS data including the night-time period (i.e., 9 p.m. to 6 a.m.) in a day was 2.79 (SD = 5.22). In total, 13 of the 33 participants met the threshold of sufficient GPS data to extract mobility features. Step data was collected from 35 (87.5%) participants. The mean number of step days collected was 17.06 (SD = 9.43). The mean completion rate (% days recorded step data) was 60.92%. More information regarding passive monitoring completeness is shown in Supplemental Table 5 and Supplemental Figure 2.

Acceptability

Acceptability was evaluated using the QFQ (see Supplemental Table 6 for a summary). Thirty-two participants completed the QFQ at the four-week follow-up. Overall, the YouXin app was acceptable, with a QFQ total score of 47.4 (SD = 13.6, range 26–83). Participants reported the app was easy to use and they felt comfortable tracking their symptoms using the app. Negative ratings related to the burden of answering questions (e.g., "Did answering the questions take up a lot of time?"), the difficulties of operating the devices (e.g., "Was using the touchscreen difficult?"), and distress induced by the questions (e.g., "Did filling in the questions make you feel worse?") were minimal.

Qualitative feedback was collected from 10 participants after the four-week assessment. All the participants found the YouXin app to be acceptable and easy to use. Benefits from using the app were expressed, including being more aware of their mental health problems, using the insights gained from the app to self-manage thoughts, emotions, and behaviours, and improving their ability to communicate their mental health problems with family members and mental health professionals based on the monitoring results. Participants reported the volume of the alerts sent by the app was too low, causing them to miss some of the scheduled notifications. The GPS monitoring consumed a significant amount of phone battery, which meant participants found they had to charge the phone more than usual; this was a major barrier to the effective use of the app and phone. A small number of participants said they felt uncomfortable giving 'truthful' responses to some ASM items for fear of their doctor finding out about changes in symptom patterns,

potentially indicating a deviation from anticipated recovery progress. A summary of illustrative quotes is shown in Supplemental Table 7. A more detailed analysis of qualitative interviews is reported elsewhere (Zhang et al., *in preparation*).

Safety and clinical outcomes

During the four-week study period, no AEs or SAEs were recorded. As the primary objective of the study was to assess feasibility, we report descriptive results to demonstrate the trend of changes in clinical outcomes between baseline and four-week follow-up. As shown in Table 3, no deterioration was found in both symptoms and functional outcomes, suggesting that using the app did not result in a worsening in symptoms and functioning.

Reliability, validity, and instability of active monitoring

Table 4 shows the summary statistics for active monitoring scores and corresponding clinical assessment scores at one-week follow-up, and the results of reliability, validity, and instability. Clinical assessment scores of guilt feelings, grandiosity, anxiety, somatic concern, hopelessness, and depression were generally low and less varied between participants, as most participants scored 1 on these items (0 on the hopelessness item). Cronbach's alpha was high (> 0.7) and showed good reliability, expect for the grandiosity (0.59) and hopelessness (0.62) ASM items. In terms of validity, hallucinations, suspiciousness, guilt feelings, delusions, and grandiosity items in active monitoring showed moderate Spearman's correlations (> 0.35) with the corresponding PANSS items. However, anxiety, somatic concern, hopelessness, and depression were not significantly correlated with their corresponding PANSS or CDS items. Instability was measured by MSSD and ICC; a visual description of the scores across time is shown in SupplementaL Figure 3. There was a lot of variation between participant's scores; many reported very little variation both within and between days, while some saw much greater moment-to-moment changes. Somatic concern was the most unstable item, followed by anxiety, hopelessness, and suspiciousness. In contrast, grandiosity, delusions, and guilt feelings were relatively more stable across time compared to other items.

Association between passive monitoring indicators and PANSS

The GPS indicators varied considerably between participants (see Supplemental Table 8). The time spent at home extracted from GPS data was not significantly correlated with self-reported time spent at home from the active monitoring item (rho = −0.06, p = .575). Supplemental Figure 4

Table 3. Clinical measures at baseline, 1 week and 4 weeks.

Measure	Baseline (n = 40)		1 week (n = 37)		4 weeks (n = 33)		Mean difference between baseline and 4 weeks[a]	95% CI of mean difference	P value	Cohen's D
	Mean	SD	Mean	SD	Mean	SD				
PANSS positive	12.7	6.4	12.3	6.4	10.7	5.0	−1.58	−2.56 to −0.59	0.003	−0.57
PANSS negative	15.6	6.4	15.8	6.5	15.2	6.6	−0.15	−0.76 to 0.46	0.616	−0.09
PANSS general	25.6	5.6	24.5	5.6	22.7	5.2	−2.12	−3.14 to −1.10	0.000	−0.74
PANSS total	53.8	14.9	52.5	15.7	48.6	14.4	−3.85	−5.70 to −2.00	0.000	−0.74
CDS total	1.6	1.9	1.1	1.7	0.9	1.9	−0.39	−1.22 to 0.43	0.336	−0.17
PSP	69.1	9.5	-	-	73.3	11.5	3.79	0.96 to 6.62	0.010	0.47
EQ-5D-5L VAS	69.0	23.4	-	-	75.8	17.9	5.52	-0.43 to 11.46	0.068	0.33

Measure	Number	%			Number	%	Wilcoxon signed-rank test		P value	Wilcoxon r (95% CI)
PHEs	-	-	-	-	-	-	42		0.393	0.16 (0−0.47)
Blackout	3	7.5	-	-	2	6.1	-		-	-
Arousal	9	22.5	-	-	4	12.1	-		-	-
Adhesion	23	57.5	-	-	21	63.6	-		-	-
Eudaimonic project	5	12.5	-	-	6	18.2	-		-	-

aPaired t-test.
PANSS: Positive and Negative Syndrome Scale; CDS: Calgary Depression Scale for Schizophrenia; PSP: Personal and Social Performance; EQ-5D-5L VAS: EuroQoL 5D 5 level version visual analogue scale; PHE-s, Patient Health Engagement Scale.

shows the correlation matrix between passive monitoring indicators and the negative symptom subscale of PANSS. Due to the missingness of the GPS data, the sample size available for calculating the correlations was small (n = 10). Therefore, the correlation coefficients reported here show the trend of the relationships between the variables; results should be interpreted with caution.

Discussion

This study examined the validity, feasibility, acceptability and safety of the YouXin app for symptom self-monitoring for people with psychosis in China using a non-randomised study design. The completeness of active monitoring met our pre-defined threshold (>80%); however, the data availability of passive monitoring is suboptimal with only a

small number of participants recording sufficient data to be able to extract mobility indicators. The validity of ASM was demonstrated with several domains achieving moderate correlation with corresponding gold-standard clinical assessment. However, some domains were not significantly correlated with clinical assessment, suggesting further refinement of ASM items is needed. Both the quantitative and qualitative evaluations showed high acceptability of YouXin among participants. Clinical measurements indicated no symptom and function deterioration. No adverse reactions to using the app or study procedures were reported, suggesting the app is safe to use. Taken together, the findings demonstrate YouXin is feasible, acceptable, and safe for symptoms self-monitoring for people with psychosis in China.

Retention in the study was high, with 82.5% of participants completing the four-week follow-up, which is

Table 4. Summary statistics for active monitoring YouXin scores and corresponding clinical assessment scores, and the results of reliability, validity, and instability (week 1 data).

Item	YouXin scores		Clinical assessment scores		Chronbach's alpha	Spearman's rho	MSSD (SD)	ICC
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)				
Hallucinations	1.86 (1.24)	1.24 (1.00–2.43)	1.92 (1.57)	1 (1–3)	0.92	0.58***	0.85 (1.72)	0.81
Suspiciousness	2.19 (1.42)	1.53 (1.20–2.54)	2.46 (1.73)	1 (1–3)	0.93	0.57***	0.90 (2.11)	0.82
Guilt feelings[a]	1.81 (1.42)	1.07 (1.00–1.73)	1.16 (0.55)	1 (1–1)	0.96	0.55***	0.63 (1.51)	0.88
Delusions	2.44 (1.32)	2.01 (1.39–3.14)	2.51 (1.68)	3 (1–4)	0.94	0.49**	0.60 (1.19)	0.86
Grandiosity	2.26 (1.12)	1.97 (1.38–2.58)	1.65 (1.23)	1 (1–1)	0.59	0.35*	0.59 (1.04)	0.79
Anxiety	2.46 (1.70)	1.61 (1.16–3.83)	2 (1.20)	1 (1–1)	0.93	0.25	1.41 (2.65)	0.81
Somatic concern[a]	2.52 (1.53)	2.08 (1.31–3.57)	1.38 (0.79)	1 (1–1)	0.79	0.15	2.04 (5.18)	0.70
Hopelessness[a, b]	2.16 (0.94)	2.00 (1.48–2.53)	0.14 (0.35)	0 (0–0)	0.62	0.14	1.35 (2.33)	0.61
Depression	2.81 (1.17)	2.67 (1.93–3.76)	1.35 (0.82)	1 (1–1)	0.78	0.13	0.72 (0.85)	0.75

aThese items only have two scores (i.e., binary variable), so we used Point-Biserial Correlation rather than Spearman's Correlation for these three items.
bThis item corresponds to the CDS item hopelessness.
*** $p < .001$, ** $p < .01$, * $p < .05$.

comparable to previous remote symptom monitoring studies conducted in people with psychosis.[16,18] Although none of the socio-demographic and clinical characteristics and device attributes predicted study drop-out, it should be noted that this sample is relatively young, had mild symptom severity and high digital and mental health literacy at baseline. Previous findings regarding the impacts of age and baseline symptoms on engagement have been mixed; some studies have reported that people with psychosis of older age[39,40] or with more severe symptoms[21,41] might have more difficulty using digital health technologies. The recruitment strategy is feasible in general, as we met our recruitment target (four participants per week on average). However, we did not collect clinical and demographic characteristics from those who declined participation, and we were not able to collect the overall proportion of clinician caseload recruited, which limited our understanding of recruitment feasibility. Of the seven potential participants who declined participation, one individual refused taking part due to a family member's concerns that the study would be triggering for the potential participant. As family members are important carers of people with psychosis in China,[42] further consideration of the views of family members in digital symptom monitoring research and clinical implementation is needed.

The completion rate of active monitoring is acceptable; 60% of participants met the threshold of completing at least 33% of the prompts. However, the completion rate is lower than in previous studies.[16,18] This might be because the feedback in the app was minimal, and we did not provide active human support in this study, which has been shown to be helpful in increasing engagement in DHIs in psychosis.[18,43]

Although passive monitoring required less effort from participants, the completion rate of passive monitoring was not largely different from active monitoring (mean completion rate 66.78% and 60.92% for GPS recording and step count respectively). We did not set any pre-defined threshold for the feasibility of passive monitoring due to the exploratory nature of this analysis. The completion rates (i.e., % days recorded passive data) we reported here were broad since we aimed to understand the quantity and quality of passive data collection in our study sample. Setting completion rate criteria is needed for future studies with specific clinical purposes to assess the feasibility of passive monitoring in a specific clinical context. Moreover, data missingness is common in passive monitoring.[44–46] There are many factors that could impact the collection of passive data, including types of sensor selected, frequency of sensing, participants turning off or incorrectly

setting data permissions, and smartphone operating systems turning off data collection for battery or performance reasons.[44,47] We allowed participants to turn on and off data permissions during the study. Also, some participants mentioned in qualitative interviews that the GPS monitoring consumed a significant amount of phone battery. This might explain the high missingness of passive data observed in this study. Additionally, more comprehensive training on using the app and smartphone device itself and close monitoring of data completeness may be needed to increase data collection as proposed by Currey and Torous.[44]

The acceptability of YouXin was demonstrated by both the quantitative and qualitative evaluations. The QFQ total score of YouXin is comparable to a previous study evaluating smartphone-based assessment for psychosis (i.e., QFQ total score = 53.0).[25] Qualitative interviews indicated that although some participants encountered some practical barriers (e.g., low alarm volume and charging smartphone more often), most participants found the app helpful, which is consistent with previous DHI studies.[48,49] Overall, participants felt the app was easy to use. These findings are in line with previous studies reporting that remote symptom monitoring is well tolerated by people with psychosis.[16,18]

Regarding the validity of ASM items, hallucinations, suspiciousness, guilt feelings, delusions, and grandiosity items in active monitoring showed moderate Spearman's correlations (> 0.35) with their corresponding PANSS items. Anxiety, somatic concern, hopelessness, and depression were not significantly correlated with their corresponding PANSS or CDS items. These findings were not in line with the initial ClinTouch study.[21] One potential reason could be due to the relatively mild symptom severity of the sample and the lack of variance of some clinical assessment items, as ASM domains corresponding with higher clinical assessment score variation (e.g., hallucinations) demonstrated better correlation compared to domains corresponding with lower clinical assessment score variation (e.g., hopelessness). Another reason could be the phrasing of the ASM items. We referenced the Chinese version of the PANSS and CDS when translating the ClinTouch items to ensure we used accurate Chinese terms to describe symptoms. However, these terms may be difficult to understand by participants as both the PANSS and CDS are designed to be used by trained mental health professionals. Some participants from the qualitative interviews reported that some terms in the ASM were hard to understand, which lends support to this interpretation and may explain the low validity. Additionally, qualitative interviews showed that some participants felt uncomfortable responding to prompts because they did not want their treating doctor to notice their symptoms fluctuations, which may also explain the low validity between some ASM and clinician-administered items. Furthermore, the low validity may also be explained by cultural factors, as cultural beliefs and behaviours may influence interpretations of schizophrenia pathology.[50]

All the ASM items showed relatively high MSSD and ICC, suggesting the items were sensitive to both within-participant symptom fluctuation and between-participant symptom differences. This finding supports the advantage of digital remote symptom monitoring of recording real-time symptom changes,[15] which cannot be achieved by conventional clinic-based face-to-face assessment due to practical (e.g., accessibility) and methodological (e.g., ecological validity, recall bias) challenges.[14]

Potential associations between mobility indicators extracted from passive data and clinical assessments were observed, with some passive monitoring indicators showing non-significant moderate correlations with some PANSS negative items. For example, distance travelled, number of unique locations, and number of places visited negatively correlated with emotional withdrawal and passive/apathetic social withdrawal, and time spent at home positively correlated with emotional withdrawal and passive/apathetic social withdrawal. These findings align with previous studies that reported an association between reduced GPS mobility and higher severity of negative symptoms.[51–53] However, due to the small sample size caused by missing data, no definitive conclusion can be drawn about the associations between passive monitoring indicators and negative symptoms in this sample. High levels of data completeness are crucial for ensuring the accuracy of features derived from passive monitoring data.[44] A more comprehensive investigation of the relationship between passive monitoring indicators and negative symptoms in people with psychosis in China is needed.

To our knowledge, this is the first study to develop and evaluate a digital remote monitoring tool specifically designed for people with psychosis in China. There were notable strengths. We used a mixed-methods design, which allowed us to have a more comprehensive understanding of the feasibility and acceptability of YouXin by combining quantitative and qualitative methods. In the development of the YouXin app, we implemented a co-production approach, involving input from both clinicians and service users, enabling us to optimise the app's alignment with end-user needs. Nevertheless, there are some limitations of the study. The study was primarily designed to understand the feasibility of YouXin; therefore, we did not include elements to promote engagement with either the prompts or app usage, which may have affected our ability to calculate validity parameters and to extract passive monitoring mobility indicators. A future study specifically designed for testing validity between ASM and gold-standard clinical assessment is needed after the refinement of the app to provide a more concrete conclusion. Despite the broad eligibility criteria, our sample was relatively young and had less severe symptoms, which limits our understanding of the feasibility of the Youxin app and our study procedures in an older more symptomatic sample in China. Additionally, the use of PANSS might have introduced a

floor effect in the assessment of people at risk of psychosis. An instrument specifically developed for people at risk of psychosis is needed when studying the feasibility and validity of the YouXin app in this population more specifically in the future. Moreover, as the participants had the option of switching off the passive monitoring, it could have increased the probability of missing data and introduced selection bias, as participants more concerned about privacy might choose to switch off passive monitoring. Furthermore, we recruited participants with high digital literacy, which may have led to selection bias with people who are more familiar with digital technology. With a single group design, we were not able to examine the clinical efficacy of the app.

Conclusions

In summary, this study provided an initial evaluation of the validity, feasibility, acceptability, and safety of a remote symptom self-monitoring smartphone app in people with psychosis in China. YouXin was found to be feasible, acceptable and safe among participants. Before conducting a future randomised controlled study, refinements are needed to improve the validity of active monitoring items. Strategies to increase passive monitoring data completeness are also required. Furthermore, more comprehensive feedback and a blended approach involving human support may be considered to improve participant engagement. Bearing these considerations in mind, as our feasibility parameters were met, a future powered efficacy study is warranted.

ORCID iDs: Xiaolong Zhang ⓘD https://orcid.org/0000-0003-2964-0485
Sandra Bucci ⓘD https://orcid.org/0000-0002-6197-5333

References

1. Rössler W, Joachim Salize H, van Os J, et al. Size of burden of schizophrenia and psychotic disorders. *Eur Neuropsychopharmacol* 2005; 15: 399–409.
2. Charlson FJ, Ferrari AJ, Santomauro DF, et al. Global epidemiology and burden of schizophrenia: findings from the global burden of disease study 2016. *Schizophr Bull* 2018; 44: 1195–1203.
3. Huang Y, Wang Y, Wang H, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study. *Lancet Psychiatry* 2019; 6: 211–224.
4. Ma C, Yu S, Huang Y, et al. Burden of mental and substance use disorders — China, 1990−2019. *China CDC weekly* 2020; 2: 804–809.
5. Wu Q, Zhao L and Ye X-C. Shortage of healthcare professionals in China. *Br Med J* 2016; 354: i4860.
6. Zhang X, Lewis S, Firth J, et al. Digital mental health in China: a systematic review. *Psychol Med* 2021; 51: 1–19.
7. Zhang X, Lewis S, Chen X, et al. Mental health professionals views and the impact of COVID-19 pandemic on implementing digital mental health in China: a nationwide survey study. *Internet Interv* 2022; 30: 100576.
8. Choi WT, Yu DK, Wong T, et al. Habits and attitudes of video gaming and information technology use in people with schizophrenia: cross-sectional survey. *J Med Internet Res* 2020; 22: e14865.
9. Tan Y, Teng Z, Qiu Y, et al. Potential of mobile technology to relieve the urgent mental health needs in China: web-based survey. *JMIR Mhealth Uhealth* 2020; 8: e16215.
10. Rodriguez-Villa E, Rauseo-Ricupero N, Camacho E, et al. The digital clinic: implementing technology and augmenting care for mental health. *Gen Hosp Psychiatry* 2020; 66: 59–66.
11. Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry* 2021; 20: 318–335.
12. Trull TJ and Ebner-Priemer U. Ambulatory assessment. *Annu Rev Clin Psychol* 2013; 9: 151–176.
13. Trifan A, Oliveira M and Oliveira JL. Passive sensing of health outcomes through smartphones: systematic review of current solutions and possible limitations. *JMIR Mhealth Uhealth* 2019; 7: e12649.
14. Ben-Zeev D, Brian R, Wang R, et al. Crosscheck: integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. *Psychiatr Rehabil J* 2017; 40: 266–275.

15. Torous J, Wisniewski H, Bird B, et al. Creating a digital health smartphone app and digital phenotyping platform for mental health and diverse healthcare needs: an interdisciplinary and collaborative approach. *J Technol Behav sci* 2019; 4: 73–85.

16. Lewis S, Ainsworth J, Sanders C, et al. Smartphone-enhanced symptom management in psychosis: open, randomized controlled trial. *J Med Internet Res* 2020; 22: e17019.

17. Eisner E, Bucci S, Berry N, et al. Feasibility of using a smartphone app to assess early signs, basic symptoms and psychotic symptoms over six months: a preliminary report. *Schizophr Res* 2019; 208: 105–113.

18. Gumley A, Bradstreet S, Ainsworth J, et al. The EMPOWER blended digital intervention for relapse prevention in schizophrenia: a feasibility cluster randomised controlled trial in Scotland and Australia. *The Lancet Psychiatry* 2022; 9: 477–486.

19. Carter H, Araya R, Anjur K, et al. The emergence of digital mental health in low-income and middle-income countries: a review of recent advances and implications for the treatment and prevention of mental disorders. *J Psychiatr Res* 2021; 133: 223–246.

20. Zhang X, Lewis S, Carter L-A, et al. A digital system to facilitate self-management in people with psychosis in China (YouXin): protocol for a non-randomised validity and feasibility study with a mixed-methods design. *JMIR Res Protoc*, 202312: 1–14. https://research.manchester.ac.uk/en/publications/a-digital-system-to-facilitate-self-management-in-people-with-psy (2023, accessed 4 August 2023).

21. Palmier-Claus JE, Ainsworth J, Machin M, et al. The feasibility and validity of ambulatory self-report of psychotic symptoms using a smartphone software application. *BMC Psychiatry* 2012; 12: 1–10.

22. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Br Med J* 2016; 355: i5239.

23. Miller S, Ainsworth B, Yardley L, et al. A framework for analyzing and measuring usage and engagement data (AMUsED) in digital interventions: viewpoint. *J Med Internet Res* 2019; 21: e10966.

24. Bell IH, Eisner E, Allan S, et al. Methodological characteristics and feasibility of ecological momentary assessment studies in psychosis: a systematic review and meta-analysis. *Schizophr Bull* 2023; sbad127.

25. Ainsworth J, Palmier-Claus JE, Machin M, et al. A comparison of two delivery modalities of a mobile phone-based assessment for serious mental illness: native smartphone application vs text-messaging only implementations. *J Med Internet Res* 2013; 15: e60.

26. Bucci S, Barrowclough C, Ainsworth J, et al. Using mobile technology to deliver a cognitive behaviour therapy-informed intervention in early psychosis (actissist): study protocol for a randomised controlled trial. *Trials* 2015; 16: 1–10.

27. Si T, Shu L, Dang W, et al. Evaluation of the reliability and validity of Chinese version of the mini-international neuropsychiatric interview in patients with mental disorders. *Chin Mental Health J* 2009; 23: 493–503.

28. Zheng L, Wang J, Zhang T, et al. Reliability and validity of the Chinese version of scale of psychosis-risk symptoms. *Chin Mental Health J* 2012; 26: 571–576.

29. He Y and Zhang M. The Chinese norm and factor analysis of PANSS. *Chin J Clin Psychol* 2000; 8: 65–69.

30. Zhang H and Xiao W. Measuring depression for schizophrenia - Calgary depression scale for schizophrenia. *J Int Psychiatry* 2006; 33: 8–15.

31. Zhang Y, Graffigna G, Bonanomi A, et al. Adaptation and validation of a Chinese version of patient health engagement scale for patients with chronic disease. *Front Psychol* 2017; 8: 104.

32. Tianmei S, Liang S, Yun'ai S, et al. The Chinese version of the personal and social performance scale (PSP): validity and reliability. *Psychiatry Res* 2011; 185: 275–279.

33. The EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16: 199–208.

34. R Core Team. R: A Language and Environment for Statistical Computing, https://www.R-project.org/ (2021).

35. Müller SR, Bayer JB, Ross MQ, et al. Analyzing GPS data for psychological research: a tutorial. *Adv Methods Pract Psychol Sci* 2022; 5: 25152459221082680.

36. Braun V and Clarke V. *Successful qualitative research: A practical guide for beginners*. Los Angeles, London, New Delhi, Singapore, Washington DC: Sage, 2013.

37. QSR International Pty Ltd. NVivo (Version 12), https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home (2018).

38. National Health Commission of the People's Republic of China. Mental Health Indicators Survey and Evaluation Programme, http://www.gov.cn/gzdt/2010-03/08/content_1550552.htm (2010, accessed 29 October 2021).

39. Bonet L, Llácer B, Hernandez-Viadel M, et al. Differences in the use and opinions about new eHealth technologies among patients with psychosis: structured questionnaire. *JMIR Ment Health* 2018; 5: e51.

40. Firth J, Cotter J, Torous J, et al. Mobile phone ownership and endorsement of "mHealth" among people with psychosis: a meta-analysis of cross-sectional studies. *Schizophr Bull* 2016; 42: 448–455.

41. Meyer N, Kerz M, Folarin A, et al. Capturing rest-activity profiles in schizophrenia using wearable and mobile technologies: development, implementation, feasibility, and acceptability of a remote monitoring platform. *JMIR Mhealth Uhealth* 2018; 6: e188.

42. Tsui MCM and Tsang HWH. Views of people with schizophrenia and their caregivers towards the needs for psychiatric rehabilitation in urban and rural areas of mainland China. *Psychiatry Res* 2017; 258: 72–77.

43. Killikelly C, He Z, Reeder C, et al. Improving adherence to web-based and mobile technologies for people with psychosis: systematic review of new potential predictors of adherence. *JMIR Mhealth Uhealth* 2017; 5: e94.

44. Currey D and Torous J. Increasing the value of digital phenotyping through reducing missingness: a retrospective review and analysis of prior studies. *BMJ Ment Health* 2023; 26: 1–8.

45. De Angel V, Adeleye F, Zhang Y, et al. The feasibility of implementing remote measurement technologies in psychological treatment for depression: mixed methods study on engagement. *JMIR Ment Health* 2023; 10: e42866.

46. Kiang MV, Chen JT, Krieger N, et al. Sociodemographic characteristics of missing data in digital phenotyping. *Sci Rep* 2021; 11: 15408.

47. Zhang H, Ibrahim A, Parsia B, et al. Passive social sensing with smartphones: a systematic review. *Computing* 2023; 105: 29–51.

48. Bucci S, Barrowclough C, Ainsworth J, et al. Actissist: proof-of-concept trial of a theory-driven digital intervention for psychosis. *Schizophr Bull* 2018; 44: 1070–1080.

49. Palmier-Claus JE, Rogers A, Ainsworth J, et al. Integrating mobile-phone based assessment for psychosis into people's

everyday lives and clinical care: a qualitative study. *BMC Psychiatry* 2013; 13: 1–12.

50. Khan A, Yavorsky C, Liechti S, et al. A Rasch model to test the cross-cultural validity in the positive and negative syndrome scale (PANSS) across six geo-cultural groups. *BMC Psychol* 2013; 1: 5.

51. Depp CA, Bashem J, Moore RC, et al. GPS mobility as a digital biomarker of negative symptoms in schizophrenia: a case control study. *npj Digit Med* 2019; 2: 108.

52. Mow JL, Gard DE, Mueser KT, et al. Smartphone-based mobility metrics capture daily social motivation and behavior in schizophrenia. *Schizophr Res* 2022; 250: 13–21.

53. Raugh IM, James SH, Gonzalez CM, et al. Geolocation as a digital phenotyping measure of negative symptoms and functional outcome. *Schizophr Bull* 2020; 46: 1596–1607.