Measuring Fear of Recurrence in Psychosis

Arann Ronan Rowe, B.Sc. (Hons), MSc.

Submitted in partial fulfilment of the requirements for the degree of
Doctorate in Clinical Psychology

School of Health and Wellbeing
College of Medical, Veterinary and Life Sciences
University of Glasgow

August 2025

Table of Contents

Appendices

Table of Tables

Chapter One: Systematic Review

Chapter Two: Major Research Project

Table of Figures

Chapter One: Systematic Review

Chapter Two: Major Research Project

Acknowledgements

For my parents, thank you for all you have done for me, without your unending support and faith in me none of this would have been possible.

To my supervisor Andrew, thank you for your warm guidance, patience, good humour and support.

To my partner, without whom I would not have made it this far, thank you for your incredible support.

My thanks to the members of the Psychosis Research Group for their guidance and advice, especially to the PPI members for their input in designing the Cognitive Interviewing study. Also to the Reference Group for contributing their time and expertise to this project.

Finally to the participants of the Cognitive Interviewing study, for sharing their lived experience of psychosis and fear of recurrence, without which this study would not be possible.

Chapter 1

A Systematic Review of the Fear of Recurrence Scale: Psychometric Properties and Clinical Associations

Prepared in accordance with the author requirements for Frontiers In Psychiatry (https://www.frontiersin.org/journals/psychiatry/for-authors/author-guidelines)

Abstract

Background and Aims: Schizophrenia (SCZ) spectrum disorders are a group of mental health conditions with hallucinations, delusional thinking, cognitive impairment and negative symptoms, associated with high levels of distress for both the person and their carers. Remission and recovery are common but for 65% of individuals are followed by relapse within six years. Fear of Recurrence in Psychosis is an understudied clinical construct that captures a persistent sense of fear of a return to a state of psychosis. To date, one measurement tool has been developed to measure Fear of Recurrence in Psychosis, the Fear of Recurrence Scale (FoRSe). This systematic review aimed to review and summarise the evidence relating to its psychometric properties and its associations with clinical outcomes. **Methods:** A systematic review of five electronic databases (Embase, Medline, PsycInfo, CINAHL and Scopus) with forward and backward searches conducted via Web of Science was conducted to identify relevant studies published between 2005 and 2023. The MMAT was used to assess studies that reported on clinical associations with the FoRSe, while the COSMIN Guidelines for Systematic Reviews of Patient Reported Outcome Measures were used to assess studies that reported on its psychometric properties. A narrative synthesis was then produced. **Results:** A total of 159 studies were screened with 7 meeting inclusion criteria. One study reported on the psychometric properties of the FoRSe. The FoRSe scores were found to be associated with Post Psychotic Post-Traumatic Stress Disorder (PP-PTSD), relapse and significant distress in people who have experienced an episode of psychosis. The available evidence on its psychometric properties was found to be insufficient to draw conclusions on its use without further research, per COSMIN guidelines. Areas of concern regarding measurement properties were in Content Validity, Structural Validity and Internal Consistency. While Reliability was found to meet sufficient criteria, the risk of bias ratings mean that further research is required. **Conclusions:** The Fear of Recurrence Scale remains the only tool to measure this important construct. While it has been shown to have clinical utility, evidence for its psychometric properties has only been examined in one study. COSMIN criteria suggest these results cannot be used to draw firm conclusions regarding recommendations for its use without further research. As qualitative work in this area suggests that the current version of

the FoRSe does not capture some aspects of Fear of Recurrence, this review recommends that the scale is revised to improve its content validity prior to further psychometric work.

1. Introduction

Schizophrenia (SCZ) spectrum disorders are typically characterised by three groups of symptoms: positive, such as auditory or visual hallucinations and delusional thinking; negative, such as anhedonia, avolition, asociality, blunted affect and alogia; and cognitive impairments (Kirkpatrick et al., 2006, American Psychiatric Association, 2013). It is estimated to have a lifetime prevalence rate of 0.7%; however, differences in diagnostic and regional approaches to mental health care mean this may be an underestimation (Solmi et al., 2023, Simeone et al., 2015). Prevalence is similar for both males and females (Charlson et al., 2018), with onset typically occurring in adolescence/early adulthood. Causal factors in the development of SCZ are complex and multifactorial, with strong evidence for environmental factors, such as early childhood adverse experiences, migration, social adversity and interpersonal trauma (Bebbington et al., 2004, Van Os et al., 2010), as well as complex genetic involvement (Owen et al., 2023). Schizophrenia is associated with a significant impact upon functioning with increased levels of disability, negative impacts on social functioning, socioeconomic disadvantage, functional skills deficits, poor physical health outcomes and an increase in all-cause mortality in comparison to the general population, resulting in a 15 to 20 year reduction in life expectancy (Harvey et al., 2019, Correll et al., 2022). Additionally, individuals experience both increased rates of stigma and discrimination (Hazell et al., 2022, Lauber, 2008). As a result, it is common for individuals to experience significant disability resulting in challenges around independent accommodation, vocational occupation, and reduced life expectancy.

The course of schizophrenia is associated with premorbid impairments in functioning followed by a prodromal phase that can involve a range of affective, negative and non-specific cognitive symptoms, as well as sub-threshold positive symptoms (Benrimoh et al., 2024), sometimes referred to as an at-risk mental state (Nieman and McGorry, 2015). These occur prior to the onset of an acute episode of psychosis, usually resulting in contact with services (Tandon et al., 2024) with subsequent assessment and treatment with anti-psychotic medication and often hospitalisation. While many individuals come to feel that coercive treatment can be necessary or beneficial (Plahouras et al., 2020), it is also commonly

experienced as a violation of human rights, traumatic and further compounding the trauma of experiencing psychosis itself (Shaw et al., 1997, Meyer et al., 1999).

Remission of symptoms, defined as improvement in core symptoms to at least a mild level of intensity where they no longer interfere significantly with behaviour, is common, with 76% experiencing remission six months following a first episode (Andreasen et al., 2005). Recovery, defined as experiencing only mild symptoms as well as improvement both clinically and socially for at least 2 years from a first episode is also common, occurring in 38% of people, and maintained for 32% of individuals at 6 year follow-up (Lally et al., 2017). This indicates that a subgroup of individuals do not experience a relapsing/remitting course, but for others relapse is a common occurrence. There remains a lack of consensus on the definition of relapse (Gleeson et al., 2024) with hospitalisation frequently used as a proxy. However, it is estimated that within 5 years of a first episode 81% of people will experience a relapse (Robinson et al., 1999).

Relapse tends to emerge over the course of several weeks, beginning with lower level psychotic symptoms and affective changes, sleep disturbance and increased suspiciousness, followed by the subsequent emergence of more severe psychotic symptoms (Birchwood et al., 1989, Gleeson et al., 2024). Relapse is associated with damaging effects upon a person's social and occupational functioning, relationships (Taylor et al., 2015, Johansen et al., 2020), increased risk of post psychotic depression (Jäger et al., 2007) and risk of progression of symptoms to persistent non-remitting psychosis (Emsley et al., 2013). Relapse also accounts for the majority of the costs of care, with direct costs mostly accounted for by unscheduled care/rehospitalisation, while indirect costs experienced by persons with schizophrenia and their caregivers are larger (Kotzeva et al., 2023, Ascher-Svanum et al., 2010). When combined, this points to substantial personal, societal and financial costs associated with schizophrenia placed upon the individual, their wider family and the health care system (Pennington and McCrone, 2017).

Psychologically, relapses are distressing experiences for individuals due to the traumatic nature of psychosis itself (Shaw et al., 1997, Meyer et al., 1999) and the increased likelihood

of coercive treatment, both of which have been found to increase risk of meeting criteria for PTSD (Buswell et al., 2021). Relapse is also associated with increased depressive symptomology with associated appraisals of defeat, entrapment, self-blame and uncertainty, thought to be rooted in experiences of loss of autonomy, social/occupational role and experiences of coercive treatment and psychosis itself (Birchwood et al., 2000, Rooke and Birchwood, 1998, Charlson et al., 2018, White and Gumley, 2009).

Understandably, it is common for people who have experienced psychosis to be aware of the consequences of relapse and to fear the possibility of its occurrence and progression (Sandhu et al., 2013, Brookmann, 2020). Fear of Recurrence in Psychosis (FoRP) is a long recognised though understudied construct (Zukowska et al., 2022). A cognitive interpersonal model of relapse that centres on FoRP has been theorised to explain both the phenomenology and mechanisms involved in risk and course of relapse in psychosis (Gumley et al., 2020). Within this model, initial lower level psychotic symptoms trigger FoRP which drives feelings of anxiety, shame, demoralisation, and fear observable as an increase in affective symptoms, prompting individuals to engage in coping strategies to regulate emotional distress and defensive behaviours prompted by feelings of suspiciousness/paranoia. These behaviours can include avoidance, worry, hypervigilance and decreased therapeutic engagement in attempt to avoid confirmation of FoRP. Traumatic experiences of both treatment and psychosis itself may also play a role in increasing distrust in services, inhibiting help seeking and interacting with therapeutic relationships (Gumley et al., 2014). In the context of an interpersonal cycle, care providers may interpret these changes as increased risk of relapse, resulting in emotional (empathy, concern, uncertainty, anxiety and suspiciousness) and behavioural changes (increased monitoring and risk orientated care) on their part. As these service responses are associated with an impending relapse, it is theorised that they act as confirmation of this to the person with psychosis, increasing FoRP and negative expectations of services, leading to an increase in coping behaviours, resulting in a cycle that may drive relapse itself. Recent work (Allan et al., 2023) has shown that FoRP can be a persistent, self-perpetuating experience that maintains emotional distress over time rather than solely a process that occurs in response to a possible relapse. This view is supported by evidence from lived experience perspectives (Zukowska et al., 2022). This suggests that FoRP is a multidimensional psychological feature

of SCZ that may be involved in the maintenance of the affective and functional burden of the condition. By acknowledging these perspectives as well as the nature and impact of SCZ, it seems reasonable that FoRP may be better conceptualised as an understandable response to coping with a persistent and unpredictable illness, the return of which can have significant consequences and as such causes persistent worry or distress.

Psychometrics is the field of measurement concerned with latent constructs (Streiner et al., 2024) and the properties that are required for adequate measurement of such constructs: Reliability and Validity. Reliability is concerned with the ability to measure in a stable and reproducible manner that is free from error. Validity is concerned with the extent to which the scale measures the construct it intends to measure. Both reliability and validity can be thought of as broad domains, which encapsulate specific types of reliability and validity (Appendix 1.1) that can be used to establish the overall psychometric properties of scores a scale produces. Choosing a methodology for examining psychometric properties is complex, choices being influenced by the structure and design of the scale and made further difficult by a disagreement around both acceptable statistical methodologies and criteria for 'good' measurement properties.

This makes undertaking development of a scale or establishing its psychometric properties complex (Boateng et al., 2018, De Vet et al., 2011, Streiner et al., 2024) but also integral to clinical health sciences, as measurement is increasingly used to justify diagnosis, treatment decisions and to assess outcome and progress. As such, flawed measurement can lead to dangerous outcomes for patients such as misdiagnosis, unsuitable treatment with associated suffering, morbidity and mortality (McClimans et al., 2017). Robust measurement is also vital within the development of interventions. Poorly designed or validated tools can lead to inaccurate assessments of the effect of treatments, with subsequent consequences for individuals, health care systems and wider society (Mokkink et al., 2018). Reviews have consistently found widespread use of scales with limited psychometric evidence (Crellin et al., 2015, Cassidy et al., 2018, Ghai et al., 2022), recurrent issues within the development methodologies of new scales (Morgado et al., 2018) as well as the use of unsound psychometric practices in randomised controlled trials (Marshall et al., 2000). The individual

injury of the use of psychometrically flawed scales in clinical contexts is multiplied when they are used within research contexts; it is impossible to rely on the findings, either positive or negative, of a study when the data that it examines is unreliable.

As there is evidence that FoRP presents an important clinical construct which has not received significant research or clinical attention, it is important to establish the psychometric properties of scales available to measure it. To the author's knowledge, there is only one scale specifically designed to measure FoRP. The Fear of Recurrence Scale (FoRSe) was developed by Gumley et al. (2006) to measure FoRP. It measures three constructs: Awareness, Fear of Relapse and Intrusiveness. Its content was derived from the concerns reported by participants in a randomised controlled trial of cognitive behavioural therapy for relapse prevention (Tait et al., 2002, Gumley et al., 2003). To date, no systematic review has considered its psychometric properties or associations with clinical outcomes.

2. Aims

This review aimed to critically appraise and synthesise the literature available on the Fear of Recurrence Scale (FoRSe), including its associations with clinical variables and outcomes, as well as psychometric properties. The review aimed to establish:

1. What are the psychometric characteristics of the FoRSe?

2. What is the evidence for its associations with clinical outcomes in people with psychosis?

3. Method

3.1. Protocol and registrations

This review was registered with the International Prospective Register of Systematic Reviews (PROSPERO) in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement Guidelines (Page et al., 2021). Registration and protocol amendments were registered with PROSPERO (CRD42023439580). Modifications were made

to include the use of the COSMIN Guideline for Systematic Reviews of Patient-Reported Outcome Measures (Mokkink et al., 2024).

3.2 Eligibility Criteria

The review adopted a broad inclusion criterion in order to ensure that all studies that used the FoRSe were included. The study included randomised trials, non-randomised trials, cross-sectional studies, pilot studies and feasibility studies that used the FoRSe in a population of people who have experienced psychosis. No formal diagnosis was required for inclusion due to the changing landscape of diagnostic definitions over time. Instead, evidence of psychosis was identified via the use of structured assessment tools, reporting of relevant symptoms or having received care from a relevant service (e.g., community mental health team or specialist service such as an Early Intervention in Psychosis team). Reviews, conference abstracts and articles not published in English were excluded.

3.3 Search Strategy

Eligible studies were identified using a pre-determined search strategy (Appendix 1.2) across five databases: EMBASE, MEDLINE, PsychINFO, CINAHL and SCOPUS. Forward and backward searches were conducted via the Web of Science platform. A manual search of the reference lists of included papers was completed to identify potential additional records. As no MESH or Subject heading apply to FoRP, a search was developed focusing on references to the Fear of Recurrence Scale in Title/Abstract/Full Text searches and searching by instrument where available (PsycInfo and CINAHL).

The search was performed in June 2023 and limited to articles published after 2005 to that date, as the FoRSe was published in 2006. A second search was performed in February 2025, limited to articles published after the previous search.

3.4 Screening

Results were exported into EndNote X9 and duplicates removed. Articles were reviewed via title/abstract screening and remaining results were full text reviewed. Forward and backward citation searches were conducted following full text screening. Reasons for exclusion were

documented. All papers were screened by the primary researcher and a second reviewer repeated the screening process. Disagreements were resolved by discussion.

3.5 Data Extraction

The following data were extracted by the primary researcher from each study: 1) study design, 2) study description, 3) participant characteristics, 4) outcomes associated with the FoRSe. For studies examining the psychometric properties of the FoRSe, further data were extracted following the COSMIN Guideline for Systematic Reviews of Patient-Reported Outcome Measures (Mokkink et al., 2024).

3.6 Quality Assessment

The Mixed Methods Appraisal Tool (MMAT) (Hong et al., 2018) was used for all studies as it allows for the review of multiple study types.

For studies examining the psychometric properties of the FoRSe, the procedures outlined in the COSMIN Guideline (Mokkink et al., 2024) were applied (Figure 1.1) .

The Cosmin Risk of Bias Tool applies a set of criteria to rate the methodological qualities of studies examining the psychometric properties of a measurement tool. Ratings can be described as ‘very good’, ‘adequate’, ‘doubtful’, ‘inadequate’ or, in the cases where a property is not assessed or the standard is not relevant, ‘not applicable’. The overall quality of a study on a given property is rated using a ‘worst score counts’ method.

The data on each measurement property is then graded using criteria defined by the COMSIN Criteria for Good Measurement Properties. Ratings can be sufficient (+), insufficient (-), or indeterminate (?). A modified Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach can then be used to provide an overall rating of the evidence for the psychometric properties of a measurement tool.

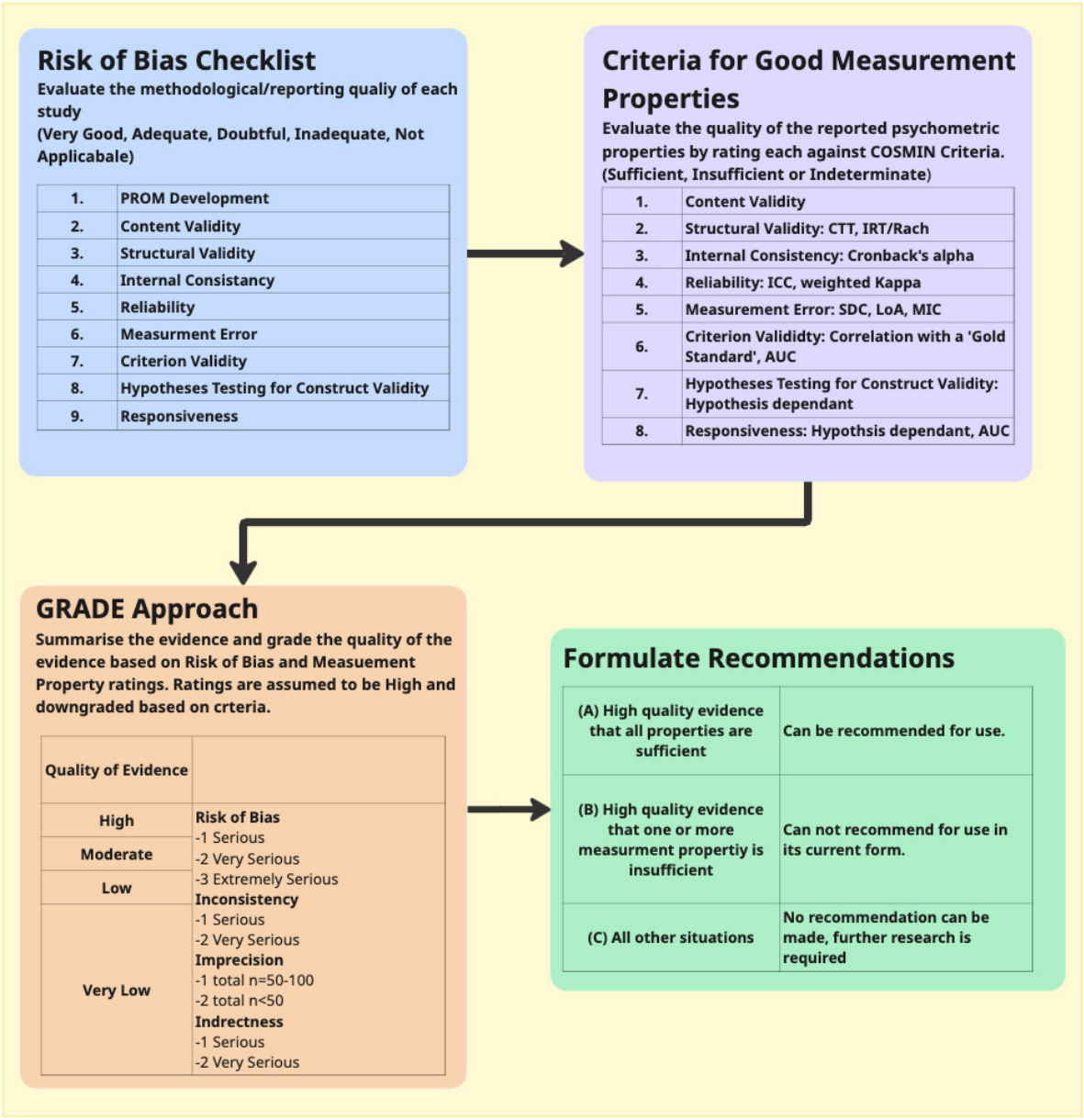

Fig 1.1: Overview of the COSMIN Systematic Review methodology, adapted from Wehr et al. (2024).

All papers were independently reviewed by a second rater for both methodologies with disagreements being resolved via consensus.

Inter-rater reliability was calculated via Cohen's Kappa for both the MMAT and COSMIN ratings prior to discussion of divergent ratings (MMAT K = 0.91, COSMIN K = 0.64).

3.7 Synthesis

Data were synthesised and presented in a narrative synthesis following the methodology outlined by the Economic and Social Research Council (ESRC) Methods Program (Popau et al 2006). An initial synthesis of the data was created via textual descriptions of the eligible studies and grouping by study type. Then, relationships between and within the studies were explored, after which the robustness of the synthesis was assessed.

4. Results

Figure 1.2 outlines the search process. The search identified 159 records, with 29 duplicates. Screening identified nine potentially eligible papers which were retrieved via University of Glasgow Institutional access. Forward and backward citation searches were conducted using Web of Science, identifying no additional papers. Full text review led to the exclusion of two records: Allan et al. (2023) did not use the full FoRSe and Eisner et al. (2019) did not report data on the FoRSe. Seven records proceeded to data extraction. A repeat of the search identified a further 11 records, with four duplicates. Title/abstract screening identified no additional eligible records.

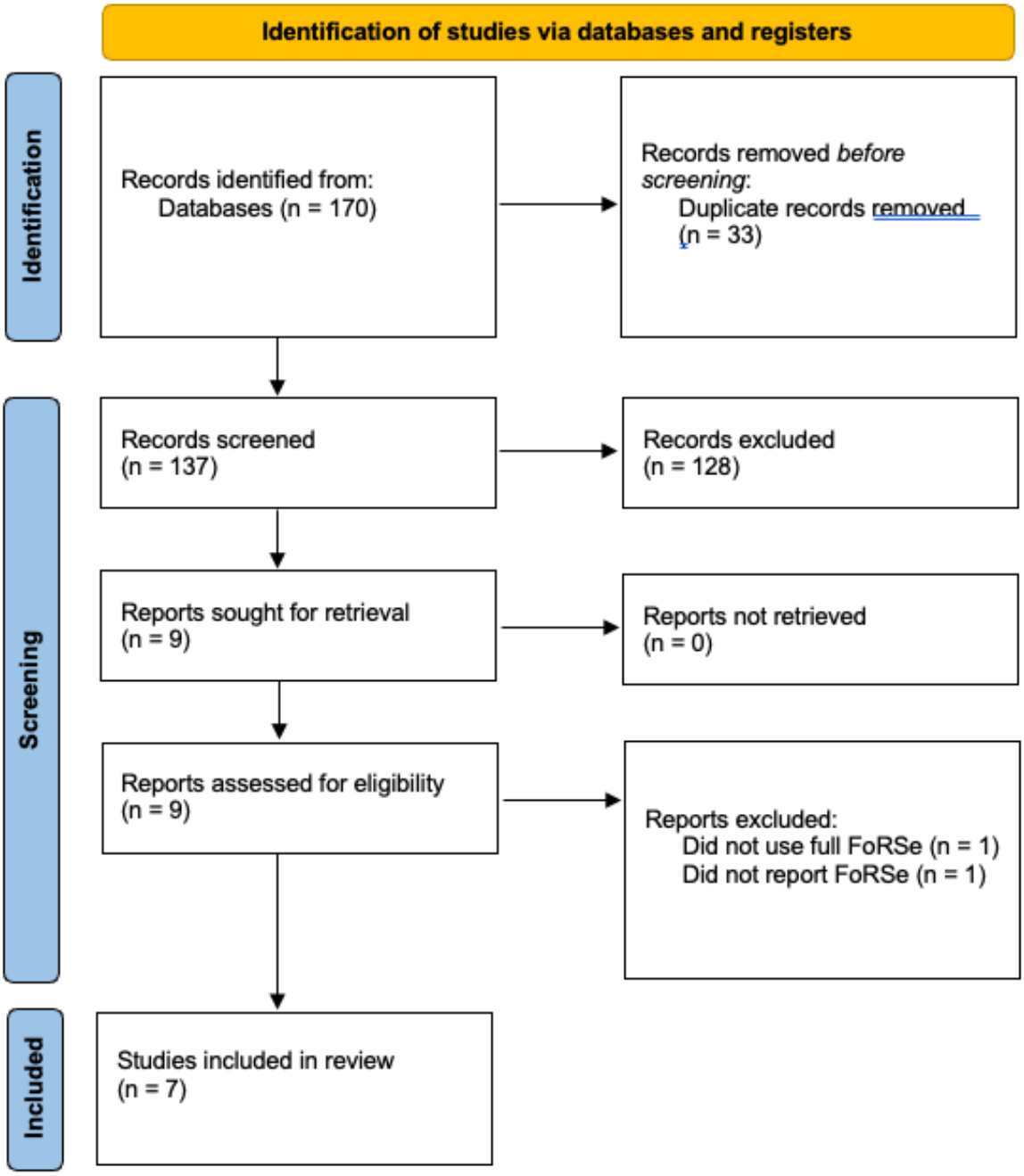

Figure 1.2: PRISMA 2020 flow diagram for new systematic reviews which included searches of databases and registers only

4.1 Study Characteristics

Study characteristics are reported in Table 1.1 (see Appendix 1.3 for additional details). Three cross-sectional studies, two randomised trials, one non-randomised pilot study and one cluster randomised feasibility study were identified. Only one study reported on the psychometric properties of the FoRSe.

The majority of studies were conducted in the United Kingdom, one conducted in the Republic of Ireland and one multi-site study conducted in both the United Kingdom and Australia. A total of 436 (184 female) participants who had experienced psychosis were included across all studies, with a mean age of 40.59 (SD=8.11). Ethnic diversity of the sample was reported as 129 White, four Asian, three Black, two mixed ethnicity. One study (Gumley et al., 2022) used two different ethnicity reporting structures that do not fit within the above reporting structure and two studies (Gumley et al., 2015, Ryan et al., 2021) did not report ethnicity. 364 participants had a diagnosis of a psychotic disorder. 33 participants self-identified as having a psychotic disorder for which treatment or a diagnosis had been received and 39 reported a diagnosis of a mental health difficulty with self-identified experience of psychosis.

Table 1.1 Study Characteristics

Study	Design	Population	Location	N (n = female)	Mean age (SD)*	Ethnicity
White and Gumley (2009)	Cross-sectional between groups	Diagnosis (DSM-IV) of Schizophrenia, with ongoing distress due to traumatic memories of psychosis	United Kingdom	27 (7)	38.87 (9.66)	All participants were White
Braehler et al (2013)	Randomized open-label, blinded end point evaluation design	ICD-10 Diagnosis of schizophrenia-spectrum disorder or biololar disorder with psychotic features	United Kingdom	40 (18)	41.76 (13.32)	All participants were White
Gumley et al. (2015)	Randomised controlled trial	ICD-10 Diagnosis of schizophrenia-spectrum disorder or related disorder, considered relapse prone	United Kingdom	169 (48)	41.48 (9.87)	Not reported
Jamalamadaka et al. (2020)	Cross-sectional questionnaire design	Diagnosis of a mental health difficulty and self-identifying experience of psychosis	United Kingdom	39 (29)	33.05 (11.30)	Asian = 2 Black = 1 White = 35 Other = 1
Ryan et al. (2021)	Pilot study	Diagnosis of a psychotic disorder	Republic of Ireland	55 (22)	36.04 (12.78)	Not Reported
Sired et al. (2021)	Cross-sectional between-groups questionnaire design	Self-report experience of Psychosis for which a diagnosis or treatment has been received and considering self to be in recovery	United Kingdom	33 (24)	34.44 (10.46)	Asian = 2 Black = 2 White = 27 Mixed background = 2
Gumley et al. (2022)	Cluster randomised controlled feasibility study	Diagnosis of Schizophrenia or related diagnosis and experienced a relapse in the last two years or received crisis input in the last two years	United Kingdom and Australia	73 (36)	42.37 (10.73)	UK Site Scottish = 37 Other British = 2 Other White = 1 Mixed = 2 Pakistani = 2 Indian = 1 African = 3 Unknown = 1 Australia Site Born in Australia = 19 Born Elsewhere = 5 Aboriginal = 1

4.2 Assessment of studies on psychometric properties of the FoRSe using the COSMIN Guidelines.

4.2.1 COMSIN Risk of Bias

One study on the Psychometric properties of the FoRSe was identified (Gumley et al., 2015). The methodological quality of this study as rated by the COSMIN Risk of Tool is described in Table 1.1.

Table 1.2: Cosmin Quality and Measurement Property Ratings

Measurement Property	Cosmin Risk of Bias				Criteria for Good Measurement			GRADE*
	Very Good	Adequate	Doubtful	Inadequate	Sufficient	Insufficient	Indeterminate	
Content Validity			X				X	
Structural Validity		X					X	
Internal Consistency	X						X	
Reliability			X		X			Low
Measurement Error				X			X	
Hypotheses Testing for Construct Validity – Convergent Validity			X				X	

Notes: *GRADE approach is not used when Criteria for Good Measurement is rated as Indeterminate

4.2.2 Content Validity

The methodological quality of the Content Validity study was rated as Doubtful due to the method of data collection used for PROM development, lack of clarity on the analysis approach used and lack of clarity on methodological aspects of the data collection.

4.2.3 Structural Validity

Methodological quality of the Structural Validity study was rated as Adequate. An Exploratory Factor Analysis was performed with a total participant number five times greater than the number of items in the FoRSe and greater than 100 participants.

4.2.4 Internal Consistency

Methodological quality of the Internal Consistency study was rated as Very Good with Cronbach's alpha having been calculated for each item.

4.2.5 Reliability

Methodological quality for the Reliability study was rated as Doubtful due to lack of clarity around identical test conditions between administrations and a lack of reporting of evidence that no systematic change between measurements has occurred.

4.2.6 Measurement Error

A relevant Measurement Error statistic was not provided by the study but was calculable from reliability statistics (see Appendix 1.4). The methodological quality for Measurement Error was rated as Doubtful. As Measurement Error is an aspect of Reliability, this was due to the issues identified under Reliability.

4.2.7 Hypotheses testing for Construct Validity

The methodological quality for Hypotheses Testing for Construct Validity was rated as Doubtful. This was due to neither the Early Signs Scale (ESS; Birchwood et al., 1989) nor the Personal Beliefs about Illness Questionnaire (PBIQ; Birchwood et al., 1993) having information available to rate them as having sufficient psychometric properties per COSMIN Guidelines.

4.3 Criteria for good Measurement Properties

The ratings of Criteria for Good Measurement Properties are described in Table 1.1.

4.3.1 Content Validity

A rating of Indeterminate was applied, due to not meeting the following criteria: evidence that least 85% of the items refer to the construct of interest and patients having been asked about the comprehensibility of the items and instructions.

4.3.2 Structural Validity

Structural Validity met the criteria for Indeterminate due to insufficient information having been reported. Factor loadings were within acceptable limits for each item (>0.30) and less than 10% of items loaded onto more than one factor. However, it was not possible to establish if the structure was in line with a theory about the construct and screen plot or kaiser criterion were measured but not reported.

4.3.3 Internal Consistency

Internal Consistency was rated as Indeterminate. Due to the Indeterminate rating of structural validity, it is there is insufficient evidence of unidimensionality which is a required assumption for calculating Cronbach's $\boldsymbol{\alpha}$.

4.3.4 Reliability

Reliability met the criteria a Sufficient rating, with $ICC_{agreement}$ and Pearsons ρ being calculated and ≥0.70.

4.3.5 Measurement Error

Measurement Error met the criteria for Indeterminate. This was due to a lack of a defined Minimum Important Change (MIC) statistic identified by the study.

4.3.6 Hypotheses testing for Construct Validity

Hypotheses testing for construct validity met the criteria for Indeterminate. The review team defined three hypotheses based on the COSMIN criteria (Table 1.2). One of these hypotheses was met while the other two hypotheses were not.

Table 1.3: Hypotheses for testing Construct Validity

Comparator Instrument	Hypothesis	Reported correlation	Hypothesis met?
Early Signs Scale	Related constructs but not identical (r=0.5-r=0.7)	0.71	Y
Personal Beliefs about Illness Questionnaire	Somewhat related constructs (r=0.3 - 0.5)	0.29	N
PANSS (Positive Symptom Scale)	Somewhat related constructs (r=0.3 - 0.5)	0.56	N

4.4 GRADE Approach

The COSMIN modified GRADE approach was applied (Table 1.1). COSMIN does not advise to apply the GRADE approach to properties that have been rated as Indeterminate, as such, only Reliability could be rated.

Reliability was graded as Low due to issues in the methodological quality of the study as rated on the Risk of Bias tool meaning that there is limited confidence in the accuracy of the results as the true measurement property could be substantially different from the estimate.

4.5 Non-psychometric studies

4.5.1 Quality Appraisal

Quality as rated by the MMAT (Table 1.3) was found to be consistent across all studies with no study meeting less than 4 of 5 criteria. The most common area of bias was in the reporting of complete outcome data, with data being missing without a reason being reported. Ryan et al. (2021) did not account for confounders within their data, however, this was a pilot study.

Table 1.4: Summary of Quality Ratings: MMAT

	Randomised Controlled Trails					Non-randomised controlled trials				
	Randomization	Comparable Groups at Baseline	Complete Outcome Data	Assessors Blinded	Adherence to Intervention	Representative Population	Appropriate Measures	Complete Outcome Data	Confounders Accounted For	Intervention Administered as intended
White and Gumley (2009)						Yes	Yes	Can't Tell	Yes	Yes
Braehler et al (2013)	Yes	Yes	No	Yes	Yes					
Gumley et al. (2015)	Yes	Yes	Can't Tell	Yes	Yes					
Jamalamadaka et al. (2020)						Yes	Yes	Can't Tell	Yes	Yes
Ryan et al. (2021)						Yes	Yes	Yes	No	Yes
Sired et al. (2021)						Yes	Yes	Yes	Yes	Yes
Gumley et al. (2022)	Yes	Yes	Yes	Yes	Yes					

4.5.2 Cross-sectional Studies

Three cross-sectional studies were included in the review.

White and Gumley (2009) conducted a cross-sectional case-control design study exploring the associations between Post Psychotic Post-Traumatic Stress Disorder (PP-PTSD), FoRSe and intolerance of uncertainty. It consisted of two groups: participants who met criteria for PP-PTSD (n=20) and a clinical control group (n=20). All participants had a diagnosis of a psychotic disorder. They found that, compared to the control group, individuals in the PP-PTSD group had significantly higher scores on the FoRSe Fear of Relapse subscale, Intrusiveness subscale and total score with large effect sizes (Cohen's D range = 1.3 – 1.6, all $p < 0.01$). Regression analysis, including the Beliefs About Paranoia Scale, FoRSe, and Intolerance of Uncertainty Scale, indicated that only FoRSe total score predicted caseness for PP-PTSD (OR 1.213, p=<0.05). A post hoc Receiver Operating Characteristic (ROC) analysis revealed that a cut-off score of >56 on the FoRSe demonstrated good sensitivity (80%) and specificity (82.4%) for PP-PTSD. CAPS-S re-experiencing scores were associated with Fear of Relapse ($r = 0.47$, p=<0.05) and Intrusiveness ($r = 0.54$, p<0.01) subscales as well as FoRSe total scores ($r = 0.54$, $p < 0.01$). CAPS-S Avoidance also had scores associated with FoRSe Fear of Relapse subscale ($r = 0.47$, $p < 0.05$), FoRSe Intrusiveness subscale ($r = 0.51$, $p < 0.05$), and FoRSe total score ($r = 0.51$, $p < 0.05$).

Jamalamadaka et al. (2020) conducted a cross-sectional between groups questionnaire-based study that examined whether people in recovery from psychosis have greater fear of illness recurrence than those recovering from common mental health problems or healthy controls. Three study populations were recruited, people in recovery from psychosis (n=39), people in recovery from other common mental health conditions (n=82) and healthy controls (n=61). Participants were recruited via NHS services and social media. The study found that those recovering from psychosis had significantly higher levels of FoRP than those recovering from common mental health problems or healthy controls (Cohen's D = −.36, p = <0.05). The hypothesized relationship between FoRSe, Mental Health Anxiety (MHA) and Mental Defeat was also found (f = 0.07). Both FoRSe and MHA were associated with higher levels of maladaptive coping behaviours (f = 0.07). The study used the FoRSe in both healthy controls

and people with common mental health conditions. The FoRSe has not been validated in either population and the study does not provide a rationale or justification for its use, nor evidence of validation. As such, the results of this study are methodologically questionable, making the results likely invalid.

Sired et al. (2021) conducted a cross-sectional between groups questionnaire-based study that aimed to explore the relationship between fear of relapse and negative appraisals of prodromal symptomology in people in recovery from psychosis (n=70), anxiety (n=70) and healthy controls (n=70). Both the anxiety and psychosis group were recruited based on self-report of diagnosis. They reported no significant differences between the psychosis group and anxiety group on FoRSe measures. FoRSe did not predict negative interpretations of psychosis symptoms in the psychosis group or anxiety symptoms in the anxiety group. The study used the FoRSe in both non-psychosis groups without providing a rationale or justification, nor evidence of validation. As such, the results of this study are methodologically questionable, making the results likely invalid.

In summary, FoRSe total score appeared to be both sensitive and specific to caseness for PP-PSTD, with a cut off score of >56. PTSD Re-experiencing and Avoidance symptoms are associated with the Intrusiveness, Fear of Relapse subscales and FoRSe Total. Studies have used the FoRSe to compare mental health anxiety in people with psychosis to people with anxiety or common mental health problems and found mixed results, however, due to the methodological issues raised by using scales in populations for which they were neither developed nor validated, these results cannot be relied upon.

4.5.3 Randomised Trials

Three randomised studies were included in the review.

Gumley et al. (2015) aimed to develop and validate the FoRSe and establish its sensitivity and specificity to relapse compared to an established measure of early warning signs of relapse, ESS in a randomised controlled trial of relapse detection in people with a diagnosis of schizophrenia or related disorders. They used a two arm randomised controlled design,

standard early signs monitoring (n=86) or FoRSe early signs monitoring (n=83). The psychometric aspects of this study are reviewed in Section 4.2. Cox proportional hazards regression was used to identify predictors of time to relapse after baseline assessment, using all subscales of the Positive and Negative Syndrome Scale, the ESS and the FoRSe. The results showed that only the Fear of Relapse subscale of the FoRSe was a significant predictor of time to relapse, with higher scores on this subscale associated with shorter time to relapse [Exp(b) = 1.20, 95% CI = 1.01–1.42, $p < .05$]. An increase of 5 points in the FoRSe showed 79% (95% CI = 52%-86%) sensitivity to relapse but 46% (95%CI = 32%-60%) specificity. There were no differences between the ESS and FoRSe in sensitivity or specificity to relapse. The all FoRSe subscales were found to be significantly correlated with the ESS and with the Positive and General scales of the PANSS. Only the Fear of Relapse scale was significantly correlated with the Negative scale of the PANSS. A significant correlation with the Calgary Depression Schizophrenia Scale (CDSS) was reported for all scales, other than for the Awareness subscale. Only the Intrusions and Fear of Relapse scales were significantly correlated with PBIQ subscales.

Braehler et al. (2013) aimed to assess the feasibility and acceptability of 16 group sessions of Compassion Focused Therapy (CFT) in people with a psychotic disorder. They used a randomised parallel group design with two groups, treatment as usual (TAU) plus CFT (n=22) and TAU alone (n=18). Total scores on FoRSe were associated with greater positive symptoms, general psychopathology, and negative illness beliefs. They found that the CFT group showed a significant reduction in FoRSe scores compared to TAU, which was maintained at follow-up. They also report statistically significant correlations between FoRSe scores and compassion scores in the CFT group, but not the TAU group. The increase in compassion in CFT group was associated with a decrease in Fear of Relapse subscale score (r=-0.74; p=0.002) and FoRSe Intrusiveness subscale score (r-0.58; p=0.022). The study concluded that the reduction in FoRSe scores in the CFT group may indicate a reduction in the perceived threat of relapse and an increase in the sense of control over intrusive thoughts and experiences.

Gumley et al. (2022) aimed to establish the feasibility of a randomized controlled trial of the effectiveness of the EMPOWER blended digital intervention versus TAU in preventing relapse

in people with a diagnosis of schizophrenia. The intervention was based on the Cognitive Interpersonal Model of relapse and consisted of daily digital monitoring of changes in wellbeing, tailored digital messages to promote self-management, telephone peer support and, in the event of possible early signs of relapse, clinical triage to local services. They used a cluster randomized design, with eight sites and a total of 86 participants, randomised into two groups, intervention (n=42) and treatment as usual (n=31). The FoRSe was used to assess if increased monitoring associated with the intervention would increase FoRP. However, it was found to be lower in the intervention group at 12-month follow up (mean difference -7.53, 95% CI -14.45 to -0.60, Cohen's D = -0.53).

In summary, the FoRSe has been shown to be sensitive to relapse but has poorer specificity with the Fears of Relapse subscale being predictive of time to relapse. Both Braehler et al. (2013) and Gumley et al. (2015) showed associations with the CDSS and PBIQ as well as positive and general symptoms on the PANSS. Increases in compassion scores, as measured by the Narrative Recovery Scale, are also associated with decreases in Fear of Relapse and Intrusiveness scores. Gumley et al. (2020) showed that the EMPOWER blended intervention resulted in decreased scores on the FoRSe at 12 months.

4.5.4 Non-randomised Studies

One non-randomised non-controlled trial was included in the review.

Ryan et al. (2021) aimed to evaluate the effectiveness of a group emotional regulation skills program in people with psychosis using an uncontrolled pilot design. Participants completed measures at four time points: baseline, pre-group, post-group and one month follow up. Fifty-five participants with a psychotic disorder diagnosis were recruited from inpatient and outpatient services based in a psychiatric hospital. The intervention consisted of emotional regulation skills derived from Dialectical Behavioural Therapy to reduce participants' FoRP. The intervention was delivered in a group format over eight weekly sessions. They reported an improvement in emotional regulation, mindfulness and recovery outcomes from pre-group to post-group and one month follow up, but no improvement in FoRP. However, only

total score was used as an outcome. There were no associations between PSYRATS-H or PSYRATS-D scores and the FoRSe.

5. Discussion

This review aimed to investigate the psychometric characteristics of the FoRSe and its associations with clinical outcomes in people with psychosis. Seven studies were identified, consisting of three randomised trials, one non-randomised trial and three cross-sectional studies. Only one study was designed as a psychometric validation of the FoRSe. Following the COSMIN guidelines, it was identified that there was neither high quality evidence that all measurement properties were sufficient, nor was there high quality evidence that one or more measurement properties were insufficient. Therefore, this review recommends that further research is required in order to draw conclusions about the psychometric properties of the scores produced by FoRSe.

Content validity is regarded by COSMIN to be the most important psychometric property, following the logic that a scale being reliable is of little use if it does not measure what it purports to. While the use of both inductive and deductive methods of item development is an improvement upon the common practice of exclusively deductive methods (Morgado et al., 2018), the data available did not provide evidence of a standard high enough to draw conclusions about the content validity of the FoRSe.

Content validity could have been improved by articulating a theoretical model for FoRP. In addition, a suitable content validity study with both professionals and people with expertise by experience would provide evidence about the comprehensiveness, comprehensibility and relevance of items and response options. This would provide evidence which could support its content validity or lead to changes to the scale.

In the case that no changes are indicated by a content validity study, it would be possible to conduct further validity work, using the current exploratory factor analytic data and performing confirmatory factor analysis (Streiner et al., 2024). The factor analytic work currently available is not sufficient for COSMIN in the absence of a confirmatory factor

analysis or a theoretical model to compare it to, and to develop a model for the purposes of supporting this data, following exploratory factor analysis, would be analogous to harking (Tackett et al., 2019). Should a confirmatory factor analysis be performed, it would then provide evidence to assess unidimensionality of the scales, which would then allow for internal consistency to be calculated. The current calculations of internal consistency cannot be used without sufficient evidence of unidimensionality (McNeish, 2018, Streiner, 2003).

Should content validity work lead to alterations to the scale, a new validation study would be required. A validation study would allow for establishing evidence for several properties at once. A new factor analytic study would need to be performed, and with a proposed theoretical model, could be used to provide evidence of unidimensionality which is required to calculate internal consistency. A confirmatory factor analysis could then be performed, in a separate group.

The reliability work conducted by Gumley et al. (2015) would also be required to be repeated in both cases. This is due to the lack of clarity about the similarity of test conditions across administrations. Differences in test conditions (i.e. location, order of administration, method of administration or instructions) can introduce variance into reliability calculations. As such, it is good practice to sufficiently describe and standardise the test conditions (Kline, 2000).

Measurement Error was unable to be rated due to there being no MIC established for the FoRSe. There are multiple ways to establish MIC, but there is no consensus on which method is most appropriate (De Vet et al., 2011, Streiner et al., 2024). Anchor based methods rely on comparing scores with an external measure, such as another measurement tool or impression of change ratings from clinicians or patients in longitudinal studies. Distribution based methods use statistical parameters such as standard deviations or standard error of measurement, without being able to define meaningful change in a given construct.

While Gumley et al. (2015) assess convergent validity via several measures, they did not provide any a priori hypotheses around the relations between subscales and the measures; evidence regarding the validity of the comparator measures; nor the expected strengths of

the hypothesized relationships. A future study should specify the rationale for the relationships between measures, including the relationships between subscales, and the expected strength of the associations.

This review is also unable to comment on the responsiveness of the FoRSe, as sensitivity and specificity are not appropriate analyses for non-dichotomous scales. While this may provide some evidence of clinical utility, a responsiveness study for scales with continuous scores would require a longitudinal study where specific a priori hypotheses can be tested, such as expected differences in mean change scores between the FoRSe and another instrument with established responsiveness that measures the same construct (i.e. a gold standard); differences in change scores after an intervention between groups; or effect size after an intervention on the construct of interest where the efficacy of the intervention is known (De Vet et al., 2011).

Designing and validating a psychometric scale is a complex process, made more difficult by disagreement around definitions of terms, acceptable methodologies, statistical approaches or even what counts as "sufficient" when considering a given metric (Boateng et al., 2018, Mokkink et al., 2024, Kottner et al., 2011, Streiner and Kottner, 2014). Theoretical approaches are also contested (McNeish, 2024, Mislevy, 2024, Sijtsma et al., 2024a, Sijtsma et al., 2024b). As such, attempts at developing standards, such as COSMIN, have not been universally accepted and continue to be an active process of refinement based on consensus (Mokkink et al., 2021, Mokkink et al., 2024). Similarly, neither the development of a scale, nor establishing evidence for its psychometric properties are singular processes, instead requiring iteration and refinement via a multiphase approach to accumulating evidence. This process is arguably one that lacks an end point, as psychometric properties are not inherent to the scores a scale produces but rather an interplay between the scale, the context in which it is used, and the population in which it is used (Streiner and Kottner, 2014). Both context and population vary as time progresses, particularly in mental health, where conceptualisations of psychological phenomena change and evolve, influenced by societal/cultural changes and research. A scale designed in the 1970s using criteria from the DSM-II will almost certainly

lack content validity in 2025, even if the developers had followed modern methodological standards of psychometric development and validation.

While no recommendation can be made about the use of the FoRSe due to the need for further evidence of its psychometric properties, this review does highlight evidence to suggest that conducting this work is justified. The Fears of Relapse subscale was found to be predictive of time to relapse, suggesting that it may have potential in relapse monitoring. It has been shown that scores on the FoRSe were associated with caseness for PP-PTSD with good sensitivity and specificity, as well as higher levels of traumatic re-experiencing symptoms being associated with increased fear of relapse. This suggests that FoRP may contribute to the persistence of feelings of current threat that characterise PP-PTSD where people report persistent fears and worries related to their experience of psychosis or its treatment. In a qualitative study of people's post-traumatic reactions to psychosis (Lu et al., 2017) participants described re-experiencing symptoms including memories of frightening hallucinations, suicidal behaviours, harm towards others, paranoia, and coercive treatment. Allen et al. (2023) also reported in a longitudinal study that fear of relapse was persistent over time and predictive of negative affect and anxiety, supporting the conceptualisation of FoRP as a persistent experience for people with schizophrenia and a potential treatment target.

The FoRSe has shown potential utility as an outcome measure, as highlighted by its use in Braehler et al. (2013), Gumley et al. (2020) and Ryan et al. (2021). Findings from Braehler et al. (2013) showed that increased compassion in CFT was associated with a decrease in FoRSe Total score suggesting that FoRP may be a modifiable treatment target that can respond to strategies designed to increase feelings of self-compassion. In addition, Gumley et al. (2022) found the use of a blended digital monitoring intervention for possible relapse led to a reduction in fears of relapse at 12 months, rather than a predicted increase. There was also a reduction in relapse rate in those randomised to the intervention compared to usual care (Hazard Ratio = 0.32, 95%CI 0.14, 0.74). This opens up the possibility that fear of relapse may be improved via targeted interventions and reduce relapse rates.

The quality of the studies within this review were rated highly on the MMAT, with no study being rated below 80%. The most common issue identified by the MMAT was missing outcome data in Braehler et al. (2013), Gumley et al. (2015), Jamalamadaka et al. (2020) and White and Gumley (2009). Additionally two studies (Jamalamadaka et al., 2020; Sired et al., 2021) used the FoRSe within populations for which the FoRSe was not developed. As psychometric properties are context and population dependant, and neither study provided or cited evidence of validation to support use of the FoRSe in their study populations, it is not possible to rely on this data.

There is evidence suggest that FoRP is a broader concept than initially conceived by Gumley and Schwannauer (2006). The FoRSe was developed based on the conceptualisation of FoRP as being part of the relapse process rather than a broader clinical feature experienced by people with schizophrenia. Several studies highlight fear of loss of autonomous decision-making as being central to FoRP, with losses concerning both fears of involuntary treatment and fears of not being able to trust one's own thoughts and its impact upon decision-making (Sandhu et al., 2013, Baker, 1995, Zukowska et al., 2022). While memories of hospital or distressing thoughts are incorporated into the FoRSe, items do not cover the traumatic nature of relapse and its potential for re-traumatisation upon its recurrence (Zukowska et al., 2022). Neither does it acknowledge the social and interpersonal nature of some of this trauma (Brookmann, 2020) nor fears of loss of social and occupational functioning (Baker, 1995). It does not fully capture the emotional experience of FoRP, which has been associated with distress, feelings of loss and entrapment (Brookmann, 2020). There is also a lack of consideration for interpersonal appraisals and beliefs about help-seeking (Zukowska et al., 2022, Brookmann, 2020) and how experiences of treatment may interact with these. Uncertainty is a recurrent theme within the research around FoRP (Sandhu et al., 2013, Baker, 1995, Baier, 1995) and has parallels in physical health illness recurrence, particularly Fear of Cancer Recurrence (Lebel et al., 2020, Simard et al., 2013). Worry and uncertainty have also been implicated in the maintenance of delusional thinking (Freeman and Garety, 2014, Freeman, 2016), and shown to be treatable via targeted psychological intervention (Freeman et al., 2016). This raises the potential for FoRP to be involved in positive symptomology, as well as being amenable to intervention through its maintenance factors.

These issues in content validity stem from the development of the FoRSe. The Fear of Recurrence Scale was originally designed as an alternative to early signs monitoring to enable greater precision in targeting a potential psychological mechanism underpinning relapse (Tait et al., 2002). As such, it is focused towards the cognitive content of early signs of relapse rather than it being developed as a measure of a distinct clinical construct.

It is important that measurement tools in health sciences have sufficient evidence to support their use. There is a large body of evidence that the standard of psychometric evidence within empirical studies is low, with the majority of measures across psychological sciences providing little evidence of validity (Evers et al., 2010, Qualls and Moss, 1996). Higgins et al. (2024) highlight the lack of psychometric rigour within clinical psychology by examining studies using or adapting the Reading the Mind in the Eyes Test (RMET) (Baron-Cohen et al., 2001). They found that 64% of studies either did not provide new validity evidence or cite evidence of previous validity work. Additionally, where modifications had been made to the RMET or where the test was used in a sample for which there had been no validity work cited, only 23% of studies provided validity evidence. Where evidence was provided, it was found to be weak other than for inter-rater reliability. Hubley et al. (2014) found that across 38 measurement tools published between 2010 and 2012, none reported evidence of content validity in the development process, with only one study reporting data from response processes such as cognitive interviewing. The widespread use of measurement tools without sufficient validation has implications for psychological science as a whole, as measurement is upstream of statistical analysis. In short, if the data cannot be relied upon, neither can the analyses. Psychology has become increasingly aware of issues in replication of study findings (Nosek et al., 2022, Tackett et al., 2019). While there is considerable focus on Questionable Research Practices such as p-Hacking and Harking, psychometricians have become concerned that poor psychometric practices (or Questionable Measurement Practices) are involved and that without addressing them, replication efforts are unlikely to be successful (Flake and Fried, 2020). Flake et al. (2022) note that fewer then 10% of replication studies provide evidence of validity for the measures they use. Without this evidence, it is not possible to differentiate if replication has failed (or succeeded) due to measurement error or theoretical implications.

This review highlights that further evidence is required before a recommendation can be made in regards to the use of the FoRSe in its current form. In light of the evidence from qualitative work that it may not account for important clinical aspects of FoRP, it is recommended that the scale is revised to improve its content validity before proceeding with further psychometric studies.

5.1 Strengths and Limitations

The strengths of this systematic review include the use of both the MMAT and COSMIN Guidelines, allowing for the assessment of quality in both psychometric and non-psychometric studies. While the MMAT allows for the review of multiple studies, the depth to which it assesses study quality is limited by this, future work may benefit from alternative approaches such as applying appraisal tools by study type (i.e. The Joanna Briggs Insitute (JBI) Critical Appraisal Tools or Critical Apprasial Skills Program (CASP) Checklists) for a more detailed examination of risk of bais. The study used a calibrated search strategy, developed with target papers from the scale's author. Papers were also screened by a second reviewer, who was independent from the review. Despite this, it is possible that records could have been missed. The study attempted to minimise this through forward/backward searching conducted on all identified studies, including the scale and its validation study. A further limitation is that only one study could contribute to the COSMIN ratings, highlighting the need for further work in establishing its psychometric properties. Professor Gumley is the scale's author and both a member of the research team and the lead author's supervisor, creating potential for bias. The study aimed to manage this via Professor Gumley providing only supervisory and editoral input and was not involved in any aspect of screening, study seletion, reviewing or rating.

6. Conclusions

FoRP is an important clinical construct. It is associated with PP-PTSD, relapse and significant distress for people who have experienced an episode of psychosis. The FoRSe is the only tool currently available to measure this construct in psychosis. Its psychometric properties have

been examined in one study but this evidence does not reach the criteria where a conclusion about its use can be drawn without further research, according to the COSMIN Guidelines.

In particular, there is no available evidence regarding content validity outside of its development. Additionally, there is evidence from qualitative work that it does not capture some aspects of FoRP and, as such, this review suggests that a revised version of the scale should be developed and its content validity is examined before proceeding to further psychometric validation.

7. Funding details and Conflicts of Interest

This study received no external funding sources. Andrew Gumley supervised the project and is the author of the FoRSe, as such he was not involved in any data extraction, rating or screening.

8. References

Allan, S., O'Driscoll, C., Mcleod, H. J., Gleeson, J., Farhall, J., Morton, E., Bell, I., Bradstreet, S., Machin, M. & 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***,** 1-15.

American Psychiatric Association 2013. *Diagnostic and statistical manual of mental disorders: DSM-5,* Arlington, Va;London;, American Psychiatric Association.

Andreasen, N. C., Carpenter, W. T., 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**,** 441-449.

Ascher-Svanum, H., Zhu, B., Faries, D. E., Salkever, D., Slade, E. P., Peng, X. & Conley, R. R. 2010. The cost of relapse and the predictors of relapse in the treatment of schizophrenia. *BMC psychiatry,* 10**,** 1-7.

Baier, M. 1995. Uncertainty of illness for persons with schizophrenia. *Issues Ment Health Nurs,* 16**,** 201-12.

Baker, C. 1995. The development of the self-care ability to detect early signs of relapse among individuals who have schizophrenia. *Arch Psychiatr Nurs,* 9**,** 261-8.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. & Plumb, I. 2001. The "Reading the Mind in the Eyes" Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *Journal of Child Psychology and Psychiatry,* 42**,** 241-251.

Bebbington, P. E., Bhugra, D., Brugha, T., Singleton, N., Farrell, M., Jenkins, R., Lewis, G. & Meltzer, H. 2004. Psychosis, victimisation and childhood disadvantage: Evidence from the second British National Survey of Psychiatric Morbidity. *The British Journal of Psychiatry,* 185**,** 220-226.

Benrimoh, D., Dlugunovych, V., Wright, A. C., Phalen, P., Funaro, M. C., Ferrara, M., Powers, A. R., Woods, S. W., Guloksuz, S., Yung, A. R., Srihari, V. & Shah, J. 2024. On the proportion of patients who experience a prodrome prior to psychosis onset: A systematic review and meta-analysis. *Molecular Psychiatry,* 29**,** 1361-1381.

Birchwood, M., Iqbal, Z., Chadwick, P. & Trower, P. 2000. Cognitive approach to depression and suicidal thinking in psychosis: I. Ontogeny of post-psychotic depression. *The British Journal of Psychiatry,* 177**,** 516-521.

Birchwood, M., Mason, R., Macmillan, F. & Healy, J. 1993. Depression, demoralization and control over psychotic illness: a comparison of depressed and non-depressed patients with a chronic psychosis. Psychol Med, 23, 387-95.

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**,** 649-656.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R. & Young, S. L. 2018. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health,* 6**,** 149.

Braehler, C., Gumley, A., Harper, J., Wallace, S., Norrie, J. & Gilbert, P. 2013. Exploring change processes in compassion focused therapy in psychosis: Results of a feasibility randomized controlled trial. *British Journal of Clinical Psychology,* 52**,** 199-214.

Brookmann, L. 2020. *A meta-synthesis of service user perspectives of shared decision-making and a qualitative investigation of lived experiences of fear of recurrence in relation to psychosis.* D Clin Psy, University of Glasgow.

Buswell, G., Haime, Z., Lloyd-Evans, B. & Billings, J. 2021. A systematic review of PTSD to the experience of psychosis: prevalence and associated factors. *BMC Psychiatry,* 21**,** 9.

Cassidy, S. A., Bradley, L., Bowen, E., Wigham, S. & Rodgers, J. 2018. Measurement properties of tools used to assess suicidality in autistic and general population adults: A systematic review. *Clinical psychology review,* 62**,** 56-70.

Charlson, F. J., Ferrari, A. J., Santomauro, D. F., Diminic, S., Stockings, E., Scott, J. G., Mcgrath, J. J. & Whiteford, H. A. 2018. Global epidemiology and burden of schizophrenia: findings from the global burden of disease study 2016. *Schizophrenia bulletin,* 44**,** 1195-1203.

Correll, C. U., Bitter, I., Hoti, F., Mehtälä, J., Wooller, A., Pungor, K. & Tiihonen, J. 2022. Factors and their weight in reducing life expectancy in schizophrenia. *Schizophrenia Research,* 250**,** 67-75.

Crellin, D. J., Harrison, D., Santamaria, N. & Babl, F. E. 2015. Systematic review of the Face, Legs, Activity, Cry and Consolability scale for assessing pain in infants and children: is it reliable, valid, and feasible for use? *Pain,* 156**,** 2132-2151.

De Vet, H. C. W., Terwee, C. B., Mokkink, L. B. & Knol, D. L. 2011. *Measurement in Medicine: A Practical Guide*, Cambridge University Press.

Emsley, R., Chiliza, B. & Asmal, L. 2013. The evidence for illness progression after relapse in schizophrenia. *Schizophrenia research,* 148**,** 117-121.

Evers, A., Sijtsma, K., Lucassen, W. & Meijer, R. R. 2010. The Dutch review process for evaluating the quality of psychological tests: History, procedure, and results. *International Journal of Testing,* 10**,** 295-317.

Flake, J. K., Davidson, I. J., Wong, O. & Pek, J. 2022. Construct validity and the validity of replication studies: A systematic review. *American Psychologist,* 77**,** 576.

Flake, J. K. & Fried, E. I. 2020. Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in methods and practices in psychological science,* 3**,** 456-465.

Freeman, D. 2016. Persecutory delusions: a cognitive perspective on understanding and treatment. *The Lancet Psychiatry,* 3**,** 685-692.

Freeman, D., Bradley, J., Waite, F., Sheaves, B., Deweever, N., Bourke, E., Mcinerney, J., Evans, N., Černis, E., Lister, R., Garety, P. & Dunn, G. 2016. Targeting Recovery in Persistent Persecutory Delusions: A Proof of Principle Study of a New Translational Psychological Treatment (the Feeling Safe Programme). *Behavioural and Cognitive Psychotherapy,* 44**,** 539-552.

Freeman, D. & Garety, P. 2014. Advances in understanding and treating persecutory delusions: a review. *Soc Psychiatry Psychiatr Epidemiol,* 49**,** 1179-89.

Ghai, V., Subramanian, V., Jan, H. & Doumouchtsis, S. K. 2022. A systematic review highlighting poor quality of evidence for content validity of quality of life instruments in female chronic pelvic pain. *Journal of Clinical Epidemiology,* 149**,** 1-11.

Gleeson, J. F., Mcguckian, T. B., Fernandez, D. K., Fraser, M. I., Pepe, A., Taskis, R., Alvarez-Jimenez, M., Farhall, J. F. & Gumley, A. 2024. Systematic review of early warning signs of relapse and behavioural antecedents of symptom worsening in people living with schizophrenia spectrum disorders. *Clinical Psychology Review,* 107**,** 102357.

Gumley, O'grady, M., Mcnay, L., Reilly, J., Power, K. & Norrie, J. 2003. Early intervention for relapse in schizophrenia: results of a 12-month randomized controlled trial of cognitive behavioural therapy. *Psychological Medicine,* 33**,** 419-431.

Gumley, A., Bradstreet, S., Ainsworth, J., Allan, S., Alvarez-Jimenez, M., Beattie, L., Bell, I., Birchwood, M., Briggs, A., Bucci, S., Castagnini, E., Clark, A., Cotton, S. M., Engel, L., French, P., Lederman, R., Lewis, S., Machin, M., Maclennan, G., Matrunola, C., Mcleod, H., Mcmeekin, N., Mihalopoulos, C., Morton, E., Norrie, J., Reilly, F., Schwannauer, M., Singh, S. P., Smith, L., Sundram, S., Thomson, D., Thompson, A., Whitehill, H., Wilson-Kay, A., Williams, C., Yung, A., Farhall, J. & Gleeson, J. 2020. Early Signs Monitoring to Prevent Relapse in Psychosis and Promote Well-Being, Engagement, and Recovery: Protocol for a Feasibility Cluster Randomized Controlled Trial Harnessing Mobile Phone Technology Blended With Peer Support. *JMIR Res Protoc,* 9**,** e15058.

Gumley, A., Karatzias, A., Power, K., Reilly, J., Mcnay, L. & O'grady, M. 2006. Early intervention for relapse in schizophrenia: Impact of cognitive behavioural therapy on negative beliefs about psychosis and self-esteem. *British Journal of Clinical Psychology,* 45**,** 247-260.

Gumley, A. I., Bradstreet, S., Ainsworth, J., Allan, S., Alvarez-Jimenez, M., Aucott, L., Birchwood, M., Briggs, A., Bucci, S., Lewis, S., Machin, M., Maclennan, G., Mcleod, H., Schwannauer, M., Singh, S. P., Sundram, S., Thompson, A., Williams, C., Yung, A. R., Farhall, J. & Gleeson, J. 2022. The EMPOWER blended digital intervention for relapse prevention in schizophrenia: a feasibility cluster randomised controlled trial in Scotland and Australia. *The Lancet Psychiatry,* 9**,** 477-486.

Gumley, A. I., Macbeth, A., Reilly, J. D., O'grady, M., White, R. G., Mcleod, H., Schwannauer, M. & Power, K. G. 2015. Fear of recurrence: Results of a randomized trial of relapse detection in schizophrenia. *British Journal of Clinical Psychology,* 54**,** 49-62.

Gumley, A. I., Taylor, H. E. F., Schwannauer, M. & Macbeth, A. 2014. A systematic review of attachment and psychosis: measurement, construct validity and outcomes. *Acta Psychiatrica Scandinavica,* 129**,** 257-274.

Harvey, P. D., Strassnig, M. T. & Silberstein, J. 2019. Prediction of disability in schizophrenia: Symptoms, cognition, and self-assessment. *Journal of Experimental Psychopathology,* 10**,** 2043808719865693.

Hazell, C. M., Berry, C., Bogen-Johnston, L. & Banerjee, M. 2022. Creating a hierarchy of mental health stigma: testing the effect of psychiatric diagnosis on stigma. *BJPsych Open,* 8**,** e174.

Higgins, W. C., Kaplan, D. M., Deschrijver, E. & Ross, R. M. 2024. Construct validity evidence reporting practices for the Reading the Mind in the Eyes Test: A systematic scoping review. *Clinical Psychology Review,* 108**,** 102378.

Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F. & Nicolau, B. 2018. Mixed methods appraisal tool (MMAT), version 2018. *Registration of copyright,* 1148552.

Hubley, A. M., Zhu, S. M., Sasaki, A. & Gadermann, A. M. 2014. Synthesis of Validation Practices in Two Assessment Journals: Psychological Assessment and the European Journal of Psychological Assessment. *In:* ZUMBO, B. D. & CHAN, E. K. H. (eds.) *Validity and Validation in Social, Behavioral, and Health Sciences.* Cham: Springer International Publishing.

Jäger, M., Riedel, M., Messer, T., Laux, G., Pfeiffer, H., Naber, D., Schmidt, L. G., Gaebel, W., Huff, W. & Heuser, I. 2007. Psychopathological characteristics and treatment response of first episode compared with multiple episode schizophrenic disorders. *European Archives of Psychiatry and Clinical Neuroscience,* 257**,** 47-53.

Jamalamadaka, T., Griffith, E., Steer, H. & Salkovskis, P. 2020. Fear of illness recurrence and mental health anxiety in people recovering from psychosis and common mental health problems. *British Journal of Clinical Psychology,* 59**,** 403-423.

Johansen, K. K., Hounsgaard, L., Frandsen, T. F., Fluttert, F. A. & Hansen, J. P. 2020. Relapse prevention in ambulant mental health care tailored to patients with schizophrenia or bipolar disorder. *Journal of Psychiatric and Mental Health Nursing*.

Kirkpatrick, B., Fenton, W. S., Carpenter, W. T., Jr. & Marder, S. R. 2006. The NIMH-MATRICS consensus statement on negative symptoms. *Schizophr Bull,* 32**,** 214-9.

Kline, P. 2000. *The handbook of psychological testing,* London, Routledge.

Kottner, J., Audige, L., Brorson, S., Donner, A., Gajewski, B. J., Hróbjartsson, A., Roberts, C., Shoukri, M. & Streiner, D. L. 2011. Guidelines for Reporting Reliability and

Agreement Studies (GRRAS) were proposed. *International Journal of Nursing Studies,* 48**,** 661-671.

Kotzeva, A., Mittal, D., Desai, S., Judge, D. & Samanta, K. 2023. Socioeconomic burden of schizophrenia: a targeted literature review of types of costs and associated drivers across 10 countries. *Journal of medical economics,* 26**,** 70-83.

Lally, J., Ajnakina, O., Stubbs, B., Cullinane, M., Murphy, K. C., Gaughran, F. & Murray, R. M. 2017. Remission and recovery from first-episode psychosis in adults: systematic review and meta-analysis of long-term outcome studies. *The British Journal of Psychiatry,* 211**,** 350-358.

Lauber, C. 2008. Stigma and discrimination against people with mental illness: a critical appraisal. *Epidemiology and Psychiatric Sciences,* 17**,** 10-13.

Lebel, S., Mutsaers, B., Tomei, C., Leclair, C. S., Jones, G., Petricone-Westwood, D., Rutkowski, N., Ta, V., Trudel, G., Laflamme, S. Z., Lavigne, A.-A. & Dinkel, A. 2020. Health anxiety and illness-related fears across diverse chronic illnesses: A systematic review on conceptualization, measurement, prevalence, course, and correlates. *PLOS ONE,* 15**,** e0234124.

Lu, W., Mueser, K. T., Rosenberg, S. D., Yanos, P. T. & Mahmoud, N. 2017. Posttraumatic Reactions to Psychosis: A Qualitative Analysis. *Frontiers in Psychiatry,* 8.

Marshall, M., Lockwood, A., Bradley, C., Adams, C., Joy, C. & Fenton, M. 2000. Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. *The British Journal of Psychiatry,* 176**,** 249-252.

Mcclimans, L., Browne, J. & Cano, S. 2017. Clinical outcome measurement: Models, theory, psychometrics and practice. *Studies in History and Philosophy of Science Part A,* 65**,** 67-73.

Mcneish, D. 2018. Thanks coefficient alpha, we'll take it from here. *Psychological Methods,* 23**,** 412-433.

Mcneish, D. 2024. Practical Implications of Sum Scores Being Psychometrics' Greatest Accomplishment. *Psychometrika,* 89**,** 1148-1169.

Meyer, H., Taiminen, T., Vuori, T., Äijälä, A. & Helenius, H. 1999. Posttraumatic Stress Disorder Symptoms Related to Psychosis and Acute Involuntary Hospitalization in Schizophrenic and Delusional Patients. *The Journal of Nervous and Mental Disease,* 187**,** 343.

Mislevy, R. J. 2024. Are Sum Scores a Great Accomplishment of Psychometrics or Intuitive Test Theory? *Psychometrika,* 89**,** 1170-1174.

Mokkink, L. B., Elsman, E. B. M. & Terwee, C. B. 2024. COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. *Quality of Life Research,* 33**,** 2929-2939.

Mokkink, L. B., Prinsen, C., Patrick, D. L., Alonso, J., Bouter, L., De Vet, H., Terwee, C. B. & Mokkink, L. 2018. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). *User manual,* 78**,** 6-3.

Mokkink, L. B., Terwee, C. B., De Vet, H. C. W., Bouter, L. M., Alonso, J. & Patrick, D. 2021. Reply to the concerns raised by McKenna and Heaney about COSMIN. *Journal of Medical Economics,* 24**,** 857-859.

Morgado, F. F. R., Meireles, J. F. F., Neves, C. M., Amaral, A. C. S. & Ferreira, M. E. C. 2018. Scale development: ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica,* 30**,** 3.

Nieman, D. H. & Mcgorry, P. D. 2015. Detection and treatment of at-risk mental state for developing a first psychosis: making up the balance. *The Lancet Psychiatry,* 2**,** 825-834.

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Kline Struhl, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D. & Vazire, S. 2022. Replicability, Robustness, and Reproducibility in Psychological Science. *Annual Review of Psychology,* 73**,** 719-748.

Owen, M. J., Legge, S. E., Rees, E., Walters, J. T. R. & O'donovan, M. C. 2023. Genomic findings in schizophrenia and their implications. *Molecular Psychiatry,* 28**,** 3638-3647.

Page, M. J., Mckenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A. & Brennan, S. E. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International journal of surgery,* 88**,** 105906.

Pennington, M. & Mccrone, P. 2017. The cost of relapse in schizophrenia. *Pharmacoeconomics,* 35**,** 921-936.

Plahouras, J. E., Mehta, S., Buchman, D. Z., Foussias, G., Daskalakis, Z. J. & Blumberger, D. M. 2020. Experiences with legally mandated treatment in patients with schizophrenia: A systematic review of qualitative studies. *European Psychiatry,* 63**,** e39.

Qualls, A. L. & Moss, A. D. 1996. The Degree of Congruence between Test Standards and Test Documentation with in Journal Publications. *Educational and Psychological Measurement,* 56**,** 209-214.

Robinson, D., Woerner, M. G., Alvir, J. M. J., Bilder, R., Goldman, R., Geisler, S., Koreen, A., Sheitman, B., Chakos, M. & Mayerhoff, D. 1999. Predictors of relapse following response from a first episode of schizophrenia or schizoaffective disorder. *Archives of general psychiatry,* 56**,** 241-247.

Rooke, O. & Birchwood, M. 1998. Loss, humiliation and entrapment as appraisals of schizophrenic illness: A prospective study of depressed and non-depressed patients. *British Journal of Clinical Psychology,* 37**,** 259-268.

Ryan, A., Crehan, E., Khondoker, M., Fell, M., Curtin, R. & Johns, L. C. 2021. An emotional regulation approach to psychosis recovery: The Living Through Psychosis group programme. *Journal of Behavior Therapy and Experimental Psychiatry Vol 72 2021, ArtID 101651,* 72.

Sandhu, A., Ives, J., Birchwood, M. & Upthegrove, R. 2013. The subjective experience and phenomenology of depression following first episode psychosis: A qualitative study using photo-elicitation. *Journal of Affective Disorders,* 149**,** 166-174.

Shaw, K., Mcfarlane, A. & Bookless, C. 1997. The phenomenology of traumatic reactions to psychotic illness. *The Journal of nervous and mental disease,* 185**,** 434-441.

Sijtsma, K., Ellis, J. L. & Borsboom, D. 2024a. Recognize the Value of the Sum Score, Psychometrics' Greatest Accomplishment. *Psychometrika,* 89**,** 84-117.

Sijtsma, K., Ellis, J. L. & Borsboom, D. 2024b. Rejoinder to McNeish and Mislevy: What Does Psychological Measurement Require? *Psychometrika,* 89**,** 1175-1185.

Simard, S., Savard, J. & Ivers, H. 2010. Fear of cancer recurrence: specific profiles and nature of intrusive thoughts. *J Cancer Surviv,* 4**,** 361-71.

Simeone, J. C., Ward, A. J., Rotella, P., Collins, J. & Windisch, R. 2015. An evaluation of variation in published estimates of schizophrenia prevalence from 1990–2013: a systematic literature review. *BMC Psychiatry,* 15**,** 193.

Sired, R., Griffith, E., Jamalamadaka, T. & Salkovskis, P. 2021. Negative interpretations of ambiguous 'psychosis-like' and 'anxiety-like' experiences in recovery from psychosis or anxiety. *British Journal of Clinical Psychology,* 60**,** 463-485.

Solmi, M., Seitidis, G., Mavridis, D., Correll, C. U., Dragioti, E., Guimond, S., Tuominen, L., Dargél, A., Carvalho, A. F., Fornaro, M., Maes, M., Monaco, F., Song, M., Il Shin, J. & Cortese, S. 2023. Incidence, prevalence, and global burden of schizophrenia - data, with critical appraisal, from the Global Burden of Disease (GBD) 2019. *Molecular Psychiatry,* 28**,** 5319-5327.

Streiner, D. L. 2003. Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. *Journal of Personality Assessment,* 80**,** 99-103.

Streiner, D. L. & Kottner, J. 2014. Recommendations for reporting the results of studies of instrument and scale development and testing. *Journal of Advanced Nursing,* 70**,** 1970-1979.

Streiner, D. L., Norman, G. R. & Cairney, J. 2024. *Health measurement scales: a practical guide to their development and use,* Oxford, Oxford University Press.

Tackett, J. L., Brandes, C. M., King, K. M. & Markon, K. E. 2019. Psychology's Replication Crisis and Clinical Psychological Science.

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**,** 141-153.

Tandon, R., Nasrallah, H., Akbarian, S., Carpenter, W. T., Delisi, L. E., Gaebel, W., Green, M. F., Gur, R. E., Heckers, S., Kane, J. M., Malaspina, D., Meyer-Lindenberg, A., Murray, R., Owen, M., Smoller, J. W., Yassin, W. & Keshavan, M. 2024. The schizophrenia syndrome, circa 2024: What we know and how that informs its nature. *Schizophrenia Research,* 264**,** 1-28.

Taylor, P. J., Hutton, P. & Wood, L. 2015. Are people at risk of psychosis also at risk of suicide and self-harm? A systematic review and meta-analysis. *Psychol Med,* 45**,** 911-26.

Van Os, J., Kenis, G. & Rutten, B. P. 2010. The environment and schizophrenia. *Nature,* 468**,** 203-212.

Wehr, S., Weigel, L., Davis, J., Galderisi, S., Mucci, A. & Leucht, S. 2024. Clinical Assessment Interview for Negative Symptoms (CAINS): A Systematic Review of Measurement Properties. *Schizophrenia Bulletin,* 50**,** 747-756.

White, R. G. & Gumley, A. I. 2009. Postpsychotic Posttraumatic Stress Disorder: Associations With Fear of Recurrence and Intolerance of Uncertainty. *Journal of Nervous & Mental Disease,* 197**,** 841-849.

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**,** 1305-1318.

Chapter 2

Measuring Fear of Recurrence in Psychosis – Development of the Fear of Recurrence Scale Revised (FoRSe-R) and examination of its Content Validity

Prepared in accordance with the author requirements for Frontiers in Psychology (Psychology for Clinical Settings) for Original Research (https://www.frontiersin.org/about/author-guidelines)

Plain Language Summary

Title: Measuring Fear of Recurrence in Psychosis – Development of the Fear of Recurrence Scale Revised (FoRSe-R) and Examination of its Content Validity

Background

Psychosis is a severe and common mental health disorder. While recovery is common, relapse is frequent and associated with poor outcomes and trauma. Fear of recurrence in psychosis may be described as catastrophic thoughts and fears about the occurrence of a relapse. Fear of recurrence is associated with post-psychotic PTSD, depression and negative self-beliefs. The Fear of Recurrence Scale (FoRSe) was developed to measure Fear of Recurrence in Psychosis. Research has suggested that the scale may not measure all aspects of Fear of Recurrance. This study aims to begin the development of a revised version.

Aims

1. Develop a model of Fear of Recurrence in Psychosis and an initial item pool.
2. Establish the content validity of the initial item pool using a Content Validity Index study with Academic and Clinical experts, removing items that do not meet pre-set cutoffs.
3. Examine the relevance, comprehensiveness and comprehensibility of the updated item pool and domains via a Cognitive Interviewing study with people who have lived experience of psychosis.
4. Using the above evidence, make any changes required (item rewording, item removal, addition etc.) and select items for continued development of the FoRSe-R.

Methods

Participants

Study One – 12 participants were recruited from the research team's professional networks.

Study Two – Four participants with lived experience were recruited via mental health charities and social media.

Study Design

This study is a mixed methods content validity study, examining the items and subscales from the perspective of professional experts (Study One) and people with lived experience of psychosis (Study Two). Study Two used Cognitive Interviewing, which involves one-to-one interviews that are recorded and then transcribed, where people with lived experience are asked questions about the items. Study Two used this method to explore how people with experience of Fear of Recurrence in Psychosis understood the items, if they found them relevant and if they thought there were important questions or areas that were not covered.

Findings and Conclusions

Study One removed 24 items due to lack of relevance, finding 51 items had sufficient relevance. In study two 14 items were removed with 27 repaired, four retained and six added. The additional items were added to improve comprehensiveness of the domains. Further development work is suggested, examining the items that have been added and the response validity of the item pool before further quantitative psychometric work.

Abstract

Background and Aims: Psychotic disorders are severe mental health conditions, with a course that is often characterised by a cycle of recovery followed by relapse. Fear of Recurrence in Psychosis describes catastrophic thoughts and fears about the occurrence of a relapse. Fear of recurrence has been associated with relapse itself, post-psychotic PTSD, increased rates of depression, feelings of entrapment, shame, and self-blame. To the author's knowledge, only one measurement tool of fear of recurrence in psychosis exists, the Fear of Recurrence Scale. Review of the evidence of its psychometric properties has highlighted potential issues, including its content validity. Qualitative work also suggests that it does not fully capture this unique clinical construct. As such this study aimed to: 1). Develop an extensive item pool and proposed theoretical model underpinning Fear of Recurrence in Psychosis. 2). Establish the initial content validity of the initial item pool using a Content Validity Index study with Academic and Clinical experts, with removal of items that do not meet a priori outcomes. 3). Establish the evidence of the relevance, comprehensiveness and comprehensibility of the updated item pool and subscales via a Cognitive Interviewing study with experts by experience. 4). Using the above evidence, make any further changes required (item rewording, item removal, addition etc.) and select items for continued development of the FoRSe-R. **Methods**: A mixed methods approach was used to examine the content validity of the pilot item pool of FoRSe-R. Study One examined Content Validity using a qualitative agreement study with professional experts. Items that met a priori criteria for retention were then examined in Study Two using a Cognitive Interviewing methodology to examine the relevance, comprehensiveness and comprehensibility of the items and domains with experts by experience. This data was analysed using collaborative review to inform the retention, repair, removal or addition of items and domains. **Results**: 12 participants consented to participate in Study One. 51 items met criteria for retention and 24 items met criteria for removal. Four participants consented to participate in Study Two. Analysis resulted in the repair of 27 items and six domains, removal of 14 items, retention of four items and the addition of six items. **Conclusions**: This study has developed the initial item pool for the FoRSe-R and provided evidence as to the relevance, comprehensibility and comprehensiveness of the items and domains. Further developmental work is required to

establish the content validity of the repaired and added items. Response process work will then be required, followed by exploring the new scales validity and reliability through quantitative methods.

1. Introduction

Fear, defined as “a transient feeling that emerges when a person experiences an impending threat to one’s survival” (Ridenour et al., 2025) is a common feature of the subjective experience of psychosis (Karon, 1992, Sweeney et al., 2015, Fusar-Poli et al., 2022). Fear can be distinguished from anxiety, which is “a more diffuse feeling of worry, apprehension and rumination that is associated with a perceived but non-existent threat” (Gross and Canteras, 2012, pp. 656). Fear of psychosis itself is commonly reported (Sweeney et al., 2015). Lived experience accounts of first episode psychosis (FEP) describe the experience as terrifying, with disintegration of inner and outer boundaries, loss of agency, ego-dissolution and reality disruption, highlighted by experts by experience as core to the subjective experience (Fusar-Poli et al., 2022). These accounts echo psychoanalytic theories about psychosis, where fears of self-disintegration or annihilation of self are theorised as a central feature (Frosch, 1983, Davidsen and Rosenbaum, 2012).

Hallucinations, perceiving stimuli that are not present in the environment via any sensory modality, are described as confusing, fearful, demonic, frightening or terrifying (Sweeney et al., 2015, Alderson-Day and Ward, 2022) with feelings of pervasive terror, existential fear and entrapment described in first person accounts (Fusar-Poli et al., 2022). Delusional beliefs, intense societally and/or culturally inconsistent beliefs held with conviction and subjective certainty in the face of proof or evidence to the contrary, are sometimes described as bringing clarity to the incomprehensible and confusing experience of acute psychosis as well as personal meaning (Fusar-Poli et al., 2022). They are also characterised by feelings of distrust, hostility and threat (Ritunnano et al., 2022). The addition of the interpersonal nature of these fears results in an experience of existence in both an internal and external world that is characterised by persistent fear (Boyd and Gumley, 2007). It is unsurprising to consider that the process of receiving treatment, often coercively, while experiencing such an existential state can be highly distressing and traumatic. Contact with services, which can involve involuntary admission and detainment in hospital, physical or chemical restraint, seclusion and forced administration of medication, while in a psychotic crisis can echo delusional or

persecutory beliefs that an external agency is intending to inflict harm upon the person in crisis (Seed et al., 2016).

While many people recover from FEP, the possibility of the return of acute psychotic symptoms, and the consequences of such a relapse is a common concern for service users, carers and health care systems (Bassett et al., 2009, Lally et al., 2017, Estrade et al., 2022). There is strong evidence to justify this concern as both understandable and valid as 81% of people who have experienced FEP relapse within 5 years (Robinson et al., 1999). The impacts of FEP are extensive for both the individual and their wider network. Relapse has been shown to further impact individuals with increased risks of depression, further damage to social and occupational functioning, increased rates of all-cause mortality and risk of progression of symptoms to persistent, non-remitting psychosis (Jäger et al., 2007, Taylor et al., 2015, Johansen et al., 2020, Emsley et al., 2013). Psychologically, recovery from a relapse is associated with increased demoralization, feelings of entrapment and increased risk of Post Psychosis-Post Traumatic Stress Disorder (PP-PTSD) due to increased exposure to the trauma of experiencing psychosis its self as well as traumatic experiences associated with coercive treatment (Wiersma et al., 1998, Gumley et al., 2004), 1998, Gumley 2004). Indirect costs upon the individual and their carers are also extensive (Kotzeva et al., 2023). For health care systems, relapse is responsible for the majority of the direct costs associated with psychotic illnesses (Ascher-Svanum et al., 2010).

Considering the inherently fearful nature of psychosis and the impacts of a relapse, it is unsurprising that fear of a relapse is a common experience (Bassett et al., 2009). Fear of Recurrence in Psychosis (FoRP) describes catastrophic thoughts and fears about the occurrence of a relapse. FoRP has been associated with increased risk of relapse itself, PP-PTSD, increased rates of depression, feelings of entrapment, shame, and self-blame. Fear of relapse has been theorised in a cognitive interpersonal model (Gumley et al., 2020), where it is hypothesised that the experience of fear of relapse drives the affective surge observed during early warning signs (EWS) including feelings of anxiety, shame, demoralisation, and fear. Emotional dysregulation in this model may be driven by traumatic memories of previous psychotic episodes, both from the experience of psychosis and from iatrogenic harms

associated with treatment. To avert or minimise the possibility of relapse, individuals then engage in cognitive coping strategies such as hypervigilance, rumination and worry, and cognitive avoidance or behavioural coping strategies such as avoidance, social withdrawal and withdrawing from help-seeking. These coping strategies may in turn be interpreted by care staff as EWS and prompt clinical interventions that may enable prompt and helpful treatment or confirm individuals' fears and drive further disengagement and catastrophic expectations.

FoRP may also be understood in comparison to fear of cancer recurrence (FCR) (Simard et al., 2013), and other chronic physical health conditions (Lebel et al., 2020). FCR is considered a complex multidimensional experience that includes a range of concerns, some of which may overlap with FoRP - such as fears of the negative effects of treatment and loss of function (Lebel et al., 2020, Almeida et al., 2019). Despite the long-standing awareness of FoRP, research interest has been limited. A mixed methods systematic review found only nine studies that met criteria for inclusion (Zukowska et al., 2022), highlighting the existence of two scales, the Mental Health Worries Questionnaire (MHQW) and the Fear of Recurrence Scale (FoRSe), which was the only scale designed specifically for use in psychosis. The FoRSe was initially developed by Gumley and Schwannauer (2006) to measure FoRP. It measures three constructs, Awareness, Fear of Relapse, and Intrusiveness. Its content was derived from the concerns reported by participants in a randomised controlled trial of cognitive behavioural therapy for relapse prevention (Tait et al., 2002, Gumley et al., 2003).

Qualitative research has suggested that while the FoRSe measures some important aspects of FoRP, it may be missing important phenomena. Baker (1995) interviewed people with a diagnosis of schizophrenia and their family members, focusing on uncertainty of illness, highlighting loss of stability, the fear of previous distress/torment (PP-PTSD) and fears of the consequences of relapse. Nagle et al. (2002) explored FoRP in terms of occupational barriers, finding fears of relapse presented barriers to attaining and maintaining occupation of both leisure and monetary nature. Sandhu et al. (2013) highlighted how the perceived destructive power of relapse increased people's sense of fear of its recurrence and associated it with feelings of powerlessness, lack of control and associated social withdrawal. A qualitative study by Brookmann (2020) examined phenomenological experiences of people who experience

fear of relapse using Interpretative Phenomenological Analysis. This identified a broader range of fears such as fears of loss, fears of self-harm, fears of embarrassment, fears of letting others down, feelings of entrapment, fears of forced treatment and disclosure anxiety including fears of the consequences of disclosure of fears/symptoms (i.e. forced treatment, hospitalisation, loss of occupational and social autonomy) and stigma. Allan et al. (2020) explored patient, staff and carer experiences of EWS monitoring for relapse prevention, highlighting service users' fears of disclosure due to potential consequences and that these responses to EWS were based on their experiences of treatment and personal meanings of relapse. A systematic review by Zukowska et al. (2022) also suggested the inclusion of additional constructs to measure FoRP. Synthesizing four qualitative studies, they identified fears of losing social and occupational functioning, fear of the loss of ability to make autonomous decisions, trauma/re-traumatisation, and fear of relapse effects on wellbeing as potential areas for construct development.

Additionally, in the previous chapter (Chapter One pp 08 - 48), a COSMIN systematic review methodology was used to examine the evidence for the psychometric properties of the FoRSe. This review concluded that further work is required to establish the evidence for the psychometric properties of the scale. It highlights issues within its development and a lack of content validity work and suggests that as there is evidence within the current literature that supports FoRP being a broader concept than the scale's authors originally theorised; that a revised version of the scale should be developed in order to improve its content validity before proceeding to establishing other aspects of its psychometric properties.

The development and use of psychometric measurement tools, that is, those that are concerned with measuring latent constructs (Streiner et al., 2024), is central to the practice of evidence-based medicine. Psychometric measurement tools are used to establish the evidence for treatments, to inform diagnosis and treatment decisions as well as to assess the progress of patients. This highlights the importance of establishing the reliability and validity of the scores of the measurement tool as poor measurement can lead to dangerous outcomes for patients (McClimans et al., 2017). Within research, scores produced by measurement tools are 'upstream' of statistical analysis, highlighting flawed measurement as a potential culprit

in the replication crisis, with considerable, consistent evidence of poor psychometric methodological practices being noted across research (Ghai et al., 2022, Marshall et al., 2000, Morgado et al., 2018). Content validity is defined as the "the degree to which the content of a measurement instrument is an adequate reflection of the construct to be measured" (Mokkink et al., 2010). Content validity is considered by some to be the most important psychometric property (Mokkink et al., 2024), following the thesis that a highly reliable scale has little value if it does not measure its intended construct.

Content Validity can be evidenced through a variety of methods. Typically input from two groups are required, experts by profession and experts by experience. Experts by profession provide content validation evidence via their academic and/or clinical knowledge of the subject area while experts by experience contribute evidence via their lived experience of the condition/construct of interest and as the end users of the measurement tool. Both qualitative and quantitative methodologies can be used with either group. The most common method used with experts by profession are quantitative methods of agreement, such as content validity ratio, content validity index or Cohen's coefficient kappa, while qualitative methods including the Delphi method (Dalkey and Helmer, 1963) can also be used, but are less common. Qualitative methods are more common when evidencing content validity with experts by experience, with Cognitive Interviewing being the most frequently used methodology (Boateng et al., 2018, Willis, 2015).

Cognitive Interviewing is a qualitative methodology employed in questionnaire design that focuses on how respondents interpret, understand and respond to questions (Willis, 2015, Miller et al., 2014), not to be confused with a different procedure of the same name common in legal settings used to enhance reliability of eye witness testimony (Fisher and Geiselman, 1992). Cognitive Interviewing arose out of the field of survey design with the introduction of cognitive psychology to the field in the 1980s by the Cognitive Aspects of Survey Methodology (CASM) approach (Jabine., 1984). Based on the assumption that "the respondent's cognitions drive the survey response, an understanding of cognition is central to designing questions and to understanding and reducing sources of response error" (Willis, 2015, pp. 23), the CASM approach moved the focus from the question writer to the respondent and their internal

processes as central to reducing sources of error. CASM articulates that question response involves four stages: comprehension, retrieval, judgement, and response (Tourangeau, 1984, Tourangeau et al., 2000) and that investigation into these stages allows understanding of the complexity of the response process and through this reduces response error (Miller et al., 2014, Willis, 2015). CASM is credited with both establishing a methodology for conducting scientific enquiry into question evaluation and introducing the importance of establishing and evidencing theoretical prospectives for both the response process and the methodologies used to explore them, which had previously been ignored (Sudman et al., 1996, Miller et al., 2014).

Cognitive interviews involve conducting extensive interviews with a representative sample of end users of a questionnaire to explore aspects of survey functioning and, in order to understand how respondents interpret items, relate them to their own experiences and understandings and so produce responses. The process involves collecting verbal reports from respondents who are shown the questions of interest and commonly asked to answer them but are also then asked for additional information. Respondents are usually asked to think aloud, that is to detail their thinking process as they answer the question, or verbally probed, asked specific questions intended to target the processes of interest to the investigator. This data is then collected and analysed in various ways depending on the objective and theoretical orientation of the investigator. Objectives in cognitive interviewing can be roughly divided into two approaches: Reparative and Descriptive. The Reparative approach is concerned with improving survey functioning to reduce response error and is the most common use of cognitive interviewing (Willis, 2015). In contrast, the Descriptive approach is concerned with describing how a question functions as a measure of the construct of interest. This can be thought of as a continuum rather than a dichotomy, with studies often including elements of both. Analysis methodologies are varied, drawing from sociology, linguistics, social psychology and particularly from qualitative research such as Grounded Theory.

This study proposed to develop a revised version of the Fear of Recurrence Scale, the Fear of Recurrence Scale – Revised (FoRSe-R). The initial development proposed using a deductive

process, informed by up-to-date research about FoRP, as well as information from FCR, and Mental Health Anxiety literature. This was followed by a review by a stakeholder group consisting of clinicians, academics and experts by experience, with the view to develop an extensive item pool and proposed theoretical model. Two content validity studies were then conducted, one with academic and clinical experts and a second using Cognitive Interviewing with experts by experience to examine the content validity of the scale and select items for further development of the scale.

1.1 Aims

1. Develop an extensive item pool and proposed theoretical model underpinning Fear of Recurrence in Psychosis.

2. Establish the initial content validity of the initial item pool using a Content Validity Index study with Academic and Clinical experts, with removal of items that do not meet a priori outcomes.

3. Establish the evidence of the relevance, comprehensiveness and comprehensibility of the updated item pool and subscales via a Cognitive Interviewing study with experts by experience.

4. Using the above evidence, make any further changes required (item rewording, item removal, addition etc.) and select items for continued development of the FoRSe-R.

2. Methods

2.1 Development of the FoRSe-R

The original FoRSe is a 23-item scale developed by Gumley and Schwannauer (2006) contained three subscales measuring the constructs of Intrusiveness, Awareness and Fear of Relapse. Items are anchored by a five-point Likert scale (strongly agree, agree, neither agree nor disagree, disagree or strongly disagree). The scale content was informed by the individualised early warning signs provided by participants in a trial of cognitive behavioural therapy (CBT) for relapse prevention (Gumley et al., 2003, Tait et al., 2002).

The draft item pool for the Fear of Recurrence Scale – Revised (FoRSe-R) was developed by exploring the current literature on FoRP, as well as referring to the relevant literature on fear of cancer recurrence, mental health anxiety, and anxiety about experiencing mental ill health. The research team used this to develop an exhaustive draft item pool and a proposed theoretical model. This draft item pool and theoretical model were then evaluated by a stakeholder group independent of the initial process. This group consisted of members with relevant expertise via clinical, academic and/or lived experience, recruited from the research team's professional networks. The group was asked to provide general feedback as well as feedback focusing on the relevance, comprehensibility and comprehensiveness of the scale, subscales, items and proposed theoretical model. This feedback was collated and reviewed by the research team and the FoRSe-R updated.

The item pool resulting from this process was then evaluated using a two-phase mixed-methods approach. First a quantitative Content Validity study (Study One) was utilised to assess relevance of items from a group of experts, with academic or clinical expertise in psychosis or FoRP. Following this a Cognitive Interviewing study (Study Two) was utilised to provide content validity evidence from experts by experience in regards to relevance, comprehensiveness and comprehensibility of the subscales, items and response options. Study Two was reviewed by the Psychosis Research Group's Patient and Public Involvement Group at the design stage which provided input into methodological choices as well as adaptations to improve the accessibility for people who have experienced psychosis.

2.2 Study One - Content Validity with Experts by Profession

2.2.1 Participants

Participants were recruited via the Psychosis Research Group's professional networks. Inclusion criteria were status as an academic and/or clinical expert by profession, with a specific interest in either psychosis or fear or recurrence in psychosis. Participants with additional expertise by experience were also encouraged to participate.

2.2.2. Procedures

The study was approved by the University of Glasgow MVLS Ethics Committee (Reference: 200210206) (Appendix 2.2). The study was hosted online using the Qualtrics system.

Once the study was accessed, participants were provided with a brief introduction to the study followed by a privacy notice and participant information sheet (Appendix 2.9). Following this, participants were required to provide their informed consent (Appendix 2.10). Eligibility criteria were assessed via a list of screening questions, based on the inclusion criteria outlined above. Following informed consent, baseline demographics were collected (age, gender, ethnicity, profession/source of expertise, country of location).

Participants then proceeded to the study. Participants were shown each of the eight subscales of the FoRSe-R, alongside a definition of the subscale and each of the items contained within that subscale and asked to rate each item in terms of relevance. Ratings were on a four point Likert scale (not relevant, slightly relevant, relevant, very relevant). Upon finishing rating one subscale, participants were shown the next subscale and its items to rate until all items and subscales had been rated.

2.2.3 Analysis

Data were analysed using Microsoft Excel version 16.98 and SPSS version 29.0.2.0. Item Level Content Validity Index (I-CVI) was calculated for each subscale to assess agreement between participants. Pre-defined criteria for the retention of items, deletion of items and appropriate statistical analysis were defined a priori (Table 2.1). Items with an I-CVI ≥ 0.78 (range 0–1.0) were agreed to be acceptable (Polit et al., 2007). Items with a I-CVI lower than 0.78 were considered for rewording or removal. As I-CVI calculations do not take into account chance agreement with less than ten participants, it was agreed that, should this occur, K*-Modified (k) would be calculated instead, which is able to account for chance agreement at this participant number, with a cut off of k>0.74 (range 0-1.0) required for item retention (Polit et al., 2007, Cicchetti and Sparrow, 1981, Fleiss, 1981). As it was anticipated that further revisions to the scale would be required following Cognitive Interviewing, Scale level Content Validity Index (S-CVI) was not calculated. Items that were retained were taken forward to the second study - Cognitive Interviewing with people who have experienced psychosis.

Table 2.1: A Priori Analysis Plan

Recruitment outcome	Actions
Participants ≥ 10	Calculate I-CVI
Participants ≤ 9	Calculate K*-Modified (k)
Results	
I-CVI ≥ 0.78 or k > 0.74	Item retained
I-CVI ≤ 0.77 or k < 0.74	Item removed

2.3 Study Two – Cognitive Interviewing with Experts by Experience

2.3.1 Participants

Participants were recruited via two pathways. Initially the research team made contact with mental health organisations who work with individuals who have experienced serious mental illness within the Glasgow area. Additional recruitment took place via social media (BlueSky). Inclusion criteria were defined as: ≥16 years of age, who self-identify as having experienced psychosis and are currently in the United Kingdom.

Participants were offered reimbursement for their participation of £25 per hour of interview, up to a total of £75, in line with NIHR Involve rates (NIHR, n.d.). Reimbursement was available via bank transfer, cash or voucher at participant preference. Approval for the rate and methods of participant reimbursement were sought and approved by the MVLS University of Glasgow MVLS Ethics Committee (Reference: 200210206 Amendment 4) (Appendix 2.6).

2.3.2 Procedures

The study was approved by the University of Glasgow MVLS Ethics Committee (Reference: 200210206) (Appendix 2.2). The recruitment portal for the study was hosted on the Qualtrics system.

Once the study recruitment portal was accessed, participants were provided with a brief introduction to the study followed by a privacy notice and participant information sheet (Appendix 2.11). Following this, participants were required to provide their informed consent (Appendix 2.12). Eligibility criteria were assessed via a list of screening questions, based on the inclusion criteria outlined above.

Following informed consent, contact information was collected so that an initial meeting with the research team could be arranged with the participant to answer any questions, develop rapport, confirm eligibility and arrange the study interviews. This data was only used to contact the participant and not retained for analysis.

Study interviews were conducted either online via video call or in-person according to participant preference and geographic practicality. Interviews were expected to last three hours in total, and as such took place across several occasions based on participant preference. A topic guide was developed (Appendix 2.13) using a semi-structured prompt methodology (Willis, 2015). Adaptations to the interview to facilitate participation were encouraged in line with guidance from Gupta et al. (2024), considering common issues for people with schizophrenia, such as cognitive impairment, alogia and alternate communication styles, while maintaining a focus on collecting information about comprehensibility, relevance and comprehensiveness. Demographic information (age, gender, ethnicity, years of education completed, employment status, residential status) was also collected.

All interviews were conducted by the principal investigator, who is a final year Trainee Clinical Psychologist with previous experience of conducting semi-structured interviews with people who have experienced psychosis for research purposes. Interviews were recorded via MS Teams, with a second digital recorder used as a backup. Transcription was conducted by a member of the research team and reviewed by the principal investigator before being finalised.

2.3.3 Analysis

Analysis was undertaken using NVivo (version 14.24.3). This study derived its analytic approach from a theme coding method proposed by Knafl et al. (2007), which was based on procedures developed by Miles and Huberman (1994). Theme coding sits within an Interpretivist Grounded Theory stance, initially introduced to cognitive interviewing by Miller et al. (2011b) and developed further by Chepp and Gray (2014). The Interpretivist perspective involves a stance that meaning is constructed socially by the perceiver, within a social environment, with all interpretations being valid rather than seeking to establish correct and

incorrect perspectives. It is often descriptive in nature, that is, its focus is on describing the function of an item, rather than reparative, where the focus is to find issues with items in order to fix them or reduce the number of items . The methodology employed by Knafl et al. (2007), used bottom-up theme coding to allow for a reparative stance. Following transcription, participant quotes were assigned to each subscale and item and appended with their demographic data for context. The researchers then summarized participant comments and applied codes that were inductively defined. At the end of this process, the codes applied to each subscale and item were reviewed and decision rules were developed and used to inform modifications to the scale. Outcomes of this process included item removal, item re-wording and item retention. This method was adapted in the following ways. This study's interviewing approach was broader than Knafl et al. (2007) in that it explored comprehensiveness, comprehensibility, and relevance (as per COSMIN Guidance). Top-down coding was introduced based on these COSMIN defined constructs, resulting in a mixed coding approach. During this process, the lead researcher (AR) undertook the initial coding. Each transcript was then checked by a second coder (AG). In order to track this process and improve clarity of coding, elaborations to the COSMIN definitions were added. This was undertaken participant by participant. As with Knafl et al. (2007) participant quotes were applied to illustrate these codes but did not assign participant demographic to protect confidentiality. A collaborative analysis method was used (Miller et al., 2011a) which involved the research team jointly examining the whole body of data, using data displays, coding matrices and pattern coding methods, to then make decisions regarding each item and subscale which involved either repair, retention, deletion or addition of items (Knafl et al., 2007, Willis, 2015). The principal investigator and supervisor were involved in all analysis, including coding.

2.3.4 Transparency and Reflexivity

The principal investigator is a doctoral researcher and Trainee Clinical Psychologist who has worked in the field of mental health clinically for seven years and is experienced in delivering assessment, formulation and psychological intervention with adults who experience mental health difficulties, including those which fall across the psychosis spectrum. They have worked prior to this as a researcher within several mental health settings, with a primary interest in

psychosis research. The researcher, as such, has awareness of the power dynamics involved within both research and clinical care for people who have experienced psychosis and acknowledges that this awareness is positioned within their own experience as a clinician/researcher rather than that of an expert by experience. Through this, they anticipated that these dynamics may inhibit the responses of participants, especially as experiences of coercive care are common for people who have experienced psychosis. With this awareness they sought to facilitate interactions that acknowledged the value of participants knowledge and their autonomy as individuals. As part of their Doctoral training in clinical psychology, they were also aware that they have been trained to view and analyse experiences in terms of the biopsychosocial model of mental illness. They do not have lived experience of psychosis and their understanding of the phenomenology of psychosis is informed first by clinical and academic literature and training, which they are aware is produced primarily from the perspective people who do not have lived experience, and secondly from experiences, insights and perspectives shared with them during research and therapeutic encounters with people with lived experience.

3. Results

3.1 Development of the FoRSe-R

The research team developed an initial pool of 73 items divided into eight subscales (Hypervigilance, Early Signs, Worry and Uncertainty, Perceived Controllability, Fears of Relapse, Help Seeking, Interpersonal Expectations and Intrusive Memories) and proposed a model of FoRP (Appendix 2.14). This was provided to the reference group who consisted of eight members with a mix of academic (n=6), clinical (n=6) and lived experience expertise (n=2) in psychosis (Total >8 as some members hold multiple identities). This resulted in an extensive review of item wording, improving consistency in wording, reduction of technical language and clarifying items where they were felt to be ambiguous. Instructions were clarified and reworded to normalise the experiences being measured as part of an understandable response to the threat of relapse. No concerns were raised about relevance. One item was removed as part of the process and an additional eight items were added to improve comprehensiveness. Theoretical feedback resulted in the division of the subscales

into separate Threat and Mitigation scales. A description of the group, process and changes is available in Appendix 2.15 and Appendix 2.16.

The resulting theoretical model (Fig 2.1) proposed that an external or internal trigger results in: the activation of a Threat network associated with intrusive memories of traumatic past experiences of psychosis/treatment; activation of hypervigilance for mental/emotional changes as signs of relapse; and worries about relapse or specific fears of the consequences of a relapse. The response to these Threat experiences are then influenced by factors within the Mitigation aspect of the model related to beliefs about the controllability of relapse, understanding/knowledge of early signs and beliefs about help seeking and expectations of responses from others. Where an individual has positive beliefs/expectations in these areas, they are more able to mitigate or ameliorate their FoRP and engage in behaviours that reduce the chance of relapse occurring. Negative expectations in these areas are proposed to lead to increased threat, resulting in the individual engaging behaviours, such as avoidance, and reduced help seeking, resulting in increased distress and risk of relapse. Both the Threat and Mitigation aspects of the model can be co-maintained, with increased threat levels leading to further activation of attempts to mitigate the threat of relapse, and mitigation attempts not allowing for threat to become deactivated - meaning that fears of recurrence can become a persistent and ongoing problem over time.

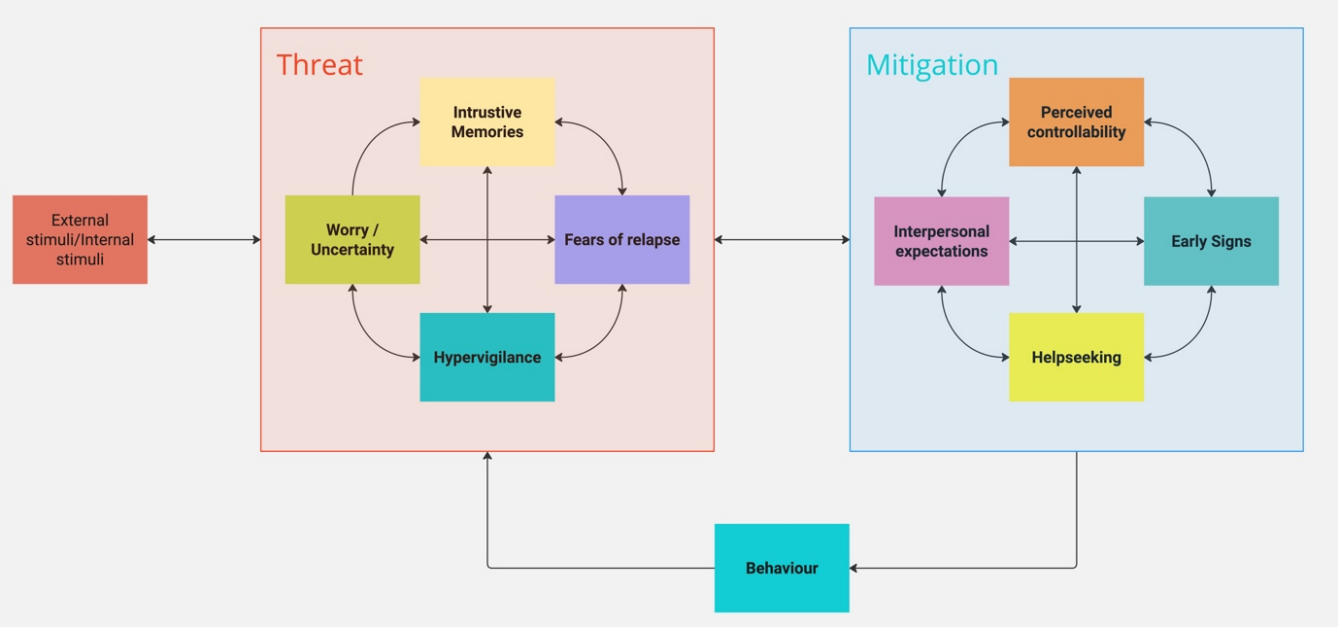

Figure 2.1. Fear of Recurrence in Psychosis Model

The items and structure of the pilot FoRSe-R are available within Appendix 2.17. It contained 80 items that are rated on a 5-point Likert scale from 0-4 (strongly disagree, disagree, neither agree nor disagree, agree, strongly agree) with higher scores indicating increased difficulty in that domain. As a result, several items are reverse scored. It had a timeframe of the past two weeks. It consisted of two primary subscales, Threat and Mitigation, which are proposed as a bi-factor hierarchical factor structure. The structure of the scales is defined below.

The Threat scale construct consists of Hypervigilance, Worry and Uncertainty, Intrusive Memories and Fears of Relapse subscales.

Hypervigilance (9 items) – This subscale includes items designed to measure the extent to which people experience heightened feelings of threat in regard to their psychosis returning, and the associated vigilance for signs of deterioration of their mental health.

Worry and Uncertainty (13 items) - This subscale includes items designed to measure the extent to which people experience uncertainty and worry in relation to fears of recurrence.

Intrusive Memories (9 items) – This subscale includes items designed to measure the extent to which people experience distressing, aversive or intrusive memories in relation to their previous experience of psychosis.

Fears of Relapse (10 items) – This subscale includes items designed to measure the extent to which people have specific fears about the consequences of having a relapse.

The Mitigation construct consists of Early Signs, Perceived Controllability, Help Seeking and Interpersonal Expectations.

Early Signs (9 items) – This subscale includes items designed to measure the extent to which people feel they can utilize their knowledge of early warning signs to recognise and act to prevent a relapse.

Perceived Controllability (9 items) – This subscale includes items designed to measure the extent to which people feel that they have control over a relapse occurring.

Help Seeking (11 items) – This subscale includes items designed to measure the extent to which people feel that they can seek help from others in the context of a relapse.

Interpersonal Expectations (10 items) – This subscale includes items designed to measure people's expectations about the way in which others will respond to a potential relapse.

The initial item pool and subscales were then included into the first Content Validity Study with Experts by Profession.

3.2 Study One – Content Validity with Experts by Profession

3.2.1 Participant Characteristics

A total of 18 individuals were contacted via the Psychosis Research Group's professional networks by email and invited to participate. A total of 15 responses were recorded by the Qualtrics system. Of these 12 identified themselves as eligible to participate in the study.

Table 2.2 shows the participant characteristics. Six (50%) were female, and one (8%) was non-binary/non-conforming. All 12 (100%) participants identified as being White. Eight (67%) identified their expertise as being clinical, nine (75%) as academic and two (17%) had additional expertise by experience in addition to either clinical or academic expertise. Nine (75%) participants were based in the United Kingdom and three (25%) participants were based outside the United Kingdom.

Table 2.2: Study One Participant Characteristics

Variable		Frequency(N=x)	Percentage (%)
Gender	Male	5	41.7
	Female	6	50.0
	Non-Binary/Non-conforming	1	8.3
Age	25-34	2	16.7
	35-44	5	41.7
	45-54	3	25.0
	55-64	2	16.7
Ethnicity	White	12	100
Source of Expertise	Academic	9	66.7
	Clinical	8	75.0
	Lived Experience	2	16.7
Country of Location	United Kingdom	9	75.0
	Other	3	25.0

3.2.2 Content Validity Index

As more than 10 participants were recruited, Item Level Content Validity Index (I-CVI) was calculated as per the a priori analysis plan with a cut off ≥ 0.78 I-CVI defined as acceptable for item retention (Table 2.1). I-CVI scores per item are presented in Table 2.3. In total 28 items were removed due to insufficient agreement. Four items were removed from the Hypervigilance Subscale; five items were removed from the Worry and Uncertainty subscale, two items were removed from the Intrusive Memories subscale, four items were removed from the Fears of Relapse subscale, five items were removed from the Early Signs subscale, three items were removed from the Perceived Controllability subscale, three items were removed from the Help Seeking subscale and two items were removed from the Interpersonal Expectations subscale resulting in 51 items meeting criteria for retention. An additional item in the Fears of Relapse scale was removed due to an administration error in the Qualtrics system, resulting in no data being collected. A decision was made to exclude this item in the cognitive interviewing study, resulting in 51 items proceeding to the second study.

Table 2.3: I-CVI Results

Subscale	Item	I-CVI	Retained (Y/N)	Subscale	Item	I-CVI	Retained (Y/N)
Hypervigilance	I constantly feel the need to monitor my mental health	**0.91**	**Y**	Worry and Uncertainty	I worry about the possibility of relapse	**1.00**	**Y**
	I am afraid of any changes in my mental health	**0.91**	**Y**		Worrying about relapse causes me distress	**0.91**	**Y**
	I am alert for any signs that my mental health might be getting worse	**0.91**	**Y**		I can't stop worrying about relapse	**1.00**	**Y**
	I am constantly checking if my mental health is okay	**0.66**	**N**		The thought of relapse makes me feel very uncertain and worried	**0.91**	**Y**
	I feel scared if I notice changes in my mental health	**0.83**	**Y**		Worrying about relapse makes me feel like I'm becoming unwell	**0.66**	**N**
	I feel scared that changes in how I am feeling will mean that I will relapse	**1.00**	**Y**		I don't believe I will ever be able to stop worrying about relapse	**0.91**	**Y**
	I do not feel I need to constantly monitor changes in my mental health	**0.41**	**N**		When I think about relapse, I start to panic	**0.66**	**N**
	I do not need to be alert for changes in my mental health.	**0.33**	**N**		I feel able to manage my worries about relapse	**0.83**	**Y**
	I check that my view of reality fits with others' views	**0.16**	**N**		I spend time worrying about what will happen to me if I relapse	**0.91**	**Y**
					I feel overwhelmed by worries about relapsing	**0.91**	**Y**
					Becoming emotionally upset may cause me to relapse	**0.50**	**N**
					There's no point worrying about relapse	**0.25**	**N**
					If I worry too much, it may cause me to relapse	**0.50**	**N**
	Total Items Retained/Removed		**5/4**				**8/5**

Table 2.3 [Contd.]: I-CVI

Subscale	Item	I-CVI	Retained (Y/N)	Subscale	Item	I-CVI	Retained (Y/N)
Intrusive Memories	I have been remembering previous episodes of relapse	**0.75**	**N**	Fears of Relapse	The thought of having a relapse frightens me	**1.00**	**Y**
	I have been remembering past experiences of being in hospital for my mental health	**0.83**	**Y**		I fear that I will end up in hospital if I have a relapse	**0.91**	**Y**
	I have vivid and distressing memories of relapse	**1.00**	**Y**		I fear I will let my friends/family down if I have a relapse	**0.83**	**Y**
	I find it hard to put memories of relapse out of my mind	**1.00**	**Y**		If I have a relapse, it will be my fault	**0.75**	**N**
	Memories of previous relapses keep popping into my mind	**1.00**	**Y**		The thought of relapsing overwhelms me	**0.58**	**N**
	I am distressed by memories of how I was treated during previous relapses	**0.91**	**Y**		I worry if I relapse, I may not recover	**0.83**	**Y**
	I am distressed by memories of how I behaved/what I believed during previous relapses	**0.91**	**Y**		I am afraid how I might act if I relapse	**0.83**	**Y**
	I am distressed by memories of what other people did to me when during previous relapses	**0.83**	**Y**		I am afraid of what I might do if I relapse	**0.75**	**N**
	I am distressed by not being able to remember previous experiences of relapse	**0.41**	**N**		I fear what will happen to me if I relapse		
					I am afraid of what others will do to me if I relapse	**0.75**	**N**
	Total Items Retained		**7/2**				**5/4**

Table 2.3 [Contd.]: I-CVI

Subscale	Item	I-CVI	Retained (Y/N)	Subscale	Item	I-CVI	Retained (Y/N)
Early Signs	I feel confident at recognising signs I may be having a relapse	**1.00**	**Y**	Perceived controllability	There's nothing I can do to prevent a relapse	**0.91**	**Y**
	I know what to do when I notice signs of a possible relapse	**0.91**	**Y**		I am confident that I can prevent myself having a relapse	**0.75**	**N**
	I feel confident that I can manage my wellbeing in response to signs of a possible relapse	**0.83**	**Y**		I have no control over whether I relapse or not	**0.91**	**Y**
	I can safely share details of signs of a relapse with other people	**0.58**	**N**		There are things I can do to stop myself from having a relapse	**1.00**	**Y**
	I have a positive plan of what to do if I experience signs of a possible relapse	**0.75**	**N**		I can recognize and act on the early signs of a relapse	**0.91**	**Y**
	Other people know how to support me if I experience signs of a possible relapse	**0.66**	**N**		If I start to relapse, I have strategies to stay well	**0.91**	**Y**
	I feel confident that by recognizing signs of relapse, I can prevent it happening	**0.83**	**Y**		If my symptoms return, I have coping skills that help me self-manage	**0.91**	**Y**
	Other people understand signs that I may be having a relapse and what to do to support me	**0.75**	**N**		I have skills that help me cope with and manage emotional distress	**0.75**	**N**
	If I experience signs of a possible relapse I just ignore them	**0.75**	**N**		In order to avoid relapse, I have to repeatedly check things	**0.25**	**N**
	Total Items Retained		**4/5**				**6/3**

Table 2.3 [Contd.]: I-CVI

Subscale	Item	I-CVI	Retained (Y/N)	Subscale	Item	I-CVI	Retained (Y/N)
Help Seeking	I feel confident to seek help from others if I'm worried about relapse	**1.00**	**Y**	Interpersonal Expectations	Other people will be caring and understanding if I have a relapse	**0.83**	**Y**
	I find myself seeking reassurance from others about my mental health if I'm worried about relapse	**0.66**	**N**		Other people will feel disappointed and let down if I have a relapse	**0.91**	**Y**
	I would not look for help from others if I felt like I was about to relapse	**0.83**	**Y**		Other people will force me into treatment if I have a relapse	**0.91**	**Y**
	If I'm worried about relapse, I play down my fears to others	**0.83**	**Y**		Other people will over-react if I share with them my worries about relapse	**0.75**	**N**
	I feel secure enough to tell others about my fears around relapse	**0.83**	**Y**		Other people will not take my worries about relapse seriously	**0.58**	**N**
	I feel I can tell my mental health team/provider about my fears about relapse	**0.91**	**Y**		Other people will help me if I relapse	**1.00**	**Y**
	If I tell my mental health team/provider about my fears of relapse, they will over-react	**0.58**	**N**		I will lose contact with friends or family if I relapse	**0.91**	**Y**
	I avoid talking about my fears about relapse at all costs	**0.83**	**Y**		Other people will play down my worries about relapse	**0.83**	**Y**
	I am able to ask for help from my friends/family when I am worried about relapse	**1.00**	**Y**		Other people will panic if I tell them I'm having a relapse	**0.83**	**Y**
	If I tell my mental health team/provider about my fears of relapse, they will not take me seriously	**0.58**	**N**		Other people will take away my control if I'm having a relapse	**0.83**	**Y**
	If I'm worried about relapse I just try to keep it to myself	**0.91**	**Y**				
	Total Items Retained		**8/3**				**8/2**

3.3 Study Two – Cognitive Interviewing with Experts by Experience

3.3.1 Participant Characteristics

Six participants were recruited. One did not respond to initial contact attempts following informed consent and one did not respond to further contact following the first interview. 4 participants participated in the study until completion. Recruitment occurred via Mental Health Charities (n=1), and the study recruitment portal (n=3). To reduce barriers to participation, demographics were collected after completion of interviews (Table 2.4).

Table 2.4: Study Two Participant Characteristics

Variable		Frequency (N=x)	Percentage (%)
Gender	**Male**	**2**	**50**
	Female	**1**	**25**
	Non-Binary/Non-conforming	**1**	**25**
Age	**25-34**	**1**	**25**
	35-44	**2**	**50**
	45-54		
	55-64	**1**	**25**
Ethnicity	**White**	**2**	**50**
	Mixed or multiple ethnic group	**1**	**25**
	Asian	**1**	**25**
Years of Education Completed	**0-11**	**1**	**25**
	12-17	**1**	**25**
	18+	**2**	**50**
Employment Status	**Unemployed**	**1**	**25**
	Part-time	**2**	**50**
	Full-time	**1**	**25**

3.3.2 Interview Characteristics

Interviews took place in person (n=1) or via video call (n=3) depending on geographic practicality and participant preference. All participants resided in the United Kingdom at the time of participation. The average length of interview was 245.43 minutes and were completed in an average of three sessions. Over four participants a total of 981.7 minutes (16 hours) of interview data were collected.

3.3.3 Analysis

A first pass of coding was conducted to apply codes for Relevance (+/-), Comprehensibility (+/-) and Comprehensiveness (+/-). During this pass, data that could relate to more complex bottom-up coding was identified and marked for review. A second review of the data was then conducted, reviewing the highlighted issues and developing codes from the data to process them into reparative themes within interviews. These themes were then compared across interviews and developed into themes that applied across the whole of the data. The themes developed from this process were: 1) Issues with Conceptualisation, 2) Epistemic Issues, 3) Exclusion of some respondents, 4) Clarity/Specificity Issues, 5) Complexity, 6) Language that evokes themes of Psychosis, 7) Items that produce a difficult response, 8) Utility issues and 9) Cultural issues.

The content of these codes was then summarised and incorporated into data displays. The research team collaboratively analysed and reviewed these using the data displays, coding matrices as well as assessing the verbatim data directly. This process was conducted across three meetings which were recorded. Each item and domain was reviewed and discussed until consensus was reached as to retention, rewording, removal or if an additional item was required. A summary of the outcome of this process is presented in Table 2.5. A detailed item by item analysis is available in Appendix 2.18.

Table 2.5: Cognitive Interviewing Analysis Outcome

Subscale	Item	*Summary of Issues	Decision	Updated Item
Hypervigilance			Retained	
HV1	I am afraid of any changes in my mental health	Scale	Updated	I have been afraid of any changes in my mental health.
HV2	I am alert for any signs that my mental health might be getting worse.	Rel Comp Complex Unrepairable	Removed	
HV3	I constantly feel the need to monitor my mental health	Rel C/S ToP Scale	Updated	I have been hypervigilant for signs of relapse.
HV4	I feel scared if I notice changes in my mental health	Epist Scale	Updated	I have felt scared when I notice signs of a relapse.
HV5	I feel scared that changes in how I am feeling will mean that I will relapse	Scale Auth	Updated	I have felt scared that changes in how I am feeling mean that I will relapse.

*Summary of issues: Rel= Relevance, Comp=Comprehensibility, Hensiv= Comprehensiveness, C/S=Clarity/Specificity, Cons=Conceptualisation, Epist=Epistemic, Miss=Misses some groups, Complex=Complexity, DiffR=Item evokes a difficult response, Scale=Scaling issue, Domain=Domain level Feedback, ToP = Evokes Themes of Psychosis

Table 2.5: Cognitive Interviewing Analysis Outcome [Contd.]

Subscale	Item	*Summary of Issues	Decision	Updated Item
Worry and Uncertainty			Updated	Worry Domain - defined as concerning the extent that people experience worry in relation to having a relapse.
WU1	I can't stop myself worrying about relapse	DiffR Scale	Updated	I have not been able to stop worrying about relapse
WU2	I don't believe I will ever be able to stop worrying about relapse	C/S Complex Unrepairable	Removed	
WU3	I feel able to manage my worries about relapse	Domain Scale	Updated	I have felt able to manage my worries about relapse
WU4	I feel overwhelmed by worries about relapsing	Comp Scale	Updated	I have felt distressed by worries about relapse
WU5	I spend time worrying about what will happen to me if I relapse	Comp C/S Unrepairable	Removed	
WU6	I worry about the possibility of relapse	Domain Complex Scaling	Updated	Worrying about relapse has caused me distress
WU7	The thought of relapse makes me feel very uncertain and worried	Cons C/S Unrepairable	Removed	
WU8	Worrying about relapse causes me distress	Scale	Updated	Worrying about relapse has been causing me distress

*Summary of issues: Rel= Relevance, Comp=Comprehensibility, Hensiv= Comprehensiveness, C/S=Clarity/Specificity, Cons=Conceptualisation, Epist=Epistemic, Miss=Misses some groups, Complex=Complexity, DiffR=Item evokes a difficult response, Scale=Scaling issue, Domain=Domain level Feedback, ToP = Evokes Themes of Psychosis

Table 2.5: Cognitive Interviewing Analysis Outcome [Contd.]

Subscale	Item	*Summary of Issues	Decision	Updated Item
Intrusive Memories		Hensiv C/S Miss	Updated	Aversive Memory Experience – This subscale includes item designed to measure the extent to which people have memories or a lack of memory related to previous experiences during psychosis that are experienced as distressing, intrusive or aversive
IM1	I am distressed by memories of how I behaved/what I believed during previous relapses	Miss Complex DiffR Scale	Updated	I have been distressed by memories of previous experiences of psychosis
IM2	I am distressed by memories of how I was treated during previous relapses	C/S Miss Domain Scale	Updated	I have been distressed by memories of how services treated me during previous experiences of psychosis
IM3	I am distressed by memories of what other people did to me when during previous relapses	Miss Scale	Updated	I have been distressed by memories of what other people did to me during previous experiences of psychosis
IM4	I find it hard to put memories of relapse out of my mind	Rel C/S Miss Complex	Removed	
IM5	I have been remembering past experiences of being in hospital for my mental health	Rel Comp Miss Auth	Removed	
IM6	I have vivid and distressing memories of relapse	Miss Scale	Updated	I have experienced vivid and distressing memories of psychosis
IM7	Memories of previous relapses keep popping into my mind	Rel Miss Complex	Removed	
			Added	I have been distressed by not being able to remember previous experiences during psychosis

*Summary of issues: Rel= Relevance, Comp=Comprehensibility, Hensiv= Comprehensiveness, C/S=Clarity/Specificity, Cons=Conceptualisation, Epist=Epistemic, Miss=Misses some groups, Complex=Complexity, DiffR=Item evokes a difficult response, Scale=Scaling issue, Domain=Domain level Feedback, ToP = Evokes Themes of Psychosis

Table 2.5: Cognitive Interviewing Analysis Outcome [Contd.]

Subscale	Item	*Summary of Issues	Decision	Updated Item
Fears of Relapse		Hensiv	Updated	Additional Items added
FoR1	I am afraid how I might act if I relapse		Retained	
FoR2	I fear I will let my friends/family down if I have a relapse	DiffR Cult	Updated	I fear I will disappoint others if I relapse
FoR3	I fear that I will end up in hospital if I have a relapse	C/S Miss Domain	Updated	I fear I will be forced into treatment if I have a relapse
FoR4	I worry if I relapse, I may not recover.	Comp Epist DiffR Unrepairable	Removed	
FoR5	The thought of having a relapse frightens me	Comp	Updated	Thinking about having a relapse frightens me
			Added	I am afraid of the impact a relapse would have on my life
			Added	I fear I will lose my autonomy if I relapse

*Summary of issues: Rel= Relevance, Comp=Comprehensibility, Hensiv= Comprehensiveness, C/S=Clarity/Specificity, Cons=Conceptualisation, Epist=Epistemic, Miss=Misses some groups, Complex=Complexity, DiffR=Item evokes a difficult response, Scale=Scaling issue, Domain=Domain level Feedback, ToP = Evokes Themes of Psychosis

Table 2.5: Cognitive Interviewing Analysis Outcome [Contd.]

Subscale	Item	*Summary of Issues	Decision	Updated Item
Early Signs		Hensiv Cons	Updated	Additional Items added
ES1	I feel confident at recognizing signs I may be having a relapse.	Complex Auth	Updated	I can recognise signs that I may be having a relapse
ES2	I feel confident that I can manage my wellbeing in response signs of a possible relapse.	Rel Comp	Removed	
ES3	I feel confident that by recognizing signs of relapse, I can prevent it happening.	Rel Cons Epist	Removed	
ES4	I know what to do if I notice signs of a possible relapse.	C/S Epist	Retained	
			Added	I feel confident that other people can help me recognise my signs of relapse
			Added	I feel confident that other people know my signs of relapse

*Summary of issues: Rel= Relevance, Comp=Comprehensibility, Hensiv= Comprehensiveness, C/S=Clarity/Specificity, Cons=Conceptualisation, Epist=Epistemic, Miss=Misses some groups, Complex=Complexity, DiffR=Item evokes a difficult response, Scale=Scaling issue, Domain=Domain level Feedback, ToP = Evokes Themes of Psychosis

Table 2.5: Cognitive Interviewing Analysis Outcome [Contd.]

Subscale	Item	*Summary of Issues	Decision	Updated Item
Perceived controllability		Rel Cons Epist Miss	Further development required	Significant revision, See Preceved Controllability in Domain and Item Analysis, p91
PC1	I can recognize and act on the early signs of a relapse.	Rel Complex	Further development required	
PC2	I have no control over whether I relapse or not	Rel Cons Epist	Further development required	
PC3	If I start to relapse, I have strategies to stay well.	Comp C/S Cons Domain	Further development required	
PC4	If my symptoms return, I have coping skills that help me self-manage.	Complex, Miss	Further development required	
PC5	There are things I can do to stop myself from having a relapse.	Cons Epist	Further development required	
PC6	There's nothing I can do to prevent a relapse.	Cons, Epist, DiffR Rel	Further development required	

*Summary of issues: Rel= Relevance, Comp=Comprehensibility, Hensiv= Comprehensiveness, C/S=Clarity/Specificity, Cons=Conceptualisation, Epist=Epistemic, Miss=Misses some groups, Complex=Complexity, DiffR=Item evokes a difficult response, Scale=Scaling issue, Domain=Domain level Feedback, ToP = Evokes Themes of Psychosis

Table 2.5: Cognitive Interviewing Analysis Outcome [Contd.]

Subscale	Item	*Summary of Issues	Decision	Updated Item
Interpersonal Expectations		Hensiv Auth	Updated	Additional Items added
IE1	I will lose contact with friends or family if I relapse	C/s Cons Miss	Updated	I worry my relationships will be damaged if I have a relapse
IE2	Other people will be caring and understanding if I have a relapse	Comp C/S	Updated	People around me will be caring if I have a relapse
IE3	Other people will feel disappointed and let down if I have a relapse	Diffr Cult	Removed	
IE4	Other people will force me into treatment if I have a relapse		Retained	
IE5	Other people will help me if I relapse	Epist	Updated	Other people will be helpful if I relapse
IE6	Other people will panic if I tell them I'm having a relapse	Rel Cons	Updated	Other people will overact if I tell them I think I'm having a relapse
IE7	Other people will play down my worries about relapse	Complex	Updated	Other people will minimise my concerns about relapse
IE8	Other people will take away my control if I'm having a relapse	C/S Auth	Removed	
			Addition	Other people will take me seriously if I say I'm having a relapse

*Summary of issues: Rel= Relevance, Comp=Comprehensibility, Hensiv= Comprehensiveness, C/S=Clarity/Specificity, Cons=Conceptualisation, Epist=Epistemic, Miss=Misses some groups, Complex=Complexity, DiffR=Item evokes a difficult response, Scale=Scaling issue, Domain=Domain level Feedback, ToP = Evokes Themes of Psychosis

Table 2.5: Cognitive Interviewing Analysis Outcome [Contd.]

Subscale	Item	*Summary of Issues	Decision	Updated Item
Help Seeking			Retained	
HS1	I am able to ask for help from my friends/family when I am worried about relapse	C/S Miss Auth	Updated	I am able to ask for help from people around me when I think I'm relapsing
HS2	I avoid talking about my fears about relapse at all costs	Comp Complex Utility Auth	Removed	
HS3	I feel confident to seek help from others if I'm worried about relapse	Domain	Updated	I would feel confident to seek help from others if I thought I was having a relapse
HS4	I feel I can tell my mental health team/provider about my fears about relapse	Comp Miss Complex Auth	Updated	I feel I can tell my doctor when I think I'm relapsing
HS5	I feel secure enough to tell others about my fears around relapse	Comp Complex	Updated	I feel safe to tell others if I think I'm relapsing
HS6	I would not look for help from others if I felt like I was about to relapse.		Retained	
HS7	If I'm worried about relapse I just try to keep it to myself	Complex Auth	Updated	If I'm worried I am having a relapse, I try to keep it to myself
HS8	If I'm worried about relapse, I play down my fears to others	Comp Complex	Removed	

*Summary of issues: Rel= Relevance, Comp=Comprehensibility, Hensiv= Comprehensiveness, C/S=Clarity/Specificity, Cons=Conceptualisation, Epist=Epistemic, Miss=Misses some groups, Complex=Complexity, DiffR=Item evokes a difficult response, Scale=Scaling issue, Domain=Domain level Feedback, ToP = Evokes Themes of Psychosis

3.3.4 Domain and Item Analysis

Hypervigilance

Hypervigilance was found to be relevant to three participants and had mixed relevance for one, who discussed hypervigilance of carers being important, reducing the relevance of the domain. No participants identified comprehensibility issues. It was understood as about a state of vigilance, heightened awareness or constantly scanning one's environment for signs, thoughts or experiences that could indicate a relapse. One participant described this experience as checking one's mind for signs of relapse and analogised this to body scanning in fear of cancer recurrence. Another understood the domain as feeling a heightened state of awareness for signs of relapse. Participants raised comprehensiveness concerns regarding day-to-day changes such as emotional changes or potential psychotic symptomology. Conceptual issues were raised by one participant who felt that hypervigilance can be also be centred in carers. Two participants highlighted that the normalisation of some level of hypervigilance is important for staying well and as part of a normal response to a condition with serious impacts.

During the collaborative review, this domain was found to be relevant, comprehensible and comprehensive. While it is likely that FoRP could create hypervigilance in carers, this was agreed to be outside of the scope of this scale.

Review of the items resulted in: one item, HV2, being removed due to issues with comprehensibility, relevance and complexity that were found to be unrepairable. Four items were repaired: HV1 due to scaling issues identified by the authors, HV3 due to relevance, clarity/specificity issues, wording that evoked themes of psychosis and author identified scaling issues, HV4 due to epistemic issues and scaling issues identified by the authors and HV5 due to scaling issues identified by the authors.

Worry and Uncertainty

This domain was relevant for all participants. Comprehensibility was mixed due to the use of both "Worry" and "Uncertainty", all participants found that their use together was hard to

understand, stating that they were distinct but important concepts. Participants raised comprehensiveness concerns, one felt that worries about being taken seriously when requesting help were a common experience but absent, one felt that worries about the practical impacts of relapse, such as loss of housing or financial implications were important to this area and another felt that worry about the impacts of distressing worry on day-to-day functioning were not assessed. Three participants raised conceptual issues with combining "Worry" and "Uncertainty" into one scale. "Worry" was viewed as unambiguously relevant, while "Uncertainty" was viewed with mixed relevance, two participants finding it relevant and one relating it to rationality with negative relevance.

During the collaborative review, the conceptualisation of "Worry" and "Uncertainty" as one domain was found to lack relevance, with strong evidence for the relevance of "Worry" as a domain, while "Uncertainty" was found to have insufficient evidence of relevance to be retained. The construct was redefined as "concerning the extent that people experience worry in relation to having a relapse."

The review of items resulted in: five items being repaired: WU1 due to provoking a difficult response and scaling issues identified by the authors, WU3 due to domain level feedback and scaling issues identified by the authors, WU4 due complexity and scaling issues identified by the authors, WU6 due to domain level feedback, complexity issues, wording that evokes themes of psychosis and scaling issues identified by the authors and WU8 due to scaling issues identified by the authors. Three items were removed: WU2 due to clarity/specificity issues and complexity issues that were found to be unrepairable and WU7 due to clarity/specificity issues and conceptualisation issues which were found to be unrepairable and WU5 due to complexity issues that were unrepairable.

Intrusive Memories

This domain was relevant for all participants. It was understood as about negative, intrusive or uncontrollable memories of psychosis or treatment experiences. Participants raised comprehensiveness concerns, one participant highlighted experiences of distressing absence of memory during psychosis, one found that treatment experiences were covered narrowly

and one regarding non-treatment experiences with services such as the police or social work. Utility issues were raised by one participant regarding the importance that the scale was not used when people are unwell. Issues around missing some people's experiences were raised by one participant in relation to the experience of memory blackouts, two felt that while the items covering delusional beliefs and behaviours during psychosis were relevant, other traumatic experiences of psychosis itself, such as thought disorder and hallucinations, were not covered. A conceptual issue was raised by one participant regarding memory being described as intrusive and thus having agency, which did not fit with their own conceptualisation of memory.

During the collaborative review, the conceptualisation of the domain was reworded to describe memories as being experienced as intrusive to address conceptualisation concerns and to cover a broader range of memory experiences. Feedback relating to comprehensiveness about the range of treatment, non-treatment and psychosis experiences covered was addressed in the review of items IM2 and IM3. The domain description was reworded to "This subscale includes items designed to measure the extent to which people have memories or a lack of memory related to previous experiences during psychosis that are experienced as distressing, intrusive or aversive." The domain was renamed to "Aversive Memory Experiences" to better cover non-intrusive memory experiences such as blackouts, and an additional item was added to the domain "I have been distressed by not being able to remember previous experiences during psychosis".

The review of items resulted in four items being repaired: IM1 due to excluding some respondents, complexity issues, provoking a difficult response and scaling issues, IM2 due to clarity/specificity issue, excluding some respondents, domain level feedback and scaling issues, IM3 due to excluding some respondents and scaling issues and IM6 due to excluding some respondents and scaling issues. Three items were removed from the scale: IM4 due to relevance, clarity/specificity issues, excluding some respondents and complexity, IM5 due to relevance and complexity issues, excluding some respondents and additional issues identified by the authors, and IM7 due to relevance and complexity issues as well as excluding some respondents.

Fears of Relapse

This domain was relevant for all participants. It was understood as about fears related to a relapse occurring and about fearing the consequences of a relapse. No comprehensibility issues were identified. Participants raised comprehensiveness concerns, three regarding the practical impacts of relapse, such as fears of loss of accommodation, occupation or financial impacts, one regarding fears related to the impact of relapse on life course/trajectory. Two participants raised that hospital was the only treatment fear specifically in the domain, one reporting that this can sometimes act as a proxy for relapse but items that cover other experiences would be useful and another raised that non-hospital service experiences were missing, such as interactions with the police or community mental health services. Exclusion of some groups was raised by one participant regarding the use of hospital and friends and family, echoing feedback from other domains.

During the collaborative review, this domain was found to be relevanant and comprehensible. Comprehensiveness issues raised were addressed by the addition of an item "I am afraid of the impact a relapse would have on my life" and in the review of FoR3. The domain itself was retained unaltered.

Review of the items resulted in: one item, FoR1 being retained unaltered. Three items were repaired: FoR2 due to provoking a difficult response and cultural issues, FoR3 due to clarity/specificity issues, missing some respondants and domain level feedback and FoR5 due to comprehensibility issues. An additional item was added when repairing FoR3. One item, FoR4, was removed due to comprehensibility and epistemic issues as well as evoking a difficult response, these issues were found to be unrepairable.

Early Signs

This domain was relevant for all participants. It was understood as concerning a person's knowledge of their idiosyncratic early signs of relapse by one participant and as being about knowing the early signs of relapse and how to respond to them by three participants. No comprehensibility issues were identified. Comprehensiveness issues were raised by one participant regarding a person's confidence in others responding to their early signs. Conceptualisation issues were raised by two participants, regarding awareness of early signs

being possible to centre in their social/care network, due to lack of personal awarness, echoing comprehensiveness issues. Epistemic issues were raised by one participant regarding the framing of relapse as preventable or completely controllable and implied total control of relapse was possible, centring that locus of control/responsibility within the respondent. They found a framing of relapse as influenceable or modifiable with the acknowledgment that aspects will be outside the control of individuals more acceptable.

During the collaborative review, the domain was found to be relevant and comprehensible. An item about belief in appropriate responses from others to signs/help seeking was relevant, but more related to Interpersonal Expectations. The issues raised of epistemic injustice were agreed to be unacceptable and not fit with the intended conceptualisation of the domain and were addressed in item ES4. Two additional items were added to address comprehensiveness issues, “I feel confident that other people can help me recognise early signs of relapse” and “I feel confident that other people recognise my signs of relapse”.

Review of the items resulted in: one item, ES4, being retained. One item, ES1, was repaired due to complexity and author identified issues. Two items were removed: ES2 due to relevance and comprehensibility issues and ES3 due to relevance, conceptual and epistemic issues.

Perceived Controllability

This domain had mixed relevance. Two participants found the domain relevant, one found its relevance was mixed and one did not find it relevant. Comprehensibility was found to be positive by two participants and two did not feed back on its comprehensibility. No comprehensiveness issues were identified. Conceptualisation and epistemic issues were raised by one participant who found the framing of relapse as preventable in this domain reduced relevance and did not fit with their understanding of relapse. They found that changing this to how much a person feels they can influence the risk, impact or extent of a relapse addressed these issues and better related to self-efficacy. One participant raised that some groups would be missed in items due to referring to the return of symptoms, as people often live with enduring symptoms without being acutely unwell.

During the collaborative review, the domain was agreed to require repair due to relevance and conceptualisation issues. The definition was found to have created issues around epistemic injustice due to framing relapse as preventable. A review of the the domain and items resulted in a conceptualisation of this domain as relating to the extent to which a person believes that relapse is modifiable, i.e. its risk, severity, length, course or consequence can be influenced via their or others influence, related to self-efficacy, autonomy and empowerment. As part of the review, it was agreed that the domain would require further evaluation and development work before finalisating any repair. When considering the items within this domain a large number were initially cadidates for removal. In light of the highlighted need for further development work within this domain, it was agreed that the data and items should be included in that process rather than drawing conclusions at this stage.

The issues identified in the items where as follows: PC1 was found to have to relevance, conceptualisation and epistemic issues, PC2 was found to have conceptualisation, relevance and epistemic issues, PC3 was found to have issues with comprehensibility, clarity/specificity, conceptalisation as well as domain level feedback, PC4 was found to have complexity issues and excluding some respondents, PC5 was found to have conceptualisation, epistemic and author identified issues and PC6 was found to have relevance, conceptualisation and epistemic issues as well as provoking a difficult response. Multiple potential items were generated as part of this review (Appendix 2.19) but were not added to the domain and will be brought forward alongside the domain and items for further development and evaluation.

Help Seeking

This domain was relevant for all participants. It was understood to refer to if people think they would be able to seek help from others in a potential relapse, but not about the quality of that help. No comprehensibility issues were identified. Comprehensiveness issues where raised, one participant raised concerns that being listened to or having your view taken seriously should be addressed here or in Interpersonal Expectations, one felt that the items covered the network around the respondent unequally, one raised that the stage in which a person would seek help was absent. Conceptualisation issues were raised regarding the

interaction between this domain and insight, as people will only seek help for something if they perceive it to be a problem.

During the collaborative review, the domain was found to be relevant and comprehensible. Comprehensiveness was improved via an item to assess if a person feels they will be listened to or have their view taken seriously was agreed to be more relevant to Interpersonal Expectations. At what point in the process of a relapse a person would seek help was agreed to be outside of the conceptualisation of this domain. The feedback regarding insight was acknowledged to be important and was agreed to be a potential avenue for hypothesis testing in later psychometric work. The domain was retained as worded.

Review of the items resulted in the following, one item, HS6, was retained unaltered. Five items in this domain were repaired: HS1 due to clarity/specificity issues, excluding some respondents and author identified issues, HS3 due to domain level feedback, HS4 due to comprehensibility and complexity issues, HS5 due to comprehensibility and complexity issues and HS7 due to complexity and author identified issues. Two items were removed: HS2 due to comprehensibility, complexity and utility issues as well as additional issues identified by the author and HS8 due to comprehensibility issues and missing some respondants.

Interpersonal Expectations

All participants found this domain to be relevant. It was understood by as concerned with expectations of how others will treat you during a relapse or if they believe one is possible and as how safe/confident people feel in their relationships in the context of a relapse. No comprehensibility issues where identified. Comprehensiveness issues were raised, one participant regarding having ones view taken seriously or being listened to and one regarding expectations of how relationships will be following recovery.

During the collaborative review, the domain was found to be both relevant and comprehensible. It was agreed that expectations of people's responses after recovery were outside of the scope of this domain. Other people taking worries seriously or being listened to was found to be relevant, but raised issues related to Study One, as an item related to this "Other people will not take my worries about relapse seriously" was removed due to not

meeting criteria for retention. It was agreed that this experience has clear evidence for relevance from a lived experience perspective requiring the addition of an item "Other people will take me seriously if I say I'm having a relapse".

Review of the items resulted in the following: one item, IE4, was retained unaltered. Five items in the domain were repaired: IE1 due to clarity/specificity and conceptual issues as well as excluding some respondents, IE2 due to comprehensibility and clarity/specificity issues, IE5 due to epistemic issues, IE6 due to relevance and conceptualisation issues, IE7 due to complexity issues. Two items were removed: IE3 due to provoking a difficult response and cultural issues, IE8 due to clarity/specificity issues in addition to issues identified by the authors.

Overall questionnaire

All participants reported that the questionnaire was relevant and felt to be comprehensive. One participant highlighted that they felt that it was important to try to keep the questions as independent of the health care system as was reasonable to do so and that they preferred questions that were applicable to a range of experiences rather than multiple items for individual experiences. One participant raised that the length of the questionnaire would create utility issues if it was kept at its current length.

3.3.5 Reasons for Repair, Retention, Removal or Addition

The majority (n=30) of items were found to be relevant by participants. Despite extensive review by the authors and the reference group, comprehensibility issues were expressed by participants in 30 items, with seven items removed involving issues of comprehensibility.

The Comprehensiveness of Worry and Uncertainty, Fears of Relapse and Early Signs domains were improved by the identification of experiences understood to be important by participants, resulting in either additional items (n=6) or the repair of existing items (n=2). Of the additional items, two were similar in content to items removed during Study One.

Clarity/Specificity issues were the most common issue identified by participants, affecting 26 items. The analysis found that of these 26 instances, six required repair and five were

removed. Items required repair due to using terms such as “Mental health” where this was understood as being vague, or were items had multiple objects or allowed for ambiguity to create difficulties in understanding the intended object of inquiry. These issues were closely related to Complexity, which was identified in 16 items by participants. The analysis found that eight items required repair and eight were removed. Complexity related to issues such as idiomatic language that would create issues for people who speak English as a second language, questions that contained too many parts, use of double negatives, or the use of formal or technical language.

The analysis found that the 12 items were either worded or conceptualised in a way that excluded a group of potential respondents. Wording such as “previous relapses” within the Intrusive Memories domain, unintentionally excluded those who had experienced only a single episode of psychosis. The use of hospitalisation as a proxy for coercive treatment excluded the responses of those who had experienced psychosis but not been hospitalised. The framing of the non-clinical aspect of an individual’s social network as “friends and family” was highlighted as exclusionary to individuals who have either no contact with their family or small/limited social networks, an experience that the participants felt to be common among some people who experience psychosis.

There were 14 instances where participants discussed epistemic issues in relation to the items. There were several instances where items, particularly within the Early Signs and Perceived Controllability domains, referred to relapse as preventable or controllable rather than as something that can be influenced. This framing was understood by participants to centre the responsibility of controlling relapse within the respondent, rather than acknowledging relapse as unpredictable, affected by multiple factors some of which are outside of respondent’s control. Early Signs as defined was also raised as not accounting for the experience of a condition that impairs ‘insight’ making noticing changes more complex. Some of these items also were experienced by participants as difficult, unacceptable or challenging, provoking difficult responses. In the analysis this resulted in the need to repair two items, where items were understood as judgemental or inappropriate culturally for some participants, and the removal of three items, where epistemic injustice issues such as items being understood as perpetuating stigma of psychosis via associations with negative tropes

of those who have experienced psychosis suffering from a neurodegenerative condition with a homogenous course. Some items where also felt to produce needless anxiety.

Not all epistemic issues were felt to be negative. Participants raised that some items addressed injustice issues by acknowledging a person's autonomy, were relevant to lived experience of invalidation or navigated the potentially conflicting perspectives of service users and services regarding forced treatment in a way they felt to be positive.

Participants experienced some items as difficult, unacceptable or challenging, mostly in relation to epistemic issues or conceptualisation issues. Some items were understood as perpetuating epistemic injustice, such as the framing of relapse as controllable or where the wording or conceptualisation of items were understood by participants as infantilising, inappropriate culturally or implying judgement. Similarly, some wording was understood as evoking themes of psychosis itself and thought to possibly increase anxiety or suspiciousness by some participants. Analysis resulted in the repair of two items to remove the use of language such as "manage" and "monitor", both of which were found not to be required for item relevance or comprehensibility and reduced their utility.

4. Discussion

This study aimed to begin the development process of a revised measurement tool for FoRP. It followed best practice using guidance from the literature on modern psychometric development procedures (Boateng et al., 2018, Swan et al., 2023, Streiner et al., 2024, De Vet et al., 2011, Mokkink et al., 2024).

Study One found that the majority of items (n=51) met *a priori* standards (I-CVI≥0.78) for retention, with 28 items being rejected for not reaching a level of acceptable agreement. The population of this study consisted of a mixture of clinical and academic experts, two of whom also had additional lived experience of psychosis. The retained items were then included in the item pool of Study Two.

Study Two explored the evidence for the relevance, comprehensibility and comprehensiveness of the domains and retained items from Study One via Cognitive

Interviewing. The methodology was developed from the Interpretivist perspective (Miller et al., 2014), that is the study did not seek to understand the cognitive processes of how a respondent understands an item or generates a response, but rather the study sought to understand how respondents interpreted the items and related them to their own lived experience. Collaborative analysis employed to make decisions about retention, rewording, deletion or addition of items and to assess the item pool content validity.

This process resulted in six of the domains requiring repair and two domains being retained unaltered. Of the repaired domains, three required revision to their definition or conceptualisation and three required the addition of items to improve comprehensiveness without such alterations. One domain, Preceived Controllability, was found to require further development and evaluation work rather than repair, due to the extensiveness of revision required and small sample size of this study. In addition to this, 27 items were repaired, four items were retained unaltered, and 14 items were removed. An additional six items were added, of which two had content similar to items removed due to lack of agreement as to their relevance in Study One. See Figure 2.2 for the process flowchart for the study.

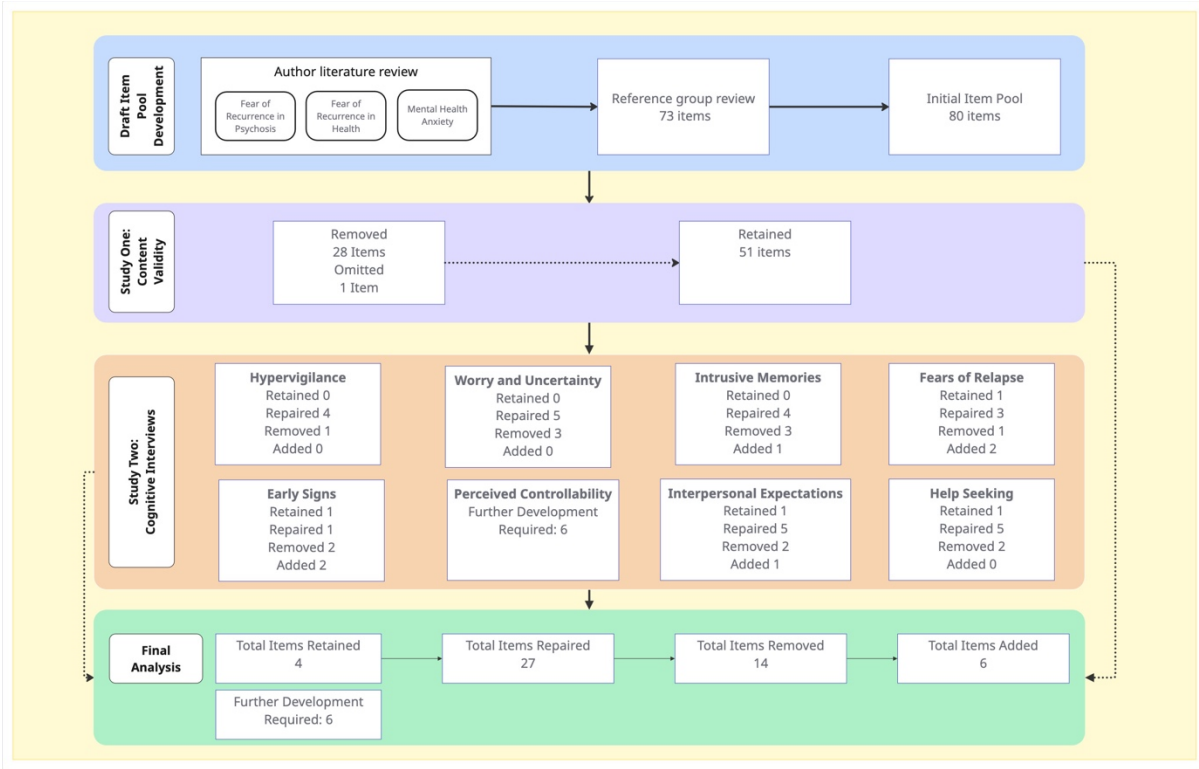

Figure 2.2 Study Process Flowchart

4.1 Contradictory Findings

There were instances of contradictory data between the studies. The experience of fearing that efforts to seek help from services in the context of a potential relapse would be met with inappropriate responses such as dismissal, neglect or a lack of urgency was raised by multiple participants. They shared lived experience of services not making referrals to mental health teams, clinicians not recognising or treating their concerns as valid and inappropriate or untimely responses. In Study One, the item "Other people will not take my worries about relapse seriously" was removed due to lack of agreement as to its relevance to the Interpersonal Expectations domain (I-CVI=0.57) and another "If I tell my mental health team/provider about my fears of relapse, they will not take me seriously" was also removed due to lack of agreement regarding its relevance to Help Seeking (I-CVI=0.58). In Study One participants were not asked to provide reasons for their rating, nor assign the item to the domain that they felt most relevant, due to the practicalities of recruiting experts and the length of the task. As such, it is possible that both items could have achieved acceptable agreement in another domain, or that they were perceived to not be relevant due to other factors such as comprehensibility. Analysis of the data from Study Two raised sufficient evidence of issues in the comprehensiveness of Interpersonal Expectations to justify the inclusion of an item with content similar to those removed in Study One.

Two participants raised that their early signs confidence/knowledge was centred in others in their lived experience, two items "Other people know how to support me if I experience signs of a possible relapse" (I-CVI=0.66) and "Other people understand signs that I may be having a relapse and what to do to support me." (I-CVI=0.75) were removed from the Early Signs scale during the first study. Collaborative analysis of the data from cognitive interviewing raised sufficient evidence of a lack of comprehensiveness of the Early Signs domain due to the absence of items relating to the experience of early signs knowledge being viewed as centred in a person's social network, sometimes due to the difficulties of being aware of signs that an insight impairing condition is relapsing. As a result, an item similar to those removed was added to be examined in further development work.

An item regarding blackouts/distressing experiences related to an absence of memory was removed from the Intrusive Memories domain, "I am distressed by not being able to

remember previous experiences of relapse" (I-CVI=0.41). During the collaborative analysis, which resulted in the reconceptualization of the domain, it was agreed to reintroduce an item similar to this for examination in further development work.

Two additional items presented conflicts between the studies but were not readded. "If I have a relapse, it will be my fault" (I-CVI=0.75) was not readded despite evidence of the relevance of fearing experiencing of feelings of guilt and shame in response to a potential relapse. Similarly, "I check that my view of reality fits with others' views" (I-CVI=0.16) was not readded despite relevance of data for this experience from one participant. In both cases, it was agreed that the data were too limited to justify their reintroduction, however further developmental work may provide further evidence to do so.

4.2 Major Issues

The evidence from the second study suggested that a reconceptalisation of the Perceived Controllability domain is required. The cognitive interviews raised that the use of control within the name and description raised epistemic and relevance issues for participants. The epistemic issues centred around the framing of relapse as preventable and the centring of the ability and responsibility of this within the individual who experiences psychosis, to the exclusion of the system around them, which was not the intended framing of the domain. Collaborative analysis resulted in a reframing of this domain as intended to assess how modifiable a person believes the course, risk or extent of a potential relapse is, acting as a bridge between Early Signs, Help Seeking and Interpersonal Expectations. At its core, it intends to assess "Do you believe that there is anything that can be done to change the likelihood or severity of relapse?" with links to empowerment, self-efficacy and autonomy, while also acknowledging that the health/social/care network around the individual has power and influence over the course of a relapse. The use of self-efficacy/autonomy in this scale is not meant to centre responsibility within the person, but instead to understand "Do you believe that relapse is modifiable?" and "Do you think that you can modify it?" either through their own actions or by seeking help from others. Multiple items were developed with this updated framing but further work to explore how this concept is defined and relates to FoRP is required before the domain can be examined in further development work. Considering the sample size of this study and the need for further development work on the

Perceived Controllability domain, the current items will also be included in this work prior to decisions regarding their repair or removal.

The Worry and Uncertainty domain was also revised significantly. Worry as a domain had strong evidence for relevance to FoRP, which allowed us to straightforwardly accept a repair of the domain to be concerned with Worry. Uncertainty's evidence was less straightforward. Participants were less clear around the relevance of uncertainty, referring to it as essential to the experience of the condition or 'part of the furniture' but finding it more nuanced than worry. Uncertainty was spoken about as being protective, some level of uncertainty being conducive to staying well but an excessive amount being unbearable. Collaborative review resulted in the domain being repaired by the removal of Uncertainty from the domain's name and definition, and a potential conceptualisation of uncertainty within FoRP more fitting with the idea of "uncertainty experienced as/of a degree that is intolerable" could be more relevant but resulted in the conclusion that further qualitative work would be necessary to understand what role uncertainty plays in FoRP.

The evidence from this process highlights the importance of content validity developed from the population for which a scale is intended as well as professional experts, a process considered best practice but rarely achieved in the literature, where expert opinion alone continues to be the main evidence used to evidence content validity (Morgado et al., 2018). This is despite clear guidance that evidence from the intended population of use is essential and forms the basis for content validity (Clark and Watson, 1995) and is a key component of evidencing validity, the most important of psychometric properties (Swan et al., 2023). Despite this, factor analysis is the most often cited evidence of validity for psychometric scales designed through Classical Test Theory, with Hubley et al. (2014) finding none of the 59 studies reviewed in *Psychological Assessment* and the *European Journal of Psychological Assessment* provided evidence of content validity as part of their validity work, a practice considered methodologically unacceptable if one uses Classical Test Theory as the basis of their design (McNeish, 2024). This project has aimed to navigate the complex literature and guidance available regarding the development of psychometric tools. In doing so, it is hoped that it has begun the initial, iterative process of producing a tool for measuring an important

latent construct commonly experienced by those who have experienced psychosis and live with a sense of fear at the idea of its return.

4.3 Future Directions

This study has further developed the initial item pool for the FoRSe-R, providing evidence as to the relevance, comprehensibility and comprehensiveness of the items and domains. The repair of the items, as well as the repair and issues indentifed in the Perceived Controllability scale and Worry scale merits further content validation work with both professional and lived experience experts. Perceived Controllability and Locus of Control have been raised as important mediators of illness distress in multiple health conditions including psychosis (Birchwod et al., 2003, Shelley and Pakenham., 2004), considering this and the views and experiences of the participants of this study, further development work in collaboration with experts by experience as well as experts by profession is warranted in addition to further evaluation work, to ensure that changes to this domain and its items are well considered and representative.

The scaling, instructions and time frame should also be examined as part of this work. Of note, this study has not provided evidence of the response process, another essential element of validity work both required for the use of Classical Test Theory and commonly found to be absent in the development of psychometric measurement tools (Hubley et al., 2014). This work requires conducting further cognitive interviews focusing on how responses to the items with their associated scales are produced, usually using think aloud methodologies (García and Baena, 2014, Willis, 2015)

Following this stage, should no further content validity work be required, validation via quantitative methods will be required including the examination of reliability, responsiveness and factor analytic work. This process may necessitate further revisions to the scale and should be considered part of its development process until a final measure can be shown to have sufficient evidence of such a high methodological quality that the scores it produces can be considered valid and reliable within the context and population of its intended use.

With further development the measure has potential to provide clear annunciation of tractable psychological targets that are amenable to change via psychological therapy for people with FoRP. The addition of the FoRSe-R to routine clinical care for people who have experienced psychosis may help to identify indviduals for whom FoRP presents a clinically significant problem and help establish areas that may be suitable for intervention. It is intended that the Forse-R will help with the development of target treatment modules designed to address those specific domains, such as interpersonal therapy (MacBeth et al., 2019) where scores are raised in Interpersonal Expectations or Help Seeking domains, trauma focused therapy (van den Berg et al.,2018) for people who have raised scores on Intrusive Memories, Hypervigilance or Fears of Relapse, individualised early signs monitoring (Gumley et al., 2022) may be of benefit where scores on Early Signs are raised and compassion focused therapy (Braehler et al., 2013) may be of benefit for those who have raised scores of Hypervigilance or Worry. This would address an unmet clinical need for people who experience FoRP, and potentially a reduction in the severity and frequency of relapse itself.

4.4 Strengths and Limitations

This study aimed to document its processes explicitly and to fit an open science model so as to allow for replication, meeting methodological considerations regarding psychometric scale design and as part of the principle of credibility in qualitative research and cognitive interviewing (Miller et al., 2014). It aimed to use a methodology that fits with the requirements of developing a psychometric tool within Classical Test Theory, examining the evidence of content validity from the perspectives of both experts and people with lived experience, a practice that is commonly absent in psychometric design (Hubert et al 2014). The methodology for cognitive interviewing was informed by lived experience expertise, to ensure it was accessible for participants and to achieve as representative a sample as possible. It also adopted advice from Gupta et al. (2024) on conducting qualitative research with people who have experienced psychosis, including adapting interviews where required to meet individuals' accessibility needs. This allowed for the collection of data from individuals who would otherwise may not have been able to participate in Cognitive Interviewing. This further increased the representativeness of the sample.

The study is not without limitations. The design of Study One means that while the items retained can be evidenced as relevant to their domain, the reasons for item removal are difficult to discern. Other methodologies, such as a Delphi study, interviews or focus groups with experts may have allowed us to examine their perspective in more depth. However, the methodology used provided a practical baseline for examining relevance with limited resources and in a way that experts found an acceptable level of participation burden. Additionally, the sample was recruited via the study team's professional networks and word of mouth, which may have limited the range of perspectives and views of the expert pannel. Future developmental work on the item pool developed from this study should consider a qualitative avenue to develop a deeper understanding of its content validity from the expert perspective as well as using broader recruitment such as an open call via social media/directly contacting professionals via their institution.

This study did not recruit from NHS services, but all participants reported having used NHS services and made reference to this within the interviews. Future studies recruiting via non-NHS sources may benefit from collecting clinical demographics such as diagnosis, types of services used (i.e. CMHT, Early Intervention in Psychosis Service, Inpatient Services), number of hospitalisations and number of relapses to allow for better description of clinical aspects of the sample. Further purposeful sampling from NHS services, including FEP services, could also improve the power and representativeness of the study sample in order to help further develop and refine candidate scales and items.

While a sample of four participants is considered acceptable for content validity work by the COSMIN guidelines, a larger sample size may have allowed for the development of further themes and perspectives on the items and evidence to support or refute larger scale changes, such as within the Perceived Controllability domain. However the limiting of the sample to four participants allowed for the initial exploration of the content validity of the items and reduction in item numbers required at this stage of development and will allow for further content validation work to be conducted in a larger sample in shorter interviews and reduced participant burden.

While development of robust psychometric measurement tools is critically important to clinical mental health, the scale development is complex and resource intensive. The depth of knowledge required in multiple fields (the construct of interest, psychometrics and qualitative research) presents a major barrier to development. Psychometrics is a complex field and despite great efforts, there remains no clear "best path" methodology. As such, scale developers are presented with multiple, often conflicting methodological choices that are debated, often without agreement on terminology. This is made further difficult by disagreements over theory as to which approaches are valid or useful. This project has situated its choices within Classical Test Theory (CTT), the most common methodology used in clinical health scale design. Item Response Theory (IRT) presents an alternate approach, focusing on item level information rather than test level information with "harder" assumptions than those required by CTT. Despite IRT becoming the favoured theoretical approach in psychometrics, CTT remains the most used methodology within clinical health scale development, where constructs are often complex and multifaceted and the large sample sizes required in IRT are more difficult to recruit. This study has attempted to navigate an array of guidance on scale design (Streiner et al., 2024, Mokkink et al., 2024, De Vet et al., 2011, Boateng et al., 2018) to produce a scale using CTT that can meet such standards. This presents a strength, but it must be acknowledged that another developer may have evaluated that guidance and arrived at different methodological conclusions or explored the development via another psychometric tradition.

5. Acknowledgements

Funding was received from the University of Glasgow for transcription services and participant payments. No other source of funding was received for the project.

6. References

Alderson-Day, B. & Ward, T. 2022. The Sound of Fear. *Voices in Psychosis: Interdisciplinary Perspectives [Internet].* Oxford University Press.

Almeida, S. N., Elliott, R., Silva, E. R. & Sales, C. M. 2019. Fear of cancer recurrence: a qualitative systematic review and meta-synthesis of patients' experiences. *Clinical psychology review,* 68**,** 13-24.

Ascher-Svanum, H., Zhu, B., Faries, D. E., Salkever, D., Slade, E. P., Peng, X. & Conley, R. R. 2010. The cost of relapse and the predictors of relapse in the treatment of schizophrenia. *BMC psychiatry,* 10**,** 2.

Baker, C. 1995. The development of the self-care ability to detect early signs of relapse among individuals who have schizophrenia. *Arch Psychiatr Nurs,* 9**,** 261-8.

Bassett, M., Sperlinger, D. & Freeman, D. 2009. Fear of madness and persecutory delusions: Preliminary investigation of a new scale. *Psychosis,* 1**,** 39-50.

Birchwood, M., Mason, R., MacMillan, F. and Healy, J., 1993. Depression, demoralization and control over psychotic illness: a comparison of depressed and non-depressed patients with a chronic psychosis. Psychological medicine, 23(2), pp.387-395.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R. & Young, S. L. 2018. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Frontiers in Public Health, 6, 149.

Boyd, T. & Gumley, A. 2007. An experiential perspective on persecutory paranoia: A grounded theory construction. *Psychology and Psychotherapy: Theory, Research and Practice,* 80**,** 1-22.

Brookmann, L. 2020. *A meta-synthesis of service user perspectives of shared decision-making and a qualitative investigation of lived experiences of fear of recurrence in relation to psychosis.* D Clin Psy, University of Glasgow.

Charmaz, K. 2006. *Constructing grounded theory: A practical guide through qualitative analysis*, sage.

Chepp, V. & Gray, C. 2014. Foundations and new directions. *Cognitive interviewing methodology***,** 7-14.

Cicchetti, D. V. & Sparrow, S. A. 1981. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American journal of mental deficiency,* 86**,** 127-137.

Clark, L. A. & Watson, D. 1995. Constructing Validity: Basic Issues in Objective Scale Development. *Psychological assessment,* 7**,** 309-319.

Dalkey, N. & Helmer, O. 1963. An experimental application of the Delphi method to the use of experts. *Management science,* 9**,** 458-467.

Davidsen, K. A. & Rosenbaum, B. 2012. Fear of annihilation in subjects at risk of psychosis: A pilot study. *Psychosis,* 4**,** 149-160.

De Vet, H. C. W., Terwee, C. B., Mokkink, L. B. & Knol, D. L. 2011. *Measurement in Medicine: A Practical Guide*, Cambridge University Press.

Emsley, R., Chiliza, B., Asmal, L. & Harvey, B. H. 2013. The nature of relapse in schizophrenia. *BMC psychiatry,* 13**,** 50.

Estrade, A., Salazar De Pablo, G., Zanotti, A., Wood, S., Fisher, H. L. & Fusar-Poli, P. 2022. Public health primary prevention implemented by clinical high-risk services for psychosis. *Translational Psychiatry,* 12**,** 43.

Fisher, R. P. & Geiselman, R. E. 1992. *Memory enhancing techniques for investigative interviewing: The cognitive interview*, Charles C Thomas Publisher.
Fleiss, J. L. 1981. Balanced incomplete block designs for inter-rater reliability studies. *Applied psychological measurement,* 5**,** 105-112.
Frosch, J. 1983. *The Psychotic Process*, International Universities Press.
Fusar-Poli, P., Estradé, A., Stanghellini, G., Venables, J., Onwumere, J., Messas, G., Gilardi, L., Nelson, B., Patel, V. & Bonoldi, I. 2022. The lived experience of psychosis: a bottom-up review co-written by experts by experience and academics. *World Psychiatry,* 21**,** 168-188.
Ghai, V., Subramanian, V., Jan, H. & Doumouchtsis, S. K. 2022. A systematic review highlighting poor quality of evidence for content validity of quality of life instruments in female chronic pelvic pain. *Journal of Clinical Epidemiology,* 149**,** 1-11.
Gross, C. T. & Canteras, N. S. 2012. The many paths to fear. *Nature Reviews Neuroscience,* 13**,** 651-658.
Gumley, A., Bradstreet, S., Ainsworth, J., Allan, S., Alvarez-Jimenez, M., Beattie, L., Bell, I., Birchwood, M., Briggs, A. & Bucci, S. 2020. Early signs monitoring to prevent relapse in psychosis and promote well-being, engagement, and recovery: protocol for a feasibility cluster randomized controlled trial harnessing mobile phone technology blended with peer support. *JMIR research protocols,* 9**,** e15058.
Gumley, A., O'grady, M., Mcnay, L., Reilly, J., Power, K. & Norrie, J. 2003. Early intervention for relapse in schizophrenia: results of a 12-month randomized controlled trial of cognitive behavioural therapy. *Psychological Medicine,* 33**,** 419-431.
Gumley, A., O'grady, M., Power, K. & Schwannauer, M. 2004. Negative beliefs about self and illness: a comparison of individuals with psychosis with or without comorbid social anxiety disorder. *Australian & New Zealand Journal of Psychiatry,* 38**,** 960-964.
Gumley, A. & Schwannauer, M. 2006. *Staying well after psychosis: A cognitive interpersonal approach to recovery and relapse prevention*, John Wiley & Sons.
Gupta, V., Syrett, S., Gumley, A., Hardy, A., Ward, T. & Allan, S. 2024. Identifying problems and solutions to qualitative research privileging “rich” data within psychosis research.
Hubley, A. M., Zhu, S. M., Sasaki, A. & Gadermann, A. M. 2014. Synthesis of Validation Practices in Two Assessment Journals: Psychological Assessment and the European Journal of Psychological Assessment. *In:* ZUMBO, B. D. & CHAN, E. K. H. (eds.) *Validity and Validation in Social, Behavioral, and Health Sciences.* Cham: Springer International Publishing.
Jabine T., S., M., Tanur J., & Tourangeau, R. 1984. *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines,* Washington, DC, The National Academies Press.
Jäger, M., Riedel, M., Messer, T., Laux, G., Pfeiffer, H., Naber, D., Schmidt, L. G., Gaebel, W., Huff, W. & Heuser, I. 2007. Psychopathological characteristics and treatment response of first episode compared with multiple episode schizophrenic disorders. *European Archives of Psychiatry and Clinical Neuroscience,* 257**,** 47-53.
Johansen, K. K., Hounsgaard, L., Frandsen, T. F., Fluttert, F. A. & Hansen, J. P. 2020. Relapse prevention in ambulant mental health care tailored to patients with schizophrenia or bipolar disorder. *Journal of Psychiatric and Mental Health Nursing*.
Karon, B. P. 1992. The fear of understanding schizophrenia. *Psychoanalytic Psychology,* 9**,** 191.

Knafl, K., Deatrick, J., Gallo, A., Holcombe, G., Bakitas, M., Dixon, J. & Grey, M. 2007. The analysis and interpretation of cognitive interviews for instrument development. *Research in Nursing & Health,* 30**,** 224-234.

Kotzeva, A., Mittal, D., Desai, S., Judge, D. & Samanta, K. 2023. Socioeconomic burden of schizophrenia: a targeted literature review of types of costs and associated drivers across 10 countries. *Journal of medical economics,* 26**,** 70-83.

Lally, J., Ajnakina, O., Stubbs, B., Cullinane, M., Murphy, K. C., Gaughran, F. & Murray, R. M. 2017. Remission and recovery from first-episode psychosis in adults: systematic review and meta-analysis of long-term outcome studies. *The British Journal of Psychiatry,* 211**,** 350-358.

Lebel, S., Mutsaers, B., Tomei, C., Leclair, C. S., Jones, G., Petricone-Westwood, D., Rutkowski, N., Ta, V., Trudel, G., Laflamme, S. Z., Lavigne, A.-A. & Dinkel, A. 2020. Health anxiety and illness-related fears across diverse chronic illnesses: A systematic review on conceptualization, measurement, prevalence, course, and correlates. *PLOS ONE,* 15**,** e0234124.

Marshall, M., Lockwood, A., Bradley, C., Adams, C., Joy, C. & Fenton, M. 2000. Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. *The British Journal of Psychiatry,* 176**,** 249-252.

Mcclimans, L., Browne, J. & Cano, S. 2017. Clinical outcome measurement: Models, theory, psychometrics and practice. *Studies in History and Philosophy of Science Part A,* 65**,** 67-73.

Mcneish, D. 2024. Practical Implications of Sum Scores Being Psychometrics' Greatest Accomplishment. *Psychometrika,* 89**,** 1148-1169.

Miles, M. B. & Huberman, A. M. 1994. *Qualitative data analysis: An expanded sourcebook*, sage.

Miller, K., Chepp, V., Willson, S. & Padilla, J. L. 2014. *Cognitive Interviewing Methodology,* Newark, UNITED STATES, John Wiley & Sons, Incorporated.

Miller, K., Fitzgerald, R., Padilla, J. L., Willson, S., Widdop, S., Caspar, R., Dimov, M., Gray, M., Nunes, C., Prüfer, P., Schöbi, N. & Schoua-Glusberg, A. 2011a. Design and Analysis of Cognitive Interviews for Comparative Multinational Testing. *Field methods,* 23**,** 379-396.

Miller, K., Mont, D., Maitland, A., Altman, B. & Madans, J. 2011b. Results of a cross-national structured cognitive interviewing protocol to test measures of disability. *Quality & Quantity,* 45**,** 801-815.

Mokkink, L. B., Elsman, E. B. M. & Terwee, C. B. 2024. COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. *Quality of Life Research,* 33**,** 2929-2939.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M. & De Vet, H. C. 2010. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of clinical epidemiology,* 63**,** 737-745.

Morgado, F. F. R., Meireles, J. F. F., Neves, C. M., Amaral, A. C. S. & Ferreira, M. E. C. 2018. Scale development: ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica,* 30**,** 3.

Nagle, S., Cook, J. V. & Polatajko, H. J. 2002. I'm Doing As Much As I Can: Occupational Choices of Persons with a Severe and Persistent Mental Illness. *Journal of Occupational Science,* 9**,** 72-81.

Padilla García, J. L. & Benítez Baena, I. 2014. Validity evidence based on response processes.
Polit, D. F., Beck, C. T. & Owen, S. V. 2007. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health,* 30**,** 459-467.
Ridenour, J. M., Hamm, J. A., Leonhardt, B. L., Buck, B. & Lysaker, J. T. 2025. Facing Fear in the Psychotherapy for Psychosis. *Journal of Contemporary Psychotherapy*.
Ritunnano, R., Kleinman, J., Oshodi, D. W., Michail, M., Nelson, B., Humpston, C. S. & Broome, M. R. 2022. Subjective experience and meaning of delusions in psychosis: a systematic review and qualitative evidence synthesis. *The Lancet Psychiatry,* 9**,** 458-476.
Robinson, D., Woerner, M. G., Alvir, J. M. J., Bilder, R., Goldman, R., Geisler, S., Koreen, A., Sheitman, B., Chakos, M., Mayerhoff, D. & Lieberman, J. A. 1999. Predictors of Relapse Following Response From a First Episode of Schizophrenia or Schizoaffective Disorder. Archives of General Psychiatry, 56, 241-247.
Sandhu, A., Ives, J., Birchwood, M. & Upthegrove, R. 2013. The subjective experience and phenomenology of depression following first episode psychosis: A qualitative study using photo-elicitation. Journal of Affective Disorders, 149, 166-174.
Seed, T., Fox, J. R. & Berry, K. 2016. The experience of involuntary detention in acute psychiatric care. A review and synthesis of qualitative studies. Int J Nurs Stud, 61, 82-94.
Shelley, M. and Pakenham, K.I., 2004. External health locus of control and general self-efficacy: Moderators of emotional distress among university students. Australian journal of psychology, 56(3), pp.191-199.
Simard, S., Thewes, B., Humphris, G., Dixon, M., Hayden, C., Mireska ndari, S. & Ozakinci, G. 2013. Fear of cancer recurrence in adult cancer survivors: a systematic review of quantitative studies. Journal of Cancer Survivorship, 7, 300-322.
Streiner, D. L., Norman, G. R. & Cairney, J. 2024. *Health measurement scales: a practical guide to their development and use,* Oxford, Oxford University Press.
Sudman, S., Bradburn, N. M. & Schwarz, N. 1996. *Thinking about answers: The application of cognitive processes to survey methodology*, Jossey-Bass.
Swan, K., Speyer, R., Scharitzer, M., Farneti, D., Brown, T., Woisard, V. & Cordier, R. 2023. Measuring what matters in healthcare: a practical guide to psychometric principles and instrument development. *Frontiers in Psychology,* 14**,** 1225850.
Sweeney, A., Gillard, S., Wykes, T. & Rose, D. 2015. The role of fear in mental health service users' experiences: a qualitative exploration. *Social psychiatry and psychiatric epidemiology,* 50**,** 1079-1087.
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**,** 141-153.
Taylor, P. J., Hutton, P. & Wood, L. 2015. Are people at risk of psychosis also at risk of suicide and self-harm? A systematic review and meta-analysis. *Psychological Medicine,* 45**,** 911-926.
Tourangeau, R. 1984. Cognitive sciences and survey methods. *Cognitive aspects of survey methodology: Building a bridge between disciplines,* 15**,** 73-100.
Tourangeau, R., Rips, L. J. & Rasinski, K. 2000. The psychology of survey response.

Wiersma, D., Nienhuis, F. J., Slooff, C. J. & Giel, R. 1998. Natural Course of Schizophrenic Disorders: A 15-Year Followup of a Dutch Incidence Cohort. *Schizophrenia Bulletin,* 24**,** 75-85.

Willis, G. B. 2015. *Analysis of the Cognitive Interview in Questionnaire Design,* Oxford, UNITED STATES, Oxford University Press, Incorporated.

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**,** 1305-1318.

Appendices

Appendix 1.1 COSMIN Definitions of Psychometric Terminology

https://osf.io/gnkxw/files/osfstorage/689629abb217e80fb31b5417

Appendix 1.2 Example Search Strategy

Search terms were adapted according to the database searched. An example of the PsychINFO search is provided below.

#	Query	Results
1	fear of recurrence scale.ti,ab,tm.	27
2	FoRSe.ti,ab,tm	32
3	1 or 2	58
4	limit 3 to english language	39

Appendix 1.3 - Systematic Review Study Characteristics

https://osf.io/gnkxw/files/osfstorage/68961b2fb220775ab4f62e62

Appendix 2.1: Study Check List

The COSMIN reporting guideline for studies on measurement properties of PROMs

Item Code	Item name	Item description	Page Number
Report section: Title			
T1	Title	Identify the report as a study of one or more measurement properties of a specific PROM to measure a specified construct in a specified population.	50
Report section: Abstract			
A1	Objectives	Provide the specific objective(s) of the research, specifying (1) the name (and version, if relevant), and construct(s) of the PROM, (2) the measurement properties being evaluated, and (3) relevant study characteristics.	53
A2	Design	Specify (details of the) study design used to evaluate the measurement properties.	53
A3	Methods	Specify the methods for evaluating each measurement property.	53
A4	Results	Provide the main results for all measurement properties evaluated.	53
A5	Discussion/Conclusions	Provide a brief statement of the implications of the findings in the context of existing evidence on the PROM.	53
Report section: Introduction			
I1	PROM	Specify the name and, if relevant, the version, and construct(s) of the PROM.	57
I2	Target population & context of use	Specify the target population and context of use that the PROM was designed for.	57
I3	State of knowledge & Rationale	knowledge & Rationale Provide a description of the current scientific knowledge (what is known and not known) regarding the measurement properties of the PROM. Explain why the new study is necessary. Provide citations for the original development paper(s).	55-61
I4	Objectives	Provide the specific objective(s) of the research, specifying (1) the name (and version, if relevant) of the PROM, (2) the measurement properties being evaluated, and (3) relevant study sample characteristics.	61
Report section: General Methods			

GM1	Study design	Specify (details of the) study design used to evaluate the measurement properties.	61-67
GM2	Participants	Specify how the study participants were selected. Specify the inclusion and exclusion criteria	62,64
GM3	PROM details	Provide details about the original version of the PROM as well as of the PROM version being studied, specify the conceptual framework (reflective/formative model), details on the structure (the number of items and subscales), the language, response scale, recall period, direction of scoring, and scoring algorithm of the PROM. Specify how the PROM was administered (e.g., in what setting, mode of administration (e.g. paper, electronic) what instructions were given), including the country in which it is administered	57, 67-70
GM4	Additional data collection	Describe why and how other data was collected (e.g., construct and measurement properties of the comparator instruments, characteristics of groups being compared, and rationale for choosing groups), including mode of administration (e.g., paper, electronic).	N/A
GM5	Time points procedures	Provide all time points of all measurements.	N/A
GM6	Justification for sample size	Provide a rationale for the sample size for all measurement properties analyses (including subgroups).	62-64
GM7	Statistical analyses	Describe the statistical analyses corresponding to all objectives (see measurement properties specific boxes). Describe the criteria for good measurement properties. Name the statistical package used and the version.	63-66
GM8	Missing data	Describe approaches for dealing with missing data.	N/A
GM9	Unplanned analysis	Specify analyses that were unplanned and their rationale.	N/A
Report section: General results			
GR1	Participant characteristics	Provide study participants' characteristics, specified per subgroup if applicable.	71, 76

GR2	Sample size	Provide the total number of participants included in the study and the sample size for each analysis.	71, 76
GR3	Missing data	Provide amount of (proportion or count) and reasons for missing data for each analysis for the PROM, and for any analyses of other outcome measurement instruments.	N/A
GR4	Results	Describe the results corresponding to all objectives	70-75, 77-95
Report section: Discussion/conclusions			
DC1	Measurement property evidence	Provide the main findings and if each measurement property is sufficient or insufficient and why.	95-99
DC2	Practical relevance	Discuss the practical relevance of the findings in terms of recommendations for (not) using the PROM.	N/A
DC3	Strengths and limitations	Discuss strengths and limitations of each study. For example, discuss if there were any potential biases in the study that could have impacted the results.	101-103
DC4	Generalizability	Discuss generalizability of the results. For example, discuss whether the results could be generalized to other populations given the sample studied.	101-103
DC5	Instrument changes	Discuss what modifications are needed to the existing PROM.	70-103
DC6	Future research	Describe new research questions or hypotheses generated from these findings, and provide/describe the research needed to answer those questions.	100-103
DC7	Conclusions	Provide the overall conclusions for the use of the PROM.	N/A
Report section: Other information			
O1	Conflict of interest	State any conflict of interest you may have related to the PROM. This may include any involvement in the development of the PROM or any commercial funding or profit.	103

Appendix 2.2: University of Glasgow MVLS Ethics Committee Approval Letter

University of Glasgow MVLS Ethics Committee Approval letter removed due to confidentiality issues.

Appendix 2.3: University of Glasgow MVLS Ethics Committee Amendment Approval

University of Glasgow MVLS Ethics Committee Amendment Approval removed due to confidentiality issues.

Appendix 2.4: Amendment 3

https://osf.io/gnkxw/files/osfstorage/6895c267b2b1b840fe6b0d18

Appendix 2.5: Amendment 3 Approval

Amendment 3 Approval removed due to confidentiality issues.

Appendix 2.6: Amendment 4 Approval

Amendment 4 Approval removed due to confidentiality issues.

Appendix 2.7: Amendment 5 Approval

Amendment 5 Approval removed due to confidentiality issues.

Appendix 2.8: Amendment 6 Approval

Amendment 6 Approval removed due to confidentiality issues.

Appendix 2.9: Study One PIS

https://osf.io/gnkxw/files/osfstorage/686507d8dd6d7cc17c9fef50

Appendix 2.10: Study One Consent

https://osf.io/gnkxw/files/osfstorage/686507d45684dfd88a49f0b6

Appendix 2.11: Study Two PIS

https://osf.io/gnkxw/files/osfstorage/6865074744661ef24803ab82

Appendix 2.12: Study Two Consent

https://osf.io/gnkxw/files/osfstorage/68650753dba2e48fb849efbf

Appendix 2.13: Cognitive interviewing Topic Guide

https://osf.io/gnkxw/files/osfstorage/68650748c459b599d69fef6c

Appendix 2.14: Initial Development Item Pool

https://osf.io/gnkxw/files/osfstorage/6895de505fd2716dbacd912a

Appendix 2.15: Reference Group

https://osf.io/gnkxw/files/osfstorage/6895dc910ad619924db069c4

Appendix 2.16: Feedback Summary

https://osf.io/gnkxw/files/osfstorage/6895dc9c5fd2716dbacd90ae

Appendix 2.17: Pilot FoRse-R

https://osf.io/gnkxw/files/osfstorage/6895deba8b4f559152cd8ecd

Appendix 2.18: Item-by-item Review

https://osf.io/gnkxw/files/osfstorage/6895f5253b6e5a6987f62d09

Appendix 2.19: Potential Items for Revised PC Scale

https://osf.io/gnkxw/files/osfstorage/6895de0e766d57a512b06c7e