Clinical Psychology Review 107 (2024) 102357

Contents lists available at ScienceDirect

Clinical Psychology Review

journal homepage: www.elsevier.com/locate/clinpsychrev

Review

Systematic review of early warning signs of relapse and behavioural antecedents of symptom worsening in people living with schizophrenia spectrum disorders

J.F. Gleeson [a,*], T.B. McGuckian [a], D.K. Fernandez [a], M.I. Fraser [a], A. Pepe [a], R. Taskis [a], M. Alvarez-Jimenez [b,c], J.F. Farhall [d], A. Gumley [e]

[a] Healthy Brain and Mind Research Centre, School of Behavioural and Health Sciences, Australian Catholic University, Fitzroy, VIC, Australia
[b] Orygen, Parkville, VIC, Australia
[c] Centre for Youth Mental Health, The University of Melbourne, Parkville, VIC, Australia
[d] Department of Psychology and Counselling, La Trobe University, Melbourne, VIC, Australia
[e] Glasgow Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK

ARTICLE INFO

Keywords:
Early signs
Prodrome
Relapse
Schizophrenia
Psychosis
Ecological momentary assessment

ABSTRACT

Background: Identification of the early warning signs (EWS) of relapse is key to relapse prevention in schizophrenia spectrum disorders, however, limitations to their precision have been reported. Substantial methodological innovations have recently been applied to the prediction of psychotic relapse and to individual psychotic symptoms. However, there has been no systematic review that has integrated findings across these two related outcomes and no systematic review of EWS of relapse for a decade.

Method: We conducted a systematic review of EWS of psychotic relapse and the behavioural antecedents of worsening psychotic symptoms. Traditional EWS and ecological momentary assessment/intervention studies were included. We completed meta-analyses of the pooled sensitivity and specificity of EWS in predicting relapse, and for the prediction of relapse from individual symptoms.

Results: Seventy two studies were identified including 6903 participants. Sleep, mood, and suspiciousness, emerged as predictors of worsening symptoms. Pooled sensitivity and specificity of EWS in predicting psychotic relapse was 71% and 64% (AUC value = 0.72). There was a large pooled-effect size for the model predicting relapse from individual symptom which did not reach statistical significance ($d = 0.81$, 95%CIs = −0.01, 1.63).

Conclusions: Important methodological advancements in the prediction of psychotic relapse in schizophrenia spectrum disorders are evident with improvements in the precision of prediction. Further efforts are required to translate these advances into effective clinical innovations.

1. Introduction

Schizophrenia spectrum disorders, which affect around 20 million people globally (James et al., 2018), most often have their onset by age 20 (Solmi et al., 2022). Whilst up to 70% of consumers who experience a first episode of psychosis achieve symptomatic remission from acute symptoms within the first year of treatment (Phahladira et al., 2020), the majority experience deficits in psychological wellbeing and social functioning that persist beyond the first year (Santesteban-Echarri et al., 2017). Unfortunately, by 3–4 years follow-up just over half of consumers diagnosed with a first-episode psychosis will experience a psychotic relapse (Alvarez-Jimenez et al., 2012; Phahladira et al., 2020). In addition to being inherently distressing for consumers and their families, relapses interrupt hard-won momentum attained in psychosocial recovery (Kane, 2007; Takeuchi et al., 2019). Furthermore, each relapse conveys a risk for the progression to persistent psychotic symptoms (Emsley, Chiliza, & Asmal, 2013), post-psychotic depression (Jager, Hintermayr, Bottlender, Strauss, & Moller, 2003) and fear of relapse itself (Zukowska, Allan, Eisner, Ling, & Gumley, 2022). For health services, relapses markedly increase the cost burden of treating schizophrenia primarily due to rehospitalization (Pennington & McCrone, 2017). Therefore, the prediction and prevention of relapses are critical

* Corresponding author at: Healthy Brain and Mind Research Centre, Australian Catholic University, Level 5, Daniel Mannix Building, Young Street, Fitzroy, VIC 3065, Australia.
E-mail address: john.gleeson@acu.edu.au (J.F. Gleeson).

https://doi.org/10.1016/j.cpr.2023.102357
Received 8 March 2023; Received in revised form 3 November 2023; Accepted 14 November 2023
Available online 22 November 2023

priorities in clinical practice and research.

Theoretical accounts of relapse have ranged in their level of explanation from broad high-level integrated explanations (e.g., stress-vulnerability model (Nuechterlein, Snyder, & Mintz, 1992)) to more specific biological (e.g., the dopamine dysregulation hypothesis of relapse (Emsley, Chiliza, Asmal, & Harvey, 2013)), and psychological (e.g., cognitive subsystems model of relapse (Gumley, White, & Power, 1999)) accounts. These models have informed hypotheses regarding who is at risk and when relapse is most likely.

In terms of predicting who is most at risk for relapse beyond the first episode, established risk factors include poor premorbid adjustment, medication nonadherence, persistent substance use disorder, and exposure to unsupportive interpersonal environments (Alvarez-Jimenez et al., 2012). Fortunately, maintenance antipsychotic medications, compared to placebo, (Ceraso et al., 2020) and individual CBT and family interventions (usually added to antipsychotic medication) significantly reduce the risk of relapse (Bighelli et al., 2021; Rodolico et al., 2022). However, antipsychotic medications produce a range of aversive side-effects which reduce adherence (Schneider-Thoma et al., 2022) and the availability of effective family interventions is sparse (Eckardt, 2022). In short, in the real-world treatment context the risk of relapse and rehospitalisation remains a significant problem (Jorgensen et al., 2021).

An adjunctive strategy, typically used in combination with maintenance antipsychotic medication, is to identify when relapse is imminent so that timely preventive interventions can be implemented (Birchwood, Spencer, & McGovern, 2000). Initial empirical support for this strategy was derived from observational studies which showed that early warning signs (EWS) of relapse unfolded over the course of days and weeks, with specific signs and symptoms entailing anxiety, dysphoria, insomnia, concentration problems, and attenuated psychotic symptoms (Birchwood et al., 1989). These findings provided evidence of a window of time for the marshalling of personal and clinical responses, however, a previous review concluded that the effectiveness of EWS monitoring is limited by their modest predictive validity (median sensitivity 61%; median specificity 81%) (Eisner, Drake, & Barrowclough, 2013).

Whilst earlier research into EWS relied upon weekly to monthly repeated administration of retrospective self-report surveys and standardised clinical rating instruments, over the last decade there has been rapid growth in the deployment of novel methodologies to track EWS and other proximal behavioural antecedents of relapse in real-time. These include ecological momentary assessment (EMA) and ecological momentary intervention (EMI) using self-report in the context of daily life which reduces the risk of forgetting and biases associated with recall (Myin-Germeys, Klippel, Steinhart, & Reininghaus, 2016; Shiffman, Stone, & Hufford, 2008). In addition, the use of data collected passively from personal mobile digital smartphones and other devices has recently introduced the added benefit of real-time "passive sensing" of idiographic behavioural changes immediately preceding relapse – these methods reduce the burden on participants from completing frequent surveys and address the biases inherent in self-report (Trifan, Oliveira, & Oliveira, 2019).

These recent methodological innovations offer the prospects of major advances in both the understanding and prevention of reoccurrences of psychosis. Grounded in ecological psychology, EMA has the capacity to generate large intensive longitudinal data sets which enable a significantly enriched understanding (both nomothetical and idiographic) of fluctuations in the severity of psychosis in relation to the everyday context (Myin-Germeys et al., 2018). The capability of EMA to also provide individuals with new insights into their symptoms via feedback mechanisms has also spawned a rapid growth in the development and evaluation of personalised digital EMIs which promise to empower individuals diagnosed with psychosis to take an active role in the management of their mental health (Bell, Lim, Rossell, & Thomas, 2017).

Given these advances over the last 10 years, an updated and integrated synthesis of the evidence for EWS and other behavioural antecedents of relapse in schizophrenia spectrum disorders is timely and required. For example, these newly emerging methods allow high frequency bursts of passive sensing and self-report measurement, enabling the identification of new antecedents of subtle worsening in the severity of psychotic symptoms over the course of hours or a day. Furthermore, researchers have recently noted that these antecedents to subtle increases in symptom severity appear to overlap with EWS of full-blown relapse (Lüdtke, Moritz, Westermann, & Pfuhl, 2022). However, there has been no systematic integration of these findings to date. Therefore, we sought to synthesise the evidence pertaining to EWS of psychotic relapse (usually over the course of days to weeks) and behavioural antecedents to worsening in psychotic symptoms (usually over the course of hours or a day) which typically fall below the threshold for a full-blown relapse. Hence, the primary objective of this review was to synthesise the evidence for EWS and other relevant antecedent behaviours and changes in psychological states as determined from studies utilising intensive repeated measurement and/or EMA/EMI in people diagnosed with schizophrenia spectrum disorders. Specifically, self-report data and passive sensing data from personal digital devices (e.g., geospatial data, call and text frequency data, social media usage) were within scope.

The secondary objectives were: 1) to systematically identify investigations of EWS and other behavioural antecedents of psychotic relapse including the methods used for assessing EWS, psychotic relapse and antecedents of worsening in psychotic symptoms; and 2) to evaluate the quality of studies that have investigated EWS of relapse and other antecedents of worsening in psychotic symptoms, as measured by EMA/EMI, passive sensing, and/or routine intensive repeated measurement.

2. Method

This review was designed and conducted in accordance with the Preferred Reporting Items for Systematic Reviews (PRISMA) Statement (Moher, Liberati, Tetzlaff, Altman, & Grp, 2009) and was registered with Prospero (ID=CRD42021225532) (https://www.crd.york.ac.uk/PROSPERO/). We conducted searches of peer-reviewed original investigations in EMBASE, MEDLINE, PsychINFO, Web of Science, and SCOPUS. The database search was conducted in September 2021, with two updated searches conducted in November 2022 and July 2023 that covered the interim periods. Additional information sources included reference lists of relevant reviews, publication lists of prominent authors, and reference list of included articles. Where clarification of individual studies was required (e.g., potential overlap in data across two or more published reports), corresponding authors were contacted directly via email to request further information.

2.1. Study identification and eligibility

Eligible studies included: papers published in English in peer reviewed journals; original empirical reports of EMA/EMI and/or routine intensive repeated measurements and/or digital indicators designed to detect EWS or clinical antecedents to psychotic relapse and exacerbation in psychotic symptoms in populations diagnosed with schizophrenia spectrum disorders. Eligible study designs included intervention studies, observational studies, and case report/case study designs. Both the study design and analysis of eligible papers (e.g., time-lagged modelling of predictors) needed to allow for conclusions regarding antecedents as opposed to cross-sectional effects only. We excluded studies published in languages other than English, non-human studies, studies that were exclusively qualitative, papers that were not peer reviewed, conference abstracts, and literature reviews. The date of publication was not limited.

For clarity, we excluded studies with mixed diagnoses (schizophrenia spectrum and bipolar disorder) unless separate findings were reported for participants with schizophrenia spectrum disorders. Specific diagnoses included schizophrenia, schizoaffective disorder, delusional disorder, schizophreniform disorder, brief psychotic disorder, as

well as psychosis associated with substance use or medical conditions (APA, 2013). These disorders entail acute episodes of psychosis, therefore, studies that recruited participants with a primary diagnosis of schizotypal personality disorder were excluded. To reduce study heterogeneity, schizophrenia spectrum disorders needed to be the primary diagnosis, therefore psychosis secondary to mood disorders were not included. Samples with a putative risk of psychotic illness were also excluded (e.g., ultra-high risk for psychosis (Yung, Phillips, Yuen, & McGorry, 2004)). To preserve heterogeneity in the age range, the mean age of participant samples needed to be over 16 years and we excluded studies sampling exclusively from populations with childhood onset and from the older aged populations, i.e., with mean age $\geq$ 65 years.

2.2. Measurement features

We included studies that utilised EMA/EMI and/or routine intensive repeated measurement utilising self-report instruments and/or passive sensing data from personal digital devices. Studies that utilised either passive digital sensing or intensive self-report sampling (i.e., measurements completed at least daily) to detect worsening in psychotic symptoms needed to be at least 5 days in duration, consistent with opinions regarding the minimum duration required for assessing highly variable symptoms (Myin-Germeys et al., 2018). Other studies that used self-report instruments, repeated clinical assessments or other methods to predict and detect psychotic relapse needed to sample at least once every two weeks over a duration of at least three months to ensure an adequate likelihood of detecting EWS and relapses given the established duration of the relapse prodrome (Birchwood et al., 1989). We did not specify a maximum duration of follow up.

2.3. Main outcomes

Our primary outcomes were psychotic relapse, exacerbation of psychosis, or worsening of psychotic symptoms. Considering the diversity of definitions for psychotic relapse in the literature (Gleeson, Alvarez-Jimenez, Cotton, Parker, & Hetrick, 2010), we did not exclude studies based on any specific definition of psychotic relapse. Exacerbation of psychosis, usually measured alongside relapse, was defined as worsening in psychotic symptoms that were sub-threshold for relapse (e.g., did not meet the severity thresholds for a full-blown relapse). Worsening in psychotic symptoms was defined as any degree of increased severity in a specific psychotic symptom (e.g., hallucinations) or positive psychotic symptoms overall without reference to any threshold or criteria for a relapse.

We specified additional outcomes as antecedents of psychotic relapse, which may include early warning signs of relapse and fear of relapse, as well as any other relevant variables that are identified within the literature.

2.4. Data extraction

Data extraction included demographics, study inclusion and exclusion criteria, study recruitment and sampling process, study design, study aims, measures, intervention characteristics and treatment conditions (where applicable), outcome data at each time point, limitations, and implications.

The data extraction template was piloted independently by two reviewers (DF, JG) over three studies. Data extraction was then completed independently and with duplication by JG and DF. Reviewers successfully resolved discrepancies by consensus. We planned to contact corresponding authors to resolve any uncertainties - one author was contacted.

2.5. Quality assessment

We used a combination of tools to assess the methodological quality and risk of bias in individual studies. For studies that use EMA/EMI, the Checklist for Reporting on EMA Studies (CREMAS) was used to assess quality of reporting (Liao, Skelton, Dunton, & Bruening, 2016). We assigned ratings of 'yes', 'no', 'not reported' or 'not applicable' to each checklist item. Risk of bias was primarily assessed using the NIH quality assessment tool for observational and cross-sectional studies (https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools) (National Heart Lung and Blood Institute, 2023). As recommended by PRISMA, we did not calculate an overall risk of bias score for individual studies, but instead reported on risk of bias domains.

2.6. Data synthesis and meta-analysis

We undertook a narrative synthesis of findings. We initially planned to conduct a meta-analysis specifically in relation to investigations of psychotic relapse if there were at least four studies utilising similar study designs and measuring the same outcome (e.g., hallucinations). We subsequently loosened this criterion to at least three studies to enable analysis of effects across symptoms. Meta analyses were planned to 1) calculate effect sizes that demonstrated how specific symptom domains were related to psychotic relapse, and if possible 2) aggregate relevant statistics related to sensitivity and specificity of EWS. All analyses were conducted in R version 4.2.2.

2.6.1. Relationship between symptom domains and psychotic relapse

Separate effect sizes (Cohen's *d*) were calculated for each effect for each symptom domain (Cohen, 1992). To calculate effects sizes, mean, standard deviations, and sample sizes were utilised when available. Otherwise, effect sizes were estimated using *p* value and sample size using the *compute.se* package (Del Re, 2013). If *p* values were not reported, a *p* value of 0.50 was assumed for non-significant results and 0.05 for statistically significant results. Only one effect was calculated using an assumed *p* value. Sensitivity analysis showed that removing this outcome from analyses did not change the results, and therefore this outcome was retained in the reported results.

Using the *meta* package (Balduzzi, Rucker, & Schwarzer, 2019) we utilised a random effects multi-level meta-analysis to allow for the possible non-independence of effects with three levels, i.e., participants (level 1), nested within specific effects (level 2), nested within studies (level 3) (Fernandez-Castilla et al., 2020). This approach also enabled us to evaluate heterogeneity separately at levels 2 and 3. We also included adjustments for the small number of studies using the Hatung-Knapp-Sidik-Jonkman method to estimate pooled effects (Hartung, 1999).

We estimated the statistical heterogeneity of studies by calculating point estimates from the I^2 statistic where heterogeneity estimates of 0–40% 'may not be important', 30–60% may represent 'moderate' heterogeneity, 50–90% may represent 'substantial' heterogeneity, and 75–100% may represent 'considerable' heterogeneity (Higgins et al., 2019).

To assess the relative benefits of selecting a 3-level versus 2-level probabilistic model, we utilised the ANOVA function to compare Akaike and Bayesian information criteria outcomes. If the 3-level model provided an improved fit, the model would be retained. An overall pooled ES was calculated in addition to pooled estimates at the symptom domain level. Positive values indicated that an increase in the symptom domain was associated with relapse status. Significant pooled ESs were indicated by 95% confidence intervals that did not cross zero. The magnitude of each combined estimate was interpreted according to Cohen's (1992) descriptors of 0.3 (small), 0.5 (moderate), 0.8 (large), >1.0 (very large).

We utilised a funnel plot to assess potential study bias (Borenstein, Hedges, Higgins, & Rothstein, 2009). If sufficient studies were available, subgroup analyses were planned based on study methodology (e.g., observational vs intervention studies) or measurement type (e.g., EMA/EMI vs intensive repeated measures vs passive sensing) to reduce heterogeneity and to explore potential group differences.

2.6.2. Sensitivity and specificity of EWS

Univariate analyses were conducted to pool sensitivity, specificity, and diagnostic odds ratio outcomes from studies that reported sufficient summary data (i.e., true positive, true negatives, false positives, and false negatives)(Hartzes & Morgan, 2019; Shim, Kim, & Lee, 2019). Again, multi-level random-effects models with logit transformation were used given the non-independence of outcomes resulting from some studies testing more than one indicator of early warning signs. Further, bivariate analysis was conducted to estimate a summary receiver operating characteristic curve (sROC) for diagnostic test accuracy using the *mada* package (Doebler, Holling, & Sousa-Pinto, 2015). Insufficient data were available to conduct subgroup analyses.

3. Results

3.1. Study selection and characteristics

As depicted in Fig. 1 a total of 4129 references were imported into Covidence from the database searches. After de-duplication of 950 abstracts, 3179 studies remained for title and abstract screening. From these, 2792 abstracts were screened out as non-relevant, leaving 387 articles for full text screening. From these, 61 were eligible for extraction. A search through reference lists from eligible papers produced an additional three studies that matched our inclusion criteria (Gaebel & Riesbeck, 2014; Marder et al., 1994; Subotnik & Nuechterlein, 1988). An updated search conducted in November 2022 produced an additional 70 articles, five of which were eligible for inclusion (Daemen, van Amelsvoort, Group, & Reininghaus, 2022; Lüdtke et al., 2022; Postma et al., 2021; Radley, Barlow, & Johns, 2022; Zhou, Lamichhane, Ben-Zeev, Campbell, & Sano, 2022). A second updated search was conducted in July 2023 which produced a further 228 articles, three of which were eligible for inclusion (Allan et al., 2023; Cohen et al., 2023; Lamichhane, Zhou, & Sano, 2023). There was a final tally of seventy-two studies.

These studies entailed a total of 6903 participants with a mean sample size of 95.88 (125.93) and range 1–907. All papers were published between 1989 and 2023 and were conducted in USA (k = 29), The UK (k = 15), The Netherlands (k = 6) Germany (k = 7), Denmark (k = 2), Australia (k = 1), multiple locations (k = 9), and individual studies conducted in France, Hong Kong, and Czech Republic. The majority recruited from populations with long standing diagnoses. In addition,

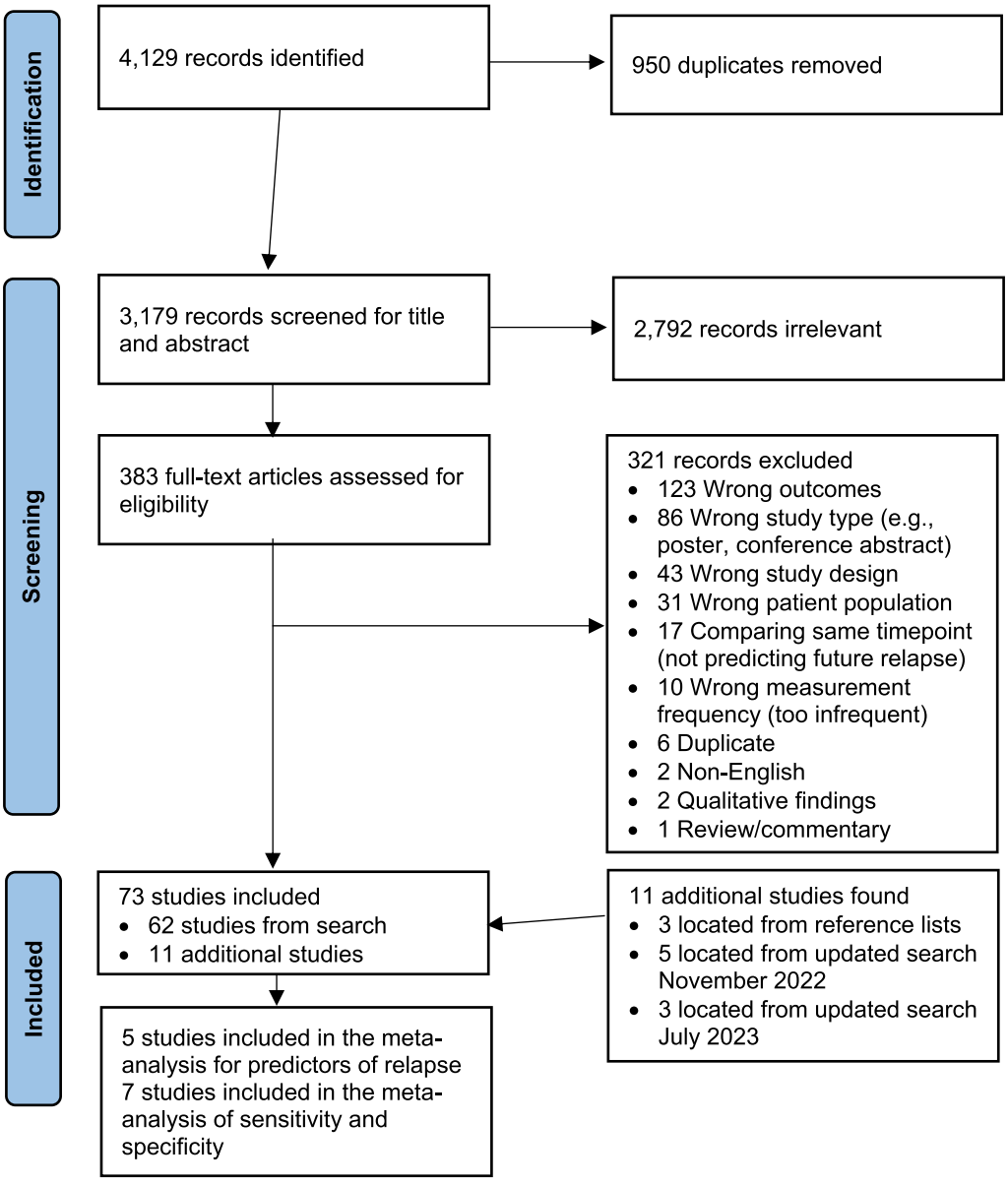

Fig. 1. PRISMA flow diagram of the study selection process.

two studies recruited exclusively from the first-episode psychosis population and two studies recruited a subgroup of participants from the first-episode population. Twenty-seven (37%) of studies reported the race of participants.

As shown in Appendix A, 17 studies utilised traditional repeated administration of standardised self-report or clinical symptom severity assessment tools (e.g., Early Signs Scale, PANSS) entailing 2550 participants, 52 studies utilised EMA, EMI or passive sensing methods with 4163 participants, and three used other approaches (e.g., passive sensing of social media use, internet search patterns) with 190 participants.

The studies varied in terms of their stated aims as can be seen in Appendix A. Twenty-five studies (35%) aimed to investigate the predictive validity of early warning signs of relapse, and four (6%) tested a specific hypothesis regarding a mechanism of relapse. Nine (13%) tested a specific hypothesis regarding worsening in psychotic symptoms broadly and ten (14%) aimed to explore predictors of symptom worsening. Sixteen (22%) tested specific hypotheses regarding paranoia or delusions, nine (13%) tested specific hypotheses regarding hallucinations and six studies (8%) had other stated aims (e.g., testing feasibility of EMI methods).

In relation to outcomes, 32 studies measured relapse, psychotic exacerbation or clinical deterioration as the primary outcome and 40 measured worsening of symptoms as an outcome. For the relapse studies the duration of follow-up ranged from 3 to 30 months, however, it was not clearly specified in four studies. A majority of the studies that predicted worsening in psychotic symptoms entailed one measurement wave over 6–7 days (k = 25), one study included 2 waves of 6 days, and two studies included 3 waves of 6–10 days. The maximum follow-up duration for the symptom worsening studies was 12 months (k = 2). The frequency of surveys per day in the ESM studies that predicted symptom worsening ranged from <1 to 12 with a mode (k = 22) of 10 surveys per day. Next, we describe the study features and report the pattern of findings for the relapse outcome studies followed by the symptom worsening studies.

3.2. Study features and findings: relapse studies

Appendix A contains the detailed outcomes for all studies and Table 1 maps the measurement methods, relapse criteria specified, and the outcomes in the subset of 32 studies for a range of predictors of psychotic relapse.

In terms of measurement methods in the relapse studies, 16 utilised repeated administration of standardised clinical assessment tools (i.e., self-report questionnaires or clinically administered tools), 11 studies used EMA methods, 11 used passive sensing, and three used analysis of internet search patterns. Nine studies deployed both EMA and passive sensing methodologies. Of note, the relapse studies that incorporated EMA deployed these methods over long periods with the total duration comprising a year in nine of the eleven studies.

Twenty-one studies included changes on standardised symptom measurement instruments in their definition of relapse, however, seven relied upon hospitalization alone, three relied upon clinical judgment, and one relied on changes to medication to determine a relapse.

In terms of study outcomes utilising measures of EWS, the Early Signs Scale (ESS) was predictive of relapse in five studies however, the Early Signs Questionnaire (ESQ) did not predict relapse in three studies and the other methods of measuring prodromal symptoms predicted relapse in one study, failed to predict relapse in a second study, and produced uncertain findings in a third study.

Four studies proposed specific hypotheses regarding a mechanism of relapse. Each of these hypotheses was supported, specifically, that fear of recurrence (Gumley et al., 2015), individual residual symptoms (Saito et al., 2020) changes in social behavior (Buck et al., 2019) and negative affect and aberrant salience would predict subsequent relapse (Lüdtke et al., 2022).

Sensitivity of EWS in predicting relapse ranged widely from 10% (Gaebel et al., 2000) to 95% (Gaebel & Riesbeck, 2014) and specificity ranged from 38% (Gaebel & Riesbeck, 2007) to 93% (Gaebel et al., 2000). One challenge for the field has been large trade-offs between sensitivity and specificity, and, as has previously been observed, there were heterogenous approaches to specifying thresholds (often applied retrospectively) on these scales and varying timeframes for analyses (Eisner et al., 2013). In summary, EWS questionnaires have produced mixed results in relation to predicting psychotic relapse.

Changes in clinician administered measures of positive psychotic symptoms (i.e., The BPRS and PANSS) predicted relapse in four studies. In other words, there is evidence that small changes in positive psychotic symptoms including hallucinations, suspiciousness, and conceptual disorganization detected by standardised clinical symptom severity interviews is predictive of full-blown relapse.

In terms of the findings from the EMA and passive sensing data, in eight instances the combination of EMA self-report with passive sensing data provided the best predictor of relapse. These predictors reflected statistical anomalies in the idiographic time series (e.g., complex combinations of changes in self-reported mood and changes in physical and geospatial activity) or clustered features of changes. In term of studies that only used EMA methods, changes in self-reported mood (k = 2), paranoia (k = 3) and aberrant salience (k = 1) were predictive of relapse.

Of note, combining EMA and passive sensing data to detect paired anomalies in the data has produced a combined sensitivity and specificity of 89% and 75%, respectively (Henson, D'Mello, Vaidyam, Keshavan, & Torous, 2021).

3.3. Study features and findings: Symptom worsening studies

Table 2 maps the findings specifically for the symptom worsening studies. The symptom worsening studies fell into three categories of outcomes: 1) delusions/persecutory ideation/paranoia (k = 25); 2) hallucinations (k = 12); or 3) positive symptoms measured as a cluster/dimension of symptoms (k = 12). In terms of study methods, 36 used EMA, 7 used passive sensing, 2 used real-time cognitive assessment, and 1 used self-report surveys. Six of the EMA studies also used passive sensing.

Next, we map statistically significant and non-significant findings in relation to specific predictors of worsening against the three symptom outcome categories.

3.3.1. Prediction of worsening in delusions/paranoia/ suspiciousness

There were 25 studies that examined predictors of worsening in delusions/paranoia/suspiciousness. Negative affect predicted worsening in four studies, sadness/depression predicted worsening in two studies, reduced positive affect predicted worsening in one study and increased mood predicted worsening in one study. Negative affect did not predict worsening in two studies and sadness and depression failed to predict worsening in two studies. Anxiety predicted worsening in one study but failed to predict worsening in three studies. Worry and rumination was predictive in two studies and did not predict worsening in one study. Emotion regulation/ experiential avoidance predicted worsening in two studies, in one study it did not predict worsening, and in one study variable findings were produced.

Poor sleep predicted worsening in four studies and change in sleep predicted worsening in one study. Stress and stressful events were predictive across three studies.

Low self-esteem was predictive of worsening in two studies; feeling confident did not predict relapse in one study. Attachment insecurity predicted worsening in two studies.

Auditory or visual hallucinations predicted worsening in two studies, increased EWS predicted worsening in one study, and suspiciousness in one study. Conversely, feeling supported did not predict worsening in one study. Use of substances and thought control each predicted

Table 1
Methods, relapse criteria, and relapse predictor outcomes of included studies.

	Method				Relapse criteria				Predictor											
Study	RCA	EMA	PS	O	H	Δmeds	CJ	ΔSS	ESS	BS	ESQ	EWSQ	OPS	FoR	BPRS	PANSS	MDDPD	EMAQ	SMPD	IS
Adler et al., 2020; USA		✓	✓		✓	✓		✓									X			
Bak, Drukker, Hasmi, & Van Jim, 2016; UK		✓				✓												✓		
Barnett et al., 2018; USA		✓	✓		✓												✓	✓		
Ben-Zeev et al., 2017; USA		✓	✓		✓												✓	✓		
Birchwood et al., 1989; UK	✓				✓		✓	✓	✓											
Birnbaum et al., 2019; USA				✓	✓														✓	
Birnbaum et al., 2020; USA				✓	✓															✓
Birnbaum et al., 2020; USA				✓	✓															✓
Buck, Scherer, et al., 2019; USA			✓		✓	✓	✓	✓									✓			
Buck et al., 2021; USA		✓			✓	✓		✓										✓		
Cohen et al., 2023; USA and India		✓	✓		✓	✓	✓	✓									✓	✓		
Eisner et al., 2019; UK	✓				✓	✓		✓	✓	✓										
Gaebel & Riesbeck, 2014; Germany	✓							✓					?							
Gaebel et al., 2000; Germany	✓				✓			✓			X									
Gaebel & Riesbeck, 2007; Germany	✓				✓			✓			X									
Gumley et al., 2015; UK;	✓							✓	✓					✓						
Henson et al., 2021; USA		✓	✓		✓		✓	✓									✓	✓		
Jorgensen, 1998a; Denmark	✓				✓			✓	✓											
King & Shepherd, 1994; UK;	✓						✓		✓											
Lahti, Wang, Pei, Baker, & Narayan, 2021; USA		✓	✓			✓		✓									✓			
Lamichhane et al., 2021; USA		✓	✓				✓										✓	✓		
Lamichhane et al., 2023; USA			✓		✓			✓									✓			
Lüdtke et al., 2022; Germany	✓				✓		✓	✓										✓		
Marder et al., 1991; USA;	✓							✓			X				✓					
Marder et al., 1994; USA	✓							✓					X							
Saito et al., 2020; USA	✓				✓		✓	✓							✓					
Spaniel et al., 2018; Czech Republic	✓				✓							✓								
Subotnik and Nuechterlein (1988); USA	✓							✓							✓					
Tait, McNay, Gumley, & O'Grady, 2002; UK	✓							✓					✓							
Wang et al., 2018; USA	✓				✓			✓								✓				
Wang et al., 2020: USA		✓	✓		✓	✓		✓									✓	✓		
Zhou et al., 2022; USA		✓	✓		✓	✓	✓	✓									✓	✓		

Note: Ticks = statistically significant finding, crosses = a non-significant finding;? = unable to determine from array of results; RCA = Repeated clinical assessment; EMA = Ecological momentary assessment; PS = passive sensing; O = other; H = hospitalization; Δmeds = change to medication; CJ = clinician judgment; ΔSS = symptom score change; ESS = The Early Signs Scale; BS = basic symptoms; ESQ = Early Symptom Questionnaire; EWSQ = Early Warning Signs Questionnaire; OPS = other prodromal symptoms; FoR = fear of recurrence; BPRS = Brief Psychiatric Rating Scale; PANSS = The Positive and Negative Syndrome Scale; MDDPD = mobile digital device passive data; EMAQ = Ecological momentary assessment questions; SMPD = social media passive data; IS = internet search.

Table 2
Methods, outcomes measured, and outcomes for predictors for symptom worsening studies.

	Method				Outcome				Predictor by Outcome		
Study	Survey	EMA	PS	RtCOGAx	D/Pi/P	PosSx	H		D/Pi/P	PosSx	H
Allan et al., 2023; Australia and UK		✓				✓			SleepΔ ↑ Anx X FoR X NA X Conf X Supp X		
Jorgensen, 1998b; Denmark	✓					✓			EWS ↑		
Ben-Zeev, Ellington, Swendsen, & Granholm, 2011; USA		✓				✓			Sad/De ↑ Anx ↑ SU ↑		
Ben-Zeev, Morris, Swendsen, & Granholm, 2012; USA		✓				✓			Sad/De X VHS/Ahs ↑ Anx X		
Ben-Zeev, Frounfelker, Morris, & Corrigan, 2012; USA		✓					✓			PA X NA X Self-Stigma X	
Daemen et al., 2022; The Netherlands		✓				✓	✓		LowSE ↑		
Buck et al., 2019; USA		✓	✓			✓			Sad/De ↑ VHS/AHs ↑ Str/Se ↑ LdTrav ↑ FeSoc/SocSh ↑ PSleep ↑ LTsPhone ↑ TsSit ↑ LTsVeh ↑		
Dupuy et al., 2021; France				✓			✓			TsComTests ↑	
Geraets et al., 2020; The Netherlands		✓				✓			Sad/De X Anx X Alone X Unsafe X Sus/D/Tc ↑		
Hartley, Haddock, Sa, Emsley, & Barrowclough, 2014; UK		✓				✓		✓	W/Rum ↑		W/Rum ↑
Hartley, Haddock, Sa, Emsley, & Barrowclough, 2015; UK		✓				✓		✓	Tc ↑		Tc ↑
Hays, Keshavan, Wisniewski, & Torous, 2020; USA		✓		✓			✓			Sad/De ↑ Anx ↑ PSleep ↑ RtCogAx X	
Henquet et al., 2010; The Netherlands		✓						✓	THC/SU X		THC/SU ↑
Hermans et al., 2020; Belgium, The Netherlands		✓				✓		✓	PA ↓ NA X		PA X NA ↑
Kammerer, Mehl, Ludwig, & Lincoln, 2021; Germany		✓	✓			✓			PSleep ↑^		
Klippel et al., 2018; data from 6 studies		✓				✓			Str/Se ↑ Alone X		
Klippel et al., 2021; data from 6 studies		✓			✓					Str/Se X	
Lüdtke et al., 2021; Germany and Switzerland		✓			✓	✓	✓		NA ↑ PSleep ↑ Cb X W/Rum ↑	NA X W/Rum ↑	NA X PSleep ↑ Cb X W/Rum X
Ludwig, Mehl, Schlier, Krkovic, & Lincoln, 2020; Germany		✓			✓				NA ↑ ERS/EA X		
Meyer et al., 2021; UK		✓			✓	✓	✓		PSleep ↑	PSleep ↑	PSleep ↑
Mulligan, Emsley, Haddock, Neil, & Kyle, 2016; UK		✓	✓		✓		✓		PSleep ↑		PSleep ↑
Myin-Germeys, Nicolson, & Delespaul, 2001; The Netherlands		✓			✓				PP/Cont ↓^ Loc X Δcom X Act ↓ Act X***		
Nittel et al., 2018; Germany		✓			✓				W/Rum X NA ↑^ FeSoc/SocSh X ERS/EA ↑ ERS/EA X**		
Oorschot et al., 2012; The Netherlands		✓			✓				M ↑		
Oorschot, Lataster, Thewissen, Wichers, & Myin-Germeys, 2012; The Netherlands and Belgium		✓					✓				PA ↓ NA ↑^ Sus/D/Tc ↑
Postma et al., 2021; UK		✓				✓				LowSE ↑	

(*continued on next page*)

Table 2 (*continued*)

Study	Method				Outcome			Predictor by Outcome		
	Survey	EMA	PS	RtCOGAx	D/Pi/P	PosSx	H	D/Pi/P	PosSx	H
Radley et al., 2022; UK		✓				✓			Str/Se ↑	
Raugh et al., 2020; USA		✓	✓			✓			Δd X	
Sa, Wearden, Hartley, Emsley, & Barrowclough, 2016; UK		✓				✓			PP/Cont X	
Sitko, Varese, Sellwood, Hammond, & Bentall, 2016; UK		✓			✓			AtacSec/In ↑		
So et al., 2021; Hong Kong		✓					✓			NA ↑
Swendsen, Ben-Zeev, & Granholm, 2011; USA		✓				✓			THC/SU ↑	
Torous et al., 2018; USA		✓	✓			✓			DataQual X	
Tseng et al., 2020; USA		✓	✓				✓			PhLogs/ PhAct ↑
Udachina, Varese, Myin-Germeys, & Bentall, 2014; UK		✓			✓			ERS/EA ↑ LowSE ↑		
Vaessen et al., 2019: Europe		✓			✓			Se ↑		
Varese, Udachina, Myin-Germeys, Oorschot, & Bentall, 2011; UK		✓					✓			Dis ↑ Sus/D/Tc ↑ EA X
Wang et al., 2016; USA			✓		✓		✓	LdTrav ↑ PhLogs/ PhAct ↑* PhLogs/ PhAct ↓*		LdTrav ↑ PhLogs/ PhAct ↑* PhLogs/ PhAct ↓* SleepStart ↑
Wigman et al., 2015; The Netherlands		✓			✓			AtacSec/In ↑ NA↑		
Bell et al., 2018: Australia		✓					✓			Act ↑ EngAHs ↑

Notes: Ticks indicate that method/outcome was measured in the study; Arrows indicate direction of associations; EMA = Ecological momentary assessment; PS = passive sensing; RtCOGAx = real-time cognitive assessment; D/Pi/P = delusions/persecutory ideation/paranoia; Pos Sx = positive symptoms; H = hallucinations; EWS = Early warning signs; Sad/De = sadness/depression; VHS/AHS = Visual hallucinations or auditory hallucinations; Anx = anxiety; W/Rum = worry/rumination; PA = positive affect; NA = negative affect; M = mood; PSleep = poor sleep, Dis = dissociation; SMp = social media posting; Str/Se = stress/stressful events; SMv = social media viewing; LdTrav = less distance travelled; FeSoc/SocSh = feeling social/social sharing; LTsPhone = less time on phone; TsSit = time sitting; LTsVeh = less time in vehicles; TsComTests = time to complete tests; Alone = lonely/aloneness; PP/Cont = people present/contact; AtacSec/In = attachment insecurity/insecurity; Sus/D/Tc = suspicious/delusions/suspiciousness/thought control; Tc = thought control; THC/SU = cannabis/psychoactive substance use; Cb = cognitive bias; ERS/EA = emotion regulation strategies/experiential avoidance; Act = activity; Loc = location/GPS; Accel = accelerometer; PhLogs/PhAct = phone logs/phone activity; LowSE = low self-esteem; SleepStart = sleep start time; EngAHs = engaging with voices; Δcom = change in company; DataQual = data quality metrics; Δd = distance change; FoR = fear of relapse; Conf = feeling confident; Supp = feeling supported.
* the direction of phone use parameters varied, e.g., more phone calls but using the phone less overall; ** the effect of emotion regulation strategies varied, e.g., increased expressive suppression predicted paranoia, reappraisal did not; *** the effect of activity parameters varied, e.g., reduced activity predicted increased delusions, transition to leisure did not. ^ = the predictor was significant in one of various models/analyses but was non-significant in others.

worsening in one study.

Feeling social and social sharing failed to predict worsening in one study. Feeling alone or lonely failed to predict worsening in two studies. Fear of relapse and cognitive biases each failed to predict worsening in one study.

There were several studies that examined variables generated by passive sensing as predictors of suspiciousness. Idiographic changes in phone logs and activity (increase or decrease depending on the activity) predicted worsening, and less time on phone predicted worsening in one study. Location accelerometer data, time sitting, and less time in a vehicle each predicted worsening in one study. Activity predicted worsening in one study but not in a second study. People present and contacts were predictive in one study, however, change in company did not predict worsening in another study.

There were sixteen studies that proposed specific hypotheses in relation to worsening in paranoia. In nine studies the hypotheses were supported and in seven studies there was mixed support for the hypotheses. Specifically, as hypothesised paranoia was predicted by attachment insecurity, suspiciousness, thought control, sleep fragmentation and sleep quality, low self-esteem, prior anxiety, and prior sadness, worry and rumination, and negative affect.

In summary, changes in mood and sleep have been investigated the most frequently and have most consistently been shown to predict worsening in suspiciousness.

3.3.2. *Prediction of worsening in hallucinations*

In relation to hallucinations, we identified 13 studies that predicted worsening. Poor sleep was a significant predictor in two studies and change in sleep start time predicted worsening in one study.

Negative affect predicted worsening in two studies and failed to predict worsening in one study, whilst reduced positive affect predicted worsening in one study but not in another study. Emotion regulation failed to predict worsening in one study. Worry and rumination predicted worsening in one study but failed to predict in another study. Suspiciousness predicted worsening in hallucinations in three studies and engaging with voices was a predictor in one study.

Activity, dissociation, thought control, and substance use each emerged as predictors in single studies whilst cognitive bias failed to predict worsening in one study.

In relation to passive sensing data, phone logs and phone actions predicted worsening in two studies and less distance travelled in one study.

There were nine studies that tested a specific hypothesis in relation to worsening in hallucinations and each of these was supported. Specifically, as hypothesised, hallucinations were predicted by delusional intensity, reduced positive affect, dissociation, negative affect, sleep quality, sleep duration, sleep fragmentation, sleep efficiency, thought control, cannabis use, and worry and rumination.

In summary, fewer studies examined worsening in hallucinations, than delusions/paranoia/suspiciousness. Sleep and suspiciousness emerged as the most frequent predictors, and findings were mixed for changes in emotions.

3.3.3. Prediction of worsening in positive symptoms

Taken as a symptom cluster, worsening in positive symptoms was investigated across 12 studies. Poor sleep was a predictor across two studies. Sadness and depression predicted worsening in one study, but in two studies negative affect did not predict worsening and in one study positive affect did not predict worsening. Anxiety, worry and rumination, stress, and low self-esteem each predicted worsening in one study. Use of cannabis predicted worsening in one study.

Real time cognitive assessment and time to complete tests each failed to predict worsening in one study. Self-stigma, data quality metrics, contact with relatives, and change in distance travelled each failed to predict worsening in one study.

Nine of these twelve studies tested a specific hypothesis regarding symptom worsening and in six studies the hypotheses were supported. These related to parenting stress, self-esteem, sleep quality and duration, time to complete a color word test, use of psychoactive substances, and worry.

In summary, sleep once again emerged as the most frequent predictor of worsening in psychotic symptom as a cluster. There were a range of other predictors that have not been subject to replication. There were several non-significant findings in relation to mood as a predictor.

3.4. Meta-analysis outcomes

To estimate meaningful pooled effects, we grouped studies that measured similar predictors of relapse (e.g., suspiciousness). We grouped these studies regardless of the criteria for relapse that were applied, whether EMA or standardised measurement tools were utilised, and regardless of the duration of follow-up. It was possible to pool effects from a subset of five relapse outcome studies for three predictors, namely, conceptual disorganization (three instances), hallucinations (four instances), and suspiciousness (three instances) (Buck et al., 2021; Gaebel & Riesbeck, 2007; Saito et al., 2020; Subotnik & Nuechterlein, 1988; Wang, Gopal, Baker, & Narayan, 2018).

We found significant heterogeneity with moderate heterogeneity at the study level (38.50% of variance) and substantial heterogeneity at the effects level (61.45%), justifying a three-level random effects model (Q (6) = 35,299.4, $p < 0.0001$).

As depicted in Fig. 2, effect sizes ranged from moderate effects for conceptual disorganization to large effects for hallucinations and suspiciousness with an overall large pooled-effect size for the model which did not reach statistical significance ($d = 0.81$, 95%CIs = −0.01, 1.63). The result was a consequence of imprecision in the estimates and heterogeneity, most notably in relation to suspiciousness where estimates ranged in magnitude from small negative to large positive effects.

To test if the effects varied across subgroups, we conducted two additional analyses with study design and measurement methodology entered as moderators. Study design (i.e., observational (three predictors from two studies), RCT (four predictors from two studies), and other intervention design (three predictors from one study)) was a significant moderator (F(2, 7) = 5.343, $p = 0.039$). Specifically, the RCT design produced a stronger effect ($d = 1.92$, 95%, CIs = −0.83, 4.67) than observational ($d = 0.18$, 95%, CIs = −1.23, 1.60) or other intervention studies ($d = 0.67$, 95%CIs = 0.25, 1.08). Measurement type (i.e., EMA (four predictors from two studies) or repeated clinical assessment (six predictors from three studies)) was not a significant moderator (F (df = 1, 8) = 0.048, $p = 0.833$).

A total of seven outcomes were pooled for assessment of diagnostic test accuracy. In some studies, multiple estimates of sensitivity and specificity were produced by the inclusion of more than one measure (e.g., ESS and Fear of Recurrence). The pooled sensitivity was 0.71 (95% CIs = 0.64, 0.77), specificity was 0.64 (95%CIs = 0.47, 0.78), and diagnostic odds ratio was 4.21 (95%CIs = 2.18, 8.14) (see supplementary material). The summary ROC curve yielded a high AUC value (AUC = 0.72).

3.5. Quality assessment

Appendix B displays the outcome in relation to the quality assessment items for each study. It shows that the most significant shortcomings were the lack of sample size justification (k = 64), followed by recruitment from the same population ($n = 15$), and a lack of reporting of the rate of eligible participants recruited (k = 55). In 14 studies there was no report of an attempt to control for potential confounding

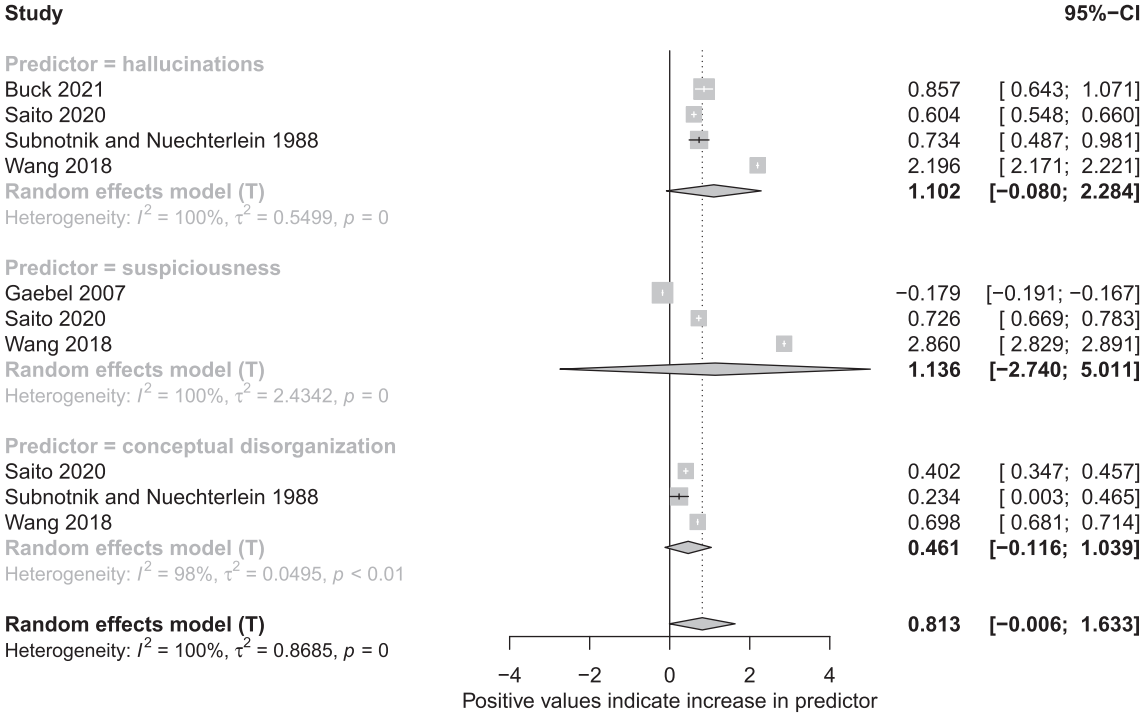

Fig. 2. Forest plot.

variables in the analyses (e.g., medication) and in eight studies the population was not clearly defined. On the positive side, study aims were specified in all studies.

4. Discussion

This is the first systematic review and meta-analysis to integrate findings in relation to the proximal predictors of relapse as a categorical outcome and worsening in psychotic symptoms as a continuous variable. Our review incorporated traditional EWS studies as well as newer real time EMA/EMI, passive sensing and digital phenotyping studies. Our findings indicate that changes in early signs questionnaires alone have produced mixed results as statistical predictors of relapse. Our review produced a higher pooled sensitivity (71% versus median 61%) and lower pooled specificity (66% versus 81%) than a previous review (Eisner et al., 2013). However, changes in EMA items combined with digital passive data show promise in relation to improving upon the prediction of psychotic relapses. Specifically, a recent promising development has been the integration of EMA and passive sensing data for the detection of anomalies in the time series of individual patient data which has resulted in a possible improvement in combined sensitivity and specificity of early warning signs (Henson et al., 2021). In relation to individual symptoms as predictors of relapse, our meta-analysis indicated large overall effects for the prediction of imminent relapse from a subset of three predictors with large effects for hallucinations and suspiciousness.

In relation to the prediction of symptom worsening as a continuous variable over the course of a day, sleep and changes in mood (especially increased negative affect) emerged as the most consistent predictors of worsening in the delusion domain from EMA and passive sensing studies. In the hallucination domain, sleep, and suspiciousness emerged as consistent predictors. Poorer sleep also predicted worsening in positive symptoms overall measured as a continuous variable.

In terms of integrating findings across the relapse and symptom worsening studies, first it is notable that sleep, suspiciousness, and mood changes are evident in both sets of predictors. Symptoms such as poor sleep and incipient psychotic symptoms are reflected in items within EWS scales as well as in EMA surveys in symptom worsening studies. Therefore, these findings suggest a dynamical systems hypothesis whereby, through an iterative function, the process of subtle deterioration in symptoms over hours can lead to a larger scale pattern of deterioration over days and weeks leading to a full-blown relapse.

In relation to ascertaining the strengths of the relapse studies included in our review, the timeframes over which relapse have been assessed are important to ensure that adequate time has elapsed for sufficient events to occur. Therefore, it is positive that most studies have been conducted over the course of a year. Of course, the symptom worsening studies are of much shorter duration but nonetheless the intense repeated measures designs enable some conclusions to be drawn regarding the predictors of more subtle changes in symptoms.

We note four major caveats in drawing firm conclusions regarding the predictors of relapse and symptom worsening. First, we found many inconsistencies and contradictory findings pertaining to symptom worsening. Second, there were multiple instances of predictors that have been investigated in only a single study. Third, many studies were exploratory in nature as opposed to positing specific a priori hypotheses which has significantly increased the likelihood that this body of research entails chance findings. Fourth, whilst the recent studies have demonstrated the promise of machine learning methods (e.g., in the analysis of idiographic anomalies in time series data) for improving the prediction of relapse, the reliance upon very large, complex, and bespoke data sets limits the translation of these findings to the clinical context where it is extremely difficult to track and analyse this array of variables.

Our assessment of study quality indicated some uncertainty with regards to the representativeness of participants because of lack of reporting of the rates of consent. We also noted persistent problems in the literature with reliance upon changes in treatment to determine the occurrence of relapse, which, as we noted over a decade ago (Gleeson et al., 2010), runs the risk of conflating psychotic relapse with other psychosocial and clinical crises faced by consumers.

4.1. Implications

Our review raises several implications pertaining to research, and clinical practice including ethical considerations. First, there is evidence of significant and rapid recent growth in the deployment of real-time methods (i.e., EMA, EMI, and passive sensing). Many of these studies have entailed data driven empirical investigations of the antecedents of relapse utilising machine learning methods which has broadened the predictors of relapse. Given the shortcoming from our quality assessment in relation to sample size justification, some doubt remains about the adequacy of statistical power for testing and validating of algorithms. A related concern is the representativeness of samples overall in relation to reflecting population diversity, e.g., the low rate at which race was reported limiting the capacity to draw meaningful conclusions.

By contrast, a minority of studies tested hypotheses which have direct implications for the understanding of the mechanisms of relapse (Emsley, Chiliza, & Asmal, 2013; Gumley et al., 1999; Nuechterlein et al., 1992).

In relation to clinical implications, this review highlights the importance of assessment of changes in mood and suspiciousness in clinical practice. Disturbances in sleep, which have long been highlighted in early warning signs checklists, is borne out in this review as an important target of assessment. However, interventions to maintain and restore good quality sleep remain relatively overlooked as potential clinical interventions for relapse prevention in schizophrenia (Waite, Sheaves, Isham, Reeve, & Freeman, 2020). Given the findings from mood disorders that have highlighted the relationship between social rhythms and circadian rhythm changes, these processes are worthy of closer investigation in psychosis (Ehlers, Frank, & Kupfer, 1988; Frank et al., 2005).

Our review highlights welcome methodological advances from paper and pencil self-report questionnaires to the integration of EMA and passive sensing which have shown promise in improving the prediction of psychotic relapse. The promise of real-time data collection methods, including a capacity to capture the idiosyncratic differences between unique individuals, the improved predictive value of such data and the reduced demand on participants, calls for further replication studies to validate these measures and enhance our understanding of mechanisms of relapse. Further advancement can be achieved by addressing the persisting lack of international agreement on the standards for defining and operationalising psychotic relapse, which has remained elusive despite consensus for remission (Andreasen et al., 2005). This remains a high priority for the field and may significantly reduce the heterogeneity in the estimates of effects. Similarly, increased standardisation and validation of ESM methods would significantly improve the capacity for direct comparison between study findings and would facilitate the establishment of population norms. The ESM item repository, which entails expert consensus methods with psychometric validation, is an exemplar of this strategy (Kirtley et al., 2020).

There are important treatment and ethical considerations in the translation of our findings into clinical practice. The recent incremental improvements in sensitivity and specificity of early warning signs, resulting from passive sensing and machine learning methods, is very promising for early intervention to prevent relapse. However, as machine learning methods proliferate in early warning signs research, a significant new challenge for the research community will be the translation of increasingly complex algorithms into transparent psychoeducation and feasible and acceptable clinical tools that facilitate informed and active consumer and carer participation and choice.

In relation to ethical considerations, the potential benefits for

consumers, carers, and clinicians in keeping vigilant for early warning signs of relapse needs to be balanced against the risk of inadvertently increasing fear of relapse (Zukowska et al., 2022), especially in light of the remaining limitations to the specificity of early warning signs. Ensuring that the process of building awareness of early signs is situated within an empowering interpersonal context that fosters shared decision making between consumers, carers, and clinicians is the best bulwark against this risk (Allan et al., 2020). Related to this, the EMI and passive sensing framework could pose a risk of generating creeping assumptions about the need for a surveillance model that jeopardises shared decision making and privacy to attain effective prevention.

4.2. Limitations and strengths of review

There are several limitations to the current review. First, we restricted papers to those with a minimum frequency of repeated assessment which may have been overly conservative, e.g., there were several papers that were excluded because of the intervals between assessment timepoints were greater than two weeks.

In addition, the definitions of relapses were varied which may have introduced a high level of heterogeneity. However, the field is yet to establish a gold standard definition, which we assert remains a high priority (Gleeson et al., 2010). The intensive repeated measurement studies collect data in a different timeframe from the earlier EWS studies – one risk is that these sets of studies have captured fundamentally different processes, however, our view is that these studies show differences in degree rather than kind of change – an important hypothesis for direct testing in future investigations. On the other hand, we note the high level of consistency in the measurement of symptoms and antecedents across the EMA studies (Myin-Germeys et al., 2009). In addition, due to resources our quality assessment was completed by only one researcher which may have biased results.

Furthermore, we acknowledge the relatively small number of studies eligible to be included in our meta-analyses of individual symptom predictors of relapse and pooling of sensitivity and specificity, due to the relative heterogeneity of studies and appropriate reporting of summary data by the included studies. However, we feel it is important to report meta-analytic results to demonstrate the current state of literature and to guide future work. The pooling of effects has significant advantages over the calculation of raw median values in relation to the weighting of individual findings.

In relation to strengths, this is the first review to synthesise findings in relation to worsening in psychotic symptoms alongside studies of relapse thereby facilitating a more inclusive view of the phenomena and the relevant literature in schizophrenia spectrum disorders. Sleep, mood and suspiciousness have emerged as important targets of intervention for both subtle and larger scale deterioration in psychosis.

4.3. Future directions

Given that only a minority of studies referenced specific hypotheses, we argue for an increased focus on integrating the important technological and methodological advances afforded by EMS, passive sensing, and machine learning together with theory testing and building. In addition, there are recent findings pertaining to several predictors of relapse or symptom worsening that urgently require replication, e.g., targets of passive sensing such as internet search patterns. Related to this, the augmentation of data-driven machine learning methods with expert domain knowledge held by consumers, families, and clinicians provides a potential pathway to improving the prediction of relapse (Gennatas et al., 2020).

In addition, future research is needed to address the relative paucity of investigations that have assessed the predictive validity of psychological constructs such as self-esteem and fear of relapse. These constructs have the advantage of being amendable to intervention.

In relation to improving the quality of research, greater transparency is needed in relation to the reporting of recruitment and retention of participants across all stages of investigations. Further rigor can be afforded by more careful consideration of potential confounding factors such as medication adherence and the presence of co-occurring syndromes such as post-psychotic depression.

5. Conclusion

There has been a rapid advance over the preceding decade in methodological and technological innovation in the empirical investigation of the prediction of psychotic relapse in people diagnosed with schizophrenia spectrum disorders. However, there has been little translation from these advances into specific new knowledge to improve prevention of relapse.

The most appropriate path forward is to systematically utilize methodological advances to test and build theoretical models of psychotic relapse so that consumers can be readily empowered with knowledge that can be applied in their daily lives to reduce the distress and other psychosocial costs of psychotic relapse.

Role of funding sources

This work was supported by an internal grant from the Australian Catholic University.

Contributions

JG developed the key ideas for the manuscript and was the lead writer of the manuscript. DF, MF, RT completed the screening and extraction. TMcG completed the meta-analyses with support from AP. AP undertook the quality assessment. DF undertook a quality assurance check of all extracted and tabulated findings. JF, AG and MAJ provided significant contributions to the key ideas and to the drafting and editing of the manuscript.

Declaration of Competing Interest

The authors have no conflicts of interest to disclose.

Data availability

No data was used for the research described in this article.

Acknowledgements

We acknowledge ACU Librarian Jessica Dickson for execution of the database searches and Mary Barillaro and Katherine Nguyen (RAs) for undertaking preliminary screening.

Appendix A. Study characteristics and outcomes

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/ other; Race/ Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
Studies using repeated clinical assessment or standardised self-report or clinical tools							
(Birchwood et al., 1989); UK	To investigate effectiveness of detecting prodromal signs of relapse in the clinical setting.	19; schizophrenia; CATEGO program and PSE.	24.6 (NR); female: 9; NR	Observational	Early Signs Scale (ESS) observer and patient forms; 2-weekly; up to 9 months or relapse	Relapse = readmission or imminent relapse defined independently by clinician. Confirmed by PSE.	Cut-off of 30 on ESS predicted relapse v. no relapse in 74% of cases (Fisher's exact $p = 006$).
(Eisner, Bucci, et al., 2019); UK	To refine the predictive value of app-based monitoring by adding basic symptoms to conventional early signs as putative relapse predictors.	18; schizophrenia: 14 (77.8%), schizoaffective: 4 (22.2%); DSM-IV.	37.9 (9.9); male: 12 (66.7%); Asian or Asian British: 1 (5.6%), Black or Black British: 2 (11.1%), White British: 15 (83.3%).	Observational	Smartphone (ExPRESS app), 6 months; PANSS positive items, mood symptoms (CDS), basic symptoms (BSC), early signs (ESS); weekly surveys; 6 months.	Relapse = symptom ↑ for 1 week resulting in management change, medication change or ↑ observation by clinical team, including admission; Symptom increase: assessed via PANSS items.	1 week later: early signs predicted suspiciousness (b = 0.194, p = 0.016); 2 weeks later: NS; 3 weeks later: early signs predicted psychotic symptoms (b =0.192, p = 0.011), hallucinations (b = 0.283, p =0.003); basic symptoms significantly predicted psychotic symptoms (b = 0.174, $p = 0.009$) and delusions (b = 0.216, p = 0.017).
(Gaebel et al., 2000); Germany	Investigate prevalence nature, time course, and predictive value of prodromal symptoms in impending relapse.	158; schizophrenia; ICD-9 and RDC.	35 (9.1); female: 91, male: 67; NR	*Re*-analysis of German multi-center observational study of intermittent versus maintenance neuroleptic long-term treatment in schizophrenia.	Adapted Early Symptom Questionnaire (ESQ); 2-weekly, reduced to 4-weekly if stable; 2 years	Relapse = clinically defined as a psychotic deterioration of maximum intensity usually with hospitalization.	Sensitivity = 10%, specificity = 93%. PPV = 43%, NPV of 67%. Relapse prediction from prodromal symptoms no better than chance, predictions 2 weeks before relapse were successful.
(Gaebel & Riesbeck, 2007); Germany	Examine predictive validity of prodromal symptoms in relation to relapse.	364; schizophrenia; ICD-9 and RDC.	34.8 (9.3); male: 157 (46.3%); NR.	Re-analysis of German multi-center observational study of intermittent versus maintenance neuroleptic treatment.	ESQ; initially 2-weekly, then changed to four-weekly; 2 years.	Relapse = psychotic deterioration of maximum intensity usually with hospitalization with minimum change in BPRS ≥ 10, CGI-Change ≥ 6, ↓ in GAS ≥ 20.	Sum of severity assessments of all prodromal symptoms led to a sensitivity of 72%; specificity of 38%.
(Gaebel & Riesbeck, 2014); Germany	To examine and enhance the relapse predictive validity of prodrome symptoms.	135 first-episode psychosis patients; NR	31.7 (9.9); male: 57.8%, NR.	Maintenance follow up study (1 year) followed by RCT evaluating antipsychotic maintenance treatment versus intermittent	Broad spectrum of 45 unspecific and specific prodrome symptoms; every 2 weeks; 2 years.	Clinical deterioration = ↑ on PANSS positive score ≥7 if sum score is ≥ 17, ↑ on the PANSS positive score ≥ 5 if sum score ≥ 20, at	Sum score of unspecific prodromes sensitivity = 95.2%, specificity = 39.6%; sum score of specific prodromes sensitivity =

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/ other; Race/ Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
				medication treatment (1 year).		least 1 PANSS positive item ≥5, CGI change score ≥ 1 if score ≥ 6, ↑ in CGI ≥ 2 if score ≥ 5, and ↓ in GAF ≥ 15.	48.8%, specificity = 80.1%.
(Gumley et al., 2015); UK	Establish reliability and validity of a Fear of Recurrence measure.	171; schizophrenia or related disorder; ICD-10.	Early Signs Scale group: 40.74 (11.33), Fear of Recurrence group: 42.22 (10.90); male: 121; NR.	Participants randomised to one of two early signs monitoring conditions - ESS or FoRSe.	ESS or FoRSe; frequency fortnightly alongside four-weekly PANSS assessment; 6 months.	Relapse = rating of moderate or greater representing an ↑ in at least 2 points on any one of the seven PANSS items.	Cut-off of 5 on ESS produced sensitivity of 79% (95%CIs = 62,89) and specificity of 35% (95% CIs = 23,50). Cut off of 5 on FoRSe produced sensitivity of 72% (95% CIs = 52,86) and specificity of 46% (95% CIs = 32,60).
(Jorgensen, 1998b); Denmark	To identify predictors of delusion formation	Sample 1: 60, Sample 2: 71; schizophrenia; DSM-IV and ICD-10.	Sample 1: 37 (10), Sample 2: 39 (12); Sample 1: male: 39 (65%), Sample 2: male: 44 (62%); NR.	Observational. Eight items from the ESS identified in one sample and tested in a second sample.	Eight items from the ESS – the Warning Signals Scale (WSS); 2-weekly; 6 months.	Symptom worsening = rating of moderate or greater, ↑ by at least 2 scale points on PANSS delusions items.	Criterion cut-off of 5 on WSS combined a high degree of sensitivity (77%) and specificity (68%).
(Jorgensen, 1998a); Denmark	To evaluate the predictive validity and temporal link of early signs to psychotic relapse.	60 (30 in each of 2 samples); schizophrenia; DSM-IV and ICD-10.	Sample 1: 38 (9), Sample 2: 36 (10); Sample 1: male: 21 (70%), Sample 2: male: 18 (60%); NR.	Observational. Two samples	ESS; 2-weekly; 6 months.	Relapse = rating of 'moderate' or greater with at least two scale points ↑ on any of seven positive scale items of PANSS.	Criterion cut-off of ≥ 10 points compared with baseline ESS score achieved sensitivity of 74% and specificity of 79% in predicting relapse.
(King & Shepherd, 1994); UK	To present a case of the use of the ESS with an inpatient with extremely severe residual symptoms.	1; severe schizophrenia with persistent delusions and hallucinations; NR.	32, male, NR	Observational Case study.	ESS observer and self-report; 2-weekly; 18 weeks	Relapse = marked exacerbation of psychotic symptoms as observed by clinicians	ESS sensitive to symptom changes in a patient with severe persisting psychotic symptoms in period of relapse.
(Lüdtke et al., 2022); Germany	To compare the effects of negative affect and aberrant salience on subsequent psychotic symptoms between a 1-week ESM phase and a 1-year follow-up phase.	30; non-affective psychoses; Mini International Neuropsychological Interview (MINI).	Possible age range 18–65 years; NR; NR.	Observational: initial ESM plus follow-up assessments.	Smartphone based ESM: anxiety, self-esteem, sadness, negative affect, aberrant salience, paranoia, AVHs; every 2 weeks and relapse assessments every 2 months for 1 year	Relapse = hospitalization, ↑ psychiatric care and 25% ↑ of Community Assessment of Psychic Experiences (CAPE) total score or clinical deterioration.	Subsequent paranoia predicted by NA ($b = 0.184$, $p = 0.001$) and aberrant salience ($b = 0.187$, $p < 0.001$); Paranoia at one-year follow-up predicted by aberrant salience ($b = 0.336$, $p < 0.001$); AVHs at one-year follow-up predicted by NA ($b = 0.093$, $p = 0.029$).
(Marder et al., 1991); USA	To compare methods of identifying prodromal	50; schizophrenia (stabilized); DSM-III-R.	NR; NR; NR	Randomised to behavioural skills training or supportive group	Anxious-Depression symptoms BPRS, ESQ, idiosyncratic	Psychotic exacerbation = worsening of 4 points or more on	AUC for BPRS Anxious-Depression subscale (BPRS-

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/other; Race/Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
	periods in relation to predicting relapse in schizophrenia.			therapy, randomised to low dose antipsychotic or placebo when prodromal.	prodromal scale (IPS); NR; at least 8 assessments; weekly.	the sum of BPRS cluster scores for thought disturbance and paranoia or an of 3 or more on either cluster.	D) = 0.62 (p = 0.026); for IPS AUC = 0.58 (p = 0.093). ESQ not significantly better than chance at predicting exacerbations.
(Marder et al., 1994); USA	To evaluate active oral fluphenazine versus placebo during relapse prodromes and to study the validity of prodromal criteria.	80; schizophrenia; DSM-III-R.	Never randomised (n = 44): 37.9 (8.3); male: 100%; nonwhite: 70%; Drug group (n = 17): 40.4 (9.4); male: 100%; nonwhite: 71%; placebo group (n = 19): 39.1 (8.6); male: 100%; nonwhite: 79%.	RCT: placebo versus fluphenazine deconate. Randomization occurred if the patient was stabilized.	Idiosyncratic prodrome Scale based on ESQ; weekly ratings; 2 years.	Psychotic exacerbation = worsening of 4 points or more on sum scale of the BPRS cluster score for thought disturbance and paranoia or ↑ of 3 or more on either cluster.	Placebo: PPV of prodrome = 48%, drug group: PPV = 37%. PPV increased over time in the placebo group and decreased in drug group (p =0.034).
(Saito et al., 2020); USA	Examine contribution of individual residual symptom to prediction of relapse.	305; schizophrenia: 206 (68%), schizoaffective = 99 (32%) disorder; DSM-IV-TR.	38.3 (12.1); male: 218 (71%); Caucasian: 156 (51%), African American: 85 (28%), Hispanic: 58 (19%), Others: 6 (2%).	Secondary analysis of RCT: patients randomised to biweekly LAI-risperidone or daily oral second-generation antipsychotics.	BPRS, biweekly, 30 months.	Relapse = CGI-GI score of 6 (much worse) or 7 (very much worse), psychotic hospitalization, increase in level of care, continuous increase in psychotic symptoms judged by raters, or self-injury/suicidal ideation.	Emotional withdrawal scores significantly higher 8 and 2 weeks before relapse compared to the baseline value (p = 0.032 and p = 0.043, respectively).
(Spaniel et al., 2018); Czech Republic	Identify the onset of changes in health and wellbeing, behavioural symptoms and pre-psychotic symptoms, before relapse.	51; schizophrenia: 31 (60.8%), schizoaffective disorder: 16 (31.4%), other: 4 (7.8%); ICD-10.	male: 32.4 (9.0) years, female: 35.2 (8.4) years; male: 37 (73%), female: 14 (27%); NR.	Observational and single group intervention study.	Early Warning Signs Questionnaire (EWSQ) patient and family version; weekly via text, 20 weeks.	Relapse = hospitalization.	Gradual increase pattern began 8 weeks before relapse in patient- and family-reported EWSQ sum scores.
(Subotnik & Nuechterlein, 1988); USA	To examine prodromal signs and symptoms of relapse in schizophrenia using a systematic and carefully controlled research design.	50; schizophrenia: 41, schizoaffective disorder: 9; Research Diagnostic Criteria.	23.4(3.4); men: 39, women: 11; Caucasian: 45, Hispanic: 4, Mixed heritage: 1.	Longitudinal follow up study: 6-week period prior to relapse period was compared with a 6-week period not preceding relapse for relapsing patents.	Brief Psychiatric Rating Scale (BPRS): every 2 weekly; total duration not specified (mean interval from admission to relapse was 18.3 months).	Relapse = rating of 6 or 7 on the Unusual Thought Content, Hallucinations, or Conceptual Disorganization items of BPRS.	Within relapsing patients: BPRS anxiety-depression factor (p < 0.003), BPRS thought disturbance (p < 0.009), BPRS thought disturbance (p < 0.02) elevated in prodrome period. Between patients: BPRS hostile-suspiciousness and thought disturbance classified 58.8% of prodromal

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/other; Race/Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
							periods, 88.2% of comparison periods.
(Tait et al., 2002); UK	Investigate if individualised early signs monitoring system effectively predicts relapse.	20; schizophrenia spectrum disorder; DSM-IV.	39 (NR), age range 25–66 years, male: 14 female: 6; NR.	Data from intervention arm of RCT: Individuals randomised to treatment as usual or cognitively oriented intervention.	Idiosyncratic early signs monitoring (ESM) questionnaire; 2-weekly; 10 months	Relapse = 50% ↑ in the total score of the positive scale of the PANSS.	75% of relapses predicted by ESM system (n = 3).
(Wang et al., 2018); USA	To identify individual items of PANSS that changed the most prior to relapse.	907; schizophrenia; DSM-IV-TR.	Relapse group: 38.4 (10.8), Non-relapse group: 38.5 (11); relapse group: male: 158 (59.2%), non-relapse group: male: 404 (63.1%); White: relapse group: 178 (66.7%), non-relapse group: 404 (63.1%), Black: relapse group: 41 (15.4%), non-relapse group 94 (14.7%), Asian: relapse group: 29 (10.9%), non-relapse group 56 (8.8%), Other: relapse group: 19 (7.1%), non-relapse group: 86 (13.4%).	Data pooled from three RCTs, double-blind, placebo-controlled withdrawal studies.	PANSS data every 4 weeks.	Relapse = psychiatric event (i.e., hospitalization, suicidal/homicidal ideation, or aggressive behavior), or significant ↑ in at least one PANSS items, or significant ↑ in PANSS total score.	PANSS items had on average 1-point ↑ 0.3–1.2 days before relapse: delusions, suspiciousness, hallucinations, anxiety, excitement, tension, conceptual disorganization.
EMA, EMI and Passive Sensing Studies							
(Adler et al., 2020); USA	Develop algorithm to predict specific days of symptom exacerbation before relapse using exclusively passive sensing data.	62; schizophrenia: 26, Schizoaffective disorder: 25, psychosis NOS: 9; chart diagnosis.	Relapse participants: 33 (NR), range = 23–47, Non-relapse participants: 40 (NR), range = 26–50; Relapse: female: 8 (44%), Non-relapse: female: 17 (40%); NR.	RCT: randomised to smartphone arm for passive sensing data collection or to treatment as usual. Data from smartphone arm only.	Smartphone android app: "crosscheck" - acceleration, app use, call logs, conversations, location, screen activity, sleep, text message activity, self-reported positive and negative symptoms; continuous data stream and self-report every 2–3 days; 12 months	Relapse = hospitalization, ↑ in care, ↑ medication, additional medication plus 25% ↑ in BPRS, suicidal/homicidal ideation, self-injury or violent behavior.	Anomaly detection system achieved a median sensitivity of 0.25 (IQR 0.15–1.00) and specificity of 0.88 (IQR 0.14–0.96; a 108% increase in anomalies near relapse.
(Allan et al., 2023); Australia and UK	To conduct time series on EMA dataset from people who have experienced a relapse within the	25; schizophrenia spectrum disorders; ICD-10.	43.3 (12.0); female: NR (52%); NR.	RCT: randomised to EMPOWER or treatment as usual. Data from EMPOWER arm.	Smartphone: EMPOWER platform, mood, anxiety, coping, psychotic experiences, self-	Symptom worsening.	Sleep change predicted paranoia (r = 0.08, 95%CIs = 0.03, 0.14).

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/ other; Race/ Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
	2 years using network analysis.				esteem, fear of relapse; 1 daily survey, 1 year.		
(Bak et al., 2016); The Netherlands	To investigate degree to which selected symptoms co-occur and co-vary over a year.	1; schizophrenia, paranoid type; DSM-IV.	46 (N/R); female; NR.	Case study, observational study.	Mobile device; questions regarding mood, loss of control, paranoia, hearing voices, relaxed; 4 days/week with 10 beeps per day; 1 year.	Relapse = increase clozapine dose to 450 mg/day; impending relapse = increase dose to 400 mg/day due to moderate increase in symptom severity.	Mood and paranoia fuelled each other, During relapse, symptom levels and clustering between symptoms ↑.
(Barnett et al., 2018); USA	Explore feasibility and utility of smartphone digital phenotyping for relapse prediction.	17; schizophrenia; NR	NR; NR; NR.	EMA observational study.	Personal smartphone, Beiwe app, medication adherence, mobility, sociability features, symptom surveys; 1 min every 10 min (passive) and self-report survey 2/ week; 3 months	Relapse = Hospitalization, increase in levels of psychiatric care	Rate of anomalies detected in the passive data streams in the 2 weeks prior to relapse was 71% higher than the rate of anomalies detected further away from relapse. Significant anomalies in all data streams 9 days prior to hospitalization in 1 patient with available data.
(Bell et al., 2018); Australia	To develop brief coping-focused intervention for distressing voices	1;. Schizophrenia; NR.	38, male, NR.	Single case illustration of EMI	Smartphone (RealLife app); emotions, surroundings, activities, coping strategies; 10 daily surveys; 1 EMA wave for 6 days, 2 EMI waves for 10 days each.	Symptom worsening (e.g., voice intensity).	Voices more intense following times when he was doing something important to him and when engaging with voices.
(Ben-Zeev et al., 2011); USA	To examine the prospective relationships predicted by a cognitive model of persecutory ideation.	199; schizophrenia, schizoaffective disorder; DSM-IV.	46.5 (11.16); male: 61%; white: 60%, African American: 15%, Hispanic: 14%, other: 11%.	EMA observational study.	Personal digital assistant, anxiety, sadness, external events, anomalous experiences, conviction, distress, substance use; 4 surveys per day; 7 days.	Symptom worsening, e.g., more intense persecutory ideation	Prior anxiety (b =0.28, P < 0.01) and prior sadness (b = 0.23, *P* < 0.01) significant positive relationship with the log-odds of subsequent persecutory ideation.
(Ben-Zeev, Morris, et al., 2012); USA	To examine if negative emotional states predict delusion subtypes.	199; schizophrenia, schizoaffective disorder; DSM-IV.	46.2 (11.24); male: 59%; white: 59%, African American: 15%, Hispanic: 14%, other: 12%.	EMA observational study.	Personal digital assistant, Purdue Momentary Assessment Tool; Anxiety, Sadness, Hallucinations, Delusions; 4 surveys per day; 7 days.	Symptom worsening in delusions	Hallucinations predicted delusions of control (OR = 4.63, 95%CIs = 2.55, 8.41) and reference (OR = 2.18, 95%CIs = 1.15,4.12) over subsequent hours, anxiety, sadness not predictors.

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/other; Race/Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
(Ben-Zeev, Frounfelker, et al., 2012); USA	To examine the relationship between self-stigmatizing beliefs, contextual factors and symptoms.	24; schizophrenia or schizoaffective disorder; DSM-IV.	44.87 (9.27); male: 71%; African-American: 79%.	EMA observational study	Personal digital assistant (PALM™), self-stigma, positive affect, negative affect, 6 surveys per day for 1 week.	Symptom worsening	No previous (time t-1) predictors were significantly related to change in symptoms.
(Ben-Zeev et al., 2017); USA	To test a multimodal data collection system for continuous remote monitoring and identification of indicators of psychotic relapse.	5; psychosis not otherwise specified, schizophrenia, schizoaffective disorder; chart diagnosis.	NR; female: 80%; African American: 40%, Hispanic: 40%, American Indian: 20%, white: 20%.	EMA observational study.	Smartphone Samsung Galaxy S5 (CrossCheck); self-report, Multimodal behavioural sensing (i.e., physical activity, geospatials activity, speech frequency, and duration) and device use data.	Relapse = emergency room visits for psychiatric reasons or psychiatric hospitalization.	Participants had unique digital indicators of their psychotic relapse evident from self-report or behavioural sensing data trends
(Buck, Scherer, et al., 2019); USA	To evaluate if smartphone-collected digital measures of social behavior provide early indication of relapse.	61; schizophrenia: 26 (42.62%), schizoaffective disorder: 26 (42.62%), psychosis NOS: 9 (14.75%).	37.11(13.85); male: 36 (59.02%), female: 25 (40.98%); White/Caucasian: 22 (36.07%), Black/African-American: 18 (29.51%),. Pacific Islander 4 (6.56%), American Indian or Alaskan Native 1 (1.64%), Asian-American 1 (1.64%), Multiracial 13 (21.31%), Missing/declined 2 (3.28%)	Data from RCT – (CrossCheck versus treatment-as usual). Data from the CrossCheck arm of the study only.	Smartphone Samsung Galaxy S5 (CrossCheck); included: sensing of speech frequency and duration, incoming/outgoing SMSs and phone calls over 12 months.	Relapse = psychiatric hospitalization, significant increase in psychiatric care, increased medication plus either an increase of 25% from baseline BPRS total score, suicidal or homicidal ideation that was clinically significant, deliberate self-injury, or violent behavior.	Outgoing call duration: 3 significant effects (ranging from $\beta = -0.009$, $p = 0.005$ to $\beta = -0.019$, $p = 0.030$), Outgoing calls: 3 significant effects (ranging from $\beta = -0.046$, $p = 0.002$ to $\beta = -0.209$, $p = 0.043$), Incoming SMS: 3 significant effects, (ranging from $\beta = -0.819$, $p = 0.009$ to $\beta = -2.228$, $p = 0.0141.080$, $p = 0.017$), Outgoing SMS: 3 significant effects (ranging from $\beta = -0.986$, $p = 0.003$ to $\beta = -2.435$, $= 0.0191.070$, $p = 0.031$).
(Buck, Hallgren, et al., 2019); USA	Quantify between - and within -person variability in paranoia and identify passively sensed indicators of paranoia over 1 year.	62 (45 completed the study); schizophrenia, schizoaffective disorder, or psychosis not otherwise specified; NR.	NR; NR; NR.	EMA observational study.	Smartphone - Samsung Galaxy S5, Crosscheck; self-report paranoia, other mental health signs, functioning, multimodal behavioural sensors via passive sensing; 3 days per week; 1 year	Symptom worsening in paranoia.	EMA predictors: depression ($\beta = 0.29$), stress ($\beta = 0.34$), hearing voices ($\beta = 0.41$), seeing things ($\beta = 0.50$), feeling social ($\beta = 0.07$), sleeping well ($\beta = 0.06$). Passive sensing predictors: time sitting ($\beta = 1.35$), distance travelled ($\beta = -0.20$), time in vehicles ($\beta = -3.92$), time in speech ($\beta = -0.65$) and phone calls ($\beta = -6.08$).

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/ other; Race/ Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
(Buck et al., 2021); USA	To determine whether a brief report of individual symptoms assessed via EMA detects changes occurring before, during, and after psychiatric relapses.	61; schizophrenia spectrum disorders; NR.	NR; NR; NR.	EMA observational data drawn from intervention arm of an RCT.	Smartphone -Samsung Galaxy S5 (Crosscheck); negative mood, anxiety, sleep; 1 survey per day, 3 days per week for 1 year.	Relapse = psychiatric hospitalization, significant ↑ in psychiatric care, ↑ medication plus either an ↑ of 25% from baseline BPRS total score, suicidal or homicidal ideation that was clinically significant, deliberate self-injury, or violent behavior.	Significant and steady increases (d = 0.05 per week) in persecutory ideation and hallucinations over the 100-day period preceding relapse.
(Cohen et al., 2023); USA and India	To explore the feasibility of digital phenotyping for relapse prediction across different regions, cultures, and languages	76; schizophrenia; DSM-5.	NR; NR; NR.	EMA observational with passive sensing	Smartphone – mindLAMP; active: sleep, sociability, psychosis, medication adherence; passive: accelerometer, GPS, screen state; 6/day; variable across settings mean of 66 days to mean of 195 days)	Relapse = 1) 25% increase in participant's PANSS score, 2) psychiatric hospitalization, 3) suicidal attempt or significant and sudden increase in suicidal ideation, 4) significant/ sudden increase in psychosis symptoms requiring clinical intervention.	Anomalies 2.12 more frequent in month preceding relapse and 2.78 times more frequent in month preceding and following a relapse compared to intervals without relapses.
(Daemen et al., 2022); The Netherlands and Belgium	To investigate whether fluctuations in self-esteem are associated with psychotic experiences in daily life.	147; non-affective psychotic disorder; DSM-IV	34.3 (8.2); women: 48 (32.7%); men: 99 (67.3%); Caucasian: 131 (89.1%), non-Caucasian: 16 (10.9%).	Observational: ESM	PsyMate digital device: thoughts, feelings, activity, social context, location, affect, self-esteem, psychotic symptoms; 10/day; 6 days	Symptom worsening in paranoia and psychotic experiences.	Self-esteem ($\beta = -0.07, p < 0.001$) predicted paranoia.
(Dupuy et al., 2021); France	To investigate role of momentary fluctuations in cognitive performance and experience of positive symptoms.	33; schizophrenia; DSM-IV-TR.	33.9 (10.0); male: 24 (73%); NR.	Observational with real-time assessment of cognitive performance and positive symptoms.	Smartphone - Samsung Galaxy; Real-time cognitive performance, psychotic symptoms; 5 surveys per day; 1 week.	Symptom worsening	Time to complete color-word test predicted ↑ in psychotic symptoms (Coeff = 0.06, SE = 0.02, $p < 0.05$, Odds Ratio = 1.07).
(Geraets et al., 2020); The Netherlands	To examine effects of VR-CBT for paranoia on affective states and on interplay between affective states and paranoia.	116; schizophrenia: 79, schizoaffective disorder: 5, not-otherwise specified psychotic disorder: 7; DSM-IV.	39.5 (10.1); male: 63 (69%); Dutch origin: 61 (67%).	Multi-centre RCT Intervention with randomization to VR-CBT or TAU. The baseline associations are relevant to this review.	PsyMate - electronic momentary assessment device; suspicious, dislike, hurt, negative affect (e.g., anxious, down, unsafe, lonely); 10 surveys per day; 6–10 days, 3 waves – baseline, 3 months, 6 months.	Symptom worsening	At baseline significant autocorrelations for suspiciousness, VR-CBT 0.19, p-< 0.01.TAU 0.17 p- < 0.01.

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/ other; Race/ Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
(Hartley et al., 2014); UK	To investigate if rumination and worry are associated with persecutory delusions, auditory hallucinations and the associated distress.	32; schizophrenia: 15, psychotic disorder NOS: 14, schizoaffective disorder: 2, acute psychotic disorder: 1; NR.	33 (10.7); male: 22; all but one were white.	Observational	Palm computer + programmed watch (ESP Software); Rumination, Worry; 10 surveys per day, 6 days.	Symptom worsening	Worry predicted delusions β = 0.332 (95%CIs 0.188–0.475) and hallucinations β = 0.206 (95%CIs 0.039–0.374),. Rumination also predicted delusions (β = 0.203, 95%CIs 0.072–0.334) and hallucinations (β = 0.202,95%CIs 0.051–0.354).
(Hartley et al., 2015); UK	To examine the associations between thought control and the experience of persecutory delusions and auditory hallucinations.	36; schizophrenia: 15, psychotic disorder NOS: 14, schizoaffective disorder: 2, acute psychotic disorder: 1; NR.	33 (10.7); male: 22; All but one were white.	Observational	Palm computer + programmed watch (ESP Software); thought control; 10 surveys per day, 6 days.	Symptom worsening	Thought control predicted severity (b = 0.140, 95%CIs = 0.015–0.266) and distress (b = 0.307, 95% CIs = 0.082–0.532) for persecutory delusions and distress for hallucinations (b = 0.371, 95% CIs = 0.196–0.545).
(Hays et al., 2020); USA	To find unique interactions of schizophrenia symptoms as experienced on a moment-by-moment basis.	47; schizophrenia; NR.	38.09 (14.64); male: 23 (54.8%); American Indian or Alaskan Native: 2 (4.8%), Asian: 0 (0%), Black or African American: 15 (35.7%), Multiracial or other: 2 (4.8%), White: 23 (54.8%).	Observational	Smartphone (mindLAMP app); mood, sleep, social functioning, anxiety, cognitive functioning, psychosis; 5/7 days per week; 90 days	Symptom worsening	Transition probabilities included anxiety-inducing psychosis (0.204, p = NR), mood-inducing psychosis (0.162, $p < 0.001$), and sleep-inducing psychosis (0.189, $p < 0.189$).
(Henquet et al., 2010); The Netherlands	To examine the effects of cannabis on psychotic symptoms and mood.	48; schizophrenia: 10, schizoaffective disorder: 28, unspecified functional psychosis: 4; RDC.	36.1 (9.3); male: 31; female: 11; NR.	Observational	Digital wristwatch, paper-and-pen ESM booklet; cannabis use, mood, and psychotic symptoms; 12 surveys per day; 1 wave for 6 days.	Symptom worsening	Cannabis use associated with hallucinations (β = 0.08, 95%CIs = 0.03–0.13, p = 0.002) and auditory hallucinations (β = 0.11, 95%CIs 0.04–0.17, p = 0.003).
(Henson et al., 2021); USA	To utilize smartphone digital phenotyping to predict clinical relapse.	83; schizophrenia; NR.	36.45 (14.96); female: 24 (38.1%), male: 35 (55.6%), other: 4 (6.3%); American Indian or Alaskan Native: 4 (6.6%), Asian: 1 (1.6%), Black or African	Observational: EMA and passive sensing.	Smartphones (mindLAMP and Beiwe); anxiety, medication, depression, mobility, sociability, cognition, screen time; sleep; 1 or 2 surveys per day (either twice each day or five times	Relapse = psychiatric hospitalization, 25% increase in PANSS from baseline, CGI change score of 6 or 7, exacerbation in symptoms requiring immediate	Paired anomalies in passive sensing and EMA survey response had sensitivity of 89%, specificity of 75%, positive predictive value (PPV) 60%, and negative predictive value (NPV) of 94% in

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/other; Race/Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
			American: 18 (29.5%), Multiracial or Other: 5 (8.2%), White Caucasian: 32 (52.5%), Native Hawaiian or Pacific Islander: 1 (1.6%).		each week); 1 wave of 3–12 month.	clinical management.	predicting relapse.
(Hermans et al., 2020); Belgium, The Netherlands	To elucidate temporal dynamics of suspiciousness and hallucinations in the early stages of psychosis.	48; first episode psychosis; multiple diagnostic systems using the CASH	22.8 (5.1); male: 19, female: 15; NR.	Observational	Digital wristwatch and daily self-assessment forms collected in a booklet (PREVENT study) or a personal digital assistant, PsyMate™(iThink study); positive and negative affect; 10 daily surveys; 6 days.	Symptom worsening = suspicious and hallucinatory experiences	↓ in positive affect before episode of suspiciousness ($\beta = -0.43$, $p < 0.05$); Negative affect ↑ before start of hallucinatory episodes ($\beta = 0.54$, $p < 0.05$).
(Kammerer et al., 2021); Germany	To determine if sleep parameters predict next-day persecutory symptoms.	77; schizophrenia: 47, schizoaffective disorder: 16, delusional disorder: 4; DSM-5.	38.04 (12.29); male: 40 (59.7%); NR.	Observational: ESM and actigraphy data from an RCT	Smartphones (movisensXS ESM app), wrist-worn actigraphs (Actiwatch 2); Sleep measures (objective and self-report); 10 daily surveys; 6 days.	Symptom worsening	Neither objective nor subjective sleep measures predicted next-day persecutory symptoms. When controlling for medication, decreased sleep efficiency significantly predicted persecutory symptoms (b = −0.00560, 95% CIs −0.0109, −0.0003, p =0.039).
(Klippel et al., 2018); data from 6 studies	To examine dynamic interplay between daily stress, momentary affect/thoughts, psychotic experiences, and other daily life contexts.	245; psychotic disorder; various diagnostic criteria across 6 studies.	35.3 (10.8); male: 111 (46%), female: 132 (54%); NR.	Observational	Diary and a wristwatch; daily stress, aloneness; 10 daily surveys; 5–6 days.	Symptom worsening	Significant associations between stress and suspiciousness ($\beta = 0.051$, $p < 0.000$); No relationship between aloneness and suspiciousness.
(Klippel et al., 2021); data from 6 studies	To investigate effects of momentary stress and affective disturbance on psychotic symptoms.	245; psychotic disorder; various diagnostic criteria across 6 studies.	35.3 (10.8); male: 111 (46%), female: 132 (54%); NR.	Observational	Diary and a wristwatch; daily stress, negative affect; 10 daily surveys; 5–6 days.	Symptom worsening	No moderated mediating effect of NA on psychotic experiences; no effects for stress on psychotic experiences.
(Lahti et al., 2021); USA	To investigate feasibility of using wearable devices and self-reported technologies to identify symptom	40; schizophrenia or schizoaffective disorder; DSM-5.	median age = 40.3; men: 63%; Black/African American: 29 (73%), White: 11 (28%).	Observational	Smartphone (The Ginger app/REDCap surveys) and wristband; sleep-wake, activity count, light exposure,	Relapse = moderately severe, very severe, or extremely severe on PANSS positive items of	In 1 patient day-to-day variation in mobility was very high before relapse. Disrupted sleep noted.

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/other; Race/Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
	exacerbation and relapse.				ambulatory activity; bidaily for question set 1, weekly for question set 2; 120 days.	2 negative items; exacerbation = required change in antipsychotic medication.	
(Lamichhane et al., 2021); USA	To investigate a machine learning based schizophrenia relapse prediction model using mobile sensing data.	75; schizophrenia; NR.	37.2 years (range 18–65 years); NR; NR.	Observational: data from clinical trial	Smartphone (Crosscheck); 10 EMA items, accelerometer, light levels, distance travelled, call duration, sound levels, conversation duration; EMA obtained items 3 times per week; up to 1 year.	Relapse = an acute increase of schizophrenia symptoms and degrading mental health.	Naive Bayes based classification gave best classification performance (F2 = 0.083). Distance travelled most relevant for relapse prediction followed by EMA items and call duration.
(Lamichhane et al., 2023); USA	To investigate a supervised personalised deep learning model for relapse prediction using mobile sensing data.	63; schizophrenia; NR.	37.2 (13.7); female: 36, male: 27; NR.	Observational: re-analysis of passive sensing from CrossCheck data. Used a long short-term model named RelapsePredNet to predict relapse in individual patients.	Smartphone Samsung Galaxy S5 (CrossCheck); Multimodal behavioural sensing (i.e., light exposure, volume, conversation, distance travelled, accelerometer, total screen usage; continuous; up to 1 year	Relapse = based on criteria such as psychiatric hospitalization, the need for increased clinical care, increased BPRS scores, etc.	Best F2 score was 0.21 when using social functional data to define patient similarity and 0.52 in the sample with a relapse. The F2 for personalised model was superior to non-personalised models.
(Lüdtke et al., 2021); Germany and Switzerland	To investigate if sleep problems or worrying, predict symptom variability.	124; non-affective psychotic disorder; NR.	Delayed access: 40.88 (9.84), Immediate access: 42.34 (10.85); Delayed access: female: 37, male: 29, Immediate access: female: 38, male: 20; NR.	RCT: delayed access versus immediate access group.	Web-based tool (EviBas): 14 items: worry; NA; anxiety; self-esteem, cognitive biases, sleep; Immediate access group: 2 times/6 days; Delayed access: once per week.	Symptom worsening	Worry predicted psychotic symptoms (b = 0.156, pFDR = 0.030) and paranoia (b = 0.116, p = 0.009), NA predicted paranoia (b = 0.058, p = 0.013), sleep predicted paranoia (b = 0.104, p = 0.012) and AVH (b = 0.087, p = 0.006). Cognitive bias did not predict symptoms.
(Ludwig et al., 2020); Germany	Investigate pathway from negative affect to paranoia.	80; schizophrenia: 51, schizoaffective disorder: 16, delusional disorder: 4; DSM 5.	37.80 (12.15); male: 57.8%; NR.	Observational	Smartphones (movisensXS); negative affect, emotion regulation strategies; 10 daily surveys, 6 days.	Symptom worsening	NA significant predictor of paranoia (b = 0.076, p < 0.001, 95%CIs 0.029, 0.122) - moderated by awareness of emotions (−ve) and rumination (+ve). No strategies predicted paranoia.

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/other; Race/Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
(Meyer et al., 2021); UK	Investigate the temporal relationship between sleep and psychopathology.	36; schizophrenia or schizoaffective disorder; NR.	40.7 (9.7); male: 24 (67%), female: 12 (33%); NR.	Observational	Smartphone (Sleepsight); self-reported sleep quality, sleep duration, psychopatholgy; daily; 12 months	Symptom worsening	Sleep quality negatively predicted psychosis symptoms (ES = −0.567), paranoia (ES = −0.502, and hallucinations (ES = −0.329). Sleep duration negatively predicted psychosis symptoms (ES = −0.501) and hallucinations (ES = −0.332) but not paranoia. Relationship between sleep duration and quality and psychosis symptoms was mediated by NA and cognitive symptoms.
(Mulligan et al., 2016); UK	To conduct a prospective examination of relationship between sleep and next-day functioning and psychotic symptoms.	24; nonaffective psychosis: 8, schizophrenia: 13, schizoaffective disorder: 1; ICD-10 and DSM-5.	37.4 (10.4); male: 13, female: 9; White British: 19, Black British: 3.	Observational	PRO-diary and sleep diary (PRO-Diary), wrist actigraphy (CamNtech); Sleep efficiency (SE), sleep fragmentation (SF), total sleep time, mood, psychotic symptoms, functioning; 5 daily surveys, 7 days.	Symptom worsening	↑ objective Sleep Fragmentation predicted ↑ AHs (β = 0.0127, p = 0.002) and paranoia (β = 0.0128, p = 0.009). ↑ objective and subjective Sleep Efficiency predicted ↓AHs. ↑ subjective Sleep Quality predicted ↓ AH, paranoia, thought control.
(Myin-Germeys et al., 2001); The Netherlands	To investigate delusions at the level of everyday functioning.	64; schizophrenia: 34, atypical psychosis: 2, delusional disorder: 1, schizoaffective disorder: 1; DSM-III-R.	35 (7); male: 28, female: 20; NR.	Observational	Digital wristwatch and physical booklet; activity, people present, mood states, delusions; 10 daily surveys, 6 days.	Symptom worsening: delusional moments (DMs).	Persons present (β = −0.715, p < 0.05) and activity (β = 1.127, p < 0.05) predicted probability of DMs; location did not predict subsequent DMs.
(Nittel et al., 2018); Germany	To explore the association between emotional instability, emotional regulation strategies, and paranoia	32; schizophrenia: 23, schizoaffective disorder: 7, schizotypal personality disorder: 1, delusional disorder: 1; DSM-V.	35.87 (11.05); male: 14, female: 18; NR.	Observational	iPod Touch (iDialogPad); Emotion regulation (ER) strategies; 10 daily surveys, 6 days.	Symptom worsening	Expressive suppression (β = 0.03, p < 0.0.05) predicted paranoia. NA (β = 0.06, p < 0.05) predicted paranoia if rumination was included as a covariate. Rumination, reappraisal, acceptance, distraction, social

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/other; Race/Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
							sharing, and reflection did not predict paranoia.
(Oorschot, Lataster, Thewissen, Wichers, & Myin-Germeys, 2012); The Netherlands	To illustrate temporal associations between mood states and paranoia.	64; schizophrenia/psychotic disorder: 56, schizoaffective disorder: 8; NR.	Paranoid patients: 38.1 (10.7), Nonparanoid patients: 36.0 (11.6); Paranoid patients: male: 87%, female: 13%, Nonparanoid patients: male: 76%, female: 24%; NR.	Observational	Digital wristwatch and assessment booklet; anxiety, irritation, relaxation; 10 daily surveys, 6 days.	Symptom worsening	At least one mood state temporally related to paranoia in 54% of patients (n = 36), not related to paranoia in 23% of patients (n = 15); did not vary in 23% of patients (n = 15).
(Oorschot, Lataster, Thewissen, Bentall, et al., 2012); The Netherlands and Belgium	To investigate phenomenology of AHs and VHs in daily life and their temporal relation to emotions and delusions.	193; schizophrenia: 144, schizoaffective disorder: 16, other psychotic disorders: 24; DSM-III-R and DSM-IV.	Mean age ranged over four groups: 36.2–29.6; male: 71%; NR.	Observational	Digital wristwatch and assessment booklets; positive and negative affect; 10 daily surveys, 6 days.	Symptom worsening	AH-onset preceded by ↑ delusional intensity (β = 0.35, $p < 0.001$). VH-onset preceded by ↓ PA (β = −0.40, $p <$ 0.01) and ↑ NA (β = 0.24, p < 0.05) and delusional intensity (β = 0.37, $p < 0.001$).
(Postma et al., 2021); UK	To examine associations between momentary self-esteem, and psychotic experiences in daily life in individuals with first-episode psychosis.	59; schizophrenia: 15 (31.3%), delusional disorder: 3 (6.3%), schizoaffective disorder: 3 (6.3%), manic psychosis: 7 (14.6%), depressive psychosis: 7 (14.6%), psychotic disorder NOS: 13 (27.1%); OPCRIT.	28.3(8.6); women: 23 (45.1%), men: 28 (54.9%), white British: 14 (27.5%); Black African: 17 (33.3%), Black Caribbean: 11 (21.6%), Asian: 1(2%), White other: 4 (7.8%), Other: 4 (7.8%).	Observational: ESM	PsyMate digital device: self-esteem, psychotic experiences; 10 daily surveys, 6 days	Symptom worsening	Self-esteem (β = −0.09, p = 0.000), fluctuations in self-esteem instability (β = 0.01, p = 0.182), variability in self-esteem (β = 0.06, p = 0.000) predicted psychotic symptoms.
(Radley et al., 2022); UK	To investigate the role of stress of parenting in the exacerbation of psychotic symptoms by using ESM to measure daily fluctuations in both.	35; primary diagnosis of any psychotic disorder (excluding postpartum psychosis, drug-induced psychosis, organic psychosis); NR	41 (6.49); female: 28 (80%), male: 7 (20%); White British: 21 (60%), Asian/Asian British: 7 (20%), Black/Black British: 4 (11.4%), White other: 2(5.7%), Mixed ethnicity: 1 (2.9%).	Observational: ESM	Smartphone (mobile interface); negative affect, positive psychotic symptoms, activity stress, event stress, social stress; 6 daily surveys, 10 days.	Symptom worsening	Parenting event stress (β = 0.09, p = 0.005), parenting activity stress (β = 0.16, $p < 0.001$), parenting social stress (β = 0.04, p = 0.016) predicted psychotic symptoms.
(Raugh et al., 2020); USA	Evaluate psychometric properties of a passive digital phenotyping method, geolocation.	51; schizophrenia; DSM-5.	39.59 (12.64); male: 13 (31.7%); African American: 13 (31.7%), Biracial: 3 (7.3%), Caucasian: 23	Observational	Blu Vivo 5R smartphone; geolocation – distance change; EMA context and symptom questions; 8 daily surveys, 1 week.	Symptom worsening	PANSS positive symptoms not associated with distance change (β = −1.25, $p >$ 0.05).

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/ other; Race/ Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
			(56.1%), Hispanic/ Latino: 2 (4.9%).				
(Sa et al., 2016); UK	To investigate whether interactions with relatives are related to symptoms in everyday life.	21; first episode psychosis: 7, schizophrenia: 8, schizoaffective: 1, psychotic disorder NOS: 2, unspecified non-organic psychosis: 3; NR.	26 (median, range 19–51); female: 6, male: 15; White: 21.	Observational	Palm device (Tungsten E2) & Digital wristwatch (TimexIron Man); contact with relative, BCIs, EE; 10 daily surveys; 6 days.	Symptom worsening	Contact not related to symptoms. EE did not moderate relationship between contact with relative and symptoms.
(Sitko et al., 2016); UK	To investigate if elevated attachment insecurity precedes the occurrence of paranoia.	20; schizophrenia: 6, psychosis NOS: 4, schizoaffective: 2, paranoid schizophrenia: 8: NR.	41.05 (12.53); male: 16; NR.	Observational	Palm pilot (Tungsten E2) or paper diaries with google calendar reminders; attachment insecurity; 10 daily surveys; 6 days.	Symptom worsening	↑ paranoia predicted by preceding elevated level of attachment insecurity ($\beta = 0.173$, $p < 0.001$)
(So et al., 2021); Hong Kong	To elucidate moment-to-moment associations between intensity of AVH and NA.	54; schizophrenia spectrum disorder; DSM-IV.	43.83 (12.40); male: 15 (36.59%); NR.	Observational	Smartphone; NA; 10 surveys per day; 6 days.	Symptom worsening	NA significantly associated with subsequent AVH ($\beta = 0.130$, S.E. = 0.034, $p < 0.001$).
(Swendsen et al., 2011); USA	To examine association of substance use with psychotic symptoms.	199; schizophrenia: 144, schizoaffective disorder: 55 DSM- IV.	46.5 (11.2); female: 39.3%, male: 61%; White: 60%, African American: 15%, Hispanic: 14%, other ethnicities: 11%.	Observational	Personal digital assistant; 4 daily surveys; negative mood, perceived stress; substance use; 7 days.	Symptom worsening	Use of any psychoactive substance associated with ↑likelihood of psychotic symptoms ($\gamma = 1.092$, SE = 0.350, $t = 3.119$, $p<0.01$).
(Torous et al., 2018); USA	Explore relationship between quality of digital phenotyping data and domains of schizophrenia.	16; schizophrenia; NR.	NR; NR; NR	Observational: pilot study	Smartphone (Beiwe): passive: GPS, accelerometer, call and text logs, screen on/off status, phone battery charging status, active: mood, anxiety, sleep, psychosis, medication adherence, 3 times/week; 90 days	Symptom worsening	No significant prediction of psychosis symptoms by data quality metrics.
(Tseng et al., 2020); USA	To predict symptom trajectories of schizophrenia from passive mobile sensor data.	61; Schizophrenia; DSM IV or DSM V.	NR; NR; NR	Observational	Smartphone (CrossCheck); self-report, multimodal behavioural sensing (i.e., physical activity, geospatials activity, speech frequency, and duration) and device use data; 3 times/week; 1 year.	Symptom worsening	Behavioural rhythms improved prediction of hearing voices (Z=-3.372, p=0.005), e.g., ↑ variation in ambient sound likely to exacerbate hearing voices; ↑ deviation in light likely to exacerbate seeing things.

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/other; Race/Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
(Udachina et al., 2014); UK	Examine the role of EA in paranoid delusions.	54; schizophrenia, schizoaffective, or delusional disorder; DSM-IV-TR.	40.24 (12.95); female: 17, male: 24; NR.	Observational	Digital wristwatch and diaries; paranoia, activity related stress, EA, self-esteem, negative mood; 10 surveys per day; 6 days.	Symptom worsening	Low self-esteem ($\beta = 0.091$, SE = 0.026, $p < 0.001$) and EA contributed to the paranoia ($\beta = 0.110$, SE = 0.018, $p < 0.001$.)
(Vaessen et al., 2019); Europe	To investigate affective recovery in response to naturally occurring stressors in everyday life.	333; early psychosis: 141 (CAARMS or SPI- AV), chronic psychosis: 192 (OPCRIT or RDC).	Early psychosis: 24.88 (NR), chronic psychosis: 36.27(NR); early psychosis: female: 59, male: 68, chronic psychosis: female: 68, male: 94; NR.	Observational	Digital wristwatch and paper diaries or electronic device; NA, tension, suspiciousness, stressful events; 10 daily surveys; 6 days.	Symptom worsening	↑ suspiciousness followed stressful events (Early psychosis: $\beta = 0.244$; SE = 0.095; $p = 0.011$; Chronic psychosis: $\beta = 0.177$; SE = 0.059; $p = 0.003$).
(Varese et al., 2011); UK	Examine relationship between AHs and dissociative experiences in daily life.	54; schizophrenia spectrum; NR.	Hallucinating: 40.09 (13.56), Non-hallucinating: 40.14 (12.36); male: 24; NR.	Observational	Wristwatch; AH, paranoia, stress, dissociation, experiential avoidance, dissociation; 10 surveys per day; 6 days.	Symptom worsening	AHs predicted by dissociation (OR = 1.20, $p < 0.01$). AHs associated with greater paranoia (OR = 1.24, $p < 0.05$). Relationship between dissociation and AHs stronger for high stress (OR = 1.52, $p < 0.01$).
(Wang et al., 2016); USA	To collect passive monitoring of mental health indicators to model changes in mental health.	48; schizophrenia, schizoaffective disorder, or psychosis; DSM-IV or DSM-V.	NR; female: 17; male: 17; African American: 11; Asian: 2; Caucasian: 19; Multiracial: 1; did not disclose: 1.	Interim data from RCT – (CrossCheck versus treatment-as usual). Data from the CrossCheck arm of the study only.	Samsung Galaxy S5 Android (CrossCheck); passive sensing (activity, sleep, sociability, audio, accelerometer, light, geolocation, phone usage) and surveys (NA, stress, hallucinations, paranoia); 1 survey per day; 3 days per week; 12 months.	Symptom worsening	Higher scores associated with staying stationary more in the morning but less in the evening, visiting fewer new places, fewer conversations, making more phone calls and SMS, and using the phone less. Higher hearing voices scores associated with staying in quieter environments (all ps < 0.05).
(Wang et al., 2020); USA	To predict whether or not relapse will occur the next day from passive mobile phone and self-report data.	75; schizophrenia; NR.	female: 26; male: 35; African American: 24; Asian: 5; Multiracial: 2; Caucasian: 29, Unknown: 1.	Data from an RCT – (CrossCheck versus treatment-as usual). Data from the CrossCheck arm of the study only.	Samsung Galaxy S5 Android (CrossCheck); passive sensing (activity, sleep, sociability, audio, accelerometer, light, geolocation) and surveys (NA, stress, hallucinations, paranoia); 1 survey per day; 3 days per week; 12 months.	Relapse = 1) psychiatric hospitalization 2) increased services, 3) increased medication and 25% increase in BPRS, 4) increased risk.	Best model combined passive and EMA data (F1 = 0.274). More conversations in the morning ($\beta = 2.631$), walk more in the evening ($\beta = 2.553$), visit fewer places in the evening ($\beta = -1.952$), visit fewer

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/ other; Race/ Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
							educational (β = −2.139) travel (β = −1.597) and residential places (β = −1.506), lower score in seeing things (β = −1.650), more time responding to EMAs (β = 1.876) more likely to relapse.
(Wigman et al., 2015); The Netherlands	To examine psychopathology from a network perspective.	263; psychosis; varied diagnostic criteria across 8 studies.	Psychosis: 35.5 (11.0); males: 179 (68%); NR.	Observational: combined data across EMA studies to undertake network analysis	Wristwatch and hardcopy booklet; 10 surveys per day; 5–6 days.	Presence of symptoms: suspiciousness.	Network analysis in psychosis group showed suspiciousness predicted by insecurity and feeling down.
(Zhou et al., 2022); USA	To develop clustering models to obtain behavioural representations from continuous multimodal sensing data to predict impending relapse.	75; schizophrenia; NR	37.2 (13.7); men: 27 (43%), women: 36 (57%); NR.	Observational: passive sensing and EMA.	Smartphone Samsung Galaxy S5 (CrossCheck); 10 EMA self-report items, Multimodal behavioural sensing (features included accelerometer magnitude, ambient light, distance travelled, call duration, sound level, conversation duration, screen use); > 12 months per patient	Relapse = psychiatric hospitalization, increased frequency or intensity of services, increased medications or dosages, ≥ 25% change in BPRS score, suicidal ideation, homicidal ideation, self-injury and violent behavior resulting in harm to self, or others.	Highest F2 score = 0.23 including baseline and clustering features, significantly higher than a random classification baseline (average F2 score =− 0.042).
Other methodologies							
(Birnbaum et al., 2019); USA	To identify and predict early relapse warning signs in social media activity in individuals receiving psychiatric care.	110; schizophrenia: 34 (66.66%), schizoaffective disorder: 13 (25.49%), unspecified schizophrenia spectrum disorder: 4 (7.84%); diagnosed by mental health professional.	23.96 (4.59); male: 70.58%; Asian: 5 (9.80%), African American: 28 (54.90%), Caucasian: 11 (21.56%), Other/Mixed: 7 (13.72%).	Observational using social media posts from Facebook.	Facebook posts; All Facebook activity 1 month prior to hospital admission records.	Relapse= hospitalization.	Individual-centric classifier achieved specificity of 0.71 in predicting psychotic relapse using linguistic and behavioural data in month prior to relapse.
(Birnbaum, Wen, et al., 2020); USA	To explore the feasibility of utilising online search archives as a tool to identify emerging psychiatric symptoms.	36; Schizophrenia: 16(15%), Schizophreniform: 8 (8%), Schizoaffective: 1 (1%), Brief Psychotic Disorder: 2(2%), Unspecified SSD: 9 (9%).	23.11 (3.3); male: 22(61%); African American/ Black: 16 (44%), Asian: 5 (13.9%), Caucasian: 12 (33.3%), Mixed Race/Other: 3 (8.3%), Hispanic: 9 (25%).	Observational	Google search history archive: all historical search activity including frequency, timing, content of search queries 52 weeks prior to first hospitalization.	Relapse = Hospitalization.	Participants search less (p = 0.030) during mornings (12 am–6 am); Content: less punctuation (p = 0.003), less search for terms related to seeing (p = 0.010), anger (p = 0.023), negative emotions (p = 0.040), perception (p =

(*continued on next page*)

(*continued*)

Author and country	Population			Study Design			Main Outcomes
	Primary Aim	Sample size; diagnoses; diagnostic criteria	Demographics: Age years: Mean (SD); Gender: no. female/male/other; Race/Ethnicity: %	Design	Measurement details (name of instrument); items measured; frequency; duration	Outcome criteria	
(Birnbaum, Kulkarni, et al., 2020); USA	To develop computational algorithms based on internet search activity designed to support relapse identification.	44; Schizophrenia spectrum disorder; NR.	NR separately for schizophrenia group; overall range 15–35 years; NR for schizophrenia group.	Observational: google search 4 weeks before and 4 weeks after relapse.	Google search history archive: all historical search activity including Length of queries, usage of linguistic inquiry and word count, variance in word frequency, Increased/Reduced usage of specific word features.	Relapse = Hospitalization.	0.030), and death ($p = 0.040$). Change in use of search term categories: sexual (↑), health (↓), hear (↑), anger (↑), sadness (↓), perception (↑); reduction in search length and frequency; Top 20 relapse classifiers ranged from 0.0097 to 0.0688 (SVM).

CATEGO and PSE = Present State Examination (Wing, Cooper, & Sartorius, 1974); NR = Not Reported; ESS = The Early Signs Scale; PSE = Present State Examination (Wing et al., 1974); PANSS = The Positive and Negative Syndrome Scale (Kay, Fiszbein, & Opler, 1987);CDS = Calgary Depression Scale (Addington, Addington, & Matickatyndale, 1993); BSC = Basic Symptoms Checklist (Eisner et al., 2019); ICD-9 (World Health Organization, 1979); RDC = Research Diagnostic Criteria (Spitzer, Endicott, & Robins, 1978); ESQ = Early Symptom Questionnaire (Herz & Melville, 1980) FoRSe = Fear of Recurrence Scale; BPRS = Brief Psychiatric Rating Scale (Overall & Gorham, 1962); EWSQ = Early Warning Signs Questionnaire; DSM-IV (American Psychiatric Association, 1994); ESM = Idiosyncratic early signs monitoring; EMA = Ecological Momentary Assessment; EMI = Ecological Momentary Intervention; ESQ = Early Symptom Questionnaire; DSM-IV-TR (American Psychiatric Association, 2000); NOS = Not otherwise specified; ESP Software (Barrett & Feldman Barrett, 2000);CASH = Comprehensive Assessment of Symptoms and History (Andreasen, Flaum, & Arndt, 1992);DSM-5 (American Psychological Association, 2013); REDCap = Research Electronic Data Capture; RCT = Randomised controlled trial; NA = Negative Affect; pFDR = false discovery rate-corrected *p* values;MINI = Mini International Neuropsychological Interview (Lecrubier et al., 1997); ES = Effect Size; ICD -10 (World Health Organization, 1992); SE = Sleep efficiency; SF = Sleep fragmentation; DSM-III-R.(American Psychiatric Association, 1987); DMs = delusional moments; ER = Emotion regulation; AHs = Auditory hallucinations; VHs = Visual hallucinations; OPCRIT = Operational criteria system (McGuffin, Farmer, & Harvey, 1991); BCI = Behaviourally controlling interactions; EE = Expressed emotion; EA = Experiential avoidance; CAARMS = Comprehensive Assessment of At-Risk Mental State (Yung et al., 1998); SPI- AV = Schizophrenia Prediction Instrument, Adult version (Schultze-Lutter, Addington, & Ruhrmann, 2007); SVM = Support Vector Machine; NS = Not significant.

Appendix B. Quality assessment outcomes from quality assessment tool for observational and cross-sectional studies

Study	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Item 12	Item 13	Item 14
Adler et al., 2020	Yes	Yes	Yes	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Allan et al., 2023	Yes	Yes	Yes	No	No	Yes	Yes	NA	Yes	Yes	Yes	NA	No	NA
Bak et al., 2016	Yes	Yes	NA	NA	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Barnett et al., 2018	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	Yes	Yes
Bell et al., 2018	Yes	Yes	NA	NA	Yes	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Ben-Zeev et al., 2011	Yes	No	Yes	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Ben-Zeev, Morris, et al., 2012	Yes	No	Yes	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	Yes	Yes
Ben-Zeev, Frounfelker, et al., 2012	Yes	Yes	NR	NR	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Ben-Zeev et al., 2017	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	No
Birchwood et al., 1989	Yes	Yes	Yes	No	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	No
Birnbaum et al., 2019	Yes	Yes	NR	No	No	Yes	Yes	NA	Yes	NA	Yes	NA	Yes	Yes
Birnbaum, Wen, et al., 2020	Yes	Yes	NR	No	No	Yes	Yes	NA	No	NA	Yes	NA	NA	No
Birnbaum, Kulkarni, et al., 2020	Yes	Yes	NR	No	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Buck, Hallgren, et al., 2019	Yes	Yes	Yes	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	Yes	Yes
Buck, Scherer, et al., 2019	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	Yes	Yes	NA	NA	No
Buck et al., 2021	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	Yes	Yes	No	NR	Yes
Cohen et al., 2023	Yes	Yes	NR	No	No	No	Yes	NA	No	Yes	Yes	NR	NR	No
Daemen et al., 2022	Yes	Yes	NR	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	NA	No	No
Dupuy et al., 2021	Yes	Yes	Yes	Yes	No	Yes	Yes	NA	Yes	No	Yes	NA	NA	Yes
Eisner, Bucci, et al., 2019	Yes	Yes	No	Yes	Yes	Yes	Yes	NA	Yes	Yes	Yes	No	NA	Yes
Gaebel & Riesbeck, 2014	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NR	Yes
Gaebel et al., 2000	Yes	No	NR	NR	No	Yes	Yes	NA	Yes	Yes	Yes	No	NR	No
Gaebel & Riesbeck, 2007	Yes	Yes	NR	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes
Geraets et al., 2020	Yes	No	NR	Yes	Yes	Yes	Yes	NA	Yes	Yes	Yes	Yes	Yes	Yes
Gumley et al., 2015	Yes	Yes	Yes	Yes	No	Yes	Yes	NA	Yes	NA	Yes	Yes	Yes	Yes
Hartley et al., 2014	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Hartley et al., 2015	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Hays et al., 2020	Yes	Yes	NR	Yes	No	Yes	Yes	Yes	Yes	NA	Yes	NA	Yes	Yes
Henquet et al., 2010	Yes	Yes	NR	Yes	No	Yes	Yes	NA	No	NA	Yes	NA	NA	Yes

(*continued on next page*)

(*continued*)

Study	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Item 12	Item 13	Item 14
Henson et al., 2021	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	No
Hermans et al., 2020	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Jorgensen, 1998a	Yes	Yes	NR	No	No	Yes	Yes	NA	Yes	NA	Yes	NA	Yes	No
Jorgensen, 1998b	Yes	Yes	NR	No	No	Yes	Yes	NA	Yes	NA	Yes	NA	Yes	No
Kammerer et al., 2021	Yes	Yes	NR	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	NA	Yes
King & Shepherd, 1994	No	Yes	NA	NA	No	Yes	Yes	NA	No	NA	No	NA	NA	No
Klippel et al., 2018	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Klippel et al., 2021	Yes	Yes	NR	Yes	Yes	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Lahti et al., 2021	Yes	Yes	NR	Yes	No	Yes	No	NA	Yes	NA	Yes	NA	Yes	Yes
Lamichhane et al., 2021	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Lamichhane et al., 2023	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	Yes	Yes	Yes	NR	No
Lüdtke et al., 2021	Yes	Yes	NR	No	No	Yes	Yes	NA	Yes	NA	Yes	No	Yes	Yes
Lüdtke et al., 2022	Yes	Yes	NR	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	NR	NR	Yes
Ludwig et al., 2020	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Marder et al., 1991	Yes	Yes	NR	No	No	Yes	NR	NA	Yes	NA	Yes	NA	NR	Yes
Marder et al., 1994	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	Yes	Yes	Yes
Meyer et al., 2021	Yes	Yes	NR	NR	Yes	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Mulligan et al., 2016	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Myin-Germeys et al., 2001	Yes	Yes	NR	NR	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Nittel et al., 2018	Yes	Yes	Yes	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Oorschot, Lataster, Thewissen, Bentall, et al., 2012	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Oorschot, Lataster, Thewissen, Wichers, & Myin-Germeys, 2012	Yes	Yes	NR	No	No	Yes	Yes	Yes	Yes	NA	Yes	NA	NA	Yes
Postma et al., 2021	Yes	Yes	NR	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes
Radley et al., 2022	Yes	Yes	NR	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes
Raugh et al., 2020	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Sa et al., 2016	Yes	Yes	NR	Yes	Yes	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Saito et al., 2020	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	Yes	No	Yes
Sitko et al., 2016	Yes	Yes	NR	NR	Yes	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
So et al., 2021	Yes	No	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Spaniel et al., 2018	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Subotnik & Nuechterlein, 1988	Yes	Yes	NR	Yes	No	Yes	NR	Yes	Yes	Yes	Yes	No	NR	No
Swendsen et al., 2011	Yes	Yes	Yes	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Tait et al., 2002	Yes	No	NR	No	No	Yes	Yes	NA	Yes	NA	No	Yes	NA	No
Torous et al., 2018	Yes	Yes	NR	Yes	No	Yes	Yes	NA	Yes	NA	No	NA	NA	Yes
Tseng et al., 2020	Yes	No	NR	NR	No	Yes	Yes	NA	Yes	Yes	No	NA	NA	Yes
Udachina et al., 2014	Yes	Yes	NR	No	No	Yes	Yes	NA	Yes	Yes	Yes	NA	NA	Yes
Vaessen et al., 2019	Yes	Yes	No	Yes	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Varese et al., 2011	Yes	Yes	NR	NR	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Wang et al., 2016	Yes	Yes	NR	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	NA	NR	Yes
Wang et al., 2018	Yes	No	NR	No	No	Yes	Yes	Yes	Yes	NA	Yes	Yes	NA	Yes
Wang et al., 2020	Yes	Yes	NR	NR	No	Yes	Yes	NA	Yes	NA	Yes	NA	NR	Yes
Wigman et al., 2015	Yes	Yes	NA	No	No	Yes	Yes	NA	Yes	NA	Yes	NA	NA	Yes
Zhou et al., 2022	Yes	Yes	NR	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	NR	NR	Yes

Notes. Item Questions: 1. Was the research question or objective in this paper clearly stated?; 2. Was the study population clearly specified and defined?; 3. Was the participation rate of eligible persons at least 50%?; 4. Subjects recruited from the same populations? Inclusion and exclusion criteria prespecified?; 5. Was a sample size justification provided?; 6. Were the exposure(s) of interest measured prior to the outcome(s) being measured?; 7. Was the timeframe sufficient to see an association between exposure and outcome if it existed?; 8. Did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?; 9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?; 10. Was the exposure(s) assessed more than once over time?; 11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?; 12. Were the outcome assessors blinded to the exposure status of participants?; 13. Was loss to follow-up after baseline 20% or less?; 14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?.

Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2023.102357.

References

Addington, D., Addington, J., & Matickatyndale, E. (1993). Assessing depression in schizophrenia - the Calgary depression scale. *The British Journal of Psychiatry, 163*, 39–44. ://A1993MN02600006.

Adler, D. A., Ben-Zeev, D., Tseng, V. W. S., Kane, J. M., Brian, R., Campbell, A. T., … Choudhury, T. (2020). Predicting early warning signs of psychotic relapse from passive sensing data: An approach using encoder-decoder neural networks. *JMIR mHealth and uHealth, 8*(8). https://doi.org/10.2196/19962

Allan, S., Bradstreet, S., McLeod, H. J., Gleeson, J., Farhall, J., Lambrou, M., … Gumley, A. I. (2020). Perspectives of patients, carers and mental health staff on early warning signs of relapse in psychosis: A qualitative investigation [Review]. *BJPsych Open, 6*(1). https://doi.org/10.1192/bjo.2019.88. Article e3.

Allan, S., O'Driscoll, C., McLeod, H. J., Gleeson, J., Farhall, J., Morton, E., … Gumley, A. (2023). Fear of psychotic relapse: Exploring dynamic relationships with common early warning signs of relapse using electronic once-a-day self-reports. *Psychosis: Psychological, Social and Integrative Approaches.*. https://doi.org/10.1080/17522439.2022.2162955

Alvarez-Jimenez, M., Priede, A., Hetrick, S. E., Bendall, S., Killackey, E., Parker, A. G., … Gleeson, J. F. (2012). Risk factors for relapse following treatment for first episode psychosis: A systematic review and meta-analysis of longitudinal studies. *Schizophrenia Research, 139*(1–3), 116–128. https://doi.org/10.1016/j.schres.2012.05.007

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed.). American Psychiatric Association.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). American Psychiatric Association.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). American Psychiatric Association.

American Psychological Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing, Inc. https://doi.org/10.1176/appi.books.9780890425596

Andreasen, N. C., Carpenter, W. T., Jr., Kane, J. M., Lasser, R. A., Marder, S. R., & Weinberger, D. R. (2005). Remission in schizophrenia: Proposed criteria and rationale for consensus. *American Journal of Psychiatry, 162*(3), 441–449. http://www.appi.org.

Andreasen, N. C., Flaum, M., & Arndt, S. (1992). The comprehensive assessment of symptoms and history (CASH) - an instrument for assessing diagnosis and psychopathology. *Archives of General Psychiatry, 49*(8), 615–623. ://WOS:A1992JG88600003.

Bak, M., Drukker, M., Hasmi, L., & Van Jim, O. S. (2016). An n=1 clinical network analysis of symptoms and treatment in psychosis. *PLoS One, 11*(9). https://doi.org/10.1371/journal.pone.0162811

Balduzzi, S., Rucker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-Based Mental Health, 22*(4), 153–160. https://doi.org/10.1136/ebmental-2019-300117

Barnett, I., Torous, J., Staples, P., Sandoval, L., Keshavan, M., & Onnela, J.-P. (2018). Relapse prediction in schizophrenia through digital phenotyping: A pilot study. *Neuropsychopharmacology, 43*(8), 1660–1666. https://doi.org/10.1038/s41386-018-0030-z

Barrett, D. J., & Feldman Barrett, L. (2000). The Experience-Sampling Program (ESP). Available http://www.experience-sampling.org/.

Bell, I. H., Fielding-Smith, S. F., Hayward, M., Rossell, S. L., Lim, M. H., Farhall, J., & Thomas, N. (2018). Smartphone-based ecological momentary assessment and intervention in a blended coping-focused therapy for distressing voices: Development and case illustration. *Internet Interventions, 14*, 18–25. https://doi.org/10.1016/j.invent.2018.11.001

Bell, I. H., Lim, M. H., Rossell, S. L., & Thomas, N. (2017). Ecological momentary assessment and intervention in the treatment of psychotic disorders: A systematic review. *Psychiatric Services, 68*(11), 1172–1181. https://doi.org/10.1176/appi.ps.201600523

Ben-Zeev, D., Brian, R., Rui, W., Weichen, W., Campbell, A. T., Hauser, M., … Scherer, E. A. (2017). CrossCheck: Integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. *Psychiatric Rehabilitation Journal, 40*(3), 266–275. https://doi.org/10.1037/prj0000243

Ben-Zeev, D., Ellington, K., Swendsen, J., & Granholm, E. (2011). Examining a cognitive model of persecutory ideation in the daily life of people with schizophrenia: A computerized experience sampling study. *Schizophrenia Bulletin, 37*(6), 1248–1256. https://doi.org/10.1093/schbul/sbq041

Ben-Zeev, D., Frounfelker, R., Morris, S. B., & Corrigan, P. W. (2012). Predictors of self-stigma in schizophrenia: New insights using mobile technologies. *Journal of Dual Diagnosis, 8*(4), 305–314. https://doi.org/10.1080/15504263.2012.723311

Ben-Zeev, D., Morris, S., Swendsen, J., & Granholm, E. (2012). Predicting the occurrence, conviction, distress, and disruption of different delusional experiences in the daily life of people with schizophrenia. *Schizophrenia Bulletin, 38*(4), 826–837. https://doi.org/10.1093/schbul/sbq167

Bighelli, I., Rodolico, A., Garcia-Mieres, H., Pitschel-Walz, G., Hansen, W. P., Schneider-Thoma, J., … Leucht, S. (2021). Psychosocial and psychological interventions for relapse prevention in schizophrenia: A systematic review and network meta. *Lancet Psychiatry, 8*(11), 969–980. https://doi.org/10.1016/s2215-0366(21)00243-1

Birchwood, M., Smith, J., Macmillan, F., Hogg, B., Prasad, R., Harvey, C., & Bering, S. (1989). Predicting relapse in schizophrenia: The development and implementation of an early signs monitoring system using patients and families as observers, a preliminary investigation. *Psychological Medicine, 19*(3), 649–656. https://doi.org/10.1017/S0033291700024247

Birchwood, M., Spencer, E., & McGovern, D. (2000). Schizophrenia: Early warning signs. *Advances in Psychiatric Treatment, 6*, 93–101.

Birnbaum, M. L., Kulkarni, P., van Meter, A., Chen, V., Rizvi, A. F., Arenare, E., … Kane, J. M. (2020). Utilizing machine learning on internet search activity to support the diagnostic process and relapse detection in young individuals with early psychosis: Feasibility study. *JMIR Mental Health, 7*(9). https://doi.org/10.2196/19348

Birnbaum, M. L., Rizvi, A. F., Arenare, E., Van Meter, R., Kane, J. M., Ernala, S. K., & De Choudhury, M. (2019). Detecting relapse in youth with psychotic disorders utilizing patient-generated and patient-contributed digital data from Facebook. *NPJ Schizophrenia, 5*(1), 17. https://doi.org/10.1038/s41537-019-0085-9

Birnbaum, M. L., Wen, H., Van Meter, A., Ernala, S. K., Rizvi, A. F., Arenare, E., … Kane, J. M. (2020). Identifying emerging mental illness utilizing search engine activity: A feasibility study. *PLoS One, 15*(10 October). https://doi.org/10.1371/journal.pone.0240820

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-analysis [book]*. https://doi.org/10.1002/9780470743386

Buck, B., Hallgren, K. A., Campbell, A. T., Choudhury, T., Kane, J. M., & Ben-Zeev, D. (2021). mHealth-assisted detection of precursors to relapse in schizophrenia. *Frontiers in Psychiatry, 12*. https://doi.org/10.3389/fpsyt.2021.642200

Buck, B., Hallgren, K. A., Scherer, E., Brian, R., Wang, R., Wang, W., … Ben-Zeev, D. (2019). Capturing behavioral indicators of persecutory ideation using mobile technology. *Journal of Psychiatric Research, 116*, 112–117. https://doi.org/10.1016/j.jpsychires.2019.06.002

Buck, B., Scherer, E., Brian, R., Wang, R., Wang, W., Campbell, A., … Ben-Zeev, D. (2019). Relationships between smartphone social behavior and relapse in schizophrenia: A preliminary report. *Schizophrenia Research, 208*, 167–172. https://doi.org/10.1016/j.schres.2019.03.014

Ceraso, A., Lin, J. J., Schneider-Thoma, J., Siafis, S., Tardy, M., Komossa, K., … Leucht, S. (2020). Maintenance treatment with antipsychotic drugs for schizophrenia [review]. *Cochrane Database of Systematic Reviews, 2020*(8). https://doi.org/10.1002/14651858.CD008016.pub3. Article Cd008016.

Cohen, A., Naslund, J. A., Chang, S., Nagendra, S., Bhan, A., Rozatkar, A., … Torous, J. (2023). Relapse prediction in schizophrenia with smartphone digital phenotyping during COVID-19: A prospective, three-site, two-country, longitudinal study. *Schizophrenia (Heidelberg, Germany), 9*(1), 6. https://doi.org/10.1038/s41537-023-00332-5

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

Daemen, M., van Amelsvoort, T., Group, I., & Reininghaus, U. (2022). Self-esteem and psychosis in daily life: An experience sampling study. *Journal of Psychopathology and Clinical Science, 131*(2), 182–197. https://doi.org/10.1037/abn0000722

Del Re, A. (2013). Compute.es: Compute Effect Sizes. R package version 0.2–2. https://cran.r-project.org/package=compute.es.

Doebler, P., Holling, H., & Sousa-Pinto, B. (2015). *Meta-analysis of diagnostic accuracy with mada*. R Packages.

Dupuy, M., Abdallah, M., Swendsen, J., N'Kaoua, B., Chanraud, S., Schweitzer, P., … Misdrahi, D. (2021). Real-time cognitive performance and positive symptom expression in schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience*. https://doi.org/10.1007/s00406-021-01296-2

Eckardt, J. P. (2022). Barriers to WHO mental health action plan updates to expand family and caregiver involvement in mental healthcare. *General Psychiatry, 35*(2). https://doi.org/10.1136/gpsych-2022-100784. Article e100784.

Ehlers, C. L., Frank, E., & Kupfer, D. J. (1988). Social zeitgebers and biological rhythms - a unified approach to understanding the etiology of depression. *Archives of General Psychiatry, 45*(10), 948–952. ://WOS:A1988Q416800011.

Eisner, E., Bucci, S., Berry, N., Emsley, R., Barrowclough, C., & Drake, R. J. (2019). Feasibility of using a smartphone app to assess early signs, basic symptoms and psychotic symptoms over six months: A preliminary report. *Schizophrenia Research, 208*, 105–113. https://doi.org/10.1016/j.schres.2019.04.003

Eisner, E., Drake, R., & Barrowclough, C. (2013). Assessing early signs of relapse in psychosis: Review and future directions. *Clinical Psychology Review, 33*(5), 637–653. https://doi.org/10.1016/j.cpr.2013.04.001

Eisner, E., Drake, R. J., Berry, N., Barrowclough, C., Emsley, R., Machin, M., & Bucci, S. (2019). Development and long-term acceptability of ExPRESS, a mobile phone app to monitor basic symptoms and early signs of psychosis relapse. *JMIR mHealth and uHealth, 7*(3). https://doi.org/10.2196/11568. Article e11568.

Emsley, R., Chiliza, B., & Asmal, L. (2013). The evidence for illness progression after relapse in schizophrenia. *Schizophrenia Research, 148*(1–3), 117–121. https://doi.org/10.1016/j.schres.2013.05.016

Emsley, R., Chiliza, B., Asmal, L., & Harvey, B. H. (2013). The nature of relapse in schizophrenia [review]. *BMC Psychiatry, 13*. https://doi.org/10.1186/1471-244X-13-50. Article 50.

Fernandez-Castilla, B., Jamshidi, L., Declercq, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2020). The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behavior Research Methods, 52*(5), 2031–2052. https://doi.org/10.3758/s13428-020-01373-9

Frank, E., Kupfer, D. J., Thase, M. E., Mallinger, A. G., Swartz, H. A., Fagiolini, A. M., … Monk, T. (2005). Two-year outcomes for interpersonal and social rhythm therapy in individuals with bipolar I disorder. *Archives of General Psychiatry, 62*(9), 996–1004. https://doi.org/10.1001/archpsyc.62.9.996

Gaebel, W., Janner, M., Frommann, N., Pietzcker, A., Kopcke, W., Linden, M., … Tegeler, J. (2000). Prodromal states in schizophrenia. *Comprehensive Psychiatry, 41*(2), 76–85. https://doi.org/10.1016/S0010-440X(00)80012-7

Gaebel, W., & Riesbeck, M. (2007). Revisiting the relapse predictive validity of prodromal symptoms in schizophrenia. *Schizophrenia Research, 95*(1–3), 19–29.

Gaebel, W., & Riesbeck, M. (2014). Are there clinically useful predictors and early warning signs for pending relapse? *Schizophrenia Research, 152*(2–3), 469–477. https://doi.org/10.1016/j.schres.2013.08.003

Gennatas, E. D., Friedman, J. H., Ungar, L. H., Pirracchio, R., Eaton, E., Reichmann, L. G., … Valdes, G. (2020). Expert-augmented machine learning. *Proceedings of the National Academy of Sciences of the United States of America, 117*(9), 4571–4577. https://doi.org/10.1073/pnas.1906831117

Geraets, C. N. W., Snippe, E., van Beilen, M., Pot-Kolder, R., Wichers, M., van der Gaag, M., & Veling, W. (2020). Virtual reality based cognitive behavioral therapy for paranoia: Effects on mental states and the dynamics among them. *Schizophrenia Research, 222*, 227–234. https://doi.org/10.1016/j.schres.2020.05.047

Gleeson, J. F. M., Alvarez-Jimenez, M., Cotton, S. M., Parker, A. G., & Hetrick, S. (2010). A systematic review of relapse measurement in randomized controlled trials of relapse prevention in first-episode psychosis. *Schizophrenia Research, 119*(1–3), 79–88. https://doi.org/10.1016/j.schres.2010.02.1073

Gumley, A., White, C. A., & Power, K. (1999). An interacting cognitive subsystems model of relapse and the course of psychosis. *Clinical Psychology & Psychotherapy, 6*, 261–278.

Gumley, A. I., Macbeth, A., Reilly, J. D., O'Grady, M., White, R. G., McLeod, H., … Power, K. G. (2015). Fear of recurrence: Results of a randomized trial of relapse detection in schizophrenia. *The British Journal of Clinical Psychology, 54*(1), 49–62. https://doi.org/10.1111/bjc.12060

Hartley, S., Haddock, G., Sa, D. V. E., Emsley, R., & Barrowclough, C. (2014). An experience sampling study of worry and rumination in psychosis. *Psychological Medicine, 44*(8), 1605–1614. https://doi.org/10.1017/S0033291713002080

Hartley, S., Haddock, G., Sa, D. V. E., Emsley, R., & Barrowclough, C. (2015). The influence of thought control on the experience of persecutory delusions and auditory hallucinations in daily life. *Behaviour Research and Therapy, 65*, 1–4. https://doi.org/10.1016/j.brat.2014.12.002

Hartung, J. (1999). An alternative method for meta-analysis. *Biometrical Journal, 41*(8), 901–916. https://doi.org/10.1002/(sici)1521-4036(199912)41:8<901::Aid-bimj901>3.0.Co;2-w

Hartzes, A. M., & Morgan, C. J. (2019). Meta-analysis for diagnostic tests. *Journal of Nuclear Cardiology, 26*(1), 68–71. https://doi.org/10.1007/s12350-018-01485-y

Hays, R., Keshavan, M., Wisniewski, H., & Torous, J. (2020). Deriving symptom networks from digital phenotyping data in serious mental illness. *BJPsych Open, 6*(6). https://doi.org/10.1192/bjo.2020.94

Henquet, C., van Os, J., Kuepper, R., Delespaul, P., Smits, M., Campo, À. J., & Myin-Germeys, I. (2010). Psychosis reactivity to cannabis use in daily life: An experience sampling study. *The British Journal of Psychiatry, 196*(6), 447–453. https://doi.org/10.1192/bjp.bp.109.072249

Henson, P., D'Mello, R., Vaidyam, A., Keshavan, M., & Torous, J. (2021). Anomaly detection to predict relapse risk in schizophrenia. *Translational Psychiatry, 11*(1). https://doi.org/10.1038/s41398-020-01123-7

Hermans, K., van der Steen, Y., Kasanova, Z., van Winkel, R., Reininghaus, U., Lataster, T., … Myin-Germeys, I. (2020). Temporal dynamics of suspiciousness and hallucinations in clinical high risk and first episode psychosis. *Psychiatry Research, 290*. https://doi.org/10.1016/j.psychres.2020.113039

Herz, M. I., & Melville, C. (1980). Relapse in schizophrenia. *American Journal of Psychiatry, 137*(7), 801–805. ://WOS:A1980JZ00400006.

Higgins, J. P. T., Lopez-Lopez, J. A., Becker, B. J., Davies, S. R., Dawson, S., Grimshaw, J. M., … Caldwell, D. M. (2019). Synthesising quantitative evidence in systematic reviews of complex health interventions. *BMJ Global Health, 4*. https://doi.org/10.1136/bmjgh-2018-000858. Article e000858.

Jager, M., Hintermayr, M., Bottlender, R., Strauss, A., & Moller, H. J. (2003). Course and outcome of first-admitted patients with acute and transient psychotic disorders (ICD-10 : F23) - Focus on relapses and social adjustment. *European Archives of Psychiatry and Clinical Neuroscience, 253*(4), 209–215. https://doi.org/10.1007/s00406-003-0435-4

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., … Murray, C. J. L. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: A systematic analysis for the global burden of disease study 2017. *Lancet, 392*(10159), 1789–1858. https://doi.org/10.1016/s0140-6736(18)32279-7

Jorgensen, K. T., Bog, M., Kabra, M., Simonsen, J., Adair, M., & Jonsson, L. (2021). Predicting time to relapse in patients with schizophrenia according to patients' relapse history: A historical cohort study using real-world data in Sweden. *BMC Psychiatry, 21*(1). https://doi.org/10.1186/s12888-021-03634-z. Article 634.

Jorgensen, P. (1998a). Early signs of psychotic relapse in schizophrenia. *The British Journal of Psychiatry, 172*, 327–330. https://doi.org/10.1192/bjp.172.4.327

Jorgensen, P. (1998b). Schizophrenic delusions: The detection of warning signals. *Schizophrenia Research, 32*(1), 17–22. https://doi.org/10.1016/S0920-9964(98)00033-4

Kammerer, M. K., Mehl, S., Ludwig, L., & Lincoln, T. M. (2021). Sleep and circadian rhythm disruption predict persecutory symptom severity in day-to-day life: A combined actigraphy and experience sampling study. *Journal of Abnormal Psychology, 130*(1), 78–88. https://doi.org/10.1037/abn0000645

Kane, J. M. (2007). Treatment strategies to prevent relapse and encourage remission [review]. *The Journal of Clinical Psychiatry, 68*(Suppl. 14), 27–30. https://www.scopus.com/inward/record.uri?eid=2-s2.0-38349099605&partnerID=40&md5=18748b05d6396e538749ad9b66b6a12c.

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin, 13*(2), 261–276.

King, C., & Shepherd, G. (1994). Early signs monitoring through a period of relapse in a highly symptomatic chronic schizophrenic patient. *Behavioural and Cognitive Psychotherapy, 22*(2), 147–152. https://doi.org/10.1017/S1352465800011929

Kirtley, O. J., Hiekkaranta, A. P., Kunkels, Y. K., Eisele, G., Verhoeven, D., Van Nierop, M., & Myin-Germeys, I. (2020, October 1). *The Experience Sampling Method (ESM) Item Repository*. https://doi.org/10.17605/OSF.IO/KG376

Klippel, A., Schick, A., Myin-Germeys, I., Rauschenberg, C., Vaessen, T., & Reininghaus, U. (2021). Modelling the temporal interplay between stress and affective disturbances in pathways to psychosis: An experience sampling study. *Psychological Medicine*. https://doi.org/10.1017/S0033291720004894

Klippel, A., Viechtbauer, W., Reininghaus, U., Wigman, J., van Borkulo, C., … Wichers, M. (2018). The cascade of stress: A network approach to explore differential dynamics in populations varying in risk for psychosis. *Schizophrenia Bulletin, 44*(2), 328–337. https://doi.org/10.1093/schbul/sbx037

Lahti, A. C., Wang, D., Pei, H., Baker, S., & Narayan, V. A. (2021). Clinical utility of wearable sensors and patient-reported surveys in patients with schizophrenia: Noninterventional, observational study. *JMIR Mental Health, 8*(8). https://doi.org/10.2196/26234

Lamichhane, B., Ben-Zeev, D., Campbell, A., Choudhury, T., Hauser, M., Kane, J., … Sano, A. (2021). Patient-independent schizophrenia relapse prediction using Mobile sensor based daily behavioral rhythm changes. In *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST, 362 LNICST* (pp. 18–33). https://doi.org/10.1007/978-3-030-70569-5_2

Lamichhane, B., Zhou, J., & Sano, A. (2023). Psychotic relapse prediction in schizophrenia patients using a personalized mobile sensing-based supervised deep learning model. *IEEE Journal of Biomedical and Health Informatics, 27*(7), 3246–3257. https://doi.org/10.1109/JBHI.2023.3265684

Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., … Dunbar, G. C. (1997). The Mini international neuropsychiatric interview (MINI). A short diagnostic structured interview: Reliability and validity according to the CIDI. *European Psychiatry, 12*(5), 224–231. https://doi.org/10.1016/s0924-9338(97)83296-8

Liao, Y., Skelton, K., Dunton, G., & Bruening, M. (2016). A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: An adapted STROBE checklist for reporting EMA studies (CREMAS). *Journal of Medical Internet Research, 18*(6). https://doi.org/10.2196/jmir.4954. Article e151.

Lüdtke, T., Moritz, S., Westermann, S., & Pfuhl, G. (2022). *Aberrant salience predicts fluctuations of paranoia two weeks in advance during a 1-year experience sampling method study in people with psychosis* (Vol. 248, pp. 361–362). Elsevier B.V.

Lüdtke, T., Pfuhl, G., Moritz, S., Ruegg, N. L., Berger, T., & Westermann, S. (2021). Sleep problems and worrying precede psychotic symptoms during an online intervention for psychosis. *The British Journal of Clinical Psychology, 60*(1), 48–67. https://doi.org/10.1111/bjc.12270

Ludwig, L., Mehl, S., Schlier, B., Krkovic, K., & Lincoln, T. M. (2020). Awareness and rumination moderate the affective pathway to paranoia in daily life. *Schizophrenia Research, 216*, 161–167. https://doi.org/10.1016/j.schres.2019.12.007

Marder, S. R., Mintz, J., Van Putten, T., Lebell, M., Wirshing, W. C., & Johnston-Cronk, K. (1991). Early prediction of relapse in schizophrenia: An application of receiver operating characteristic (ROC) methods. *Psychopharmacology Bulletin, 27*(1), 79–82. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0025781617&partnerID=40&md5=3d0b055665cf7af36bff0fe8f9b329d5.

Marder, S. R., Wirshing, W. C., Vanputten, T., Mintz, J., McKenzie, J., Johnstoncronk, K., … Liberman, R. P. (1994). Fluphenazine vs placebo supplementation for prodromal signs of relapse in schizophrenia. *Archives of General Psychiatry, 51*(4), 280–287. ://WOS:A1994NF26500003.

McGuffin, P., Farmer, A., & Harvey, I. (1991). A polydiagnostic application of operational criteria in studies of psychotic illness - development and reliability of the OPCRIT system. *Archives of General Psychiatry, 48*(8), 764–770. ://WOS:A1991GA24600013.

Meyer, N., Joyce, D. W., Karr, C., De Vos, M., Dijk, D. J., Jacobson, N. C., & Maccabe, J. H. (2021). The temporal dynamics of sleep disturbance and psychopathology in psychosis: A digital sampling study. *Psychological Medicine*. https://doi.org/10.1017/S0033291720004857

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Grp, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264–W264. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Mulligan, L. D., Emsley, R., Haddock, G., Neil, S. T., & Kyle, S. D. (2016). High resolution examination of the role of sleep disturbance in predicting functioning and psychotic symptoms in schizophrenia: A novel experience sampling study. *Journal of Abnormal Psychology, 125*(6), 788–797. https://doi.org/10.1037/abn0000180

Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., & Reininghaus, U. (2018). Experience sampling methodology in mental health research: New insights and technical developments. *World Psychiatry, 17*(2), 123–132. https://doi.org/10.1002/wps.20513

Myin-Germeys, I., Klippel, A., Steinhart, H., & Reininghaus, U. (2016). Ecological momentary interventions in psychiatry. *Current Opinion in Psychiatry, 29*(4), 258–263. https://doi.org/10.1097/yco.0000000000000255

Myin-Germeys, I., Nicolson, N. A., & Delespaul, P. (2001). The context of delusional experiences in the daily life of patients with schizophrenia. *Psychological Medicine, 31*(3), 489–498. https://doi.org/10.1017/S0033291701003646

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience sampling research in psychopathology: Opening the black box of daily life. *Psychological Medicine, 39*(9), 1533–1547. https://doi.org/10.1017/s0033291708004947

National Heart Lung and Blood Institute. (2023). Study Quality Assessment Tools. Available online https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools.

Nittel, C. M., Lincoln, T. M., Lamster, F., Leube, D., Rief, W., Kircher, T., & Mehl, S. (2018). Expressive suppression is associated with state paranoia in psychosis: An experience sampling study on the association between adaptive and maladaptive emotion regulation strategies and paranoia. *The British Journal of Clinical Psychology, 57*(3), 291–312. https://doi.org/10.1111/bjc.12174

Nuechterlein, K. H., Snyder, K. S., & Mintz, J. (1992). Paths to relapse: Possible transactional processes connecting patient illness onset, expressed emotion, and psychotic relapse. *The British Journal of Psychiatry, 161*(Suppl. 18), 88–96. http://www.rcpsych.ac.uk.

Oorschot, M., Lataster, T., Thewissen, V., Bentall, R., Delespaul, P., & Myin-Germeys, I. (2012). Temporal dynamics of visual and auditory hallucinations in psychosis. *Schizophrenia Research, 140*(1–3), 77–82. https://doi.org/10.1016/j.schres.2012.06.010

Oorschot, M., Lataster, T., Thewissen, V., Wichers, M., & Myin-Germeys, I. (2012). Mobile assessment in schizophrenia: A data-driven momentary approach. *Schizophrenia Bulletin, 38*(3), 405–413. https://doi.org/10.1093/schbul/sbr166

Overall, J. E., & Gorham, D. R. (1962). The brief psychiatric rating scale. *Psychiatric Reports, 10*, 799–812.

Pennington, M., & McCrone, P. (2017). The cost of relapse in schizophrenia. *Pharmacoeconomics, 35*(9), 921–936. https://doi.org/10.1007/s40273-017-0515-3

Phahladira, L., Luckhoff, H. K., Asmal, L., Kilian, S., Scheffler, F., Plessis, S. D., … Emsley, R. (2020). Early recovery in the first 24 months of treatment in first-episode schizophrenia-spectrum disorders. *NPJ Schizophrenia, 6*(1). https://doi.org/10.1038/s41537-019-0091-y. Article 2.

Postma, M. R., van Amelsvoort, T., Myin-Germeys, I., Gayer-Anderson, C., Kempton, M. J., Valmaggia, L., … Reininghaus, U. (2021). Across the continuum: Associations between (fluctuations in) momentary self-esteem and psychotic

experiences. *Schizophrenia Research, 238*, 188–198. https://doi.org/10.1016/j.schres.2021.09.010

Radley, J., Barlow, J., & Johns, L. C. (2022). Parenting and psychosis: An experience sampling methodology study investigating the inter-relationship between stress from parenting and positive psychotic symptoms. *The British Journal of Clinical Psychology, 61*(4), 1236–1258. https://doi.org/10.1111/bjc.12389

Raugh, I. M., James, S. H., Gonzalez, C. M., Chapman, H. C., Cohen, A. S., Kirkpatrick, B., & Strauss, G. P. (2020). Geolocation as a digital phenotyping measure of negative symptoms and functional outcome. *Schizophrenia Bulletin, 46*(6), 1596–1607. https://doi.org/10.1093/schbul/sbaa121

Rodolico, A., Bighelli, I., Avanzato, C., Concerto, C., Cutrufelli, P., Mineo, L., … Leucht, S. (2022). Family interventions for relapse prevention in schizophrenia: A systematic review and network meta-analysis. *Lancet Psychiatry, 9*(3), 211–221. https://doi.org/10.1016/s2215-0366(21)00437-5

Sa, D. V. E., Wearden, A., Hartley, S., Emsley, R., & Barrowclough, C. (2016). Expressed emotion and behaviourally controlling interactions in the daily life of dyads experiencing psychosis. *Psychiatry Research, 245*, 406–413. https://doi.org/10.1016/j.psychres.2016.08.060

Saito, Y., Sakurai, H., Kane, J. M., Schooler, N. R., Suzuki, T., Mimura, M., & Uchida, H. (2020). Predicting relapse with residual symptoms in schizophrenia: A secondary analysis of the PROACTIVE trial. *Schizophrenia Research, 215*, 173–180. https://doi.org/10.1016/j.schres.2019.10.037

Santesteban-Echarri, O., Paino, M., Rice, S., González-Blanch, C., McGorry, P., Gleeson, J., & Alvarez-Jimenez, M. (2017). Predictors of functional recovery in first-episode psychosis: A systematic review and meta-analysis of longitudinal studies [review]. *Clinical Psychology Review, 58*, 59–75. https://doi.org/10.1016/j.cpr.2017.09.007

Schneider-Thoma, J., Chalkou, K., Dorries, C., Bighelli, I., Ceraso, A., Huhn, M., … Leucht, S. (2022). Comparative efficacy and tolerability of 32 oral and long-acting injectable antipsychotics for the maintenance treatment of adults with schizophrenia: A systematic review and network meta-analysis. *Lancet, 399*(10327), 824–836. <Go to ISI>://WOS:000762307100023.

Schultze-Lutter, F., Addington, J., & Ruhrmann, S. (2007). *Schizophrenia proneness instrument, adult version (SPI-A)*. Giovanni.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*, 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Shim, S. R., Kim, S. J., & Lee, J. (2019). Diagnostic test accuracy: Application and practice using R software. *Epidemiology and Health, 41*. https://doi.org/10.4178/epih.e2019007. Article e2019007.

Sitko, K., Varese, F., Sellwood, W., Hammond, A., & Bentall, R. (2016). The dynamics of attachment insecurity and paranoid thoughts: An experience sampling study. *Psychiatry Research, 246*, 32–38. https://doi.org/10.1016/j.psychres.2016.08.057

So, S. H. W., Chung, L. K. H., Tse, C. Y., Chan, S. S. M., Chong, G. H. C., Hung, K. S. Y., & Sommer, I. E. C. (2021). Moment-to-moment dynamics between auditory verbal hallucinations and negative affect and the role of beliefs about voices. *Psychological Medicine, 51*(4), 661–667. https://doi.org/10.1017/S0033291719003611

Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., de Pablo, G. S., … Fusar-Poli, P. (2022). Age at onset of mental disorders worldwide: Large-scale meta-analysis of 192 epidemiological studies. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-021-01161-7

Spaniel, F., Bakstein, E., Anyz, J., Hlinka, J., Sieger, T., Hrdlicka, J., … Hoschl, C. (2018). Relapse in schizophrenia: Definitively not a bolt from the blue. *Neuroscience Letters, 669*, 68–74. https://doi.org/10.1016/j.neulet.2016.04.044

Spitzer, R. L., Endicott, J., & Robins, E. (1978). Research diagnostic criteria. *Archives of General Psychiatry, 35*, 773–782.

Subotnik, K. L., & Nuechterlein, K. H. (1988). Prodromal signs and symptoms of schizophrenic relapse. *Journal of Abnormal Psychology, 97*(4), 405–412. https://doi.org/10.1037/0021-843x.97.4.405

Swendsen, J., Ben-Zeev, D., & Granholm, E. (2011). Real-time electronic ambulatory monitoring of substance use and symptom expression in schizophrenia. *American Journal of Psychiatry, 168*(2), 202–209. https://doi.org/10.1176/appi.ajp.2010.10030463

Tait, A., McNay, L., Gumley, A., & O'Grady, M. (2002). The development and implementation of an individualised early signs monitoring system in the prediction of relapse in schizophrenia. *Journal of Mental Health, 11*(2), 141–153. https://doi.org/10.1080/096382301200041551

Takeuchi, H., Siu, C., Remington, G., Fervaha, G., Zipursky, R. B., Foussias, G., & Agid, O. (2019). Does relapse contribute to treatment resistance? Antipsychotic response in first- vs. second-episode schizophrenia. *Neuropsychopharmacology, 44*(6), 1036–1042. https://doi.org/10.1038/s41386-018-0278-3

Torous, J., Staples, P., Barnett, I., Sandoval, L. R., Keshavan, M., & Onnela, J.-P. (2018). Characterizing the clinical relevance of digital phenotyping data quality with applications to a cohort with schizophrenia. *npj Digital Medicine, 1*, 15. https://doi.org/10.1038/s41746-018-0022-8

Trifan, A., Oliveira, M., & Oliveira, J. L. (2019). Passive sensing of health outcomes through smartphones: Systematic review of current solutions and possible limitations. *JMIR mHealth and uHealth, 7*(8). https://doi.org/10.2196/12649. Article e12649.

Tseng, V. W. S., Sano, A., Ben-Zeev, D., Brian, R., Campbell, A. T., Hauser, M., … Choudhury, T. (2020). Using behavioral rhythms and multi-task learning to predict fine-grained symptoms of schizophrenia. *Scientific Reports, 10*(1). https://doi.org/10.1038/s41598-020-71689-1

Udachina, A., Varese, F., Myin-Germeys, I., & Bentall, R. P. (2014). The role of experiential avoidance in paranoid delusions: An experience sampling study. *The British Journal of Clinical Psychology, 53*(4), 422–432. https://doi.org/10.1111/bjc.12054

Vaessen, T., Viechtbauer, W., van der Steen, Y., Gayer-Anderson, C., Kempton, M. J., Valmaggia, L., … Myin-Germeys, I. (2019). Recovery from daily-life stressors in early and chronic psychosis. *Schizophrenia Research, 213*, 32–39. https://doi.org/10.1016/j.schres.2019.03.011

Varese, F., Udachina, A., Myin-Germeys, I., Oorschot, M., & Bentall, R. (2011). The relationship between dissociation and auditory verbal hallucinations in the flow of daily life of patients with psychosis. *Psychosis, 3*(1), 14–28. https://doi.org/10.1080/17522439.2010.548564

Waite, F., Sheaves, B., Isham, L., Reeve, S., & Freeman, D. (2020). Sleep and schizophrenia: From epiphenomenon to treatable causal target. *Schizophrenia Research, 221*, 44–56. <Go to ISI>://WOS:000542922400007.

Wang, D., Gopal, S., Baker, S., & Narayan, V. A. (2018). Trajectories and changes in individual items of positive and negative syndrome scale among schizophrenia patients prior to impending relapse. *NPJ Schizophrenia, 4*(1). https://doi.org/10.1038/s41537-018-0056-6

Wang, R., Aung, M. S. H., Abdullah, S., Brian, R., Campbell, A. T., Choudhury, T., … Ben-Zeev, D. (2016). CrossCheck: Toward passive sensing and detection of mental health changes in people with schizophrenia. In *UbiComp 2016 - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 886–897). https://doi.org/10.1145/2971648.2971740

Wang, R., Wang, W., Obuchi, M., Scherer, E., Brian, R., Ben-Zeev, D., … Campbell, A. (2020). On predicting relapse in schizophrenia using mobile sensing in a randomized control trial. In *18th Annual IEEE International Conference on Pervasive Computing and Communications, PerCom 2020*. https://doi.org/10.1109/PerCom45495.2020.9127365

Wigman, J. T. W., Van Os, J., Borsboom, D., Wardenaar, K. J., Epskamp, S., Klippel, A., … Wichers, M. (2015). Exploring the underlying structure of mental disorders: Cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. *Psychological Medicine, 45*(11), 2375–2387. https://doi.org/10.1017/S0033291715000331

Wing, J. K., Cooper, J. E., & Sartorius, N. (1974). *Measurement and classification of psychiatric symptoms; An instruction manual for the PSE and Catego program*. Cambridge U Press. https://search.ebscohost.com/login.aspx?direct=true&AuthType=shib&db=psyh&AN=1975-03367-000&site=ehost-live&scope=site&custid=s5501413.

World Health Organization. (1979). *The ICD-9 classification of mental and behavioural disorders*. The World Health Organization.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders*. The World Health Organization.

Yung, A. R., Phillips, L. J., McGorry, P. D., McFarlane, C. A., Francey, S., Harrigan, S., … Jackson, H. J. (1998). Prediction of psychosis - A step towards indicated prevention of schizophrenia. *The British Journal of Psychiatry, 172*, 14–20. https://doi.org/10.1192/s0007125000297602

Yung, A. R., Phillips, L. J., Yuen, H. P., & McGorry, P. D. (2004). Risk factors for psychosis in an ultra high-risk group: Psychopathology and clinical features. *Schizophrenia Research, 67*(2–3), 131–142. https://doi.org/10.1016/S0920-9964(03)00192-0

Zhou, J., Lamichhane, B., Ben-Zeev, D., Campbell, A., & Sano, A. (2022). Predicting psychotic relapse in schizophrenia with mobile sensor data: Routine cluster analysis. *JMIR mHealth and uHealth, 10*(4), Article e31006. https://doi.org/10.2196/31006

Zukowska, Z., Allan, S., Eisner, E., Ling, L., & Gumley, A. (2022). Fear of relapse in schizophrenia: A mixed-methods systematic review. *Social Psychiatry and Psychiatric Epidemiology, 57*(7), 1305–1318. https://doi.org/10.1007/s00127-022-02220-2