Using the Delphi method in psychosis research with experts by experience: A systematic review and development of a consensus definition of relapse

Alice Thomson, MA (hons) MSc PhD

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

December 2025

Table of Contents

List of Tables ***4***

List of Figures ***5***

Acknowledgements ***6***

Chapter 1: Systematic Review ***7***

Abstract **9**

Introduction **10**

Aims and Research Questions 11

Methods **12**

Scope of Review 12

Search Strategy 12

Search Terms 13

Screening 14

Data Extraction and Synthesis 14

Assessment of Study Quality 15

Results **15**

Screening and Selection 15

Engagement of EbE 23

Quality Appraisal 26

Transparency of Reporting 26

Discussion **27**

Summary 27

Strengths and Limitations 29

Recommendations 30

Wider Context 30

Conclusion 31

References **32**

Chapter 2: Major Research Project ***35***

Plain Language Summary **37**

Abstract **39**

Introduction **40**

Aims and Research Questions 42

Researcher Reflexivity Statement and Epistemological Position 42

Method **42**

Delphi Technique 42

Participants 43

Sample Size 43

Participant Identification and Recruitment 43

Materials 44

Questionnaire Development 44

Public and Patient Involvement 45

Pilot 45
Stage 1 Questionnaire 45
Data Analysis 46
Stage 2 Questionnaire Design and Administration 47
Stage 2 Data Analysis 47
Stage 3 Questionnaire Design and Administration 48
Feedback 48
Consensus 48
Ethical Considerations 48
Results 49
Participant Characteristics 49
Overview of Items per Delphi Round 50
Stage 1 Results 51
Stage 2 and 3 Results 52
Checklist for conceptualising relapse 53
Discussion 54
Strengths 56
Limitations 57
Next Steps 58
Conclusion 59
References 60
Appendix 1.1: PRISMA 2020 Checklist 64
Appendix 1.2: Spranger et al. (2022) Checklist 66
Appendix 1.3: Data Extraction Template 67
Appendix 1.4: Quality Appraisal 69
Appendix 1.5: Transparency Ratings 71
Appendix 2.1: Study Proposal 72
Appendix 2.2: Delphistar Checklist 73
Appendix 2.3: Delphi Surveys 77
Appendix 2.4: Data Analysis Plan 78
Appendix 2.5: Results for Delphi Rounds 79
Appendix 2.6: Ethical Approval 80
Appendix 2.7: Checklist for defining relapse 81

List of Tables

Table 1: The spider tool applied to the research question 12
Table 2: Characteristics of reviewed studies 22
Table 3: No. of EbE as a percentage of the total sample per Delphi round 25
Table 4: Quality ratings of studies included in the review 26
Table 5: Coding framework for open ended question responses 47
Table 6: Participant Characteristics 50
Table 7: General terms endorsed by 80% or more of participants in stages 2 & 3 52
Table 8: Domain specific items endorsed by 80% or more of participants in stages 2 & 3 53

List of Figures

Figure 1: PRISMA diagram 16
Figure 2: Flowchart of items in each round of the Delphi survey 51

Acknowledgements

I would like to thank everyone who participated in this research, this project would not have been possible without you and I'm extremely grateful to you for giving your time and energy towards taking part. Thank you very much to my supervisors Andrew and Stephanie. I feel exceptionally fortunate to have been on the receiving end of your generous sharing of knowledge, patience, wisdom, and compassion during the past four years. In addition to excellent research supervision, you've been inspiring role models for which I am truly grateful. Thank you to my second reviewer Christianne for making time to help me. Thank you for my amazing team of supporters, my Mum, Dad, Jude, Karen, Kate, Eve and Iona who made sure that my children were never short of love and attention throughout this process, as well as supporting me as an individual. Thank you to Rowan, Robin and Izzy for being yourselves and bringing the joy.

Chapter 1: Systematic Review

A Systematic Review on the use of Delphi Methodology in Psychosis Research involving Experts by Experience as Participants

Prepared in accordance with the author requirements for Clinical Psychology Review
https://www.sciencedirect.com/journal/clinical-psychology-review/publish/guide-for-authors

A note on Language

Throughout this thesis, the term 'expert by experience' (EbE) is used to identify people with lived experience of psychosis. This term was chosen as it recognises and values the experiential knowledge of people with first-hand experience of psychosis. However, we recognise that there is debate about the language used and this term will not fully capture the identities of everyone with lived experience of psychosis (McLaughlin, 2009).

Abstract

Background

The Delphi Method aims to build consensus among a group of experts. The method is increasingly being used in mental health research with multiple stakeholders including people with lived experience of mental health difficulties as experts by experience (EbE). Recent reviews have highlighted variability in the use of the Delphi method in mental health research, but no review has looked at psychosis research specifically. The aim of the current review was to investigate use of the Delphi method in psychosis research including participants with expertise through lived experience. This is considered particularly important as people with lived experience represent a marginalised group in mental health research.

Method

A systematic review of the literature was carried out in July 2025 using the databases APA Psycinfo, Medline (PubMed), and Embase. English language articles using the Delphi method including people with lived experience in the sample were reviewed. Data extraction and narrative synthesis of engagement of lived experience stakeholders, quality appraisal and transparency of reporting was conducted.

Results

Sixteen papers were included in the review. There was variation in the methodological approaches used and in transparency of reporting. Within the reviewed papers, there was under-reporting in relation to epistemological positioning, preliminary assumptions, and retention rates of stakeholders across Delphi rounds. It was difficult to gain an accurate representation of whether lived experience perspectives were adequately represented in consensus due to inconsistencies in reporting and lack of a reporting tool which takes the specific considerations of engaging EbE into account.

Conclusion

This review demonstrated that the Delphi method has been used as a consensus building technique in psychosis research with multiple stakeholders including people with expertise through lived experience to address a range of research questions. Methodological decisions and results need to be reported transparently to ensure that the views of people with lived experience are represented. This paper includes key recommendations aimed at enhancing methodological quality and transparency in this area.

Introduction

The Delphi method is a technique which aims to answer a research question through identifying a consensus view across a panel of experts (Barret & Heale, 2020; Hasson et al., 2000). The Delphi method has been used for various purposes in mental health research such as defining concepts, identifying collective values, making predictions and estimations (Jorm, 2015). The method has been used with patients, clinicians, carers and researchers. While some studies focus on a particular participant group, others include a range of participants seeking to gain consensus within a single sample. A review on the use of Delphi methodology in mental health research highlighted diversity in the way the method can be applied to suit different research questions (Jorm, 2015).

Within psychosis research, it has been found that different stakeholders such as clinicians and people with lived experience may differ in their opinions of specific concepts (Allan, 2019). A strength of the Delphi method is the ability to encompass a range of diverse viewpoints and combine these in a manner which has practical application. While there are many potential benefits to employing the Delphi method in psychosis research, these are dependent upon rigorous and transparent use of the method. Research using the Delphi method in other clinical areas has highlighted variability in the quality of study and transparency of reporting from Delphi studies (Junger et al., 2017). Poor quality studies and lack of transparent reporting limit the conclusions that can be drawn from such studies. Where multiple stakeholder groups are involved, it follows that a potential risk could be the development of consensus reported to reflect the views of varying stakeholder groups e.g. people with lived experience, which does not, in fact, adequately represent key stakeholders in a meaningful way. Recent critique of participatory approaches in mental health research have highlighted that the risk of perpetuating pre-existing epistemic injustice i.e. undermining people with lived experience's perspectives by assimilating their views into preconceived researcher-driven agendas (Veldmeijer & Van Os, 2025). This is termed 'participatory assimilation' and arguably, the Delphi-method may be particularly vulnerable to participatory assimilation due to it's emphasis on consensus building (as opposed to more explorative approaches) (Veldmeijer & Van Os, 2025).

A recent critique of participatory mental health research further highlights the risks of superficial or 'tokenistic' involvement of people with lived/living experience in mental health research (Colder Carras et al., 2022). These include potential dilution, minimisation or even undermining of the marginalised groups original goals, and thus the potential to perpetuate further marginalisation (Colder Carras et al.,

2022). The Delphi method provides opportunities for meaningful involvement for multiple stakeholders including those with lived experience and the condition of anonymity in the Delphi method can help to address power imbalances between stakeholders. However, there are also inherent risks such as potential misrepresentation of stakeholders' views, and this may disproportionately affect different stakeholder groups. Psychosis research sits within a real-world context in which there are inherent power imbalances between stakeholders. For example, clinicians may hold statutory powers over service users in their care, can detain individuals in hospital, and administer treatment against their will which can, at times, involve physical restraint. While such treatment is provided on the basis of least restriction and must be deemed to be in the best interest of the service user, the power differentials are undeniable and have the potential to cause physical and emotional harm to service users (Butterworth, Wood & Rowe, 2022). This is the clinical context in which psychosis research is situated. It follows that research participants come with assumptions relating to their role in the research processes that is influenced by past experiences e.g. in relation to how valued they feel their views are. Further, mirroring the dynamics of clinical services, research tends to be carried out *by* clinical researchers *to* service-users. While the engagement of multiple-stakeholders offers opportunities, there needs to be consideration of the inherent power imbalances and explicit steps taken to mitigate these within the research design. Where this is not possible, these should be acknowledged and transparently reported upon.

A Delphi study can provide a method to address some of these challenges. Firstly, the anonymity of the Delphi process may limit the capacity of traditionally powerful groups, such as clinicians, to steer discussions or shape consensus through professional status. This may create safer conditions for service users to express views that might otherwise be held back. Secondly, the iterative nature of Delphi rounds allows participants to reflect on group feedback independently and without pressure to respond in "real time". This can empower those whose opinions might diverge from majority or professional viewpoints, and it gives space for marginalised perspectives to be captured. Thirdly, Delphi studies may allow for equal weighting of contributions if researchers aiming to challenge epistemic injustice opt for this. If responses are considered at the level of concepts rather than attributed to individuals or stakeholder groups, then this shifts focus away from traditional hierarchies. However, the success of this depends on how the method has been implemented which speaks to the need for this review.

Aims and Research Questions

The current systematic review aimed to provide a synthesis of existing Delphi research, identify strengths and limitations in the existing psychosis literature, which has also included experts by experience in

psychosis, and highlight future directions for development. The current review aimed to answer the following research questions:

Research Question

What are the key considerations when using Delphi methodology in psychosis research which includes people who have expertise through lived experience?

In relation to:

1) meaningful involvement of people with expertise by experience (EbE) i.e. valuing this evidence at the same level as knowledge gained through other means

2) transparency of reporting.

Methods

The study was funded by NHS Scotland as part of the lead researcher's doctoral studies. The authors declare no competing interests.

Scope of Review

The SPIDER (Sample, Phenomenon of Interest, Design, Evaluation, Research type) tool was used to structure the research question and search strategy in the current review to explore methodological differences, strengths and weaknesses within the studies reviewed.

Sample	People with lived experience of psychosis (EbE)
Phenomenon of Interest	Delphi methodology
Design	Published peer-reviewed literature using the Delphi method
Evaluation	Characteristics, quality, adherence to Delphi reporting guidelines
Research Type	Qualitative, quantitative and mixed-methods peer reviewed studies

Table 1: The spider tool applied to the research question

Search Strategy

The PRISMA 2020 checklist was followed to ensure rigour and transparency in the search strategy (Appendix 1.1). The review protocol was registered on the open science framework (OSF) on 01/08/2025 (DOI **10.17605/OSF.IO/K79CT).** Minor changes to the protocol were made following pre-registration. These were:

1. Inclusion of a quality appraisal tool

2. EbE involvement was not rated using Arnstein's (1969) ladder of participation due to insufficient information for rating.

The following databases were used in the search: APA Psycinfo, Medline (PubMed), and Embase.

Search Terms

Consultation with a subject librarian was sought to help identify and refine appropriate search terms. Databases were searched for titles, abstracts, keywords and subject headings containing the search terms.

Delphi/

AND

Psychosis/ or schizophrenia/ or psychotic disorders/

AND

Patients/ or service user* or lived experience* or expert by experience*

The following inclusion criteria were used in the study.

Inclusion criteria

1. Psychosis as the main phenomena of interest in the study
2. Paper employs the Delphi Method
3. Participants in the Delphi have lived experience of psychosis i.e. EbE are included in the participant sample (though may not make up the entirety of the sample)
4. Written in English with full text available

Exclusion criteria

1. Grey literature, literature reviews
2. Participants are *exclusively* identified as experts through other avenues of gaining expertise (than living experience of psychosis) e.g. studies including only clinicians or researchers in the sample

According to the above criteria, research papers which used the Delphi method as part of the study, but did not include experts by experience in the Delphi process itself, were excluded.

Screening

Prior to screening, de-duplication was carried out using the software 'Rayyan'. Potential duplicates were identified by the software and manually screened by the lead researcher and duplicates were manually removed. Titles and abstracts were screened to determine whether they were relevant to the scope of the review. An independent reviewer independently screened a random 10% of articles at this stage. Full texts were retrieved and screened against eligibility criteria. An independent reviewer (CL) screened a further 10% at full text stage.

Data Extraction and Synthesis

Data were extracted and synthesised in line with the research questions. A quality appraisal checklist was applied to gain an overview of study quality (Khodyakov et al., 2023). A further checklist relating to transparency of reporting was used to gain a more detailed understanding of the reviewed studies (Spranger et al., 2022). Additional data relating to the level of engagement of participants with lived experience of psychosis was extracted to assess the extent to which people with lived experience were involved and whether this was transparently reported upon.

Transparency of Reporting

Data were extracted and tabulated according to the factors identified in Appendix 1.2 to assess whether the studies adhere to reporting guidelines and appraise the study quality in relation to the information reported (Spranger et al., 2022). While the Spranger et al. (2022) checklist is intended as a reporting guideline, the headings provided a framework to consider the decisions made by researchers, where reported. Appendix 1.3 contains the data extraction template used in the study which is based on the Spranger et al. (2022) checklist. All available information, including supplementary information where available was extracted.

Level of Engagement of Participants with lived/living experience of psychosis

To further assess level of engagement of experts by experience in the Delphi process, the percentage of experts by experience in the total sample per Delphi round was calculated and tabulated for each of the reviewed studies.

Assessment of Study Quality

A quality appraisal tool, the Delphi Critical Appraisal tool (DCAT) a 16-item checklist developed by the RAND Corporation in 2023 was applied to all reviewed papers (Khodyakov et al., 2023). The DCAT contains 4 “core” items and 12 “additional” items which are aimed at assessing study quality. The tool was developed by the RAND Corporation which originally developed the Delphi method in the 1950s (Khodyakov et al., 2023). The lead researcher applied the checklist to each paper reviewed to gain a measure of study quality in relation to checklist items.

Results

Screening and Selection

In total, the search carried out on 31/7/2025 yielded 263 references from the databases (APA PsycINFO 69; Medline 93; Embase101). Of these, 103 duplicates were removed. The lead researcher completed title-abstract screening and a further 134 references were removed. An independent reviewer (CL) independently screened a random 10% of articles including eligible and ineligible studies (n=16) at this stage with a 100% agreement. The remaining 26 studies were read in full and 16 were included in the final sample. An independent reviewer (CL) screened a further random 10% at this stage and agreement remained at 100%. Figure 1 provides a diagrammatic representation of the screening and selection process.

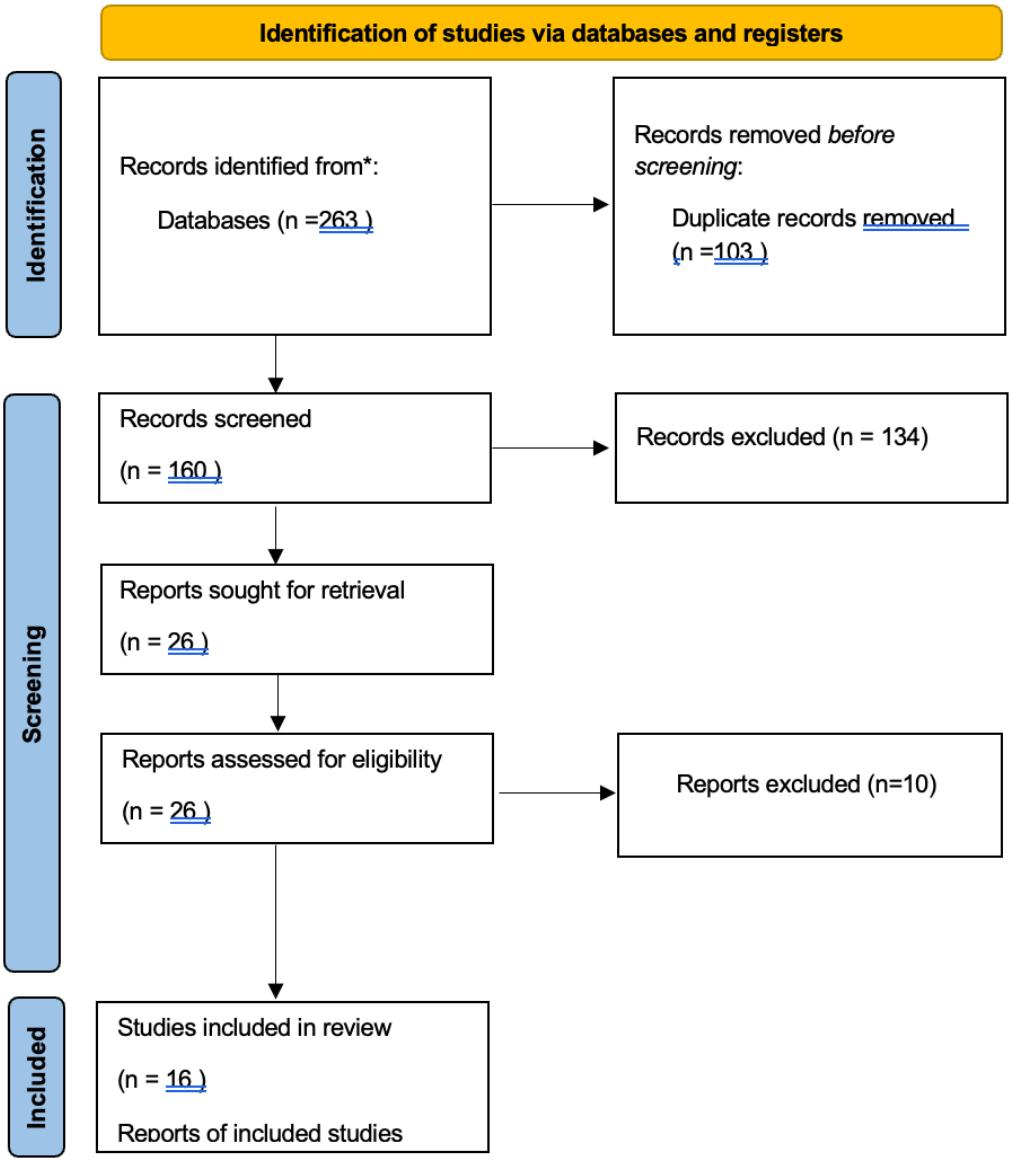

Figure 1: PRISMA diagram

Characteristics of Studies included in the review

Table 2 summarises the characteristics of the studies included in the review. The majority of studies (14/16) included more than one stakeholder, and the method was used for various purposes including development of outcome measures and guidelines, identifying priorities relating to care and defining

concepts (Table 2). Studies were carried out in a range of geographical locations and the method was also used to involve international stakeholders (Table 2).

Focus and purpose of studies

Focus and purpose of the studies was highly varied. Five studies aimed to develop mental health first aid guidelines in different geographical regions (Agrest et al., 2024; Cottrill et al., 2021; Langlands et al 2008; Li et al., 2020; Requena et al., 2024). A further six aimed to identify outcome measures for different purposes including early intervention (EI) services, schizophrenia services, diabetes management, cognitive behavioural therapy for psychosis, psychological therapy on inpatient wards, and patient reported outcome measures (Addington et al., 2005, 2012; Carswell et al., 2017; Greenwood et al., 2009; Jacobsen et al., 2024; McKenzie et al., 2022). Four studies used the Delphi method to examine priorities relating to care including service users' priorities and preferences, addressing early death in those with serious mental illness (SMI), care for immigrants with psychosis, and trauma-informed care in EI services (Byrne & Morrison, 2014, Fortuna et al., 2023; Hultsjo et al., 2011; Mitchell et al., 2021). One study used the Delphi method towards a consensus definition of recovery (Law & Morrison 2014).

Study Design

In terms of study design, five studies used the Delphi method in conjunction with other methods of data collection such as workshops, focus groups and interviews (Carswell et al., 2017; Fortuna et al., 2023; Greenwood et al., 2009; Jacobsen et al., 2024; McKenzie et al., 2022). Except for Fortuna et al. (2023), these were all studies aimed at identifying outcome measures for clinical services.

Knowledge base

All but one study included preliminary work prior to the first Delphi round (n=15, 94%), this included literature review, workshops/ meetings, EbE consultation, focus groups, previous research, translation/cultural adaptation. One study used a purely qualitative first round for the Delphi and did not report preliminary work prior to this (Mitchell et al., 2021).

Survey Instrument and Rounds

Nearly all studies (n=14, 88%) included open-ended questions as part of the Delphi survey. The exceptions were Jacobsen et al. (2024) where qualitative data was captured using interviews (but not directly in the Delphi survey) and Byrne & Morrison, 2014. All but one study (Fortuna et al., 2023) reported use of closed questions rated on a scale. In terms of scale types, the majority of studies used an odd-numbered Likert scale, the number of scale points ranged from 5 to 9 (n=14, 88%). One study used an even-numbered likert scale, forcing participants to agree or disagree with the statement (Hultsjo et al. 2011). The number

of Delphi rounds in the reviewed studies ranged from two to six, with two being the most common (n=8, 50%) (Table 2 provides details).

Feedback

The majority of studies reported that statistical feedback was provided to participants through group ratings of items (n=10, 62.5 %) (Addington et al., 2005, 2012; Byrne & Morrison 2014, Carswell et al., 2023; Cottrill et al., 2008; Fortuna et al., 2023; Greenwood et al., 2009; Hultsjo et al., 2011; Langlands et al., 2007; Mitchell et al., 2021). For studies that partially reported feedback, Fortuna et al. (2023) stated that feedback was provided but did not give details, Jacobsen et al. 2024 reported 'visual feedback', McKenzie et al. (2022) did not clearly report what feedback was provided and one study reported a leaflet summarising findings (Law & Morrisson 2014).

Consensus

All but one study reported on a consensus threshold i.e. where 'consensus' was reached through an agreedcriteria. The exception was Greenwood et al. (2009), which reported on "progression towards consensus" using a paired samples t-test. Of the studies that reported a consensus threshold, this was calculated using % endorsement or semi-interquartile range. The most common definition used was greater than 80% endorsement by the expert panel (Agrest et al., 20024; Byrne & Morrison 2014; Cottril et al., 2021; McKenzie et al., 2022; Requena et al., 2024; Li et al., 2020; Law & Morrison, 2014; Langlands et al., 2008; Mitchell, et al., 2021). Two studies used semi-interquartile range of <0.05 (Addington et al., 2005, 2012), one study used either 50% on the same ranking or 80% in either direction in a dichotimised scale (Hultsjo et al., 2011) and one study used 100% agreement (Fortuna et al., 2023). Two studies used a final consensus meeting (Carswell et al., 2017; Jacobsen et al., 2024).

Study Aims	Topic	Authors (year)	Location	Stakeholders	No. rounds	Summary of method	EbE involvement	Study outcomes
Defining concepts	Recovery in psychosis	Law & Morrisson (2014)	UK	EbE	2	Literature search, Delphi survey	Participation in Delphi.	95 recovery statements rated as essential or important by >80% sample.
Identify guidelines	MH first aid guidelines for Argentina and Chile	Agrest et al. (2024)	Argentina & Chile	healthcare professional, EbE (lived experience of psychosis or caring)	2	Translation to Spanish and cultural adaptation of survey items, two round Delphi survey	Participation in Delphi	Consensus achieved on 244 statements inc. 26 locally generated. Nearly 20% English statements not endorsed
Identify guidelines	MH first aid - redevelopment of guidelines	Cottril et al. (2021)	Australia	healthcare professional, EbE (lived experience of psychosis or caring)	3	Literature review, Delphi survey	Participation in Delphi	325 statements meeting criteria for redeveloped guidelines
Identify guidelines	MH first aid for psychosis	Langlands et al. (2008)	Australia	EbE, carers, clinicians	3	Literature review, three round Delphi survey	Participation in Delphi	89 statements endorsed
Identify guidelines	MH first aid guidelines for China	Li et al. (2020)	China	mental health professionals, EbE (carers)	2	Translation and cultural adaptation to Chinese Mandarin, two round Delphi survey	Attempted recruitment to Delphi panel	207 existing statements from existing guidelines included

								plus 8 new statements
Identify guidelines	MH first aid guidelines for Brazil	Requena et al. (2024)	Brazil	EbE, health professionals	2	Translation of English statements to Brazilian-Portuguese. Two stage online Delphi survey.	Participation in Delphi	Consensus for 257 statements
Identify outcome measures	Performance measures for EI services	Addington et al. (2005)	Canada	payer, administrative providers, clinician providers, national experts, family physicians, EbE, family members	3	Literature review to identify measures, three round Delphi survey	Participation in Delphi	24 measures identified as essential
Identify outcome measures	Performance measures for schizophrenia services	Addington et al. (2012)	Canada	schizophrenia experts, mental health clinician providers, mental health administrative providers, payer, EbE, family members	3	Literature review, three round Delphi survey	Participation in Delphi	36 measures rated essential
Identify outcome measures	COS for SMI & Type 2 diabetes	Carswell et al. (2017)	UK	EbE, carers, health & social care staff, health service managers & commissioners, academics	2	Systematic Review, multi-stakeholder and SU workshops, Delphi survey, stakeholder consensus workshop	EbE to review list of potential outcomes, participation in Delphi, 1 EbE in consensus workshop	7 outcomes identified
Identify outcome measures	CBT for Psychosis	Greenwood et al. (2009)	UK	CBTp experts, CBT therapists, EbE	2 (also included focus groups and pilot of questionnaire)	Consultation with CBT experts to develop probe areas for discussion, development of measure: focus groups, interviews, Delphi survey, pilot interviews of measure, psychometric analysis	EbE led, Participation in focus groups, participation in Delphi, pilot interviews of measure, feedback on language and clarity	Identification of outcome measure

Identify outcome measures	psychological therapy on inpatient wards	Jacobsen et al. (2024)	UK	EbE, carers, healthcare professionals, researchers, end users of research	2	Long-list of outcomes through: SR, online survey, qualitative interviews. Short-listing outcomes through E-Delphi survey. Final outcome set identified in final consensus meeting.	EbE researcher as core research team member, consultation with advisory group of service users and carers, participation in Delphi, qualitative interviews, participation in final consensus meeting	6 outcomes recommended for inclusion
Identify outcome measures	International Standard Set of Patient-Reported Outcome Measures for Psychotic Disorders	McKenzie et al. (2022)	International	EbE, clinicians, researchers	3 (modified Delphi)	Systematic Review, EbE focus groups, modified Delphi survey – videoconferencing and surveys, breakout sessions, risk adjustment factors, open review and patient validation - surveys	Involvement in working group, participation in Delphi, focus groups to identify high priority outcomes, breakout sessions narrowing lists of measures, open review and patient validation	4 outcome domains encompassing 14 outcomes important to EbE
Identify priorities relating to care	Service users' priorities and preferences for treatment	Bryne & Morrisson (2014)	UK	EbE	2	Literature review and two round Delphi survey	User-led, consultation on survey, participation in Delphi	Consensus reached for 17 statements about treatment priorities and preferences
Identify priorities relating to care	Lived experience agenda to address early death	Fortuna et al. (2023)	International	EbE, caregivers, researchers, clinician-scientists, policy makers	6	Virtual roundtable including: discussion on scientific understanding, 'story with a gap', brainstorming, draft research agenda and review, 6 round Delphi	EbE led research	7 recommendations for research priorities identified

Identify priorities relating to care	care components for immigrants with psychosis	Hultsjo et al. (2011)	Sweden	EbE, families, health-care staff	2	Developing questionnaire based on previous research, Delphi survey	Participation in Delphi	Consensus reached for being treated on equal terms, staff interest & respect. Consensus not reached for approx. half statements.
Identify priorities relating to care	Trauma-informed care in early intervention psychosis service	Mitchell et al. (2021)	UK	EbE, researchers, service providers	3	Three stage Delphi survey. Initial qualitative stage, thematic analysis to generate statement items for second and third Delphi rounds.	Participation in Delphi	16 essential principles of trauma-informed care identified

Table 2: Characteristics of reviewed studies

Engagement of EbE

User-led Research

Three studies identified as being led by experts by experience (EbE) (Byrne and Morrisson, 2014, Fortuna et al., 2023, Greenwood et al., 2009). This suggests that people with expertise by lived experience led the research design, including setting the agenda/research question to be addressed.

Delphi Study participation

All studies included information regarding the rationale regarding expert panel selection. The samples of 12 studies included multiple stakeholder groups, and in 4 studies EbE made up the entirety of participants in the Delphi survey i.e. there were no professional stakeholders such as clinicians in the Delphi (Greenwood et al., 2009; Byrne et al., 2014; Law et al., 2014; Fortuna et al., 2023). All studies provided information about how the knowledge of different stakeholder groups was to be combined. The majority of studies (n=7) included the multiple stakeholder groups in a single panel (Addington et al., 2005, 2012; Carswell et al., 2017; Hultsjo et al., 2011; Jacobsen et al., 2024; McKenzie et al., 2022; Mitchell et al., 2021). Studies of mental health first aid used multiple panels for each stakeholder group and added requirements that items must be endorsed by both panels for consensus to be met (Agrest et al., 2024; Cottril et al., 2021; Langlands et al. 2008; Li et al. 2020; Requena et al. 2024). Studies using this method required participants to specify a main identity to be grouped into a lived experience or professional panel (Agrest et al., 2024; Cottril et al., 2021; Langlands et al. 2008; Li et al. 2020; Requena et al. 2024).

A variety of methods were used for the identification and approach of participants including corresponding authors (Fortuna et al., 2023; Mitchell et al., 2021) email networks (Fortuna et al., 2023), mental health services (Addington et al., 2005, 2012; Agrest et al., 2024; Carswell et al., 2017; Hultsjo et al., 2011; Law & Morrison, 2014; Requena et al., 2024), third sector organisations (Byrne & Morrison, 2014; Carswell et al., 2017, Jacobsen et al., 2024; Langlands et al., 2008; Law & Morrison, 2014; Requena et al., 2024), social media (Agrest et al., 2024; Jacobsen et al., 2024; Law & Morrison, 2014; Requena et al., 2024), local media (Law & Morrisson, 2014), snowball sampling (Fortuna et al., 2023; Mitchell et al., 2021) previously taken part in therapy (Greenwood et al., 2009), universities (Requena et al., 2024), government representative (Addington et al., 2005, 2012), EbE in professional/advocacy roles (Langlands et al., 2008; Cottril et al., 2021), professional networks (Cottril et al., 2021).

To further assess level of engagement of EbE in the Delphi process, the percentage of EbE in the total sample for each of the reviewed studies was calculated (see Table 3). This varied from 0% due to difficulties with recruitment (Li et al., 2020) to 32.6% (Hultsjo et al., 2011). Seven studies did not report the number of EbE per Delphi round making it difficult to assess whether EbE were adequately represented during each stage of consensus building (see Table 3). One study used ambiguous language citing a 'lived experience panel' which was fully made up of carers due to difficulties in recruitment of people with lived experience of psychosis (Li et al., 2020).

Author (date)	Stage 1	Stage 2	Stage 3
	(EbE/Tot)	**(EbE/Tot)**	**(EbE/Tot)**
Addington et al. (2005)	4/25 (16%)	3/22 (13.6%)	2/20 (10%)
Addington et al. (2012)	3/20 (15%)	6/30 (20%)	6/30 (20%)
Agrest et al. (2024)	16/58 (27.6%)	NR/52	
Bryne & Morrisson (2014)*	6/6 (100%)	32 (100%)	21(100%)
Carswell et al. (2017)	23/84 (27.4%)	10/48 (20.8%)	
Cottril et al. (2021)	NR	NR	17/59 (28.8%)
Fortuna et al. (2023)*	22/28 (78.6%)	NR	NR
Greenwood et al. (2009)*	12/12 (100%)	NR	
Hultsjo et al. (2011)	14/43 (32.6%)	NR/31	
Jacobsen et al. (2024)	15/50 (30%)	NR/46	
Langlands et al. (2008)	45/157 (28.7%)	33/117 (28.2%)	31/100 (31%)
Law & Morrisson (2014)*	N/A	381/381 (100%)	100/100 (100%)
Li et al. (2020)	0/72 (0%)	0/49 (0%)	
McKenzie et al. (2022)	3/19 (15.8%)	NR	
Mitchell et al. (2021)	7/57 (12.3%)	NR/56	
Requena et al. (2024)	16/52 (30.8%)	NR/24	
Mean	**26.3%**	**20.65%**	**22.4%**
Median	**27.5%**	**20.4%**	**24.4%**

Table 3: No. of EbE as a percentage of the total sample per Delphi round

*Studies with exclusively lived experience participants in the Delphi survey and therefore excluded from mean/median analysis n.b. Fortuna et al. 2023 was excluded from mean/median calculation as this study did not identify as a multi-stakeholder study.

EbE=Expert by experience NR=not reported N/A=not applicable

Two studies reported that a final consensus meeting was held (Carswell et al., 2017; Jacobsen et al., 2024). Carswell et al. (2017) included one EbE in the final consensus meeting of ten individuals, and Jacobsen et al. (2024), included three EbE of a total nine participants. Carswell et al. (2017) provided results from the Delphi ahead of the workshop to all participants, offered a telephone call to participants with lived experience and ensured that workshop discussions started with EbE/carers sharing their views so that discussions were led from this position. They also held specific discussions for outcomes that were rated highly by people with lived experience. Jacobsen et al. (2024) held an online consensus meeting, participants presented with short-list of outcomes prior to the meeting and the importance of everyone's viewpoint was highlighted alongside use of anonymised voting software to facilitate consensus.

Quality Appraisal

The majority of studies were rated moderate-high quality using the DCAT quality appraisal tool (Khodyakov et al., 2023). Overall quality ratings are summarised in Table 4 and Appendix 1.4 provides details of full quality ratings.

Quality Rating	Studies
High	Addington et al. (2005, 2012); Cottril et al. (2021); Hulstjo et al. (2011); Langlands et al. (2008)
Moderate	Agrest et al. (2024); Bryne & Morrisson (2014); Carswell et al. (2017); Greenwood et al. (2009); Jacobsen et al. (2024); Law & Morrisson (2014), Li et al. (2020); McKenzie et al. (2022); Mitchell et al. (2021); Requena et al. (2024)
Not rated	Fortuna et al. (2023)

Table 4: Quality ratings of studies included in the review

Five studies were assessed as being of a high quality (see Table 4). Most studies (n=10) were assessed as being of moderate quality i.e. likely to provide a reasonable representation for panelists' views. The reason that these studies scored a 'moderate' quality rating is because the DCAT states that if a study does not meet one of the 'core' items, it should be downgraded from 'high' to 'moderate' quality. These studies did not meet the DCAT condition of personalised feedback. One study was not provided with an overall quality assessment due to a high volume of missing information meaning that it was not possible to rate overall study quality (Fortuna et al., 2023).

Transparency of Reporting

Appendix 1.5 details aspects of the Delphi study which were reported in the reviewed papers using the Spranger et al. (2022) checklist as a guide. Key findings from applying the Spranger checklist are detailed below.

Epistemology

None of the reviewed studies explicitly identified positioning in a specific theory of science i.e. realist or constructivist. Eleven studies (69%) commented on preliminary assumptions to some degree e.g. acknowledgement of likely differences of opinion between stakeholder groups (Addington et al., 2005, 2012; Agrest et al., 2024; Bryne & Morrisson, 2014; Greenwood et al., 2009; Hultsjo et al., 2011; Jacobsen et al., 2024; Law & Morrisson, 2014; Li et al., 2020; Mckenzie et al., 2022; Requena et al., 2024) . We also looked at whether studies formally acknowledged implicit power imbalances between stakeholder groups, this was difficult to assess due to inconsistencies in reporting. However, four studies (25%) gave a degree of acknowledgement to this (Addington et al., 2005, 2012; Carswell et al., 2017; Jacobsen et al., 2024).

Quantitative and Qualitative Evaluation

All but one study published details on the quantitative evaluation strategy through identifying the terms of consensus and item inclusion or exclusion per round. The exception was Fortuna et al. (2023) which did not report the method of quantitative data analysis. Of the studies which included open-ended questions as part of the study design, there was variability in the level of reporting data analysis. One paper (Jacobsen et al., 2024) stated that qualitative interviews with experts by experience and carers were reported in a separate paper. Six studies cited that responses were reviewed by authors for the purposes of incorporating novel ideas into the next survey round (Agrest et al., 2024; Cottril et al., 2021; Hulstjo et al., 2011; Langlands et al., 2008; Li et al., 2020; Requena et al., 2024). Thematic analysis was used to analyse open-ended data in two studies (Addington et al., 2005;Mitchell et al., 2021).

Reporting of Results

Nine studies (56%) presented the results of individual Delphi rounds (Agrest et al., 2024; Bryne & Morrisson, 2014; Carswell et al., 2017; Cottril et al., 2021; Jacobsen et al., 2024; Langlands et al., 2008; Law & Morrisson, 2014; Li et al., 2020; Requena et al., 2024). Four studies (25%) were assessed as partially reporting results to rounds (Addington et al., 2005, 2012; Hultsjo et al., 2011; Mitchell et al., 2021). This rating was used when studies reported only the endorsed items and did not include details of the removed items. Three studies (19%) did not report results for each round. Fortuna et al. (2023) used a 6 round Delphi design and reported only the final set of recommendations, Greenwood et al. (2009) used the Delphi method to assess potential items for outcome measurement as part of a wider study that included psychometric testing which formed the basis of the results section. Mckenzie et al. (2022) also used the Delphi method to develop outcomes measurement and reported on the final set rather than individual rounds to the Delphi. Eight studies (50%) provided a graphical overview of the Delphi procedure i.e. a flowchart (Agrest et al., 2024; Cottril et al., 2021; Hultsjo et al., 2011; Jacobsen et al., 2024; Langlands et al., 2008; Law & Morrisson 2014; Li et al., 2020; Requena et al., 2024).

Discussion

Summary

This study aimed to systematically review the literature using the Delphi method in psychosis research which has included people who have expertise through lived experience. Specific aims were to investigate the level of involvement of people with expertise through lived experience and assess Delphi study quality and transparency of reporting. Quality appraisal of the reviewed

studies revealed that the majority were of an acceptable standard (n=15, 94%). However, further examination revealed specific areas that could be improved upon to enhance research quality and transparency.

There was methodological variability in the reviewed papers. While the use of Delphi methodology was situated within broader studies using mixed methods in some papers (Carswell et al., 2017; Fortuna et al., 2023; Greenwood et al 2009; Jacobsen et al 2024; McKenzie et al 2022), in others the Delphi survey was reported as a standalone research paper (Addington 2005 2012, Agrest et al., 2024; Byrne & Morrison 2014; Cottril 2021, Hulstjo et al., 2011; Langlands et al., 2008; Law & Morrisson 2014; Li et al., 2020; Requena et al., 2024). Studies which utilised the Delphi technique together with other methodologies all included people with lived experience in supplementary data collection processes.

It is difficult to draw firm conclusions regarding whether the final 'consensus' views of the reported studies adequately represented the views of people with lived experience of psychosis. For studies where the sample was exclusively with people with lived experience of psychosis as participants, it is likely that the findings represented participants views given that these studies met quality appraisal criteria (Byrne & Morrisson 2014, Law & Morrisson, 2014). For multi-stakeholder studies, the results varied. At one end of the spectrum, it would not be possible to conclude that the study by Li et al. (2020) adequately included the views of people with lived experience in their consensus ratings given that no people with lived experience of psychosis took part (Li et al., 2020). The authors reflected on this in their paper and reported the attempts made to engage stakeholders with lived experience citing cultural considerations in China where the study was conducted could have been a barrier to engagement. Seven other multi-stakeholder studies did not fully report the number of people with lived experience who took part in each round making it difficult to deduce whether the final consensus adequately represented participants' perspectives (Agrest et al., 2024; Cottril et al., 2021; Hultsjo et al., 2011; Jacobsen et al., 2024; McKenzie et al., 2022; Mitchell et al., 2021; Requena et al., 2024). The two studies which used a final consensus meeting included stakeholders with lived experience and reported on the steps taken to ensure their perspectives were represented in discussions (Carswell et al., 2017; Jacobsen et al., 2024). However, experts by experience were in the minority in these meetings which may have made it difficult to express views.

Strengths and Limitations

This study had a number of methodological strengths and limitations which are important to consider when interpreting the findings. To the best of our knowledge, this is the first systematic review of published research using Delphi methodology involving EbEs in psychosis research. A strength of this study is that it drew upon research tools such as the Spranger et al. (2022) and DCAT (2023) checklists to interrogate literature. However, a limitation is that these checklists did not fully capture the aspects of research that was of interest i.e. the engagement of people with lived experience of psychosis. A recent paper highlighted the risks of participatory assimilation in research whereby the views of people with lived experiences may be distorted or minimised through the research process (Veldmeijer & van Os, 2025). This study aimed to gain a measure of the extent to which lived experience perspectives were incorporated into the Delphi process however this was difficult to do due to inconsistencies in reporting and with the available appraisal checklists, which required adaptation to assess EbE involvement. For example, the DCAT and Spranger et al. (2022) checklists did not appear to be sensitive enough to pick up on potential participatory assimilation. We attempted to partially address this through the inclusion of additional questions such as calculating the percentage of experts by experience during each Delphi round. However, some of the information was difficult to glean for papers. For example, we were interested in whether power imbalances between stakeholders were identified and mitigated however there did not appear to be standardised reporting on this in the reviewed studies. This review examined the extent of involvement of people with lived experience of psychosis, it would also have been useful to consider the extent of involvement of carers and supporters given their perspectives which may also be under-represented in the literature (Onwumere et al., 2016).

The Spranger et al. (2022) checklist was used to structure data extraction and as a checklist for transparency of reporting as this was deemed the most suitable tool available at the time of proposal development. This has subsequently been developed into a formal tool termed ‘Delphistar’ (Niederberger et al., 2024). Use of the formal ‘Delphistar’ tool may have better supported data extraction and synthesis in the current study. A limitation in the retrospective application of a reporting checklist, was that, where an item was unreported, it was not always clear whether this is because it was not carried out, or simply not reported on. Another limitation of this review is that there was limited resource available for a second independent reviewer. The available resource was used for secondary screening of articles at the title/abstract and full paper stages. Having a second reviewer at the data extraction and quality

assessment stages could have enhanced the methodological rigour of the current review by reducing the chance of researcher error.

Recommendations

Studies of co-production research in mental health have consistently highlighted power differentials and the under-valuing of experiential knowledge (Evans et al., 2025; Hopkins et al., 2024). These recommendations are provided with the aim of supporting the meaningful involvement and transparency of involvement for people with lived experience in Delphi studies.

1. Routine reporting of epistemological positioning, assumptions and researcher reflexivity.
2. Include number of stakeholders (e.g. experts by experience) per Delphi round.
3. Clear language and clarity regarding the extent of participation for those with lived experience.
4. Routine use of registries such as Open Science Framework (OSF) to enhance transparency e.g. through pre-registration of study proposal and data analysis plans, inclusion of raw materials such as questionnaires, ethical approvals, details of formal context.
5. Opportunities for people with expertise by experience to contribute during preliminary work aimed at broadening research foci and incorporation of this in the Delphi process.
6. Explicit acknowledgement of power imbalances, steps taken to mitigate this, and transparent reporting of how this has been managed in the study design.
7. Use of a reporting checklist e.g. Delphistar (Niederberger et al., 2024).
8. Follow best practice guidance for engaging experts by experience and reporting engagement (Colder Carras et al., 2023; Evans et al., 2025)

Wider Context

While the current review has focussed on Delphi studies including people with expertise through lived experience in psychosis research, there is a much wider body of literature using the method with other stakeholders e.g. clinicians and researchers. While it may not be appropriate to include lived experience expertise in some of these studies e.g. where the focus is technical expertise relating to specific aspects of professional practice, there are other studies whereby the inclusion of people with lived experience could enhance the study quality and relevance to the clinical population. It is our hope that the findings of the current review could

help to provide researchers with additional knowledge and tools to support the inclusion of expertise through lived experience using this method. There were several high-quality studies identified in this review which used the Delphi method effectively across several different topics and settings. While there were multiple examples of good practice across the reviewed papers, it should be noted that studies varied in their strengths and weaknesses, and standardisation across the field could support high quality, transparently reported research in this area. It is the current authors view that the Delphi method can be used effectively to enhance knowledge and support the inclusion of experts by experience perspectives in the knowledge base, a view which has traditionally been lacking.

Conclusion

The aim of this systematic review was to provide a description and evaluation of the use of Delphi method in psychosis research including people with expertise through lived experience as participants. Findings from the review suggest that this is a growing body of literature that can add important value to the field. However, there are specific considerations that are important for using the Delphi method with this group particularly around ethical considerations and transparency of reporting to ensure high-quality research which has value for its end-users. Recommendations for future research aiming to use the method with people who experience psychosis and their supporters have been provided.

References

Addington, D., McKenzie, E., Addington, J., Patten, S., Smith, H., & Adair, C. (2005). Performance measures for early psychosis treatment services. *Psychiatric Services*, *56*(12), 1570-1582.

Addington, D. E., Mckenzie, E., Wang, J., Smith, H. P., Adams, B., & Ismail, Z. (2012). Development of a core set of performance measures for evaluating schizophrenia treatment services. *Psychiatric Services*, *63*(6), 584-591.

Agrest, M., Tapia-Munoz, T., Encina-Zúñiga, E., Vidal-Zamora, I., Geffner, N., Ardila-Gómez, S., ... & Reavley, N. (2024). Development of mental health first-aid guidelines for psychosis: a Delphi expert consensus study in Argentina and Chile. *BMC psychiatry*, *24*(1), 113.

Allan, S., Bradstreet, S., Mcleod, H., Gleeson, J., Farhall, J., Lambrou, M., ... Gumley, A. (2019). Service Users, Carers and Mental Health Staff understandings of early warning signs of relapse in psychosis: A qualitative investigation. *BJPsych Open*, 1–7.

Arnstein, S. R. (1969). A ladder of citizen participation. *Journal of the American Institute of planners*, *35*(4), 216-224.

Barrett, D., & Heale, R. (2020). What are Delphi studies?. Evidence-based nursing, 23(3), 68-69.

Butterworth, H., Wood, L., & Rowe, S. (2022). Patients' and staff members' experiences of restrictive practices in acute mental health in-patient settings: systematic review and thematic synthesis. *BJPsych Open*, *8*(6), e178.

Byrne, R., & Morrison, A. P. (2014). Service users' priorities and preferences for treatment of psychosis: a user-led Delphi study. *Psychiatric Services*, *65*(9), 1167-1169.

Carswell, C., Coventry, P. A., Brown, J. V., Alderson, S. L., Double, K., Gilbody, S., ... & DIAMONDS Research Group. (2023). Development of a supported self-management intervention for people with severe mental illness and type 2 diabetes: theory and evidence-based co-design approach. *Journal of Medical Internet Research*, *25*, e43597.

Colder Carras, M., Machin, K., Brown, M., Marttinen, T. L., Maxwell, C., Frampton, B., ... & Jones, N. (2023). Strengthening review and publication of participatory mental health research to promote empowerment and prevent co-optation. *Psychiatric services*, *74*(2), 166-172.

Conry-Murray, C., Waltzer, T., DeBernardi, F. C., Fossum, J. L., Haasova, S., Matthews, M. S., ... & Elsherif, M. M. (2024). Validity and transparency in quantifying open-ended data. *Advances in Methods and Practices in Psychological Science*, *7*(4), 25152459241275217.

Evans, J., Lewin, C. D. C., Fabian, H., Alam, M., Calder, P., Hill, R., ... & Wykes, T. (2025). Facilitators of and barriers to patient and public involvement in mental health research within university settings: a systematic review and meta-synthesis. *Psychological Medicine*, *55*, e297.

Fortuna, K. L., Lebby, S., Geiger, P., Johnson, D., MacDonald, S., Chefetz, I., ... & Walker, R. (2023). Lived experience–led research agenda to address early death in people with a diagnosis of a serious mental illness: a consensus statement. *JAMA network open*, *6*(5), e2315479-e2315479.

Greenwood, K. E., Sweeney, A., Williams, S., Garety, P., Kuipers, E., Scott, J., & Peters, E. (2010). CHoice of Outcome In Cbt for psychosEs (CHOICE): the development of a new service user–led outcome measure of CBT for psychosis. *Schizophrenia Bulletin*, *36*(1), 126-135.

Hasson, F., Keeney, S., & Mckenna, H. P. (2000). Research guidelines for the Delphi Survey Technique Research guidelines for the Delphi survey technique, (February 2019).

Hopkins, I., Verlander, M., Clarkson, L., & Jacobsen, P. (2024). What do we know about sharing power in co-production in mental Health Research? A systematic review and thematic synthesis. Health Expectations, 27, e70014.

Hultsjö, S., Berterö, C., Arvidsson, H., & Hjelm, K. (2011). Core components in the care of immigrants with psychoses: A Delphi survey of patients, families, and health-care staff. *International journal of mental health nursing*, *20*(3), 174-184.

Jacobsen, P., Berry, K., Clarkson, L., Hiscocks, R., Hopkins, I., Morgan, C., ... & Wood, L. (2024). Development of a core outcome set for psychological therapy trials on acute psychiatric inpatient wards. *BMC psychiatry*, *24*(1), 821.

Jorm, A. F. (2015). Using the Delphi expert consensus method in mental health research. *Australian & New Zealand Journal of Psychiatry*, *49*(10), 887-897.

Jünger, S., Payne, S. A., Brine, J., Radbruch, L., & Brearley, S. G. (2017). Guidance on Conducting and REporting Delphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. *Palliative medicine*, *31*(8), 684-706.

Khodyakov, D., Grant, S., Kroger, J., & Bauman, M. (2023). *RAND methodological guidance for conducting and critically appraising Delphi panels*. Rand.

Langlands, R. L., Jorm, A. F., Kelly, C. M., & Kitchener, B. A. (2008). First aid recommendations for psychosis: using the Delphi method to gain consensus between mental health consumers, carers, and clinicians. *Schizophrenia bulletin*, *34*(3), 435-443.

Law, H., & Morrison, A. P. (2014). Recovery in psychosis: a Delphi study with experts by experience. *Schizophrenia bulletin*, *40*(6), 1347-1355.

Li, W., Jorm, A. F., Wang, Y., Lu, S., He, Y., & Reavley, N. (2020). Development of Chinese mental health first aid guidelines for psychosis: a Delphi expert consensus study. *BMC psychiatry*, *20*(1), 443.

McKenzie, E., Matkin, L., Sousa Fialho, L., Emelurumonye, I. N., Gintner, T., Ilesanmi, C., ... & Psychotic Disorders Working Group of the International Consortium for Health Outcomes Measurement. (2022). Developing an international standard set of patient-reported outcome measures for psychotic disorders. *Psychiatric Services*, *73*(3), 249-258.

McLaughlin, H. (2009). What's in a name:‘client’,‘patient’,‘customer’,‘consumer’,‘expert by experience’,‘service user’—what's next?. *British Journal of Social Work*, *39*(6), 1101-1117.

Mitchell, S., Shannon, C., Mulholland, C., & Hanna, D. (2021). Reaching consensus on the principles of trauma-informed care in early intervention psychosis services: A Delphi study. *Early Intervention in Psychiatry*, *15*(5), 1369-1375.

Morrison, A. P., & Barratt, S. (2010). What are the components of CBT for psychosis? A Delphi study. *Schizophrenia bulletin*, *36*(1), 136-142.

Niederberger, Marlen, et al. "Delphi studies in social and health sciences—Recommendations for an interdisciplinary standardized reporting (DELPHISTAR). Results of a Delphi study." *PLoS One* 19.8 (2024): e0304651.

Onwumere, J., Shiers, D., & Chew-Graham, C. (2016). Understanding the needs of carers of people with psychosis in primary care. *The British Journal of General Practice*, *66*(649), 400.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., ... & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. *A product from the ESRC methods programme Version*, *1*(1), b92

Requena, S. S., Loch, A. A., Mendes, K. N., & Reavley, N. J. (2024). Development of mental health first aid guidelines for psychosis for Brazil: A Delphi expert consensus study. *Plos one*, *19*(7), e0307064.

Spranger, J., Homberg, A., Sonnberger, M., & Niederberger, M. (2022). Reporting guidelines for Delphi techniques in health sciences: A methodological review. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, *172*, 1-11.

Veldmeijer, L., & van Os, J. (2025). “They talk the talk but don’t walk the walk”. Introducing the concept of participatory assimilation in mental health care innovation. *Psychosis*, 1-12.

Chapter 2: Major Research Project

Developing a Consensus Definition of Relapse in Psychosis: A Delphi Study

Prepared in accordance with the author requirements for Schizophrenia Bulletin
https://academic.oup.com/schizophreniabulletin/pages/information_for_authors

Note on Language

Supporter: we will use the term ‘supporter’ to describe people who have experience of supporting someone with psychosis in a personal capacity. This term has been chosen in recognition of the broad range of support that can be helpful, beyond what may have traditionally been termed 'caring' e.g. friendship, practical support, companionship etc. We recognize that this will not capture all identities and that this marks a departure from the standard terminology of ‘carer’.

Plain Language Summary

Introduction

Experiences of psychosis such as hearing things that other people do not, believing people are out to cause one harm and/or feeling very confused can be very distressing and are often associated with a need for healthcare. Some people who have experienced psychosis will go on to experience further 'episodes' of psychosis. The medical term used for this is 'relapse'. Relapse is often used as an outcome measure in research studies i.e. as a measure of whether a treatment has worked. However, there are differences in the way that 'relapse' is used and defined in different settings. Existing definitions of relapse have been created by people working in clinical and/or research settings. To the best of our knowledge, people who have experienced psychosis have not been asked about their views of 'relapse' and had these incorporated into definitions. In this study, we aimed to bring together the views of people with lived experience of psychosis, their supporters (e.g. family), clinicians, and researchers to see if we could identify consensus in defining relapse.

Method

We did this through a 'Delphi' study, which is a research method that takes place over a number of rounds (usually three) and starts with broad questions, that are narrowed down as areas of agreement and disagreement are found. This helps to reach areas of consensus through repeated surveys and the removal of items which respondents do not agree on. We used a three stage Delphi survey which asked people's views on the topic of relapse in psychosis. When more than 80% of people rated agreement, we considered this an area of consensus.

Results

The total number of participants was 119, 44 and 39 for rounds 1, 2 and 3 of the Delphi survey respectively. People with lived experience of psychosis, their supporters, clinicians and researchers were all represented during each round of the Delphi survey. Consensus was reached for 27 items relating to the concept of 'relapse' in psychosis. There was consensus that experiences of psychosis (e.g. hearing or seeing things no one else can), cognitive disorganisation (e.g. feeling confused or muddled) and mania (e.g. feeling elated and having lots of energy) are all important parts of a definition of relapse. There was also agreement to include the impact of these experiences on distress, functioning (i.e. ability to carry out the tasks of daily living) and risk in a consensus definition. There was also consensus that relapse should be defined balancing the views of different stakeholders.

Discussion

This has important implications for defining ‘relapse’ going forward, especially thinking about how to make sure definitions of relapse are relevant for the person experiencing them. We think further research that explores people with lived experiences further would be helpful in refining ‘relapse’ definitions and ensuring these are useful for the people considered to be experiencing them.

Abstract

Introduction

Relapse is common for many people who experience psychosis and relapse prevention is a common goal of evidence-based care. However, there is variation in the way that relapse is conceptualised across clinical trials and in healthcare settings. Previous research aimed at building a relapse definition in psychosis has under-represented the views of stakeholders with lived experience of psychosis and their supporters. The current study aimed to identify areas of relative consensus in defining relapse in psychosis among stakeholders comprising people with expertise by experience of psychosis, their supporters, clinicians and researchers.

Method

The Delphi method, a consensus building technique was used. Purposive sampling to recruit participants with lived experience of psychosis, their supporters, clinicians, and researchers online via social media, contacting corresponding authors and research networks was used. Questionnaires aimed at gathering respondents' views on features of relapse were administered online using the platform 'Qualtrics' in three iterative rounds. Questionnaires included rating statements using likert scales, open-ended questions and reviewing group feedback from the previous round.

Results

The total number of participants was 119, 44 and 39 for rounds 1, 2 and 3 of the Delphi survey respectively. Each stakeholder group was represented during each round of the Delphi survey and consensus was reached for 27 items relating to the concept of 'relapse' in psychosis.

There was consensus that relapse should be defined as balancing the views of different stakeholders and could be personalized to an extent. In terms of symptom profiles, there was consensus that experiences of psychosis, cognitive disorganisation and mania are all important in conceptualising relapse. There was also consensus to include levels distress, functioning and risk in each of these symptom profiles.

Discussion

This study identified areas of consensus in defining relapse among key stakeholders which were adapted into a checklist. Future research could consider how this could be utilised to enhance existing definitions. A limitation of this study is that a 'bottom up' approach to coding open-ended data was not taken. Future qualitative research aimed at further exploring the perspectives of people with lived experience of 'relapse' in psychosis could help to enhance the construct validity and utility.

Introduction

Relapse is common for many people who have experienced psychosis and is often associated with increased disability, for example through loss of important social relationships and reduced educational and occupational opportunities (Robinson et al., 1999, Scottish Intercollegiate Guidelines Network (SIGN), 2013). Relapse prevention is a common goal of evidence-based care (NICE, 2014; SIGN, 2013). However, relapse definitions vary widely across clinical trials of antipsychotic and psychosocial interventions (e.g. Leucht 2013, Kishimoto 2013, Moncrieff et. al, 2019; Cooper et al., 2019). For example, while many studies define relapse based on symptomatic increase, others use broader criteria such as psychiatric hospitalization or change in clinical management (Kishimoto 2013). This is potentially problematic as these indicators could be related more to service provision than individual factors. Furthermore, indicators of relapse used in clinical guidelines do not necessarily match those used in research settings. For instance, NICE guidelines refer to increased use of alcohol or other substances as an indicator of possible relapse:

"When a person with an established diagnosis of psychosis or schizophrenia presents with a suspected relapse (for example, with increased psychotic symptoms or a significant increase in the use of alcohol or other substances)" p.14 NICE (2014).

However, to the best of our knowledge, increased substance use is not used as an indicator of relapse in research such as clinical trials of interventions for psychosis. It is important that there is consensus around the conceptualisation, operationalisation and implementation of relapse in research and clinical settings, particularly as clinical guidelines provide recommendations based on research findings.

While the lack of a standardised definition of relapse has been noted in the literature, there continues to be ongoing debate regarding the concept (Porcelli et al., 2016). While it may not be possible to achieve a global consensus about relapse as a concept, exploring understandings of relapse for key stakeholders such as people with lived experience and clinicians, could enhance existing definitions. This would have important clinical and measurement implications as it could help to assess whether relapse has occurred. In other fields of mental health research, such as depression, such work has contributed to consistency in research and clinical practice. For example, Frank et al. (1991) examined definitions of key concepts within depression research, namely 'remission', 'relapse', 'recovery' and 'recurrence' and put forward empirically testable definitions based on the literature. This has had several benefits including

improving study design across the field, making results easier to interpret and compare; clarifying the relationships between different variables such as psychological factors involved in depression; improving clinical guidelines (Frank et al. 1991; deZwart et al., 2019). Similar work has been carried out on the concept of 'remission' in Schizophrenia by Andreasen et al. (2005). While Andreasen and colleagues put forward a 'consensus definition' based on the opinions of an 'expert working group', it is notable that their publication does not outline the inclusion criteria for membership of the working group, and it appears that the criteria may represent consensus based on Psychiatrists' views to the exclusion of other stakeholders. Subsequent research has suggested that the criteria proposed by Andreasen et al. (2005) have benefits for clinical research and practice (van Os et al., 2006).

There is a small body of existing literature focused on defining the concept of relapse in psychosis, including studies which have used the Delphi method (Burns et al., 2000, Howes et al., 2025, San et al., 2013). This has primarily focused on those with a diagnosis of 'Schizophrenia' and has emphasised clinical and research perspectives of relapse (e.g. Burns et al., 2000, Howes et al., 2025, San et al., 2013). While earlier Delphi studies of relapse included exclusively clinicians/researchers (and primarily psychiatrists) (Burns et al., 2000; San et al., 2013), a recent study cited inclusion of the perspectives of people with lived experience (Howes et al., 2025). However, the number of participants with lived experience and at which points in the study design their views were incorporated were not clearly stated. Moreover, it is unclear whether people with lived experience were given the opportunity to broaden the discourse around perspectives of relapse, or whether their views were sought only in narrowing existing criteria. This is an important distinction given that the perspectives of those with lived experience are under-represented in the existing literature (Okoroji et al., 2023). Therefore, existing definitions may exclude important indicators of relapse identified by those experiencing it, and their supporters. Further work is needed to represent the viewpoints of key stakeholders in defining relapse using methods aimed at enhancing consensus among stakeholders.

Initial research scoping definitions of relapse has identified that most clinical trials have included 1-2 of the following dimensions in defining relapse: Change in (positive) symptoms, psychiatric hospitalisation, change in clinical management, risk of harm to self/others, change in functioning, failure to respond to treatment, discontinuation of medication (e.g. Kishimoto et al., 2013). However, clinical trials have varied in the dimensions used to study outcomes, thus leading to a lack of comparability between studies. For example, while change in symptoms was the most common dimension used to measure relapse, this was not used in all studies. Additionally, some studies include a greater number of factors than others (Kishimoto et al.,

2013). A further issue is that some of the factors used to indicate relapse could represent other, unrelated, factors. For example, changes in clinical management could reflect factors relating to health service provision rather than an individual's symptoms. Despite these limitations, the factors identified by scoping work provide an initial basis for understanding the concept of relapse. It is also important to consider how these elements may combine to ascertain whether relapse has occurred. Furthermore, an increase in distress does not appear to feature within the elements of relapse identified, therefore it could be useful to consider how distress may feature within relapse as well as other features which could be missing from existing definitions.

Aims and Research Questions

The aim of the current project was to develop a consensus definition of relapse incorporating the views of key stakeholders. The following research questions were addressed:

1. What are the key components of relapse?
2. How do these combine to form a definition of relapse?
3. Is it possible to gain consensus around measurement of relapse?

Researcher Reflexivity Statement and Epistemological Position

The research team in this study are psychologists and hold a mixture of personal and professional identities in relation to experiences of psychosis and caring. The research team are passionate about promoting inclusion of people with lived experience of psychosis and their supporters in research and clinical care. The lead researcher is a trainee clinical psychologist and researcher who has worked in the field of psychosis. As a research team, we used regular research supervision to reflect and discuss emerging findings and consider how our own values and experiences shape aspects of the research, including the decisions made throughout the research process. In terms of positioning within a theory of science, a constructivist perspective was taken towards the concept of 'relapse' in psychosis whereby relapse was seen to be a social construct developed via social and cultural processes (as opposed to a 'true' construct) (Subramani, 2019).

Method

Delphi Technique

This project used a three-stage classic Delphi process (Hasson et al., 2000; Morrison & Barratt 2010; Sforzini et al., 2010), a mixed methods technique which aims to answer a research question through identifying a consensus view across a panel of experts (Barret & Heale, 2020). The Delphi method was deemed to be appropriate to answer the research question because it systematically gathers the opinions of different expert groups and can identify areas of agreement and disagreement (Niederberger et al., 2024). Three sequential rounds of

questionnaires were administered to key stakeholders via an online survey hosted on 'Qualtrics'. The number of rounds was defined in advance to be a maximum of three rounds. No outside consulting with regards to the Delphi method took place. Appendix 2.1 contains the study proposal. The study was funded by NHS Scotland as part of the lead researcher's doctoral studies. The Delphistar (2024) checklist was followed to ensure rigour and transparency in the use of the Delphi method (Appendix 2.2).

Participants

The expert panel was formed of people with expertise in psychosis with representation from four key stakeholder groups. These were people with expertise through lived experience of psychosis, lived experience of supporting someone experiencing psychosis, clinicians, and researchers in the field of psychosis.

Sample Size

While a minimum of 12 respondents is often considered sufficient in Delphi studies (Vogel et al., 2019), it was deemed appropriate to multiply this by 4 in the current study taking the sample size to 48, to allow for differences in opinions between the different stakeholder groups and ensure that each group was represented in the process of reaching consensus. To help account for attrition across three Delphi rounds, the sample size aimed for was 144 (48 x 3).

Inclusion Criteria

The following inclusion criteria were applied. Any person who self-identified as:

1. Someone with lived experience of psychosis
2. A clinician working with people who experience psychosis
3. Someone who supports or cares for someone with lived experience of psychosis
4. A researcher working in the field of psychosis

Exclusion Criteria

1. People who are under the age of 16 years old
2. People who do not read or write English

Participant Identification and Recruitment

Participant identification and recruitment was carried out by the research team using purposive sampling to ensure that individuals with the appropriate expertise were recruited. Snowball sampling was employed whereby interested individuals were asked to share the study information within their networks. Throughout recruitment, care was taken to achieve a 'balanced' sample in relation to the different participant identities through regularly checking

the sample and adjusting recruitment avenues as needed e.g. increased posting on social media groups for people with lived experience of psychosis.

Researchers and Clinicians

Researchers and clinicians who were previously involved in clinical trials with a focus on relapse prevention were emailed an invitation to participate in the study and a link to a webpage containing further information about the study. Professional networks of individuals known to be involved in relevant research e.g. study collaborators of the research team were also contacted with an invitation to participate. Social media was utilised through advertising the study on Bluesky social containing a link to the survey information.

People with lived experience of psychosis and supporters

Social media was used to advertise the study to people with expertise by experience of psychosis either personally, or through supporting a loved one. Brief information and a link to the survey containing further information was shared on relevant online community groups hosted on social media platforms specifically Facebook, Instagram and Reddit.

Materials

Data was collected using 'Qualtrics', a GDPR compliant online survey tool designed for scientific research.

Questionnaire Development

Figure 2 (p.53) provides an overview of the questionnaire at each stage of the Delphi. Stage 1 of the questionnaire was designed following scoping work previously carried out by members of the research team. This involved examination of outcome measures used in trials of treatments for psychosis (e.g. Kishimoto 2013, 2021). This work revealed the following key areas used as outcome data in clinical trials:

1. Change in (positive) symptoms
2. Psychiatric hospitalisation
3. Change in clinical management
4. Risk of harm to self/others
5. Change in functioning
6. Failure to respond to treatment
7. Discontinuation/Unplanned discontinuation of medication

These were discussed as a research team and broader themes that could be relevant were also considered, reflecting on biases in the existing literature. Our aim was to keep the initial stage of the Delphi broad to include all possible relevant domains. Following discussion as a research team, the questionnaire was structured around the following domains:

1. Change in experiences of psychosis (positive symptoms)
2. Change in cognitive disorganisation symptoms
3. Change in emotional distress
4. Change in agitation
5. Change in negative symptoms
6. Change in mania/hypomania
7. Change in level of risk to self
8. Change in level of risk to others
9. Change in functioning
10. Changes to clinical care

Public and Patient Involvement

A first version of the questionnaire was reviewed by Public and Patient Involvement (PPI) consultation consisting of people with lived experience of psychosis and/or supporting someone who has experienced psychosis. PPI consultation was organized through the University of Glasgow Psychosis Research Group PPI Group. The questionnaire was distributed to members of the University of Glasgow Psychosis Research Group PPI Group who were paid for the time they spent reviewing the study to acknowledge time, skills and expertise. Members of the group individually reviewed the questionnaire draft and provided feedback in writing and through group discussion. This resulted in the addition of another domain relating to change in behaviour as suggested by PPI group members and incorporation of the group's advice in relation to formatting the questionnaire e.g. including a progress indicator.

Pilot

Following PPI consultation, a pilot version of the questionnaire was circulated to researchers in the University of Glasgow Psychosis Research Group who provided feedback on the formatting which resulted in further minor amendments relating to accessibility/wording of statements.

Stage 1 Questionnaire

The final stage 1 questionnaire can be accessed in Appendix 2.3. This also contains the participant information sheet and consent form which participants were required to review and complete prior to beginning the survey. Prior to answering questions about relapse, participants were asked demographic questions and about which stakeholder identities they held. Participants were given the option to select as many identities as were relevant for their individual circumstances, as participants could have multiple identities in relation to the inclusion criteria. The initial questionnaire contained a 3-point rating scale. The reason for this was to keep the initial ratings broad to help consider which factors may be relevant in conceptualising relapse in psychosis. Additional open-ended questions were included to give participants the opportunity to identify the specific signs/symptoms they thought were important. Further open-ended questions were included to allow participants to reflect on broader themes beyond the domains outlined in the questionnaire and allow us to incorporate additional and novel ideas into stage 2. Stages 2 and 3 of the questionnaires were designed following analysis of the first stage, described below.

Data Analysis

A data analysis plan for each survey stage was agreed *apriori* between the research team (see Appendix 2.4). Frequencies and percentages in relation to each item were calculated. Appendix 2.5 provides responses frequencies for each item included in stages 1,2 and 3. These were also visually represented in pie charts to present back to respondents in stage 2. Responses to open-ended questions were examined by the lead researcher (AT) and content analysis was utilised to categorise responses into 4 categories that were specified *apriori* and agreed with the research team in line with good practice (Conry Murray et al., 2024). Table 5 details the categories used for coding. The lead researcher analysed comments to determine whether these represented novel ideas that were not included in the first-stage of the survey. Where there was uncertainty as to whether an idea was novel, the lead researcher brought this for discussion to the research team and a decision was made.

Category	**Applied to**	**Description**
Specific symptoms within domains	Domain specific questions	Where participants have provided information on specific symptoms/experiences as particularly important within a domain
Additional domains	General questions	Where participants have listed additional domains that should be considered in a relapse definition that have not been included in the survey
Domains/factors to exclude	General questions	Where participants have identified additional domains/factors that they feel should not be included in a relapse definition

Additional considerations	Domain specific questions, general questions	Where participants have identified considerations in conceptualising relapse that are not captured by the above 3 categories

Table 5: Coding framework for open ended question responses

Stage 2 Questionnaire Design and Administration

Following stage 1 analysis, the second stage of the questionnaire was designed (Appendix 2.3). In keeping with the data analysis plan, no domains were removed between stages 1 and 2 because none met the criteria for exclusion (i.e. <70% endorsement). While the emphasis in stage 1 was to have a broad and inclusive view on the topic of relapse in psychosis, stage 2 aimed to focus in on key areas relating to a definition of relapse, and to consider how these may relate to indicate relapse. Where stage 1 of the questionnaire included 11 potential relapse domains identified in scoping work, stage 2 encompassed follow-up questions relating to each of these domains (severity, duration, distress, impact on functioning). In addition, stage 2 included items relating to topics identified through content analysis of stage 1 open-ended responses. This comprised of:

1. Items aimed at assessing whether a broad relapse definition would be helpful versus separate relapse domains;
2. Including items relating to: different perspectives, language and utility of the relapse concept; and
3. Adding specific domains for participants to consider: changes to sleep, changes to sense of 'self agency', changes to level of 'insight'.

This resulted in a questionnaire comprising 79 closed items and 1 open question which was administered to respondents during stage 2 of the Delphi.

Stage 2 Data Analysis

Frequencies and percentages of respondent agreement relating to each item were calculated and the following consensus criteria was applied.

1. >80% agreement – consensus reached, item automatically included
2. 70-79% agreement – item re-rated in stage 3
3. <70% agreement – item excluded

Responses to the open question were analysed using the same method of content analysis as stage 1.

Stage 3 Questionnaire Design and Administration

Stage 3 focused on re-rating items for which borderline consensus was reached in stage 2 using the consensus criteria outlined above i.e. 70-79% agreement in stage 2.

This resulted in 21 items being administered in stage 3 of the questionnaire. A further item relating to cognitive disorganisation was also included due to omission through administrative error during stage 2. No additional items based on responses to open questions were added to stage 3. This resulted in a questionnaire comprising 21 closed items and 1 open item being administered to respondents in stage 3 of the Delphi.

Stage 3 Data Analysis

Frequencies and percentages were calculated, and consensus criteria was applied to stage 3 survey responses. This comprised items which had previously been rated 70-79% in stage 2 of the Delphi. For stage 3, items reaching >80% agreement by the panel were deemed to meet the consensus threshold. Responses to the open question were analysed using the same method of content analysis as stage 1 and 2.

Feedback

For rounds 2 and 3, we fed back statistical summaries of the previous round through percentages of the aggregated group's responses. This was presented visually in stages 2 and 3. We also fed back written summaries of the open-ended responses.

Consensus

Consensus was defined *apriori* as percentage agreement, whereby 80% or more of respondents rated agreement with an item consensus was deemed to be reached for that item.

Ethical Considerations

Ethical approval was sought through the University of Glasgow College of Medical, Veterinary & Life Sciences Ethics Committee for Non-Clinical Research Involving Human Participants (Appendix 2.6). A Data Protection Impact Assessment (DPIA) and data management plan was also completed.

Results

Participant Characteristics

Table 6 provides an overview of participant characteristics. The majority of participants were female and there were a range of ages represented in the sample. Participants from many different countries took part with the majority based in the UK. In terms of professional and personal identities in relation to psychosis, Table 6 shows the number of participants per round who identified as having lived experience of psychosis, experience of supporting someone, experience as a clinician, and research experience. Many participants identified as having experience in multiple areas, 34% of the sample in stage 1 (n=40), 48% in stage 2 (n=21) and 54% in stage 3 (n=21). For those with lived experience of psychosis, the majority had experienced multiple episodes of psychosis over a number of years. In terms of clinical diagnoses, 'schizoaffective disorder', 'bipolar disorder' and 'schizophrenia' were the most common diagnoses. The majority of supporters had cared for a family member over a number of years. In terms of professionals, the majority of clinicians were Clinical Psychologists and the most common research area was psychological interventions.

	Stage 1 (n=119)	Stage 2 (n=44)	Stage 3 (n=39)
Gender	female (n=72); male (n=31); non-binary/third gender (n=14); transgender (n=1); prefer not to say (n=1)	female (n=23), male (n=12), non-binary/third gender (n=5)	female (n=23), male (n=12), non-binary/third gender (n=4)
Age	18-24 (n=5), 25-34 (n=31); 35-44 (n=21); 45-54 (n=15); 55-64 (n=19); >65 (n=11)	18-24 (n=2), 25-34 (n=7), 35-44 (n=9), 45-54 (n=8), 55-64 (n=9), >65 (n=6), prefer not to say (n=1)	25-34 (n=9), 35-44 (n=5), 45-54 (n=9), 55-64 (n=10), >65 (n=6)
Ethnicity	White/Caucasian (n=101), Asian (n=6), Mixed race (n=3), Black British African (n=1), Black (n=1), Black British Caribbean (n=1), Human (n=1), Hispanic (n=1), Jewish (n=1), prefer not to say (n=1), Mestizo (n=1) Middle Eastern/North African (n=1)	White/Caucasian (n=39), Asian (n=1), Black (n=1) Human (n=1), Flemish (n=1), Irish (n=1)	White/Caucasian (n=32), Asian (n=2), Mixed race (n=1), Black British African (n=1) Human (n=2), Flemish (n=1), Irish (n=1)
Country	UK (n=57); US (n=11); Canada (n=5); Australia (n=3); Belgium (n=2); Brazil (n=1); Catalonia (n=1); Croatia (n=1); Egypt (n=1); Germany (n=2); Hungary (n=1); India (n=1); Ireland (n=1); Italy (n=1), Japan (n=1); Netherland (n=2), Peru (n=1); Taiwan (n=1)	UK (n=20), USA (n=7), Canada (n=3), Belgium (n=2), Hungary (n=1), Spain (n=1), The Netherlands (n=1), Australia (n=1), Brazil (n=1), Taiwan (n=1)	UK (n=21), USA (n=7), Canada (n=2), Belgium (n=2), Germany (1), Hungary (1), Ireland (1), The Netherlands (n=1), Australia (n=1), Brazil (n=1), Taiwan (n=1)

Expert by Experience	**n=52 (43.7%)**	**n=17 (38.6%)**	**n= 14 (35.9%)**
Diagnosis	'schizoaffective disorder' (n=10), 'schizophrenia' (n=9), bipolar disorder (n=7), psychosis NOS (n=4), Other (n=7)	'schizoaffective disorder' (n=4), 'schizophrenia' (n=2), bipolar disorder (n=4), psychosis NOS (n=2), Other (n=4)	'schizoaffective disorder' (n=3), 'schizophrenia' (n=1), bipolar disorder (n=2), psychosis NOS (n=1), Other (n=7)
No. of years since 1st episode	<1 (n=2), 1-3 (n=4), 4-10 (n=10), 10+ (n=28)	1-3 (n=1), 4-10 (n=5), 10+ (n=9)	<1 (n=1), 1-3 (n=2), 4-10 (n=3), 10+ (n=9)
No. of 'episodes'	1 (n=7), 2-3 (n=15), 3+ (n=23)	1 (n=1), 2-3 (n=5), 3+ (n=9)	1 (n=2), 2-3 (n=5), 3+ (n=8)
Supporter	**n=36 (30.3%)**	**n=18 (40.9%)**	**n=15 (38.5%)**
Relationship	parent (n=11), family member (n=10), partner (n=3), friend (n=3), other (n=1)	parent (n=6), family member (n=4), partner (n=2), friend (n=3), other (n=1)	parent (n=5), family member (n=5), partner (n=1), friend (n=2), other (n=2)
Years supported	<1 (n=1), 1-3 (n=8), 4-10 (n=5), >10 (n=14)	<1 (n=1), 1-3 (n=2), 4-10 (n=5), >10 (n=10)	<1 (n=1), 1-3 (n=4), 4-10 (n=6), >10 (n=4)
Clinician	**n=46 (38.7%)**	**n=18 (40.9%)**	**n=17 (43.6%)**
Profession	clinical psychologist (n=17), mental health nurse (n=3), psychiatrist (n=7), social worker/mental health officer (n=1), other (n=9), support worker (n=1)	clinical psychologist (n=8), mental health nurse (n=2), psychiatrist (n=2), social worker/mental health officer (n=2), trainee (n=2), support worker (n=1)	clinical psychologist (n=4), mental health nurse (n=3), psychiatrist (n=6), social worker/mental health officer (n=2), trainee (n=2)
Years	<1 (n=4), 1-3 (n=8), 4-10 (n=6), >10 (n=21)	1-3 (n=3), 4-10 (n=3), >10 (n=12)	1-3 (n=3), 4-10 (n=1), >10 (n=13)
Researcher	**n=32 (26.9%)**	**n=18 (40.9%)**	**n=17 (43.6%)**
Research area	Psychological interventions (n=21), symptom understanding & classification (n=13), pharmacological interventions(n=9), environmental & psychosocial processes (n=9), epidemiology/public health (n=6), genetics & neurobiology (n=4), other (n=2)	Psychological interventions (n=11), symptom understanding & classification (n=8), pharmacological interventions (n=5), environmental & psychosocial processes (n=5), epidemiology/public health (n=2), genetics & neurobiology (n=4), other (n=1)	Psychological interventions (n=11), symptom understanding & classification (n=7), pharmacological interventions (n=5), environmental & psychosocial processes (n=5), epidemiology/public health (n=3), genetics & neurobiology (n=4), other (n=1)
Years	4-10 (n=9), >10 (n=21)	4-10 (n=7), >10 (n=11)	1-3 (n=1) 4-10 (n=6), >10 (n=10)

Table 6: Participant Characteristics

Overview of Items per Delphi Round

Figure 2 provides a visual overview of the number of items per Delphi round. In addition to the items rated by participants using likert scales, there were an additional 14 open-ended items in stage 1, and 1 open-ended item in stages 2 and 3.

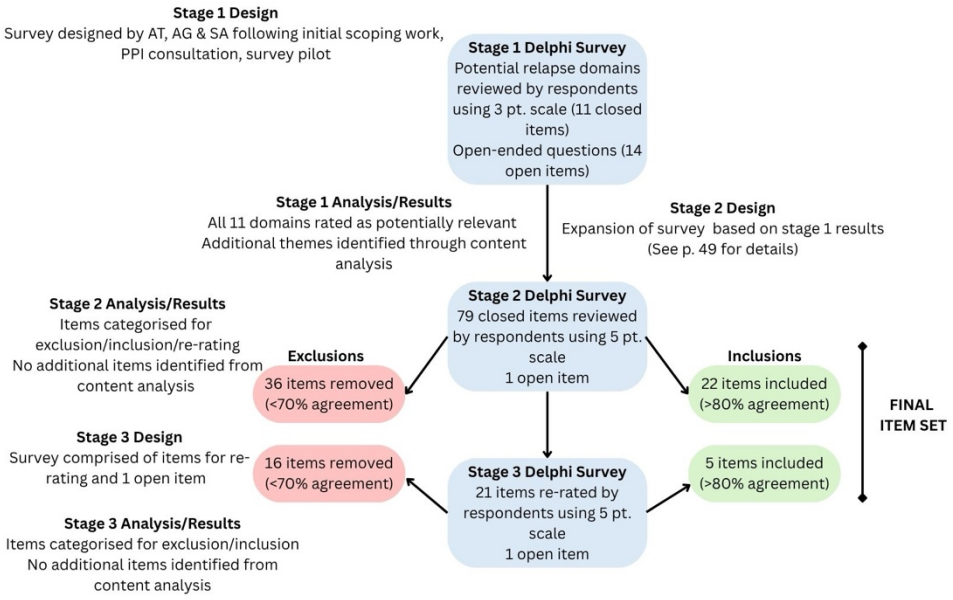

Figure 2: Flowchart of items in each round of the Delphi survey

Stage 1 Results

Appendix 2.5 provides participant responses to each of the items in stage 1 of the Delphi survey. No domains met the criteria for exclusion at this stage and therefore follow-up questions for each domain were included in stage 2.

Responses to Open-ended Questions

Following the first-stage of the survey, additional questions were added to the survey informed by participants' open-ended responses. These were:

1. 'How we think about relapse' e.g. relapse as a broad construct of many different experiences, versus separately measuring different domains.
2. 'Relapse as personal' - consideration of individuals' unique experiences and how these may be incorporated into a definition of relapse.
3. 'Diverse perspectives' - an acknowledgement that different stakeholders may hold different perspectives, and considerations of whose views are important in defining relapse in psychosis.
4. Application of the concept of 'relapse' and the language used.
5. Sleep changes.
6. Changes to 'self-agency'.
7. Changes to level of 'insight' into becoming unwell/need for additional support.

Based on participants' responses in stage 1, the questionnaire was elaborated for stage 2.

Stage 2 and 3 Results

A total of 27 items met consensus criteria following the three Delphi rounds. Table 7 details items relating to general considerations for defining relapse such as different stakeholder views and usefulness of the construct.

Item	Stage consensus reached	Percentage Agreement
It would be good practice to separately measure relapse in other domains such as mania, emotional distress etc. as separate to relapse in psychotic symptoms/experiences, for people who have experienced psychosis.	3	82%
Each individual has a personal relapse signature. It is not so much the specific experiences that matter, but the extent to which they cause the person distress and interfere with them being able to do the things they want to do.	2	85%
It would be helpful for people with lived experience to have input into a personalised definition of what relapse is for them.	2	83%
A person's own unique baseline of experiences should be taken into account in a relapse definition of psychosis.	2	85%
The views of the person experiencing psychosis are important when conceptualising relapse in psychosis.	2	92%
The views of carers are important when conceptualising relapse in psychosis.	2	86%
The views of clinicians are important when conceptualising relapse in psychosis.	2	92%
Relapse should be defined balancing the views of different stakeholders e.g. people with lived experience, carers, clinicians and researchers.	2	81%
I think the person experiencing psychosis should have the chance to use their own language to define their experiences when dealing with professionals.	2	86%
I think the term 'relapse' is useful in research settings e.g. clinical trials of medicines and therapies.	2	81%
I think the term 'relapse' is helpful in clinical settings e.g. when thinking about the support/interventions to offer someone.	2	81%

Table 7: General terms endorsed by 80% or more of participants in stages 2 & 3

Table 4 details the domain specific items that met consensus criteria. These related to: psychosis, cognitive disorganisation, mania and their impacts in terms of distress, functioning and risk.

Domain	Item	Stage Included	Percentage agreement
Psychosis	The severity of experiences of psychosis should be considered as part of a relapse definition	2	93%
	The duration of experiences of psychosis should be considered as part of a relapse definition.	3	85%

	The distress associated with psychosis should be considered as part of a relapse definition.	2	90%
	The impact on a person's functioning (i.e. ability to carry out the usual tasks of daily living) of experiences of psychosis should be considered as part of a relapse definition.	2	95%
	An increase in levels of agitation (e.g. irritability, anger, aggression) related to experiences of psychosis should be included as part of a definition of relapse of psychosis.	2	82%
	Increases of risk to self associated with experiences of psychosis should be considered as part of a definition of relapse in psychosis.	2	92%
	An increase in the level of risk to others (e.g. thoughts of harming others, behaviours, plans) associated with psychosis should be included as part of a definition of relapse in psychosis.	2	84%
	A change in behaviour (e.g. behaving in a way that is not usual for that person and is noticeable to others) associated with experiences of psychosis should be included as part of a definition of relapse.	2	81%
Cognitive Disorganisation	Cognitive disorganisation should be included as part of a relapse definition in psychosis	3	89%
	Cognitive disorganisation should be measured as a separate relapse domain.	2	80%
	The severity of cognitive disorganisation should be considered as a separate relapse domain.	2	80%
Mania	An increase in symptoms of mania/hypomania should be measured as a separate relapse domain.	2	82%
	The severity of mania/hypomania should be considered as a separate relapse domain.	3	81%
	The duration of mania/hypomania should be considered as a separate relapse domain.	2	82%
	The impact on a person's functioning (i.e. ability to carry out the usual tasks of daily living) of mania should be considered as a separate relapse domain.	3	86%
Risk	The severity of risk to self should be included as a separate relapse domain.	2	81%

Table 8: Domain specific items endorsed by 80% or more of participants in stages 2 & 3

Checklist for conceptualising relapse

Appendix 2.7 contains a checklist informed by the results of the Delphi process and endorsed items outlined in tables 3 and 4 as an output of the research. In summary, the checklist includes general considerations that were endorsed by the Delphi panel such as balancing multiple stakeholder's perspectives and taking into account an individual's unique baseline of experiences and characteristics when unwell. In addition, domain specific considerations are identified which were:

1. Experiences of psychosis (positive symptoms) and their impact e.g. on distress and functioning
2. Experiences of cognitive disorganization and impact

3. Experiences of mania and impact
4. Additional mental health domains to consider.

Discussion

This study used Delphi methodology with the aim of building consensus among key stakeholders regarding the concept of relapse in psychosis. Findings are discussed in the context of this study's strengths and limitations.

Areas of relative consensus and non-consensus among key stakeholders have been identified which could be progressed and further explored by future research. For example, there was consensus in our sample that relapse should be defined balancing the views of different stakeholders and people with lived experience should have the opportunity to input into a personalised definition. In terms of symptom profiles, there was consensus that experiences of psychosis, cognitive disorganization and mania are all important in conceptualising relapse. The sample was in agreement that it was important to consider other domains alongside psychosis. However, a broad definition of relapse (incorporating multiple domains) did not meet agreement thresholds suggesting that an overly inclusive definition of relapse was not endorsed. There was agreement that within each domain that impacts on functioning, distress and risk should be included in a definition.

There were domains e.g. 'negative symptoms' that did not meet consensus thresholds in the current study. The lack of inclusion in a relapse definition at this stage does not necessarily negate the importance of such experiences generally for someone who may experience psychosis, instead it suggests that these experiences are not regarded as central to defining relapse in psychosis. 'Emotional distress' did not meet the thresholds for inclusion as a domain in its own right, however there was agreement that distress related to experiences of psychosis should be included as part of a relapse definition in psychosis. Emotional distress is known to be closely associated with experiences of psychosis with up to 60% prevalence of depression in people diagnosed with Schizophrenia (Upthegrove et al., 2017). Therefore, it is interesting that our sample did not endorse emotional distress as a separate relapse domain. It is possible that this is related to the framing of the questions in the questionnaire which emphasised relapse in psychosis primarily, with other domains considered additionally. Factor analytic studies have shown that mood symptoms are a distinct dimension and so it is possible our sample were considering these symptoms as quite separate to relapse in psychosis (Upthegrove et al., 2017).

This would be worth exploring in future research, given that mood symptoms are often under-treated clinically for people who also experience psychosis (Gregory et al., 2017). Studies have found evidence for a close relationship between emotional distress and psychosis relapse, for example a recent review found that changes in mood were a key predictor of worsening symptoms of psychosis along with increased suspiciousness and sleep disturbances (Gleeson et al., 2024). It is possible that changes in emotional distress may in part represent an earlier stage of the relapse signature in psychosis which may account for distress not reaching consensus in the Delphi. Further research aimed at elucidating the nature of the relationship between emotional distress and psychosis relapse at different stages (e.g. pre/during/post psychosis) could help to clarify whether emotional distress may be worth considering as a key relapse domain in and of itself, in addition to distress directly associated with psychotic experiences. I

In terms of comparison to other definitions of relapse, previous research has focused on standardised definitions of relapse (e.g. Burns et al., 2000; Howes et al., 2025; San et al., 2015). In contrast, our sample endorsed items relating to personalising a definition e.g. giving people the opportunity to input into a personalised definition of what relapse may look like for them. This marks an important departure from previous work and may be in part reflective of the broader range of stakeholders included in the current sample. Future research could aim to explore this further and consider how personalising relapse definitions could be operationalised in clinical and research settings.

Previous work has focussed on relapse for those who have previously been diagnosed with 'schizophrenia' whereas our study took a transdiagnostic approach meaning that the recommendations could be relevant for a broader group e.g. those diagnosed with 'psychosis NOS', 'schizoaffective disorder and 'bipolar disorder''. A key difference between our findings, and a recent relapse definition published by Howes et al. (2025), is that where Howes et al. (2025) emphasised observable aspects of relapse e.g. service use, "significantly disturbed, new onset highly risky/dangerous behaviour" (Howes et al., 2025, Table 2), our findings suggest that the subjective experience of relapse is important to consider e.g. the distress associated with psychosis. Similar to a previous study characterising relapse in UK clinical practice, our findings suggest that an increase in symptoms/experiences of psychosis alone, are not sufficient to define relapse (Burns et al., 2000). Instead, it is important to consider the impact on the person and those around them e.g. in terms of distress, functioning and risk.

Similar to our study, Howes et al. (2025) aimed to include the perspectives of people with lived experience of psychosis into a relapse definition using the Delphi method. However, the extent of involvement of experts by experience was unclear, for example the number of participants with EbE was not reported. Epistemic injustice refers to an injustice done to someone in their capacity as an agent of knowledge (Fricker, 2007). There is a risk of epistemic injustice if the consensus views reported do not in fact adequately represent or report perspectives of people with lived experience (Colder Carras et al 2022, Veldmeijer & van Os 2025). These risks can be reduced by following good practice for engaging experts by experience e.g. through clearly defined roles, valuing of lived experience knowledge and transparently reporting extent of involvement (Colder Carras et al., 2023; Evans et al., 2025). Epistemic injustice is a particular risk when defining a concept such as relapse which has real-world implications for people assessed as having a 'relapse' (for example participants in the current study highlighted that it can affect whether one is allowed to drive their car or adopt a child). In the current study, we aimed to facilitate meaningful involvement of people with lived experience of psychosis and their supporters as much as possible, transparently report involvement, and reflect on areas for improvement. While this has meant that we cannot report fully operationalisable relapse criteria at present, we think this study makes a valuable contribution by presenting tentative areas of agreement and identifying areas for further research.

Strengths

To the best of our knowledge, this is the first study to investigate the concept of relapse in psychosis encompassing the views of multiple stakeholders throughout the Delphi process and seeking meaningful engagement of people with lived experience of psychosis and their supporters. In contrast to previous research, we used a transdiagnostic approach in the current study. Critique of a diagnostic approach towards psychosis and schizophrenia is well-documented (e.g. Bentall et al., 1988) and clinical diagnoses often do not reflect the lived experience of people receiving them. Therefore, we opted for a broader transdiagnostic approach in the current study. A strength of this study is that we followed best practice guidelines with regards to engaging multi-stakeholders including those with lived experience as much as possible (Colder Carras et al., 2023; Evans et al., 2025). For example, the initial survey design was reviewed by different stakeholders including researchers and people with lived experience of psychosis with feedback informing the final questionnaire design. We used open-ended questions during each survey round to give participants the opportunity to influence the questions asked at the next round and broaden the content. We also made efforts to ensure

that the questionnaire was accessible to a wide audience, e.g. using visuals to demonstrate concepts, using a variety of open and closed questions.

Limitations

"Participatory assimilation is the practice of involving individuals with lived experience in design and research processes...insofar as their perspectives and contributions can be integrated into or are supportive of researchers' preconceived, theory driven problem definitions and solution strategies" (Veldmeijer & van Os, 2025) p.2.

A limitation of this study is that there remains a risk of losing some of the richness and nuance of participants views, particularly those of people with lived experience of psychosis. This study used a consensus building method to gather opinions on relapse, a clinical concept which does not necessarily align with the ways people experience psychosis and would describe their experiences. Therefore, we run the risk of integrating diverse perspectives within our existing knowledge and assumptions of the concept of 'relapse'. Furthermore, the full sample was not retained throughout the three Delphi rounds, with a 67% attrition rate which is higher than previous Delphi studies of relapse (San et al., 2015; Burns et al., 2000; Howes et al., 2025) This may be accounted for, in part, by time constraints resulting in the first round of the Delphi study being 'live' for a longer period than that second and third rounds. We attempted to account for attrition in our initial sample size calculation and, given that a key aim of this study was to include the perspectives of those with lived experience of psychosis, it was important to ensure a broad sample during the early stages to include as many individual perspectives as possible. A limitation is the use of online data collection may have also risked excluding potential participants and therefore the sample may not be representative (Greer et al., 2009). It is therefore important to excise due caution when interpreting results as areas 'consensus' will not reflect the views of every person who took part in this research. Furthermore, while we used widely adopted criteria for 'consensus', it could be argued that the consensus cut-offs are somewhat arbitrary.

While we attempted to incorporate opportunities for participants to broaden the phenomena encompassed within the construct of 'relapse' through open-ended questions, it was not within the scope of this project to take a 'bottom-up' approach to the phenomenology such as grounded theory. Therefore, methodological decisions were made by AT in discussion with AG and SA based on content analysis of open ended responses e.g. which experiences to include as additional items in subsequent survey rounds, which likely did not fully capture the diversity of

experiences shared in the open-ended questions. This was reflected upon as a research team and recommendations for future research have been provided which could help to address this knowledge gap. It would have been beneficial to have had a second coder to enhance reliability of coding open-ended responses. A further limitation is that the majority of individuals who identified as clinicians were Clinical Psychologists, and it would have been beneficial to have a more diverse range of clinical professions including Nurses, Psychiatrists and Occupational Therapists given that these are a key professional groups in the provision of care and support for someone experiencing a relapse in psychosis.

Next Steps

Further research using qualitative and participatory methodologies would be beneficial in exploring the concept of relapse more broadly with people who have lived experience of psychosis. A previous Delphi study examining the concept of 'recovery' in psychosis highlighted many first-person accounts of 'recovery' and expert by experience defined measures of recovery which provided the context for the Delphi study (Law & Morrison, 2014). In contrast, there is little literature exploring the subjective experience of 'relapse' in psychosis, and no user-defined measures of relapse at the time of writing, to the best of our knowledge. Therefore, to reduce the risk of assimilation of lived experience perspectives, we suggest that further research could explore the lived experience of 'relapse' and consider how people who may be assessed clinically as experiencing 'relapse' would characterise and talk about their experiences. This could help to lessen the gap between clinical concepts and lived experience. A recent review highlighted lived experiences of people who had experienced repeated episodes of psychosis to include "*grieving personal losses, feeling split, and struggling to accept the constant inner chaos, the new self, the diagnosis and an uncertain future."* p. 168 (Fusar-Poli et al., 2022) such experiences do not neatly 'map' onto clinical constructs such as relapse.

Recommendations for defining relapse

Despite the limitations of the current research, there were areas of consensus in the current study that could be implemented into existing clinical and research operationalisations of 'relapse' in psychosis. For instance, using the person's own language as far as possible, balancing the views of multiple stakeholders and ensuring that the views of the person experiencing psychosis are included. In terms of specific domains, focusing on the experiences of psychosis, cognitive disorganisation, mania and the impact of these in terms of the associated distress, impact on functioning, and risk. Further research could further explore the views of people with lived experience using bottom-up methodology such as grounded theory. This may help to

contextualise some of the findings of the current research such as the lack of endorsement of emotional distress as a separate relapse domain.

Conclusion

'Relapse' is a clinical concept used to describe repeated experiences of psychosis. Lived experience perspectives of relapse have been under-represented in existing definitions of relapse used in clinical and research contexts, and which have serious real-world implications for people deemed to be experiencing a 'relapse'. The current study aimed to identify relative areas of consensus for relapse among multiple stakeholders including people with lived experience of psychosis and their supporters, which could be further explored by future research.

References

Allan, S., Bradstreet, S., Mcleod, H., Gleeson, J., Farhall, J., Lambrou, M., ... Gumley, A. (2019). Service Users, Carers and Mental Health Staff understandings of early warning signs of relapse in psychosis: A qualitative investigation. *BJPsych Open*, 1–7. https://doi.org/10.1192/bjo.2019.88

Andreasen, N. C., Carpenter Jr, 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*(3), 441-449.

Barrett, D., & Heale, R. (2020). What are Delphi studies?. Evidence-based nursing, 23(3), 68-69.

Beedie, S. (2019) Antipsychotic treatment for relapse prevention in psychosis: A systematic review of relapse outcomes and bias in randomised controlled trials of antipsychotic discontinuation. DCLinPsy research portfolio submission, University of Glasgow.

Bentall, Richard P., Howard F. Jackson, and David Pilgrim. "Abandoning the concept of 'schizophrenia ': Some implications of validity arguments for psychological research into psychotic phenomena." *British journal of clinical psychology* 27.4 (1988): 303-324.

Burns, T., Fiander, M., & Audini, B. (2000). A Delphi approach to characterising "relapse" as used in UK clinical practice. *International Journal of Social Psychiatry*, *46*(3), 220–230. https://doi.org/10.1177/002076400004600308

Conry-Murray, C., Waltzer, T., DeBernardi, F. C., Fossum, J. L., Haasova, S., Matthews, M. S., ... & Elsherif, M. M. (2024). Validity and transparency in quantifying open-ended data. *Advances in Methods and Practices in Psychological Science*, *7*(4), 25152459241275217.

Colder Carras, M., Machin, K., Brown, M., Marttinen, T. L., Maxwell, C., Frampton, B., ... & Jones, N. (2023). Strengthening review and publication of participatory mental health research to promote empowerment and prevent co-optation. *Psychiatric services*, *74*(2), 166-172.

Cooper, R. E., Laxhman, N., Crellin, N., Moncrieff, J., & Priebe, S. (2019). Psychosocial interventions for people with schizophrenia or psychosis on minimal or no antipsychotic medication: A systematic review. *Schizophrenia Research*, (xxxx). https://doi.org/10.1016/j.schres.2019.05.020

de Zwart, P. L., Jeronimus, B. F., & de Jonge, P. (2019). Empirical evidence for definitions of episode, remission, recovery, relapse and recurrence in depression: a systematic review. *Epidemiology and psychiatric sciences*, *28*(5), 544-562.

Evans, J., Lewin, C. D. C., Fabian, H., Alam, M., Calder, P., Hill, R., ... & Wykes, T. (2025). Facilitators of and barriers to patient and public involvement in mental health research within university settings: a systematic review and meta-synthesis. *Psychological Medicine*, *55*, e297.

Frank, E., Prien, R. F., Jarrett, R. B., Keller, M. B., Kupfer, D. J., Lavori, P. W., ... & Weissman, M. M. (1991). Conceptualization and rationale for consensus definitions of terms in major depressive disorder: remission, recovery, relapse, and recurrence. *Archives of general psychiatry*, *48*(9), 851-855.

Franklin, M., Mukuria, C., Mulhern, B., Tran, I., Brazier, J., & Watson, S. (2019). Measuring the Burden of Schizophrenia Using Clinician and Patient-Reported Measures: An Exploratory Analysis of Construct Validity. *Patient*, *12*(4), 405–417. https://doi.org/10.1007/s40271-019-00358-x

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford university press.

Fusar-Poli, P., Estradé, A., Stanghellini, G., Venables, J., Onwumere, J., Messas, G., ... & Maj, M. (2022). The lived experience of psychosis: a bottom-up review co-written by experts by experience and academics. *World Psychiatry*, *21*(2), 168-188.

Gleeson, J. F., Rawlings, D., Jackson, H. J., & McGorry, P. D. (2005). Early warning signs of relapse following a first episode of psychosis. *Schizophrenia research*, *80*(1), 107-111.

Greer, B., Robotham, D., Simblett, S., Curtis, H., Griffiths, H., & Wykes, T. (2019). Digital exclusion among mental health service users: qualitative investigation. *Journal of medical Internet research*, *21*(1), e11696.

Gregory, A., Mallikarjun, P., & Upthegrove, R. (2017). Treatment of depression in schizophrenia: systematic review and meta-analysis. *The British Journal of Psychiatry*, *211*(4), 198-204.

Gumley, A. I., Bradstreet, S., Ainsworth, J., Allan, S., Alvarez-Jimenez, M., Aucott, L., ... & 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*(6), 477-486.

Hasson, F., Keeney, S., & Mckenna, H. P. (2000). Research guidelines for the Delphi Survey Technique Research guidelines for the Delphi survey technique, (February 2019). https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x

Howes, O. D., Bukala, B. R., Chen, E. Y., Correll, C. U., Hasan, A., Honer, W. G., ... & McCutcheon, R. A. (2025). Relapse in schizophrenia: a systematic review of criteria for clinical studies and international consensus guidelines to improve them. *American Journal of Psychiatry*, *182*(11), 969-983.

Kishimoto, T., Agarwal, V., Kishi, T., Leucht, S., Kane, J. M., & Correll, C. U. (2013). Relapse prevention in schizophrenia: a systematic review and meta-analysis of second-generation antipsychotics versus first-generation antipsychotics. *Molecular psychiatry*, *18*(1), 53-66.

Kishimoto, T., Nitta, M., Borenstein, M., Kane, J. M., & Correll, C. U. (2013). Long-acting injectable versus oral antipsychotics in schizophrenia: a systematic review and meta-analysis of mirror-image studies. *J Clin psychiatry*, *74*(10), 957-965.

Kishimoto, T., Hagi, K., Kurokawa, S., Kane, J. M., & Correll, C. U. (2021). Long-acting injectable versus oral antipsychotics for the maintenance treatment of schizophrenia: a systematic review and comparative meta-analysis of randomised, cohort, and pre–post studies. *The Lancet Psychiatry*, *8*(5), 387-404.

Law, H., & Morrison, A. P. (2014). Recovery in psychosis: a Delphi study with experts by experience. *Schizophrenia bulletin*, *40*(6), 1347–1355. https://doi.org/10.1093/schbul/sbu047

Leucht, S., Cipriani, A., Spineli, L., Mavridis, D., Örey, D., Richter, F., ... & Davis, J. M. (2013). Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *The Lancet*, *382*(9896), 951-962.

Molloy, R., Brand, G., Munro, I., & Pope, N. (2023). Seeing the complete picture: a systematic review of mental health consumer and health professional experiences of diagnostic overshadowing. *Journal of Clinical Nursing*, *32*(9-10), 1662-1673.

Moncrieff, J., Crellin, N. E., Long, M. A., Cooper, R. E., & Stockmann, T. (2019). Definitions of relapse in trials comparing antipsychotic maintenance with discontinuation or reduction for schizophrenia spectrum disorders: A systematic review. *Schizophrenia Research*, (xxxx). https://doi.org/10.1016/j.schres.2019.08.035

Morrison, A. P., & Barratt, S. (2010). What are the components of CBT for psychosis? A Delphi study. *Schizophrenia bulletin*, *36*(1), 136–142. https://doi.org/10.1093/schbul/sbp118

NICE. (2014). Psychosis and schizophrenia in adults: chosis and schizophrenia in adults: prevention and management. *Www.Nice.Org.Uk*, (February 2014). Retrieved from https://www.nice.org.uk/guidance/cg178/resources/psychosis-and-schizophrenia-in-adults-prevention-and-management-pdf-35109758952133

Okoroji, C., Mackay, T., Robotham, D., Beckford, D., & Pinfold, V. (2023). Epistemic injustice and mental health research: A pragmatic approach to working with lived experience expertise. *Frontiers in psychiatry*, *14*, 1114725.

Porcelli, S., Bianchini, O., De Girolamo, G., Aguglia, E., Crea, L., & Serretti, A. (2016). Clinical factors related to schizophrenia relapse. *International Journal of Psychiatry in Clinical Practice*, *20*(2), 54–69. https://doi.org/10.3109/13651501.2016.1149195

Robinson, D., Woerner, M. G., Alvir, J. M. J., Bilder, R., Goldman, R., Geisler, S., ... Lieberman, J. A. (1999). Predictors of relapse following response from a first episode of schizophrenia or schizoaffective disorder. *Archives of General Psychiatry*, *56*(3), 241–247. https://doi.org/10.1001/archpsyc.56.3.241

San, L., Serrano, M., Cañas, F., Romero, S. L., Sánchez-Cabezudo, Á., & Villar, M. (2015). Towards a pragmatic and operational definition of relapse in schizophrenia: A Delphi consensus approach. *International Journal of Psychiatry in Clinical Practice*, *19*(2), 90–98. https://doi.org/10.3109/13651501.2014.1002501

Scottish Intercollegiate Guidelines Network (SIGN). (2013). *Management of schizophrenia. A national clinical guideline*. *Sign*. https://doi.org/10.1136/thx.2008.097741

Sforzini, L., Worrell, C., Kose, M. *et al.* A Delphi-method-based consensus guideline for definition of treatment-resistant depression for clinical trials. *Mol Psychiatry* **27**, 1286–1299 (2022). https://doi.org/10.1038/s41380-021-01381-x

Sinha, I. P., Smyth, R. L., & Williamson, P. R. (2011). Using the Delphi Technique to Determine Which Outcomes to Measure in Clinical Trials: Recommendations for the Future Based on a

Systematic Review of Existing Studies. *PLoS Medicine*, *8*(1), e1000393. https://doi.org/10.1371/journal.pmed.1000393

Subramani, S. (2019). Practising reflexivity: Ethics, methodology and theory construction. *Methodological Innovations*, *12*(2), 2059799119863276.

Taipale, H., Schneider-Thoma, J., Pinzón-Espinosa, J., Radua, J., Efthimiou, O., Vinkers, C. H., ... & Luykx, J. J. (2022). Representation and outcomes of individuals with schizophrenia seen in everyday practice who are ineligible for randomized clinical trials. *JAMA psychiatry*, *79*(3), 210-218.

Upthegrove, R., Marwaha, S., & Birchwood, M. (2017). Depression and schizophrenia: cause, consequence, or trans-diagnostic issue?. *Schizophrenia bulletin*, *43*(2), 240-244.

van Os, J., Burns, T., Cavallaro, R., Leucht, S., Peuskens, J., Helldin, L., ... & Kane, J. M. (2006). Standardized remission criteria in schizophrenia. *Acta Psychiatrica Scandinavica*, *113*(2), 91-95.

Vogel, C., Zwolinsky, S., Grif, C., Hobbs, M., Henderson, E., & Wilkins, E. (2019). A Delphi study to build consensus on the definition and use of big data in obesity research. https://doi.org/10.1038/s41366-018-0313-9

Appendix 1.1: PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Page no.
TITLE			
Title	1	Identify the report as a systematic review.	10
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	12
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	14
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	15
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	15
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	16, 18
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	16
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	17
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	17
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	17
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	17
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	17
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	N/A
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	N/A
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	N/A
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	17
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	17
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	N/A
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	N/A

Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	N/A
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	N/A
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	18, 19
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	18
Study characteristics	17	Cite each included study and present its characteristics.	22-26
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	30
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	N/A
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	27-31
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	N/A
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	27-31
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	N/A
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N/A
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	N/A
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	32
	23b	Discuss any limitations of the evidence included in the review.	33
	23c	Discuss any limitations of the review processes used.	33
	23d	Discuss implications of the results for practice, policy, and future research.	34
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	15
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	15
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	15
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	15
Competing interests	26	Declare any competing interests of review authors.	15
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Appendices

Prisma Checklist

Appendix 1.2: Spranger et al. (2022) Checklist

I.	TITLE AND ABSTRACT: This includes the identification of the study as a Delphi process in the title and abstract, and a structured summary of the study.
II.	EPISTEMOLOGY: This refers to the positioning in a specific strand of theory of science (realist or constructivist), as well as the formulation of the objective and the statement of preliminary assumptions, e.g., about possible conflicting perspectives.
III.	FORMAL CONTEXT: This includes information on the research team conducting the project and other key information on the project background (e.g., funder or objective of the project context, study protocol, ethics committee vote).
IV.	KNOWLEDGE BASE: This includes information on the consideration or integration of the current state of research, i.e., the evidence base. It should also be indicated if the designated respondents are aware of it or how it can be ensured that they are aware of necessary contexts or are informed about them and in which social context they are located.
V.	KNOWLEDGE AND KNOWLEDGE INTEGRATION: This concerns, in particular, the specification of which kind of knowledge stocks and experiences are relevant and which perspectives are necessary to answer the respective research question, as well as how the different knowledge stocks are potentially weighted and how these are to be brought together.
VI.	DELPHI VARIANT AND ROLE IN RESEARCH PROCESS: Statement and justification of which Delphi variant or which modification has been chosen and how they are situated in the research process.
VII.	SAMPLE: This includes a comprehensive description of the expert panel with regard to identification, recruitment, knowledge base, sociodemographic data, number of cases (also per round) and, if necessary, recruiting or dealing with refusers or dropouts. The handling of anonymity should also be reported.
VIII.	SURVEY INSTRUMENT: The survey instrument must be described in terms of its scope, structure, derivation and testing of questions or items, the ratio of open and closed questions, integrated scale types, and graphic design.
IX.	DELPHI ROUNDS: This includes the number of Delphi rounds, as well as disclosure of the termination criterion.
X.	FEEDBACK: Disclosure of statistical and graphic representation of feedback (also per round).
XI.	EVALUATION: This includes the definition of and dealing with consensus and disclosure of the quantitative and qualitative evaluation strategy, and also how these two aspects are weighted and combined.
XII.	RESULT: The Delphi process (response, procedure, modifications) and the results of the individual Delphi rounds are to be presented. If necessary, reasons should be given why specific elements were changed during the process. The complete procedure should also be presented graphically, e.g., in a flowchart.
XIII.	QUALITY OF DATA AND INTERPRETATION: This includes reflection on the quality of the data collection and evaluation process. Quality criteria for quantitative or qualitative research should be applied depending on the epistemological positioning and the specific Delphi variant.
XIV.	DISCUSSION AND LIMITATIONS OF THE FINDINGS: This refers to a critical reflection on the validity claim of the findings, also depending on the epistemological basis.
XV.	DISSEMINATION: Statement on how the findings will be processed or used beyond the Delphi study.

Spranger, J., Homberg, A., Sonnberger, M., & Niederberger, M. (2022).
Reporting guidelines for Delphi techniques in health sciences: a methodological review. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, *172*, 1-11.

Appendix 1.3: Data Extraction Template

Item (Spranger et al., 2022)	**Sub-items**	**Study details**
I. Title/abstract		
II. Epistemology	*realist/constructivist*	
	assumptions	
	power	
III. Formal Context	*Formal context*	
	Pre-registration	
	Ethical approvals	
iv Knowledge base	*Preliminary work*	
	Provide Participants' w information	
V. Knowledge & knowledge integration	*Sample rationale*	
	Multi vs single stakeholder	
	Combining knowledge	
VI. Sample	*Sample size*	
	Sample characteristics	
	Refusers/dropouts	
	Identification/Recruitment	

VIII. Survey Instrument	*Described*	
	Supplementary	
	Open-ended Qs	
	Scales	
IX. Delphi rounds	*no. rounds*	
	termination	
X. Feedback	*Statistical*	
	Personal	
XI. Evaluation	*Consensus def.*	
	A priori/posteriori	
	quant eval	
	qual eval	
XII. Result	*results to rounds*	
	description of process	
	graphical representation	
XIII. Quality of data & interp	*Qual criteria*	
	reflection	
XIV. Discussion and limitations		
Dissemination		
ADDITIONAL NOTES		

Appendix 1.4: Quality Appraisal

Author (date)/DCAT Items	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Overall Quality Rating
Addington et al. (2005)	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	U	Y	High
Addington et al. (2012)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	High
Agrest et al. (2024)	Y	Y	Y	N	Y	Y	Y	Y	Y	U	Y	Y	Y	U	U	Y	Moderate
Bryne & Morrisson (2014)	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	U	Moderate
Carswell et al. (2017)	U	Y	Y	Y	Y	Y	Y	U	Y	U	Y	Y	Y	U	N	Y	Moderate
Cottril et al. (2021)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	High
Fortuna et al. (2023)	U	U	U	N	Y	N	Y	U	U	U	U	U	N	U	N	U	Not rated
Greenwood et al. (2009)	Y	Y	Y	Y	Y	N	Y	Y	Y	U	Y	Y	Y	N	N	U	Moderate
Hultsjo et al. (2011)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	High
Jacobsen et al. (2024)	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Moderate
Langlands et al. (2008)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	High
Law & Morrisson (2014)	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	Moderate
Li et al. (2020)	Y	Y	Y	N	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	Moderate
McKenzie et al. (2022)	U	Y	Y	U	Y	N	Y	U	Y	U	Y	Y	U	Y	N	N	Moderate
Mitchell et al. (2021)	Y	Y	Y	N	Y	Y	Y	U	N	Y	Y	Y	N	Y	Y	U	Moderate
Requena et al. (2024)	Y	Y	Y	N	Y	Y	Y	U	Y	Y	Y	Y	U	Y	Y	U	Moderate

Key

Core Items

Item 1. Did the research team appropriately employ anonymity in the delphi study?
Item 2. Did the research team appropriately employ iteration in the delphi study?
Item 3. Did the research team appropriately employ statistical summaries of group responses in the delphi study?

Item 4. Did the research team appropriately employ controlled feedback in the delphi study?

Additional Items

Item 5. Did the research team provide an appropriate justification for the use of the delphi method in this study?
Item 6. Did the research team provide an appropriate justification for the type of delphi used in this study?
Item 7. Did the research team seek the appropriate expertise for the delphi study?
Item 8. Did the research team sufficiently retain panellists throughout the delphi study?
Item 9. Did the research team conduct preliminary work prior to the first round of the delphi study?
Item 10. Did the research team make appropriate design choices for the delphi study?
Item 11. Did the research team ask appropriate questions in the delphi study?
Item 12. Did the research team use appropriate response items in the delphi study?
Item 13. Did the research team conduct a pilot test before launching the delphi study?
Item 14. Did the research team appropriately analyse the data from the delphi study?
Item 15. Did the research team transparently report their delphi study?
Item 16. Did the research team appropriately consider ethical issues?

Reference

Khodyakov, D., Grant, S., Kroger, J., & Bauman, M. (2023). *RAND methodological guidance for conducting and critically appraising Delphi panels*. Rand.

Appendix 1.5: Transparency Ratings

https://osf.io/gak29/files/osfstorage/693f271085b65ee8e54878c0

Appendix 2.1: Study Proposal

https://osf.io/w987d/files/osfstorage/691c94a4e643abcb39770186

Appendix 2.2: Delphistar Checklist

Topic	Section	Item	Checklist Item	Location where item is reported	Exemplary answer
I Title and Abstract		1	Identification as a Delphi procedure in the title	p.40	What is a public health intervention? Results of a Delphi study.
		2	Identification as a Delphi procedure in the abstract	p.44	A Delphi procedure was selected to answer the research question.
		3	Structured abstract	p.44	e.g., background, method, results and discussion
II Context	**Formal**	4	Information about the sources of funding	p.49	The Delphi study was funded by [SOURCE].
		5	Information about the team of authors and/or researchers (e.g., discipline, institution)	p.48	The Delphi study was conducted by an interdisciplinary team with representatives from medicine, public health, and health promotion.
		6	Information about method consulting	p.49	The study group was advised by external experts from [INSTITUTION] regarding statistics. Or: No outside consulting in regard to method took place.
		7	Information about the project background	N/A	The Delphi survey was part of a mixed-methods study on [AIM].
		8	Information about the study protocol	p.49	The study protocol is available at [LINK].
	Content	9	Justification of the chosen method (Delphi procedure) to answer the research question	p.48	The Delphi method is suitable for answering the research question because it systematically gathers the judgments of different expert groups and can identity agreement and disagreement.
		10	Aim of the Delphi procedure (e.g., consensus, forecasting)	p.48	The aim of the Delphi study is to find consensus on criteria to define a public health intervention.
III Method	**Body & Integration of knowledge**	11	Identification and elucidation of relevant expertise, spheres of experience, and perspectives (e.g., theory, practice, affected groups, disciplines)	p.46-47	The experts should represent the sciences and clinical practice because [REASON].
		12	Handling of knowledge, expertise and perspectives which are missing or have been deliberately not integrated	N/A	If it is not possible to recruit experts specialized in [AREA], this is openly communicated to the other respondents during the Delphi process.

		13	Basic definition of expert[1]	Method – Inclusion criteria	A person who has been active in the area for at least [NUMBER] years is considered to be an expert.
	Delphi variations	14	Identification of the type of Delphi procedure and potential modifications (e.g., classic Delphi, real-time Delphi, group Delphi)	Method – Delphi Technique	A classic Delphi procedure was used [LITERATURE REFERENCE].
		15	Justification of the Delphi variation and modifications, including during the Delphi process, if applicable	N/A	If the willingness to participate clearly decreases between the first and second rounds, a third round will not be held.
	Sample of experts	16	Selection criteria for the experts (per round if there are different expert groups)	p.49	All of the experts who met the definition were invited to the first round. All of the experts who completed the previous round were invited to participate in the subsequent round.
		17	Identification of the experts	p.49-50	The experts were identified based on publications in [DATABASE].
		18	Information about recruiting and any subsequent recruiting of experts	p.49-50	The experts were informed about the Delphi study and invited to participate.
	Survey	19	Elucidation of the content development for the questionnaire[2]	p.50-52	The questionnaire was developed based on the results of systematic reviews [LITERATURE REFERENCE].
		20	Description of the questionnaire (content and structure)	p.52-53	The questionnaire was divided into three segments on [TOPICS]. The statements made in the questionnaire were evaluated using standardized items, with the option to comment in free-text boxes.
	Delphi rounds	21	Number of Delphi rounds	p.48	Three Delphi rounds were held.
		22	Information about the aims of the individual Delphi rounds	p.53	The first Delphi round focused on exploring relevant aspects. These aspects were then presented to the experts in the second Delphi round for standardized evaluation.
		23	Disclosure and justification of the criterion for discontinuation	p.48	The number of rounds was defined in advance to be a maximum of three rounds.
	Feedback	24	Information about what data was reported back per round	p.54	In terms of feedback, we shared the statistical results plus the summary of the open responses.
		25	Information on how the results of the previous Delphi round were fed back to the experts surveyed (e.g., via frequencies,	p.54	Mean values, standard deviations and percentage frequency distributions were reported.

			mean values, measures of dispersion, listing of comments)		
		26	Information on whether feedback was differentiated by specific groups (e.g., by field of expertise, institutional affiliation)	p.54	The feedback was aggregated across all expert groups.
		27	Information about how dissent and unclear results were handled	p.53	The results showing dissent were presented again for evaluation in the next Delphi round.
	Data analysis	28	Disclosure of the quantitative and qualitative analytical strategy	p.52	The quantitative items were descriptively analyzed. The open-ended items were analyzed using thematic analysis [LITERATURE REFERENCE].
		29	Definition and measurement of consensus	p.54	Consensus was defined as percentage agreement, meaning that agreement was assumed if at least 80% of the respondents agreed on an item.
		30	Information on group-specific analysis or weighting of experts (e.g., theory vs. practice, discipline-specific analysis)	N/A	In the analysis, the mean values for percent agreement are weighted for each expert group in terms of the number of group members.
IV Results	**Delphi process**	31	Illustration of the Delphi process (e.g., in a flow chart)	p.57	A summary of the process is illustrated in a flow chart (Figure 1).
		32	Information about special aspects during the Delphi process (e.g., deviations from the intended approach with justification)	N/A	During the Delphi procedures the political discussion mentioned climate change and the effects on health. It is possible that this influenced the experts' responses.
		33	Number of experts per round (both invited and participating)	p.56	The number of experts participating in the first Delphi round was [NUMBER], and the number of experts in the second round was [NUMBER]. This corresponds to a response rate of [NUMBER]% in the first round and [NUMBER]% in the second round.
	Results	34	Presentation of the results for each Delphi round and the final results	p.58, 59	In the first Delphi round [NUMBER]% of the respondents agreed, in the second [NUMBER]%, and in the third [NUMBER]%.
V Discussion	**Quality of findings**	35	Highlighting the findings from the Delphi study	p.60	The central findings can be summarized as follows: [STATE FINDINGS].

		36	Validity of the results (e.g., transferability of the findings)	p.63	The results are not transferable to other countries due to different legal regulations.
		37	Reliability of the results (e.g., split half, inter-rater reliability)	p.63	The responses in the free-text comments were analyzed by two independent reviewers [SPECIFY].
		38	Reflection on potential limitations (e.g., distortion, skewing, bias)	p.63	The results are to be viewed critically with regard to the composition of the panel because [REASONS].

Appendix 2.3: Delphi Surveys

Stage 1 (Includes Participant Information Sheet (PIS) and Consent Form)
https://osf.io/w987d/files/osfstorage/6931892bf28b726045aa460f

Stage 2
https://osf.io/w987d/files/osfstorage/6931892d7c769e5b1a7e466f

Stage 3
https://osf.io/w987d/files/osfstorage/6931892d671f613c2d363a5b

Appendix 2.4: Data Analysis Plan

https://osf.io/w987d/files/osfstorage/693eb92c57db7d3f9e4871f5

Appendix 2.5: Results for Delphi Rounds

https://osf.io/w987d/files/osfstorage/693ebc020d286aee82416b7a

Appendix 2.6: Ethical Approval

University of Glasgow

22nd March 2025

MVLS College Ethics Committee

Project Title: Developing a consensus definition of relapse in psychosis: A Delphi Study.
Project No: 200240278

The College Ethics Committee has reviewed your application and has agreed that there is no objection on ethical grounds to the proposed study. It is happy therefore to approve the project, subject to the following conditions:

- Project end date: End December 2025
- The data should be held securely for a period of ten years after the completion of the research project, or for longer if specified by the research funder or sponsor, in accordance with the University's Code of Good Practice in Research: https://www.gla.ac.uk/media/media_490311_en.pdf
- The research should be carried out only on the sites, and/or with the groups defined in the application.
- Any proposed changes in the protocol should be submitted for reassessment, except when it is necessary to change the protocol to eliminate hazard to the subjects or where the change involves only the administrative aspects of the project. The Ethics Committee should be informed of any such changes.
- You should submit a short end of study report to the Ethics Committee within 3 months of completion.
- For projects requiring the use of an online questionnaire, the University has an Online Surveys account for research. To request access, see the University's application procedure at https://www.gla.ac.uk/research/strategy/ourpolicies/useofonlinesurveystoolforresearch/.

Yours sincerely

Jesse Dawson
MD, BSc (Hons), FRCP, FESO
Professor of Stroke Medicine
Chair MVLS Research Ethics Committee

School of Cardiovascular and Metabolic Health
College of Medical, Veterinary & Life Sciences
Room M0.05
Office Block
Queen Elizabeth University Hospital
Glasgow
G51 4TF
Tel – 0141 451 5868
jesse.dawson@glasgow.ac.uk

Appendix 2.7: Checklist for defining relapse

Suggested Checklist for Conceptualising Relapse in Psychosis in Clinical and Research Settings

This document provides suggestions for conceptualising relapse based on the Delphi study.

General Guidance

Relapse should be defined balancing the views of multiple stakeholders i.e. the person who has experienced psychosis, their supporters, the clinical team. Relapse should *not* be defined solely on clinical judgement.

The person's own language to describe their experiences should be used as much as possible. It may be helpful to consider the following prior to assessment:

Consider the person's language	How do they talk about their experiences? How would they like you to talk about them?
Relevant mental health domains	What does the person identify as other key mental health domains relevant to them e.g. mood, anxiety, 'negative symptoms' etc.
Consider the person's support network	Who does the person consider as their key supporters?

A Priori

Prior to assessing for a potential relapse, it may be helpful to gather the following information from the person, their key supporters, and clinical team:

Consider the person's unique baseline when well *(This aims to capture the person's everyday fluctuations in their experiences)*	In relation to baseline: • Experiences of psychosis e.g. baseline voice hearing • Cognitive disorganisation e.g. tangential speech that is normal for that person • Mania e.g. fluctuations in affect • Other relevant mental health domains • Distress • Functioning • Risk
Consider the person's unique signature when unwell *(This aims to gather information about the characteristics that may indicate an episode of psychosis)*	In relation to changes in: • Experiences of psychosis • Cognitive disorganisation • Mania • Other relevant mental health domains And impact on:

	• Distress • Functioning • Risk

Assessing for Possible Relapse

Taking into account the person's own description of their experiences as defined by them and their key supporters, consider the following:

(Note that distress associated with experiences, impact on behaviour/functioning, and any associated risk should be given greater weighting than the experiences themselves)

Experiences of psychosis	• Associated distress • Impact on behaviour and functioning • Risk to self & others • Severity • Duration
Cognitive Disorganisation	• Associated distress • Impact on behaviour and functioning • Risk to self & others • Severity • Duration
Mania	• Associated distress • Impact on behaviour and functioning • Risk to self & others • Severity • Duration
Other Relevant Mental Health Domains	• Based on the person's unique baseline, past episodes of psychosis, and their current reporting Additionally consider: • Risk to self • Mania

Prior to identifying a relapse episode, ensure that the following have been consulted:

Notes detailing the person's unique baseline of experiences and signature when unwell	Details:
The person being assessed	
Key supporters of the person	

The person's clinical team	

Rating Relapse

You may wish to identify relapse if the following criterion are met:

1. The person, their key supporters and clinical team identify an increase in symptoms of: psychosis, cognitive disorganisation, mania which is out with the person's unique baseline

which is associated with either:

2. An increase in distress at a level which is out with the person's usual everyday fluctuations (identified A priori)
3. Impact on behaviour/functioning which is out with the person's usual everyday fluctuations (identified A priori)
4. Impact on risk to self and/or other which is out with the person's usual everyday fluctuations (identified A priori)

Where relapse is identified, it would also be useful to assess the other key mental health domains identified as important for the individual. It would also be useful to review the person's pre-identified 'signature' when unwell and consider whether the current difficulties are to some degree similar.

Important Considerations for Rating Potential Relapse

1. If the person is unwell to the extent that it is difficult to include their current views when rating relapse (e.g. if someone is particularly scared/confused/overwhelmed) then:
    a. Consider ways of supporting them to provide their views e.g. written materials such as diaries
    b. Consult with their key supporters
    c. Use the information they have already provided you with as outlined above
    d. Revisit at a later time to gain their perspective once they are well enough.
2. The 'unique baseline' and 'personal signature' are considered key to rating potential relapse, particularly to reduce the risk of diagnostic overshadowing i.e. when entirely unrelated personal changes risk being diagnosed as 'relapse'.
3. Relapse must be defined balancing the views of the identified stakeholders, including the person experiencing psychosis. In the event that there are conflicting perspectives then it may be helpful to consider:

a. Using supervision to identify areas of disagreement and consider what additional information may be helpful
b. Revisiting at a later time which may provide clarity regarding whether a relapse has occurred

[END OF THESIS]