

Psychometric Evaluation and Therapeutic Application of Quality-of-Life Outcome Measures in Psychiatric Inpatient Care

Sinead McLernon

A thesis submitted in partial fulfilment of the requirements for the degree of Doctorate in Clinical Psychology

Clinical Psychology Unit, Department of Psychology,

The University of Sheffield

May 2025

The results, discussions and conclusions presented herein are identical to those in the printed version. This electronic version of the thesis has been edited solely to ensure conformance with copyright legislation and all excisions are noted in the text. The final, awarded and examined version is available for consultation via the University Library

Declaration

This thesis has been prepared and submitted in partial fulfilment of the requirements for the award of Doctorate in Clinical Psychology at the University of Sheffield. I declare, this work has not been used or submitted for any other degree or institution to obtain any other qualifications. No funding has been received for this thesis, and I declare that I have no conflicts of interest. Furthermore, I declare, as the author of this thesis, that this work is my own and that I am aware of the University of Sheffield's guidance on unfair means.

Structure and Word Count

Section One: Literature Review

Word count excluding references and tables: 7975

Word count including references and tables: 20011

Section Two: Empirical Study

Word count excluding references and tables: 7970

Word count including references and tables: 12657

Total:

Word count excluding references and tables: 15945

Word count including references and tables: 32668

Lay summary

Literature Review

QoL outcome measures have become increasingly valuable in inpatient psychiatric settings. Outcome measures can enhance patient outcomes over time, and these measures should be relevant to the population they serve. To understand if the outcome measure is an appropriate and helpful measurement properties are assessed. Measurement properties encompass the various strands of scientific inquiry that establish whether a measure accurately assesses what it claims to measure and whether it maintains consistency over time.

A review of 38 studies identified 23 different QoL outcome measures that had been tested in at least a partial inpatient population. The outcome of the review suggests that no single outcome measure encompasses all measurement properties suitable for use in this setting. Several outcome measures showed promising evidence and could be adapted for this setting. Overall, it was clear that more patient and clinician involvement and feedback were needed on the content of the identified scales to make the scale relevant and valuable for adult psychiatric inpatient populations.

Empirical Study

DIALOG is a QoL and patient experience measure that has been tested in community settings. A patient may highlight areas they would like to work on when completing DIALOG, and DIALOG+ provides a method for structuring conversations and setting goals to work towards gradual change in those selected areas. DIALOG+ has been successfully tested in community settings; however, to our knowledge, it has not been tested in an inpatient setting. Therefore, the study aimed to determine if it is achievable in this setting.

The study's outcome revealed numerous challenges in applying DIALOG and DIALOG+ in this setting, including inadequate recruitment and engagement. Staff

highlighted barriers such as the need for additional training and coaching, greater team input, improved systems and leadership, as well as the unique aspects of rehabilitation settings.

Acknowledgments

To my research supervisor, Dr Claire Bone, thank you for always sharing your enthusiasm and excitement for research and the project. Your passion for research has been infectious, and your expertise and knowledge have been invaluable. Thank you for consistently offering encouragement and positive feedback when I needed it most.

I want to thank my supervisors from the University of Sheffield, Dr Jaime Delgadillo, who has provided consistent support and knowledge throughout the study. And Dr Scott Steen, who has provided invaluable feedback and support to the write-up of my review.

Thank you to all the staff and patients within the inpatient service for taking the time to engage in the research and provide feedback. I would especially like to thank Alex Fergusson, who continuously supported data collection and engagement throughout the trial.

I want to thank my parents. To my dad, who always knew that I wouldn't stop until I finished the Doctorate. To my mum, who has supported me continuously throughout this journey, I would not have had the faith and determination to keep going without her. And to my brothers, who are always just a phone call away.

Lastly, I would like to thank all the friends who supported me throughout the process with chocolate and chats. Thank you, Iain, for all the paddleboarding and adventure breaks, as well as your daily support. I also extend my gratitude to the DClinPsy gang; having a group of friends who fully understand the job we do and the effort we put into it has been invaluable throughout this journey.

Table of Contents

Part One: Literature Review *9*
Abstract *10*
Introduction *12*
Methods *15*
Results *20*
Discussion *83*
References *89*
Appendix *104*
Appendix A 104
Appendix B 105
Appendix C 105
Appendix D 112
Appendix E 113
Appendix F 114
Appendix G 115
Part Two: Empirical Study *116*
Abstract *117*
Introduction *119*
Method *122*
Quantitative Method *123*
Qualitative Method *129*
Results *136*
Discussion *155*
References *162*
Appendix *176*
Appendix A 184
Appendix B 184
Appendix C 189
Appendix D 195
Appendix E 196
Appendix F 201
Appendix G 209
Appendix H 210
Appendix I 211

Appendix J 220
Appendix K 222
Appendix L 232
Appendix M 234
Appendix N 236
Appendix O 239
Appendix P 242
Appendix Q 243
Appendix R 244
Appendix S 249
Appendix T 253

Part One: Literature Review

A COSMIN Systematic Review of the Measurement Properties of Quality-of-Life Outcome Measures used with Adult Inpatient Psychiatric Populations

Word Count: 7975

Abstract

This review aimed to identify and compare the measurement properties of quality of life (QoL) measures used in adult psychiatric inpatient settings and to assess their quality using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) manual (Mokkink et al., 2024).

Methods

Studies were included in the review if they included an adult inpatient psychiatric sample (including mixed samples) and if they evaluated the measurement properties of QoL outcome measures. Four databases were searched: MEDLINE, PsycINFO, Scopus, and CINAHL. Forward and backwards citation searching of included studies was conducted.

Each measurement property was assessed against the COSMIN criteria for good measurement properties to ascertain whether the study results were sufficient (Mokkink et al., 2024). A narrative synthesis was employed to integrate the results. The GRADE approach (Guyatt et al., 2011) was utilised to provide an overall certainty of the evidence.

Results

Twenty-three outcome measures were included in the review, from 38 studies. No outcome measure covered all nine measurement properties. Some outcome measures showed promising evidence for this population.

Discussion

Most measures demonstrated good structural validity and internal consistency, indicating some reliability and validity in this setting. There were limitations to the review, including the absence of two reviewers for measurement properties, which could potentially lead to biased results.

Key words

COSMIN, quality of life, outcome measure, measurement properties, psychiatric inpatient, reliable change index

Practitioner Points

- No outcome measures were identified that encompassed all measurement properties; clinicians should approach the identified outcome measures with caution.
- Reliable change index scores have been calculated for several outcome measures, allowing for the understanding of change over time.

Introduction

In mental health care, outcome measures are commonly used to monitor changes over time. These measures can inform care planning, track progress, evaluate services, and contribute to research (Tracy et al., 2022). When used effectively, they have the potential to enhance patient outcomes (De Jong et al., 2021).

Outcome measures are used to track improvements in quality of life (QoL) (Van Krugten et al., 2021). The World Health Organisation (WHO) defines QoL as an individual's perception of their position in life within context, culture and value system (WHOQOL Group, 1994) whilst Teoli and Bhardwaj (2023) define it as the standard of health, well-being, comfort, and happiness experienced by an individual or group. It can be considered a culturally constructed idea (Kagawa-Singer et al., 2010). Improvements in QoL can, therefore, be significant for individuals, and enhancing QoL is often a primary goal in mental health care.

It is generally agreed that the content of QoL measures should come directly from the relevant patients to enhance content validity (Auquier et al., 2003). Furthermore, Happell (2008) suggests that for outcome measures to be useful, they should be comparable across a relevant client group and service type. Using an outcome measure for a population it is not designed for may produce inaccurate or unrepresentative results. For example, measures used in a community setting may not apply to the problems faced by those in an inpatient setting. Therefore, this review will identify existing QoL measures used in an adult psychiatric inpatient setting. Research conducted in an inpatient psychiatric setting often involves multiple populations due to the dynamic nature of these services (Strathdee, 2024). Therefore, studies that include multiple populations alongside inpatient populations will be included.

QoL outcome measures in psychiatry are continually evolving. Van Krugten et al. (2021) reviewed the QoL outcome measures used in mental health settings; however, the measurement properties were not evaluated in this review. To evaluate whether an outcome measure is reliable and valid, its measurement properties are assessed and reported (Ginti, 2020). Previous reviews have evaluated the measurement properties of the EuroQoL-5D and Short Form-36 in patients with schizophrenia (Papaioannou et al., 2011). Additionally, Zuniga et al. (2024) examined the Schizophrenia QoL Scale (SQLS) and revision four (SQLS-R4) alongside their measurement properties. However, to our knowledge, no prior reviews have investigated the measurement properties of all QoL outcome measures used in an adult psychiatric inpatient setting. Therefore, this will be the focus of this review.

Domain

The domain being studied is the suitability and performance of QoL outcome measures for psychiatric inpatient populations.

Review Questions

- What are the different QoL measurement tools validated in adult psychiatric inpatient settings?
- What are their measurement properties?
- How well do they adequately measure the QoL for those in an adult psychiatric inpatient setting?

Aims

This study aims to identify the QoL measures used in adult psychiatric inpatient settings and evaluate the quality of their measurement properties using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) guideline for systematic reviews of Outcome Measurement Instruments (OMIs) (Mokkink et al., 2024). All measurement properties defined by COSMIN will be reviewed to highlight any potential

knowledge gaps. In this review, “measurement properties” refers to content validity, structural validity, internal consistency, cross-cultural validity, test-retest reliability, measurement error, criterion validity, construct validity, and responsiveness, as defined in the COSMIN manual (see Appendix A). While “psychometric properties” is the more traditional term (Swan et al., 2023), “measurement properties” will be used throughout per the COSMIN manual (Mokkink et al., 2024).

To further highlight the gaps in the knowledge, all outcome measures tested in an adult psychiatric inpatient setting will be included.

Objectives

- To identify, describe, and synthesise the outcome measures used to evaluate adult psychiatric inpatients' QoL outcomes.
- To report on the validation of outcome measures and their measurement properties using the COSMIN OMI guidelines for systematic reviews (Mokkink et al., 2024).
- To use a narrative synthesis to identify key themes and patterns between identified QoL measurement tools.

Outcomes

- A description of quality-of-life outcome measures used in inpatient psychiatric settings.
- A summary of the measurement properties of identified outcome measures and their relative performance using COSMIN OMI guidelines for systematic reviews (Mokkink et al., 2024).

Methods

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) updated COSMIN-OMI reporting guidelines (Elsman et al., 2024, Appendix B) were used to guide the review. The protocol was pre-registered with PROSPERO (https://www.crd.york.ac.uk/PROSPERO/) before the review was conducted (protocol ID: CRD420250651806).

Search Strategy

Eligibility was assessed by using the PICOS framework (population, intervention, comparator, outcome and study design) (Amir-Behghadami & Janati, 2020). The inclusion and exclusion criteria are described in Table 1, which guided the development of the search strategy by combining key terms and utilising Boolean operators. A University librarian reviewed the search strategy to ensure that no key terms were missed. The search was limited to papers published in the English language and included grey literature. Papers published in different countries were included if they were published in the English Language. There was no recommended time window for reviewing QoL outcome measures; therefore, no time limits were applied to the search strategy. The search was conducted on the 18th of April, 2025, across four databases: MEDLINE (Ovid), PsycINFO (Ovid), Scopus, and CINAHL. Searches on databases such as PsycINFO and Medline included medical subheadings (MeSH) terms and the option to "explode" specific terms, such as "quality of life" and "psychometrics". A comprehensive search strategy is outlined in Appendix C.

Furthermore, backwards citation searching was conducted by reviewing the reference lists of articles identified for inclusion from the database searches, and any relevant studies which fit the inclusion criteria were included in the review. Forward citation searching of included studies was conducted by searching Google Scholar; articles were screened against the inclusion criteria. This process helped to identify the most recent studies.

While developing the review question, a systematic review of the available QoL outcome measures used with individuals experiencing mental health issues was identified (Van Krugten et al., 2021). This review described and compared outcome measures using the O'Connell et al. (2014) framework for QoL outcomes in mental health settings. To build on this, a review of the reference list of this review was conducted, and any relevant studies that met the inclusion criteria were included.

All papers produced in the search were downloaded onto EndNote (EndNote Team, 2013). The primary reviewer selected papers according to the inclusion criteria by initially screening titles and abstracts, followed by a full-text review.

Table 1

Summary of Inclusion and Exclusion Criteria

	Inclusion Criteria	Exclusion Criteria
Population	Adult psychiatric inpatients. This refers to individuals admitted to general adult psychiatric inpatient as opposed to specialist settings. For example, studies conducted in psychiatric inpatient wards that may include a mix of diagnoses (e.g., patients with eating disorders alongside other psychiatric conditions) will be included.	Specialist inpatient settings will be excluded such as specialist eating disorder units, specialist forensic units such as secure wards, learning disability wards and drug and detox wards. These settings focus on treating specific diagnostic or legal populations outside the scope of the review which focuses on general psychiatric inpatient with a mix of diagnosis.
	Studies should report on psychiatric inpatients. Studies that include a mixed populations (for example, both inpatients and outpatients) will be included, these studies may contribute valuable comparisons between groups (for example, discriminant validity). Studies exclusively involving outpatients will be excluded.	Studies that define the population as older adult or adolescent will be excluded as the review is limited to general adult psychiatric inpatients. No specific age range was specified as countries define these categories differently. Instead, articles which defined their population as "older adult" or "adolescent" were excluded.
	As the review is exploratory to identify knowledge gaps and aims to give a broad overview of available OMIs in this population, mixed populations were included. Results were critically assessed to reflect this.	Inpatients settings focused only on physical health, such as stroke rehabilitation units or general medical wards, will also be excluded, as they do not primarily provide psychiatric care.
	Different length of stay inpatients will be included (acute, rehabilitation, and long stay patients) and different ward intensity will be included (rehabilitation, acute, PICU).	

Intervention	Papers which evaluate measurement properties of quality-of-life measurement tools, for example, validation studies.	Studies which do not report on or include measurement properties of quality-of-life outcome tool. Interventional studies, for example, those that use an outcome measure for pre- and post-intervention evaluation but do not explicitly evaluate measurement properties of the outcome measure.
Comparator	Using different measurement properties, such as reliability indices, the identified outcome measures will be compared using the COSMIN OMI guideline for systematic reviews (Elsman et al., 2024).	
Outcomes	The measurement properties of QoL measures including content validity, structural validity, internal consistency, cross-cultural validity, test-retest reliability, measurement error, criterion validity, construct validity and responsiveness.	Outcome measurement tools that do not assess QoL will not be included. For example, papers that report symptom severity outcome measures.
Study Design	Studies that evaluate measurement properties of quality-of-life outcome measure will be included. This might include longitudinal design, cross-sectional design, and repeated measures. Content validity studies typically rely on qualitative evidence, as such, relevant qualitative studies will be included.	Studies which do not evaluate or report the measurement properties of QoL outcome measures.

Analysis

The COSMIN reporting guidelines (Mokkink et al., 2024) recommend evaluating OMIs using six stages. The first two stages involve extracting data and then evaluating it. Stages three and four involve rating and summarising the study results, and at stage 5, a rating is given for the summarised results. At stage six, the quality of evidence is graded.

Stage 1: Extract

Relevant information was extracted and collated using COSMIN review management file. Data extracted included study characteristics (study setting, participants, diagnosis, design and methods, and outcome measures used), sample characteristics (age, gender, ethnicity, presenting problems, country), outcome measurement tool characteristics (domains, number of items, time to complete, common usage, related and paired measures, availability,

licensing, and scoring procedure), measurement properties (content validity, structural validity, internal consistency, cross-cultural validity, test-retest reliability, measurement error, construct validity, and responsiveness). There was no significant missing data, and no authors were contacted to retrieve any data. Information, including mean scores, standard deviations and Cronbach's Alpha, was extracted to calculate the Reliable Change Index (RCI) (Jacobson et al., 1991) for each OMI.

It is recommended that data be extracted by one reviewer and checked by a second reviewer (Mokkink et al., 2024). Due to time and resources, this was not possible for this review. A second reviewer independently assessed 13 of the 38 included studies (33%) for risk of bias. Agreement was reached across all measurement properties for 11 of these studies. In two cases, disagreements relating to construct validity were resolved through consultation of the COSMIN manual and a consensus discussion.

Stage 2: Evaluate

Each measurement property identified in each study was scored against the COSMIN risk of bias assessment (Mokkink et al., 2024; Appendix D). COSMIN risk of bias ratings are rated on a scale of Inadequate (I), Doubtful (D), Adequate (A) and Very Good (VG). A "very good" rating indicates the standard has been met using the preferred methods. An "inadequate" score indicates the standard has not been met, and the methods used are inadequate. Following COSMIN guidelines, the lowest score across all items for a measurement property determines the overall rating.

Stage 3: Rate: Study Results

Each measurement property within each study was evaluated against the COSMIN criteria for good measurement properties, which were used to determine whether the study results were sufficient or not (Mokkink et al., 2024; Appendix E). A score of (-) indicates that the measurement property is insufficient, indicating that there is evidence that the

measurement property is of poor quality. A score of (?) indicates that an indeterminant rating has been made, which means that something has been done, but there is not enough information to rate the result. A score of (+) indicates that the measurement property is sufficient, and there is evidence that it is of good quality.

Stage 4: Summarising the Results

Multiple studies that report the same outcome measure are summarised. At this stage, differences in results between studies should be noted. For example, if some studies report sufficient measurement properties and others report insufficient ones, a narrative explanation should be provided for the differences.

Stage 5: Rate: Summarised Results

Measurement properties were given a final summarised rating of sufficient (+), insufficient (-), inconsistent (+/-) or indetermined (?) for each measurement property. Content validity was evaluated based on relevance, comprehensibility, and comprehensiveness. This provided an overall conclusion for each OMI reported.

Stage 6: Grade

The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach (Guyatt et al., 2011; Appendix F) was used to assess the overall certainty of the evidence for each measurement property across the included studies, following the COSMIN guidelines. The quality of evidence was rated on a scale of “high”, “moderate”, “low” and “very low”. A “high” rating reflects strong confidence that the result is close to the actual value, while “very low” indicates that the result is likely substantially different from the actual value. The summary of the quality of the evidence regarding measurement properties was determined by the methodological quality of the studies, consistency, precision (sample size), and directness of the results (i.e., the population of interest).

Narrative Synthesis

Measurement properties were qualitatively synthesised by combining findings from various studies using words and text to explain the outcomes (Rodgers et al., 2009). Study characteristics were compared to assess potential factors influencing measurement properties.

Any final recommendations made were based on the outcome of the GRADE approach. COSMIN recommends that all nine measurement properties be assessed before making a recommendation. However, recommendations were made based on the results that produced the highest quality evidence across the most measurement properties.

If enough studies reported information about the same outcome measure and the same property, such as reliability or internal consistency, a meta-analysis of these results could have been conducted. However, the results showed that the data were insufficient for meta-analysis. In the case of outcome measures that had potential for meta-analysis, they lacked the appropriate range of data, such as unidimensionality and content validity, to produce meaningful results.

Finally, RCI (Jacobson et al., 1991) was calculated and reported by inputting means, standard deviations, and Cronbach's Alpha scores into a Shiny app developed to calculate RCI (Evans, 1998).

Results

Study selection is outlined in the COSMIN-PRISMA diagram (Figure 1). A total of 2,773 records were identified from the database searches after duplicates were removed. Duplicates were removed using an automation tool on EndNote (EndNote Team, 2013) and then reviewed by human judgment. After title and abstract screening, 130 reports were sought for retrieval. After full text screening, 23 reports were included in the review. One record by Mezzich et al. (2011) was excluded due to the ambiguity of the inpatient setting. The author

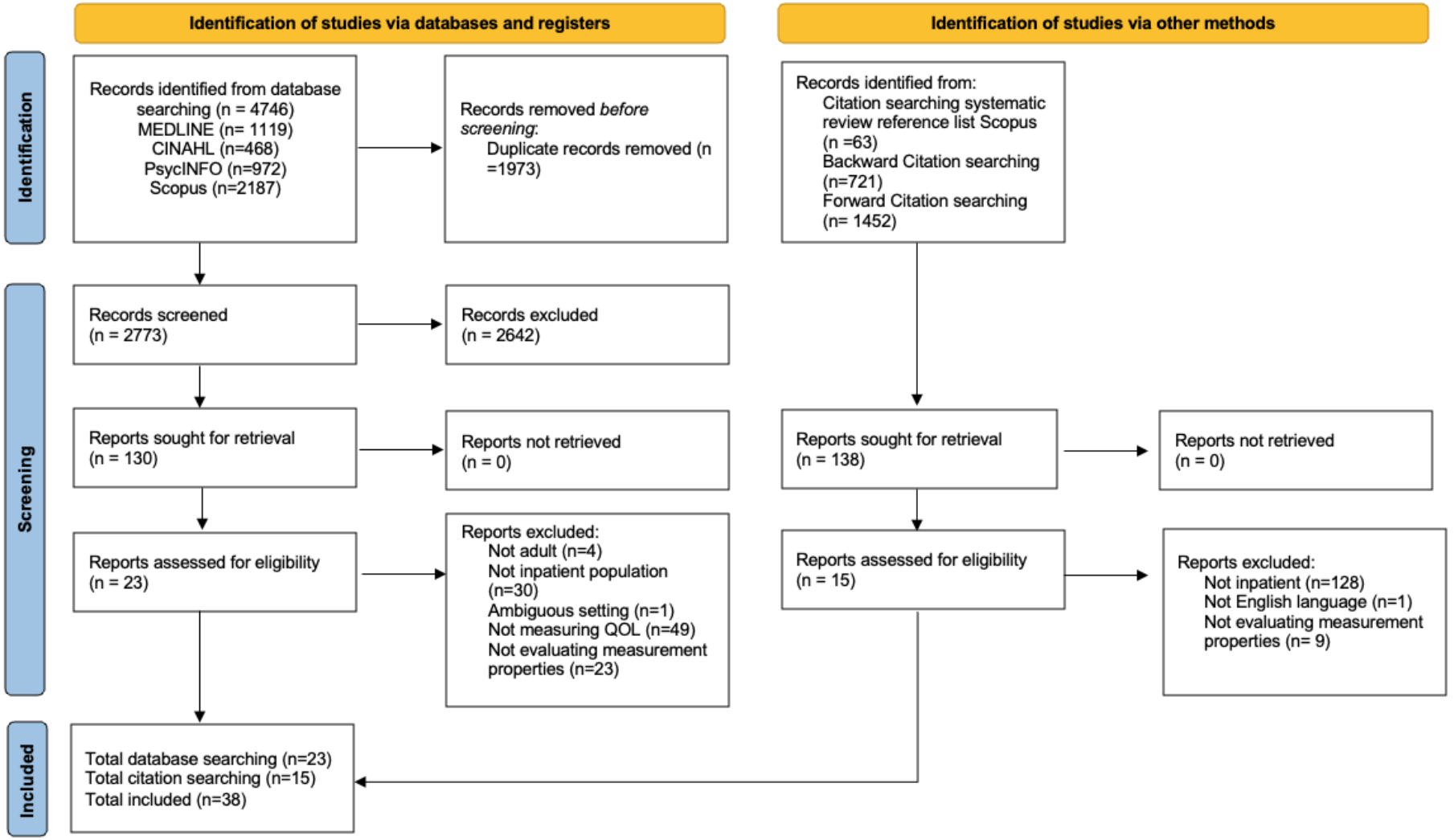
was contacted in May 2025 to confirm the setting and was given two weeks to reply. A random selection of full-text articles was independently screened by a second reviewer (10%) with 100% agreement.

A further 63 references were identified from a previous review of QoL instruments in mental health, conducted by Van Krugten et al. (2021). A total of 2,173 records were identified through forward and backwards citation searching. A total of 138 records were retrieved after title and abstract screening. After full-text screening, 15 articles were included in the review.

A total of 38 reports were included in the review. Reasons for excluding studies are included in the COSMIN-PRISMA diagram. Twenty-three QoL outcome measures were identified for review. The number of papers per outcome measure is further detailed in Table 1.

Figure 1

COSMIN-PRISMA Flow Diagram of Study Selection



OMI Characteristics

Data were extracted and presented in prespecified tables using Excel, with the COSMIN review management file (Mokkink et al., 2024) utilised. OMIs were organised together rather than listed alphabetically, as recommended by COSMIN. OMI characteristics were summarised and shown in Table 1. Studies that include more than one outcome measure are reported separately for each outcome measure. For example, a study with two outcome measures will be reported twice. As a result, a total of 42 studies are reported in Tables 1-5 to reflect those that have more than one outcome measure.

The study authors, year of publication, name of the OMI, administration time, scale domains and items per scale, response options and scoring, recall period, availability of the scale, recall method, and language versions were reported. Response options were indicated as the number of points on the scale, and most results were derived from self-report instruments. Some minor information regarding administration time was missing, and the authors were not contacted for this data.

Table 1

Outcome Measurement Instrument Characteristics

OMI	OMI	Reference	Items (N)	Administration time (minutes)	Scale Domains (N items)	Response options (Likert)	Scoring	Recall period	Measure Available (Y/N)	Patient rated? (Y/N)	Languages Versions included
1	Lehamn's Bref Quality of life interview	Anderson et al., 2002	74	16	Objective (2), subjective QoL (6) and global life satisfaction score (1)	1-7	Average, higher score indicates higher QoL	One year	Y	N-interviewer	English
2	LQoLI (full)	Russo et al., 1997 Lehman, 1988	158	45	Objective QoL: Living situation (8), frequency of family contacts (1), frequency of social contacts (1), number of leisure activities (1), work (1), frequency of religious activity, finances (2), safety (2), health (3) Subjective QoL: General life satisfaction (living, family, social, leisure, work, religious, finances, safety, health).	1-7	A higher score indicates higher QoL	Not reported (NR)	Y	N-Interviewer administered	English
3	Q-LES-Q	Bishop et al., 1999 Ritsner et al., 2002	90/93 (3 optional functioning items 83	40-45 40-45	Four of the core subscales assess the global domains of physical health (13 items), subjective feelings (14 items), leisure activities (six items), social relationships (11 items), and general activities, work, household duties, and school/course work. There are single questions addressing overall satisfaction and medication satisfaction. Items on school and coursework have been excluded from Ritsner et al. (2002)	1-5	Average score, higher score indicates greater satisfaction	Past week	Y	Y	English, Hebrew

OMI	OMI	Reference	Items (N)	Administration time (minutes)	Scale Domains (N items)	Response options (Likert)	Scoring	Recall period	Measure Available (Y/N)	Patient rated? (Y/N)	Languages Versions included
4	QLESQ-SF	Pikanen et al. 2011	16	7	Fourteen questions assess domains such as social relationships, living or housing situation, feelings and physical health. The last two questions concern medication and overall life satisfaction during the past 7 days.	1-5	Total score added, higher score indicates better QoL	Past week	Y	Y	Finish
5	Lancashire Quality of Life Profile (LQoLP) and LQoLP-EU	Ritsner et al. 2002 (27 items) Kaiser et al., 1997 Gaite et al., 2000 Hansson et al., 1998	27 100 105 100	20-25	9 domains work and education (7), leisure and participation (8), religion (4), finances (7 items), social relations (6), health 10. (Life regard index and satisfaction with life scale are additional components to the EU version.)	7 point	Average scores. Higher score indicates better QoL	Present	N	Has elements of subjectivity-completed by an interviewer	Hebrew, Dutch, Danish, English, Spanish, Italian, Swedish, German,
6	Schizophrenia Quality of Life Scale (SQOL) -18 Version 1	Boyer et al., 2010	18	NR	8 subscales (psychological well-being (3), self-esteem (2), family relationships (2), relationship with friends (2), resilience (3), physical wellbeing (2), autonomy (2), sentimental life (2). Global score of QOL (total)	1-5	0-100 higher score indicating higher QoL	Present	N	Y	French
7	SQOL-41 and Chinese version SQOL-C	Auquier et al., 2003 Chou et al., 2010	41	NR	8 domains, psychological wellbeing (10), self-esteem (6), family relationships (5), relationships with friends (5), resilience (5), physical wellbeing (4), autonomy (4), sentimental life (2)	1-5	Means 0-100 with 100 being most favourable	Present	Y	Y	English, Chinese

OMI	OMI	Reference	Items (N)	Administration time (minutes)	Scale Domains (N items)	Response options (Likert)	Scoring	Recall period	Measure Available (Y/N)	Patient rated? (Y/N)	Languages Versions included
8	Quality of Life Interview (QLI)	Angstman et al., 2009	32	NR	16 domains: health, self-esteem, goals-and-values, money, work, play, learning, creativity, helping, love, friends, children, relatives, home, neighbourhood, and community	1-6	Average responses, higher score indicates better QoL	NR	N	Y	English
9	Quality of Life in Schizophrenia (QLIS) version 1	Franz et al., 2012 Franz et al., 2012	53 (with 2 optional items relating to employment)	NR	16 subscales social contacts (5 items) appreciation by others (4 items), relationship to my family (3 items), appraisal of pharmacotherapy (6) appraisal of psyc-pathological symptoms (6) cognitive functioning (5) abilities to manage daily living (4) appraisal of accommodation/housing (5) financial situation (4) leading a normal life (3) confidence (4) global life satisfaction (3)	1-4	Means (0-10, with zero worst QoL)	NR	N	Y	German
10	The How Are You? Scale Version 5	Katsavdakis et al., 1999	55	30	5 Domains: Your health (6), how you've been feeling (20), your daily life (7), your circumstances (6), your relationships (10), overall domain (6)	0-4	Means (direction not indicated)	Present	N	Y	English
11	Quality of Life Survey (QLS)	Nair et al., 1996	63	15-20	8 domains: Material quality of life (15 items, 6, 3 and 5), knowledge and education (7 items), human relationships (4), dependency issues (2), inner experience (6), inpatient treatment (9), leisure activities (4), work and vocational rehabilitation (2).	5	Mean (direction not indicated)	Preceding month	N	Y	NR
12	Short Form (SF)-36 Health survey	Newnham et al., 2007 (only mental health subscales)	36	NR	8 domains: physical functioning (10) role-physical (4) bodily pain (2) general health (5) vitality (4) social functioning (2) role emotional (3) mental health (5)	Multiple severity scales	Mean, 0-100 with higher score indicating higher QoL	4 weeks	Y	Y	English, Japanese, Chinese

OMI	OMI	Reference	Items (N)	Administration time (minutes)	Scale Domains (N items)	Response options (Likert)	Scoring	Recall period	Measure Available (Y/N)	Patient rated? (Y/N)	Languages Versions included
		Pukrop et al., 2003 Nishiyama et al., 2009 Su et al., 2014 Tunis et al., 1999									
13	World Health Organisation Quality of Life Assessment WHOQOL-BREF	Su et al., 2014 Oliveira et al., 2016 Norholm & Bech, 2006 Chan et al., 2003 Van de Willige et al., 2005	26-28 (2 domestic on the Taiwan version)	A few minutes	4 domains: physical (5) psychological (5) level of independence (6) social relationships (3) Environment (9) spirituality and religious beliefs, overall quality of life and general health perceptions	5 point scale	Means, 0-100, higher scores refer to higher QoL	Present	Y	Y- can be administered by proxy	Chinese, Portuguese, Danish, Chinese, Dutch
14	Modular System for Quality of Life (MSQOL)	Pukrop et al., 2003	55	NR	Mood rating, sociodemographic module of objective QOL subjective QOL modules: physical health, physical vitality, personal role and social relationships, leisure activities, mental state and emotions, life in general, family, partnership, occupation	7 point scale	0-100 higher score indicating higher QoL	4 weeks	Y	Y	English
15	EuroQOL (EQ)-5D 3L and 5L (still comparable)	Pikanen et al., 2011 (descriptive scale only 3L)	5	NR	Five dimensions mobility, self-care, daily functioning, pain or discomfort, and anxiety or depression.	1-3 (3Levels) 1-5 (5Levels)	Scores are coded and a score of 1.00 indicates full health	7 days	Y	Y	Finish, Dutch

OMI	OMI	Reference	Items (N)	Administration time (minutes)	Scale Domains (N items)	Response options (Likert)	Scoring	Recall period	Measure Available (Y/N)	Patient rated? (Y/N)	Languages Versions included
		Van de Willige et al., 2005					Or as a visual analogue scale 0-100 (100 being best imaginable)				
16	Mental Health Quality of Life (MHQoL) -7D Questionnaire	Ebadi & Rezaiye, 2025	7	Quick to complete	Self-image, independence, mood, relationships, daily activities, physical health, future, psychological well-being	Level 1-4 (0-3 score)	Sum of scores 0-21	Today	Y	Y	Farsi
17	Short Form (SF)-12 health survey version 2	Soysal Gunduz et al., 2021	12	NR	8 domains: General health, physical functioning, Role physical, Bodily pain, Social functioning, Vitality, Mental health, Role-emotional.	3 point Likert and 5 point for 10 items	NR	Present - 4 weeks	Y- English online and Turkish from author	Y	Turkish
18	Schizophrenia Quality of Life Survey (SQLS) Japanese version	Kaneda et al., 2002	30	NR	Psychosocial (15), motivation/energy (7) and symptoms side effects (8)	0-4	Lower score represents better HRQOL	NR	N	Y	Japanese
19	Recovering Quality of Life (ReQoL) 10 and 20	Keetharuth et al., 2018	10 items and 20 items respectively	A few minutes	Generic overall rating of QoL, recovery, and well-being.	0-4 points 5 levels	summative	Past week	Y (NHS and public research)	Y	English
20	Schizophrenia Quality of Life Survey	Kuo et al. 2007 Kuo et al., 2009	33	NR	two domains, psychosocial (20) and vitality (13)	0-4 5 point scale	Lower score represents better HRQOL	7 days	Y- permission needed	Y	Chinese

OMI	OMI	Reference	Items (N)	Administration time (minutes)	Scale Domains (N items)	Response options (Likert)	Scoring	Recall period	Measure Available (Y/N)	Patient rated? (Y/N)	Languages Versions included
	(SQLS) revision 4	Martin & Allen, 2007 Su et al., 2017									
21	Quality of Life in Bipolar Disorder (QOLBD)	Michalak et al., 2010	56	4	a revised set of 14 domain subscales were indicated: Physical, sleep, mood, cognition, leisure, social, spirituality, finances, household, self-esteem, independence, identity, work, education	5 point scale	NR	7 days	Y	Y	English
22	QLESQ-18	Ritsner et al., 2005	18	10-12	Physical health, subjective feelings, leisure time activities, social relationships, general index.	5 point scale	Average of all 18 items, higher score is better QoL	Past week	Y	Y	English
23	Heinrich's Quality of Life Rating Scale	Simon-Abbadi et al., 1999	21	45	Domains: interpersonal relations (1-8), instrumental QoL (9-12) intrapsychic foundations (13,17,20,21) Common objects and activities (18-19)	NR	NR	NR	N	N-(clinician rated)	French

Study Characteristics

Information on study characteristics was extracted and presented in Table 2. This includes the OMI used, the first author and year of publication, the country and language of study, the study setting, participant demographics, disease characteristics, and sample size. Minor information was missing regarding the ethnicity of the study populations; the authors were not contacted to obtain this data.

The mean age of participants ranged from 31.3 to 53.4 years, with most studies reporting average ages of participants in their thirties. The proportion of male participants varied across studies, ranging from 29% to 79%, with an average of 63.45%. The duration of illness was reported, ranging from 8.5 to 26.5 years, with an average of 15.21 years. Most studies were conducted in Europe and Asia and were published in English. All studies included at least a partial inpatient population. Some studies reported a mixed sample and included additional settings such as board and care homes, community homes, community care, residential care, halfway houses, primary and secondary care and comparison groups of professionals and the public. Studies which included a mixed sample received a lower score for the quality of evidence (QOE) (Table 6) for “indirectness”.

Sample sizes ranged from 21 to 4266, with varying sample sizes used within studies for each measurement property assessed. A range of study designs were employed, including cross-sectional, longitudinal, prospective, group comparison, repeated measures, multi-trait scaling analysis, and between-subjects designs.

Table 2

Study Characteristics and Summary of Main Findings.

OMI	Reference	Country (Language, if reported)	Setting	Design	Demographic Characteristic patients			Disease characteristics		N
					Mean (SD) age	% male	Ethnicity	Diagnosis	Duration (years)	
1	Anderson et al., 2002	Canada (English)	Inpatient (93.6%) rehabilitation 3 participants were outpatient	Not reported suggestive of longitudinal/repeated measures	40.12 (8.59)	61	100% White	74% schizophrenia, 15% bipolar, 11% psychosis and 2 with major depression	3.91 hospital stay	CV[†] R[‡], N=47
2	Russo et al., 1997	USA (English)	Acute inpatient	Longitudinal	35.6 (11.4)	61.1	White (77.3%) African-American (15.7%) Asian (2.9%) and Latino (2%)	Severely ill	2 in hospital	SV[§], IC[**], CV and R, N =981
2	Lehman, 1988	USA (English)	Inpatient 21,1 %, outpatient 19.6%, and board and care homes	Not reported suggestive of repeated measures	57.5 17.5 34.5	47	74.8-90.2% White	Inpatient population schizophrenia 76.8%, Affective disorder 13.1%, Alcoholism 6.1%, drug abuse 1%, Organic brain syndrome 4%, mental retardation 9.1%, Personality disorder 5.1%	23.7	CV, N=469 Re[††], N=45 SV, N=99
2	Goodwin & Madell, 2002	UK (English)	Inpatient 23%, Community home 67%, and	Not reported, suggestive of Randomised	51	52	NR	Schizophrenia and other mental health difficulties not defined	23	CV, N=45

[††] Construct validity
[‡] Responsiveness
[§] Structural validity
[**] Internal consistency
[††] Reliability

OMI	Reference	Country (Language, if reported)	Setting	Design	Demographic Characteristic patients			Disease characteristics		N
					Mean (SD) age	% male	Ethnicity	Diagnosis	Duration (years)	
			Community own home 15%	repeated measures						
3	Bishop et al., 1999	USA (English)	Admitted to either a residential community treatment centre or a University psychiatric hospital	Not reported- suggestive of cross-sectional design	18-65	45	Ethnic information was available on 123 subjects with 61% Caucasian (*n* = 75), 29% African-American (*n* = 36), 8% Hispanic (*n* = 10) and 2% Asian (*n* = 2).	Diagnostic categories were available for 139 subjects: 34.5% were schizophrenic (*n* = 48), 16.6% were bipolar (*n* = 23), 21.6% were classified as psychosis NOS (*n* = 30), 19.4% were major depression (*n* = 27) and 7.9% were classified as other (*n* = 11).	NR	SV, IC, N=151
3	Ritsner et al., 2002 (10 items removed)	Isreal (Hebrew)	Inpatient 35.5% and outpatient 17.6% professionals 46.7%	Not reported, suggestive of a between subject's design	38.9 (10.1)	79	NR	148 patients were diagnosed with schizophrenia and the remaining 51 patients with either schizoaffective disorder (n= 33) or mood disorders (n=18)	14.1	IC, N=199 (inpatient and outpatient sample) CV, N=374 (patient and comparison group) Re, N=20 (inpatient)
4	Pitkanen et al., 2011	Finland (Finish)	Acute inpatient	Not reported, suggestive of repeated measures	38 (13)	59	NR	41% were diagnosed with schizophrenia, 29% unspecified non-organic psychotic disorder, 16% schizoaffective disorder, 6% acute and	NR	CV, N=280 SV, N= 184 IC, N=190

OMI	Reference	Country (Language, if reported)	Setting	Design	Demographic Characteristic patients			Disease characteristics		N
					Mean (SD) age	% male	Ethnicity	Diagnosis	Duration (years)	
								transient psychotic disorder, 5% persistent delusional disorder, 2% schizotypal disorder and 1% other non-organic psychotic disorder (F28)		
5	Ritsner et al., 2002 (only 27 items considered)	Isreal (Hebrew)	Inpatient 35.5% and outpatient 17.6% professionals 46.7%	Not reported, suggestive of a cross sectional comparative design	38.9 (10.1)	79	NR	148 patients were diagnosed with schizophrenia and the remaining 51 patients with either schizoaffective disorder (n= 33) or mood disorders (n=18)	14.1	IC, N= 374 CCV‡‡ N= 10 Re, N=20 CV, Patients N=199 Professionals N=175
5	Kaiser et al., 1997	England and Germany (English and German)	In (55.2%) and outpatient (54.8%)	Not reported, suggestive of cross sectional comparative design	53.4 (14)	59	NR	neuroleptics (being treated for psychosis) also depression bipolar and schizophrenia	12.32	SV, IC, CV, N=440
5	Gaite et al., 2000	Netherlands, Denmark, Uk, Italy (Dutch, Danish, English, Spanish, Italian)	Inpatient, outpatient and community	Test-retest study (longitudinal)	18-65	NR	NR	Schizophrenia, diagnosis F20-F25	NR	IC, N= 404 Re, N= 294
5	Hansson et al., 1998	Sweden (Swedish)	Inpatient	Cross-sectional (longitudinal)	31.3	51.7	NR	psychosis	NR	SV, N= 29 Re, N= 21

‡‡ Cross-cultural validity

OMI	Reference	Country (Language, if reported)	Setting	Design	Demographic Characteristic patients			Disease characteristics		N
					Mean (SD) age	% male	Ethnicity	Diagnosis	Duration (years)	
6	Boyer et al., 2010	France (French)	In 28.6% and outpatient 71.3%	Repeated measures (longitudinal)	36.5 (10.8)	70	NR	Schizophrenia	13.8	SV, IC, N= 517 Re, N=72 CV, N= 302 R, N= 28
7	Auquier et al., 2003	France (French)	In and outpatient 50%	Randomised test retest (longitudinal)	37.3 (10.9)	68	NR	Schizophrenia paranoid (46.9%), 29 disorganized (14%), 33 undifferentiated (15.9%) and 45 residual (21.7%).	13.9	SV, IC, CV, N= 207 Re, N= 53 R, N= 46
7	Chou et al., 2010	China (Chinese)	Inpatient	Prospective study	37.9	70.7	NR	Schizophrenia	NR	IC, Re and CV, N=41
8	Angstman et al., 2009	USA (English)	Inpatient	NR	37.8 (11.6)	41	88% Caucasian and 18% Native American	schizophrenia-spectrum disorders at 29%, de-pression at 23%, bipolar disorder at 20%, substance disorders at 14%, personality disorders at 11%, and other at 3%	NR	R, N=56
9	Franz et al., 2012	Germany (German)	Inpatient 67%, outpatient 31%, general population	Cross sectional	37.3-40.7 (9,8-12.9)	41-50	NR	Schizophrenic patients	25.2 - 30.7 months	ConV[§§] N= 268 Re, N= 49 (outpatients) IC, N=135
9	Franz et al., 2012	Germany (German)	Inpatient 39% and out patient	Cross sectional	37.4 (9.8)	54	NR	Schizophrenia	12	CV, N= 135
10	Katsavdakis et al., 1999 (study 2)	UK (English)	Inpatient compared to professionals	Repeated measures within subjects and between groups	33.6 (11.4)	33	NR	alcohol dependency, 14% for an anxiety disorder diagnosis, 24% for a depressive or dysthymic	16 days stay	CV, N= 117 R, N= 43

[§§] Content validity

OMI	Reference	Country (Language, if reported)	Setting	Design	Demographic Characteristic patients			Disease characteristics		N
					Mean (SD) age	% male	Ethnicity	Diagnosis	Duration (years)	
								diagnosis, 9% for a bipolar disorder diagnosis, 12% for a psychotic disorder diagnosis, 12% for an eating disorder diagnosis, and 13% for all remaining diagnostic categories.		
11	Nair et al., 1996	Singapore (Not reported)	Inpatient	Not reported suggestive of Repeated measures	49.39 (10.58)	54	Chinese (765), Malay 8% Indian 12%, other, 4%	Schizophrenia	7.68 years hospital stay	IC, CV and R, N=50
12	Newnham et al., 2007 (only subscales related to mental health)	Australia (Not stated)	Private inpatient hospital	Not reported suggestive of Repeated measures	13-87	29	NR	affective/mood (66%), and anxiety (22%) disorders, as well as schizophrenia (4%) and substance use (5%).	2 year period	CV, R, N=1830
12	Pukrop et al., 2003	Germany (Not mentioned)	Inpatient	Not reported suggestive of Between subjects longitudinal design	33.7 - 43.45(9.6-12.42)	38.9-51.6	NR	Schizophrenia (n= 91) and Depression (n=114) Healthy controls (n=346)	6.7	SV, CV, R, N=205
12	Nishiyama et al., 2009	Japan (Japanese)	Inpatient	Multi trait scaling analysis	50 (14.7)	53.6	NR	Schizophrenia (77.4%), Bipolar (4.4%), Major depressive Disorder (2.9%), Dementia (5.8%), mental retardation (7.3%)	4.8 hospital	IC, CV, N=137

OMI	Reference	Country (Language, if reported)	Setting	Design	Demographic Characteristic patients			Disease characteristics		N
					Mean (SD) age	% male	Ethnicity	Diagnosis	Duration (years)	
12	Su et al., 2014	Taiwan (Chinese)	Inpatient	Not reported suggestive of Repeated measures and cross sectional	49.1 (7.85)	66	NR	Schizophrenia	26.5	SV, Re, ME*** CV, N=100
12	Tunis et al., 1999	USA (English)	In and outpatient	Not reported suggestive of Repeated measures and cross sectional	39.29 (11.32)	70	75% White	Schizophrenia (80%) Schizoaffective disorder (20%)	24	SV, IC, R, N=1155 CV, patient sample N= 1155 general population N=2474
13	Su et al., 2014	Taiwan (Chinese)	Inpatient	Not reported suggestive of Repeated measures	49.1 (7.85)	66	NR	Schizophrenia	26.5	SV, Re, ME, CV, N=100
13	Oliveira et al., 2016	Portugal (Portuguese)	Inpatient (34%) and outpatient (66%)	Cross-sectional	43.15 (12.38)	57	NR	Schizophrenia 36.5%, Bipolar 8.2%, Anxiety 3.7%, Personality Disorders 14.1%, Depressive Disorder 20.1%, Psychotic Disorder 10.4%	NR	SV, N=403 IC, N=191 CV, N=403
13	Norholm & Bech, 2006	Denmark (Danish)	Inpatient	Cross sectional, but not specified	39(10.9)	58.9	NR	Schizophrenia	69 days hospital (+/- 64 days)	IC, CV, N=38
13	Chan et al., 2003	Hong Kong (not specified, Hong Kong version used)	Inpatient (31.3%), long stay care homes and halfway houses	Matched group comparison	49.84 (7.88)	59.7	NR	Schizophrenia	26	CV, N=207
13	Van de Willige et al., 2005	Netherlands (not described)	Inpatient and out patient	longitudinal	36 (11.2)	55	NR	Schizophrenic Disorder	11	R, N=63

*** Measurement error

OMI	Reference	Country (Language, if reported)	Setting	Design	Demographic Characteristic patients			Disease characteristics		N
					Mean (SD) age	% male	Ethnicity	Diagnosis	Duration (years)	
14	Pukrop et al., 2003	Germany (Not mentioned)	Inpatient	Between subjects longitudinal design	33.7 - 43.45(9.6-12.42)	38.9-51.6	NR	Schizophrenia (n= 91) and Depression (n=114) Healthy controls (n=346)	6.7	IC, CV, R, N=205
15	Pikanen et al., 2011	Finland (Finish)	Acute inpatient	Not reported, suggestive of cross sectional comparative design	38 (13)	59	NR	41% were diagnosed with schizophrenia, 29% unspecified non-organic psychotic disorder, 16% schizoaffective disorder, 6% acute and transient psychotic disorder, 5% persistent delusional disorder, 2% schizotypal disorder and 1% other non-organic psychotic disorder (F28)	NR	SV, N= 282 IC, N= 281 CV, N=280
15	Van de Willige et al., 2005	Netherlands (not described)	Inpatient and out patient	longitudinal	36 (11.2)	55	NR	Schizophrenic Disorder	11	CV, R, N=63
16	Ebadi & Rezaiye, 2025	Iran (Farsi)	Inpatient	Not reported suggestive of Cross sectional and Repeated measures	18+	61	NR	MDD 28.7%, BMD 27.3%, PTSD 28.7%, SCH 15.3%	NR	SV, N = 150 IC, Re, ME, N= 60
17	Soysal Gunduz et al., 2021	Turkey (Turkish)	Physical health ward and Psychiatric ward (28.4%)	Cross-sectional design	43.5 (14.4)	49		Physical and Mental health, bipolar disorder, Major depression, PTSD, OCD, Anxiety, Psychosis	NR	SV, IC, CV, N= 136
18	Kaneda et al., 2002	Japan (Japanese)	Inpatient	Not reported suggestive of Cross sectional,	46.6 (11.3)	45	NR	Schizophrenia	22.9	SV, IC,CCV, CV, N= 55
19	Keetharuth et al., 2018	UK (English)	General practices (6-27%), IAPT	Not reported suggestive of	16-65	NR	White (85-86%), Black	Common mental health disorders, Schizophrenia,	NR	Study 1= 2262 study 2= 4266

OMI	Reference	Country (Language, if reported)	Setting	Design	Demographic Characteristic patients			Disease characteristics		N
					Mean (SD) age	% male	Ethnicity	Diagnosis	Duration (years)	
			(6%), Secondary care, Inpatient (3-13%) out patient (46-57%), and Community	Cross sectional and longitudinal			and minority ethnic (9-13%)	other psychotic disorders, bipolar, personality disorders, others and missing data		R, N= 953 Re, N= 2800 CV, N=4266 SV, N=4266
20	Kuo et al. 2007	Taiwan (Chinese)	Inpatient (50%) and outpatient (50%)	Not reported suggestive of Repeated measures (longitudinal)	36.9 (8.82)	75	NR	Schizophrenia	12.5	IC, N=97 Re, N= 83 CV, N= 97 R, N= 14
20	Kuo et al., 2009	Taiwan (Chinese)	Inpatient (50%) and outpatient (50%)	Not reported suggestive of Cross sectional,	36.9 (8.82)	75	NR	Schizophrenia	12.52	SV, CV, N=100
20	Martin & Allen, 2007	UK population (Not specified)	Community (71%), day centre (13%), and inpatient (16%)	Cross sectional design	40.7 (11.43)	74	NR	Schizophrenia	24.64	SV, IC, CV, N=100
20	Su et al., 2017	Tiawan (not specified)	Inpatient	Not reported suggestive of Cross sectional,	49.2 (7.8)	66	NR	Schizophrenia	26.6	SV, CV, N=100
21	Michalak et al., 2010	USA and Canada (English)	Inpatient (58.6%) and outpatients	Longitudinal	41.03 (13.8)	31.5	NR	Bipolar type 1 and 2	NR	SV, IC, N=224 Re, CV, N= 93
22	Ritsner et al., 2005	Israel (Not reported)	Inpatient (71.8%) and outpatients	Repeated measures	38.5 (10.3)	79.8	Jewish or Arab Israelis	Schizophrenia and Schizoaffective disorders	13.7	SV, IC, CV, N= 472 outpatients' Re, N= 33
23	Simon-Abbadi et al., 1999	France (French)	Inpatient	Not reported suggestive of Repeated measures	31.5 (8.5)	63.3	NR	Schizophrenic	8.5	IC, SV, Re, CV, N=60

Risk of Bias

Each measurement property in each study was scored against the COSMIN risk of bias assessment (Mokkink et al., 2024; Table 3). Criterion validity is shaded in grey, and no further results were included for this property. This is due to the absence of a recognised gold standard QoL instrument for this population to serve as a comparator (Blome & Augustine, 2015; Burks et al., 2021; Eiser & Jenney, 2007).

A log of decision-making was kept, supporting consistency with risk of bias ratings (Appendix G). COSMIN advises that reviewers use their relevant expertise in decision-making. Therefore, decisions were guided by the primary reviewer's experience and supplemented by the COSMIN manual (Mokkink et al., 2024).

The measurement property cross-cultural validity was rated "inadequate" across all reporting studies. This was due to the lack of qualitative and quantitative methods reported to support this property. Internal consistency was rated "very good" across most studies; however, some studies were downgraded when Cronbach's Alpha was not calculated (Norholm & Bech, 2006).

Lower scores were given for construct validity and responsiveness when there was an absence of predefined hypotheses or when an inappropriate statistical analysis was used, as per COSMIN, such as when correlations or effect sizes were not calculated. Structural validity ratings were downgraded when sample sizes were insufficient. COSMIN recommends a sample of at least five times the number of items in the scale for sufficient structural validity. Furthermore, assessment of content validity often revealed a lack of details of qualitative and quantitative methods. Methodological details such as reporting saturation, use of an interview guide, and recording and transcribing interviews were often missing. Test-retest reliability was rated lower when the time between administrations was unclear and when the stability of the sample was not reported or the intraclass correlation coefficient

(ICC) was not provided. Overall, the risk of bias varied by measurement property and this assessment will contribute to the overall quality of the evidence (Table 6).

Table 3

Quality of Studies: Risk of Bias Assessment

OMI	Reference	Content Validity	Structural Validity	Internal consistency	Cross cultural validity	Reliability	Measurement error	Criterion Validity	Construct Validity		Responsiveness
									Convergent	Discriminant	
1	Anderson et al., 2002								A	A	I
2	Russo et al., 1997		A	VG					VG	VG	VG
2	Lehman, 1988	I		VG		A			VG	VG	
2	Goodwin & Madell, 2002								I	A	
3	Bishop et al., 1999		I	VG							
3	Ritsner et al., 2002 (10 items removed)			VG	I	A			A	VG	
4	Pitkanen et al., 2011		VG	VG					A	A	
5	Ritsner et al., 2002 (only 27 items considered)			VG	I	A			A	VG	
5	Kaiser et al., 1997		I	VG					A	A	
5	Gaite et al., 2000			VG	I	VG					
5	Hansson et al., 1998			VG	I	I					
6	Boyer et al., 2010		VG	VG					VG	VG	A
7	Auquier et al., 2003	A	A	VG		I			VG	VG	A
7	Chou et al., 2010			VG	I	VG			A	A	
8	Angstman et al., 2009										I
9	Franz et al., 2012	A		VG		A					
9	Franz et al., 2012								VG		
10	Katsavdakis et al., 1999	D								A	A
11	Nair et al., 1996	I		VG						D	I
12	Newnham et al., 2007									A	D

OMI	Reference	Content Validity	Structural Validity	Internal consistency	Cross cultural validity	Reliability	Measurement error	Criterion Validity	Construct Validity		Responsiveness
									Convergent	Discriminant	
12	Pukrop et al., 2003			VG					A	A	D
12	Nishiyama et al., 2009			VG					A	A	
12	Su et al., 2014		I			A	A		A	D	
12	Tunis et al., 1999		VG	VG					A	A	D
13	Su et al., 2014		I			A	A		A	D	
13	Oliveira et al., 2016		VG	VG					A		
13	Norholm & Bech, 2006			I						D	
13	Chan et al., 2003								A	A	
13	Van de Willige et al., 2005								A		D
14	Pukrop et al., 2003			VG					A	A	D
15	Pikanen et al., 2011		VG	VG					A	D	
15	Van de Willige et al., 2005								A		D
16	Ebadi & Rezaiye, 2025	D	VG	VG	I	A	A				
17	Soysal Gunduz et al., 2021		VG	VG					D	D	
18	Kaneda et al., 2002		I	VG	I				VG		
19	Keetharuth et al., 2018	A	VG	VG		A			VG	VG	VG
20	Kuo et al. 2007			VG		A			VG		A
20	Kuo et al., 2009		I				A		A		
20	Martin & Allen, 2007		I	VG						A	
20	Su et al., 2017		I						VG		
21	Michalak et al., 2010	VG	I	VG		D			A		
22	Ritsner et al., 2005		VG	VG		A			A	A	
23	Simon-Abbadi et al., 1999		I	VG		I	D		A		

Narrative Synthesis

Table 4 summarises the extracted results for each study and measurement property. Each measurement property was rated against the COSMIN criteria for good measurement properties (Mokkink et al., 2024).

A total of 23 OMIs were identified. Of those, nine reported content validity, presented in Table 5. According to COSMIN criteria for good measurement properties, 14 OMIs had sufficient evidence that structural validity was of good quality and was scored (+), four OMIs had insufficient evidence that the measurement property was of good quality and were rated (-), and three OMIs had not enough information to rate the result and were rated (?).

Fourteen studies received a score of (+) for internal consistency, indicating a Cronbach's alpha value of 0.70 or greater. Seven OMIs scored (?); in most cases, this was because there was no confirmation of unidimensionality, and seven scored (-).

Three studies reported cross-cultural validity as forward and backwards translation; however, these studies did not provide enough information about group comparisons to make any conclusive decisions using COSMIN standards, and thus, they were scored (?).

Ten studies scored (+) for test-retest reliability, that is, ICC greater than or equal to 0.70, one study scored (?) for reporting Pearson's correlations (*r*), and five studies scored (-). Three OMIs reported (+) for measurement error; in these cases, the standard error of measurement (SEM) was reported.

Convergent validity was assessed by comparing the OMI with other outcome measures or within the OMI itself. Thirty-one studies received a score of (+) for convergent validity, and two received a score of (?). This was because the expected result was unclear. Discriminant validity was assessed by comparison between groups and known groups. Twenty-five studies scored (+) for discriminant validity, and two scored (?). In some cases,

OMIs did not report a hypothesis; however, if the literature supports the outcome, it is assumed that construct validity has been met.

Twelve studies scored (+) for responsiveness; some OMI studies did not report a hypothesis. In these cases, if it was assumable from the literature, it was positively scored. It is expected that the effect size is reported. However, one study only reported means and standard deviations and was therefore scored (?), while two studies were scored (-).

Several OMIs were strongly rated against the COSMIN criteria for good measurement properties across six of the seven measurement properties, indicating their potential as robust tools. OMI 7, the SQOL-41, was reported in two studies. Auquier et al. (2003) reported on content validity and exploratory factor analysis (EFA). The EFA revealed an eight-factor structure with factor loadings ranging from 0.46 to 0.84 and explained 52% of the variance; and were deemed "adequate". Internal consistency was assessed using Cronbach's alpha (α = 0.93-0.94) (Auquier et al., 2003; Chou et al., 2010) and was deemed "very good". Chou et al. (2010) discussed the translation process of both French and English versions of the scale; however, insufficient information was provided about group comparisons to rate the scale. Test-retest reliability was assessed using ICCs (0.79); however, the time interval between testing was unclear, potentially ranging from 0 to 30 days, and measurement error was not reported, which impacted the certainty of the evidence (Auquier et al., 2003).

Convergent validity, ranging from "adequate" to "very-good", was demonstrated through significant correlations with established measures including the EuroQol (r = 0.48, p<0.001), the Quality of Life Index (QoLI) (r = 0.29-0.74, p<0.001) with one domain (finances) not significant (0.11), the SF-36 (r = 0.29-0.61, p<0.001) (Auquier et al., 2003), and significant correlations among similar subscales were found with the WHOQOL-BREF, SQLS-R4 and Rosenberg Self-Esteem Scale (RESE) (Rosenberg, 1965) (r = 0.36-0.82, p<0.01-0.05) (Chou et al., 2010).

Discriminant validity, ranging from “adequate” to “very good” against the COSMIN risk of bias checklist, was demonstrated through significant relationships with different groups, outpatients (M = 59.3, p<0.05) (Auquier et al., 2003) and those with higher numbers of hospitalisations (U = 45.06 and 48.47, p<0.01 and p<0.04) (Chou et al., 2010). Sensitivity to change was indicated by a significant effect size (d = 0.44, p < 0.05) (Auquier et al., 2003).

The SF-36 was evaluated in five studies, rated between “inadequate” and “very good” against the COSMIN risk of bias checklist. Two studies reported good structural validity. Su et al. (2014) conducted a confirmatory factor analysis (CFA) and found a standardised root mean square residual (SRMR) of 0.07, which is within the acceptable threshold (<0.08). However, the study was rated “inadequate” on the risk of bias checklist, which may impact confidence in these results. Tunis et al. (1999) reported good EFA results, with factor loadings ranging from 0.46 to 0.70 and 56% variance explained, which was rated as “very good”, strengthening confidence in these results. However, content validity and cross-cultural validity were not reported, impacting the strength of the OMI.

Internal consistency was consistently reported across studies, and rated “very good”, with Cronbach’s alpha ranging from 0.79 to 0.90 (Pukrop et al., 2003; Tunis et al., 1999), although Nishiyama et al. (2009) reported inconsistent results (0.26-0.91), this was due to the differing levels of cognitive functioning between populations in this study. The results of the internal consistency analysis did not meet the criteria for pooling using meta-analysis, as Nishiyama et al. (2009) did not report the total scale Cronbach’s alpha. Test-retest reliability was rated “adequate”, with an ICC score of 0.87 and SEM of 10.03% reported (Su et al., 2014).

Convergent validity, rated “adequate”, was demonstrated through significant correlations with established measures. The MSQoL (r = 0.61- 0.84, p<0.01) (Pukrop et al., 2003) and WHOQOL-BREF (comparative fit index (CFI)=0.982, SRMR=0.050) (Su et al.,

2014). Additionally, the SF-36 mental health component was negatively correlated with other psychiatric measures such as the BPRS (r = -0.31, P<0.001), CGI (r = -00.15, P<0.001) and MADRs (r =- 0.55, P<0.001) (Tunis et al., 1999).

Discriminant validity, ranging from “doubtful” to “adequate”, was demonstrated across different studies. Several studies have reported significant differences between patient groups and the general population (M = 524.4 ± 144.10 vs 637.51, p < 0.01) (Su et al., 2014). Comparisons also indicated significant differences between schizophrenic patients (M = 41-82.1, p<0.001, except bodily pain) and the general population (Tunis et al., 1999). Significant differences were found between the different groups at admission (F = 38.4-230.2, p < 0.001) and discharge (F = 12.4-132.9, p < 0.001) (Pukrop et al., 2003). One study also assessed the discriminative validity of the SF-36 across levels of cognitive functioning using the mini mental state examination (MMSE) (Folstein et al., 1975) thresholds, assessed by the percentage of items correlated more strongly with their scale than with others (MMSE < 24: 14-100%, MMSE <25: 29-100%, MMSE <28: 75-100%) results indicate the discriminant validity is most robust in the highest cognitive group (Nishiyama et al., 2009)**.**

Responsiveness was reported in multiple studies among depressed patients; most results showed significant change (t = 3.51-9.39, P<0.001), with two non-significant findings. In patients with schizophrenia, significant changes were found across four to five SF-36 domains (t = 2.17- 4.19, p<0.001- 0.04, mean change = 1.31- 14.23, from not significant to P<0.0001) (Pukrop et al., 2003). One author (Newnham et al., 2007) used RCI (1.96**) to measure responsiveness, reporting changes over time.

The MHQOL was reported in one study (Ebadi & Rezaiye, 2025). Most domains were rated between “adequate” and “very good”, strengthening confidence in the evidence, except that cross-cultural validity was rated “inadequate” and content validity was rated “doubtful”; these results should be interpreted with caution. EFA and CFA were conducted with factor

loadings reported (0.39-0.95) and variance explained (59.45%). Internal consistency was calculated using Cronbach's alpha (0.88). Translation took place according to the WHO guidelines; however, not enough information was reported to rate this measurement property. Test-retest reliability was assessed with a two-week interval, and ICCs were reported (ICC = 0.97) along with the SEM (0.14). Construct validity and responsiveness were not evaluated, which impacted the overall certainty of the evidence.

The ReQoL reports seven of eight measurement properties, rated between "adequate" and "very good" against the risk of bias checklist. However, content validity was conducted in an outpatient setting; it remained important to report this outcome to highlight the gap in inpatient content validity. Keetharuth et al. (2018) conducted CFA, reporting factor loadings (0.24-0.47) and explained variance (83%). Internal consistency was calculated using Cronbach's alpha (0.92-0.96). The authors conducted a translatability assessment, and differential item functioning (DIF) was calculated. Test-retest reliability was calculated over two weeks (ICC 0.85-0.9), and measurement error was not reported. Convergent validity was demonstrated through significant correlations with the SWEMWBS, CORE-10 (r = >0.8, p <0.01) and between ReQoL 10 and 20 (r = 0.98, p <0.01). Discriminant validity was demonstrated between the general population and the patient population (ReQoL 10, SES = 0.93, ReQoL 20, SES = 1.05). Finally, sensitivity to change was demonstrated between six and twelve weeks, and the effect size was calculated using Cohen's d, with values ranging from small (<0.2) to moderate (<0.5).

The SQLS-R4 was reported in four studies. SQLS-R4 was rated between "inadequate" and "very good" across measurement properties against the risk of bias checklist. Two studies reported good structural validity. Kuo et al. (2009) conducted exploratory factor analysis (EFA), reporting factor loadings (>0.3) and variance explained (68.1%). Su et al. (2017) reported a three-dimensional construct with all factor loadings

significant and 47 to 76% variance explained. Martin and Allen (2007) conducted CFA reporting (CFI) (0.95) and a slightly high root mean square error of approximation (RMSEA) (0.11). This result could be due to the small sample size of 100 participants. SQLS-R4 has 33 items, indicating this sample size is inadequate and could impact the results. Structural validity was rated "inadequate", impacting the strength of the result.

Internal consistency, rated "very good", was reported in two studies using Cronbach's alpha (0.83-0.96) (Kuo et al., 2007; Martin & Allen, 2007). Cross-cultural validity was not discussed. Test-retest reliability, rated "adequate", was reported by Kuo et al. (2007) using ICCs (0.84), and SEM was calculated (86% met criteria).

Convergent validity, rated between "adequate" and "very good", was calculated across several studies showing correlations between SF-36 (r = -0.25 to -0.65, p<0.01-0.05) with negative polarities due to different polarities of scores on the two instruments (Kuo et al., 2007). Significant negative correlations were found in the LQOLP and psychosocial domain (r = -0.46 to 0.22, p < 0.00-0.03) and vitality domain (r = -0.54 to -0.18, p < 0.00-0.07) (Kuo et al., 2009). Martin and Allen (2007) reported correlations between the Hospital Anxiety and Depression Scale (HADS) - Anxiety (A) (Zigmond & Snaith, 1983) (r = 0.89, p <0.01) and HADS-Depression (D) (r = 0.70, p <0.01). Finally, Su et al. (2017) reported a significant negative correlation with. WHOQOL-BREF (r = -0.25 to -0.69, p <0.01).

Responsiveness, rated "adequate", was reported by Kuo et al. (2007) sensitivity to change over two weeks for those who showed changes in Brief Psychiatric Rating Scale (BPRS) (Overall & Gorham, 1988) (RC = -0.09 to 0.55), however, there was no predefined hypothesis and significance was not reported, therefore, an overall rating of (?) was given. The low risk of bias rating potentially impacts the strength of these results. Cross-cultural validity was not reported.

Two OMIs reported two or fewer adequate measurement properties, indicating that these OMIs are potentially the least robust outcome measures. One author reported only responsiveness for the QLI (d = 0.34, p<0.01) (Angstman et al., 2009), rated "inadequate" against the COSMIN risk of bias checklist. One author reported two measurement properties for the How Are You? Scale (Katavdakis et al., 1999), rated "adequate". They identified discriminant validity between patient scores and normative mean scores (t = 34.13, p < 0.001). Responsiveness was reported, showing a significant effect size (0.58, p<0.001) over 16 days.

Different results were reported across studies for OMI numbers 2, 12, 13, and 20. OMI two reported different internal consistency results across studies. Russo et al. (1997) reported a Cronbach's alpha between 0.82-0.90, N= 981, and Lehman (1988) reported a Cronbach's alpha between 0.44 and 0.88, N= 99. There was a discrepancy in sample size between studies, which could have contributed to the difference in scores. The risk of bias for these studies was rated "very good", potentially strengthening confidence in the reported results.

OMI 5, LQOLP, has differing internal consistency across studies. Ritsner et al. (2002) report a low Cronbach's Alpha (α= 0.42-0.92), and several other authors report a Cronbach's Alpha in the appropriate range (α= 0.72-0.93) (Kaiser et al., 1997; Gaite et al., 2000; Hansson et al., 1998). Ritsner et al. (2002) do not report the total QoL score, which explains the difference in the results. Internal consistency was rated "very good" against the COSMIN risk of bias checklist, potentially strengthening confidence in the reported results.

With OMI-13, the WHOQOL-BREF, different results for discriminant validity were reported across studies. With three studies reporting sufficient discriminant validity (Norholm & Bech, 2006; Oliveira et al., 2016; Su et al., 2014) and one study reporting insufficient (Chan et al., 2003). Chan et al. (2003) used appropriate statistical methods showing the

relationship between different populations (inpatient, long stay care home, and halfway house residents) and certain domains of QoL, such as life satisfaction ($F = 7.74$, $p<0.001$) and environment ($F = 3.53$, $p<0.03$). It can be assumed from the literature that differences were to be expected. However, only some differences were found. Therefore, it was rated as (-) and was rated as "adequate" in terms of risk of bias.

There were also differing results reported for internal consistency across studies. Norholm and Bech (2006) reported only the Lovinger test of unidimensionality and did not report Cronbach's alpha, rated as "inadequate" against the risk of bias checklist. Oliveira et al. (2016) reported a Cronbach's alpha of between 0.68-0.85, which although high on most domains, there was no total QoL score provided and therefore the lowest domain score does not meet COSMIN criteria for internal consistency, however, this result was rated "very good" against the risk of bias checklist, strengthening confidence in this result.

Overall, a small subset of OMIs, including S-QOL-41, ReQoL, SF-36, MHQOL, and SQLS-R4, demonstrated consistently strong evidence across multiple measurement properties. However, common limitations across OMIs included poor or absent content and cross-cultural validity and lack of predefined hypotheses for construct validity and responsiveness.

Table 4

Results Table: Measurement Properties

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
1 LQoLI-BREF	Anderson et al., 2002	NR	NR	NR	NR	NR	NR	CV: intercorrelations within subjective scales: 0.37- 0.62* Objective scales independent, one significant relationship between social contact and finance: 0.55** (+) DV: No differences between patient and normative sample except: living situation: MD= 0.88-1.05** Family relationships: MD= 0.48* Family contact: MD=0.35-0.42* (+)	The scale was reported stable across time (6 weeks) effect size not calculated (-)
2 LQoLI Full	Russo et al., 1997	(+)	Factor loadings: 0.06-0.13 variance explained: 63.5% (+)	α=.82-.90 (+)	NR		NR	CV: satisfaction and functional indices assess unique aspects of QoL: ≤0.43 sharing on average less than 18% of common variance as predicted. (+) DV: significant differences reported between diagnostic groups across domains: (Wilk's λ=.93; *F* (16,2909) = 4.34*** (+)	Effect size between 0.4-0.82*** (102 days +/- 10.6) (+)
	Lehman, 1988			α= 0.44-0.88 (-)	NR	ICC= 0.29-0.98 (one week test-retest) (-)	NR	CV- correlations between two alternative forms of general life satisfaction scale: 0.63- 0.72**** (+)	

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
								DV- Between diagnosis: Depression: *r* = -. 17 to -.56* Anxiety: *r* = -.25 to -.33** (+)	
	Goodwin & Madell, 2002				NR		NR	CV- strong correlations between Lehman's QoL and QoL measures BIGALOW (Bigalow et al., 1991) and REHAB scale (five accomplishments) (Murray et al., 1998): 0.27-0.68** with QoL increasing with level of functioning. Mood and QoL: 0.42-0.5* (+) DV- QoL scores significantly differed across residences: H = 11.64-20.6, df=2*** (+)	
3 Q-LES-Q	Bishop et al., 1999	NR	Principle component analysis (PCA) revealed a four-factor model: 0.32- 0.45 Principle factor analysis (PFA) showed factors loadings: 0.35-0.50 (no variance explained) (?)	α=.89-.96 (?)					NR

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
	Ritsner et al., 2002 (10 items removed)	NR		α=.84-.95 (?)	Translated and back translated by a bilingual research assistant: Pilot tested N= 10. (?)	ICC=.47-.8 (two-week interval) (-)	NR	CV: Positive and Negative Symptom Scale (PANSS) (Kay et al., 1987):-0.24** Montgomery-Asberg Depression Rating Scale (MADRS) (Fantino & Moore, 2009): -0.46** No significant correlation with MARS-mania (0.13). (+) DV: Between groups Patients: *r* =0.63*** Professionals *r* =0.39*** (+)	NR
4 Q-LES-Q SF	Pitkanen et al., 2011	NR	Classical test theory (CTT): PCA: Factor loadings: 0.4-0.7 Variance explained: 56% Kaiser Meyer Olkin (KMO): 0.889 Bartlett's test of sphericity: χ 2 1117, df 105*** (+)	α=.89 (.40-74) (+)	NR	NR	NR	CV: EQ-5D: ρ= 0.445*** (+) DV: GAF (Global assessment of functioning): Differences in functioning: U=3357*** Diagnosis: U=4275 P=0.354 Not significant (NS) (+)	NR
5 LQOLP	Ritsner et al., 2002 (only 27 items considered)	NR		α=.42-92 (-)	Translated and back translated by a bilingual research assistant: Pilot test N= 10 (?)	ICC= .34-.82 (two week) (-)	NR	CV:PANSS: -0.19* MADRS: -0.34** MRS mania: NS (0.05). As expected from literature. (+) DV- Differences in QoL index for: Patients *r* =0.63*** Professionals *r* =0.39*** (+)	NR

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
	Kaiser et al., 1997	NR	Factor loadings: 0.51-.73 Variance explained: 15-42% Eigenvalue: > 0.50 (-)	α=.73 (.69-.82) (?)			NR	CV: BPRS: β = -0.29 to -0.33 r_2 = 0.09, F = 10.5*** BPRS total score: β = -0.25*** (+) DV: Settings: F= 0.83-6.10 NS- *p* <.001 across domains of QoL (Bonferroni corrections) (+) Demographics and QoL: <10% shared variance across domains. (+)	NR
	Gaite et al., 2000	NR		α=0.87 (0.3-0.88) (?)	Back translation completed: focus groups completed: NR: Pilot testing was completed: NR (?)	ICC=0.82 (7-15 days) (+)	NR		NR
	Hansson et al., 1998	NR		α=0.93 (0.53-0.93) (?)	Forward and back translations by first author: Pilot study N= 35 (?)	Pearsons *r* = 0.83-0.91*** (two weeks) (?)	NR		NR

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
6 S-QOL-18	Boyer et al., 2010	NR	CTT PCA: 8 factor structure Factor loadings: >0.4 Variance explained: 78% Inter-response theory (IRT): Rasch model: INFIT 0.91-1.05 (+)	α=0.88 (0.72-0.84) (+)	NR	ICC=0.73 (3-week interval) (+)	NR	CV: S-QOL-4: >0.88** QLS: >0.5** SF-36: 0.30–0.58** QoLI: 0.28–0.58** (except finances NS) with two of the three SQLS dimensions: -0.21 to -0.56* EuroQol: 0.50** PANNS: -0.28** (+) DV: Age: 13 to -24** (3 dimensions) Gender: Male 58.7 and 65.7, female 49.4 and 58.4**(2 dimensions) Education: (2 dimensions) 60.4-70.4** Patient populations: 43.4- 64.6* Duration of illness: (1 dimension) -0.21** in line with hypothesis. (+)	Effect sizes >0.02* for multiple domains and index score over 6 months. (+)
7 SQOL-41	Auquier et al., 2003	(+)	CTT EFA eight factor structure. Factor loadings: 0.46-0.84 Variance explained: 52% (+)	α= 0.94 (0.53-0.90) (+)		ICC= 0.79 (0.64-0.79) (not reported, could be 30 days) (+)	NR	CV: Intercorrelation within scale: 0.24- 0.60*** Two domains not significant: Psychological well-being and self-esteem: 0.66 Family relationships and resilience: 0.11 ($p = 0.14$) SQOL41: matching domains: psychological wellbeing, self-esteem, physical wellbeing	effect size index score 0.44* It was predicted that at least 0.2 indicated change over time. (+)

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
								EuroQol: 0.48*** QoLI: 0.29-0.74*** SF-36: 0.29-0.61*** Unique areas: autonomy, resilience and sentimental life: not significantly correlated: as expected. (+) PANSS: -0.23** Calgary Depression Scale for Schizophrenia (CDSS) (Addington et al., 1993): -0.39*** GAF: 0.27*** Clinical Global Impression of Severity (CGI) (Guy, 1976):-0.32*** As only one hypothesis deviated, overall score is (+). DV: Between groups: Outpatients score: M = 59.3 SD 17.8* Inpatients score: 52.1 SD 18.3* (only *P* reported) Length of illness: NS Education: NS Age: NS (+)	
	Chou et al., 2010			α = 0.93 (0.44-0.93) (+)	Forward translation and backward translation. (?)		NR	CV: SQLS-R4, WHOQoL-BREF, and RESE: 0.36–0.82* (+)	

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
								DV: Mann–Whitney U test showed significant difference between subgroups based on number of hospitalisations and episodes U= 45.06** and 48.47* (+)	
8 QLI	Angstman et al., 2009		NR	NR	NR	NR	NR	NR	Effect size 0.34* (average 20 days). No predefined hypothesis. (+)
9 QLIS	Franz et al., 2012	(+)	NR	α= 0.63-0.88 with CTT described (-)	NR	ICC=0.73-0.87 (7-10 days apart with stable non hospitalised patients)	NR		NR
	Franz et al., 2012		NR		NR		NR	CV: WHOQOL-BREF: 0.30- 0.72 variance: 4.2-55.8% Subjective Well-being Under Neuroleptic Treatment Scale (SWN) (Maurino, 2012): 0.14-0.83 variance: 1.6-42.3% Satisfaction scales: 0.18- 0.69 variance: 2.5-51.7% (+)	NR
10 How Are You? Scale	Katsavdakis et al., 1999 (study 2)	(+)	NR	NR	NR	NR	NR	DV: Between groups: Patients and normative sample: t=34.13*** (+)	Effect size 0.58*** (mean 16 days) (+)

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
11 QLS	Nair et al., 1996	(+)	NR	α = -1.5 - 0.8: negative value was reported for work/vocational rehab. No evidence of unidimensionality (-)	NR	NR	NR	DV: Between groups: Staff: M = 2- 3.17 SD= 0.2- 1.27 Patient scores M= 1.86- 2.49 SD= 0.35- 1.07 *P*<0.05 (only *P* reported) (+)	Responsiveness 6-month period between: Old hospital: M= 0.98-2.43 New hospital: M= 0.94-2.26 No effect size reported. (-)
12 Sf-36	Newnham et al., 2007 (only subscales related to mental health)	NR			NR			DV: Between recovery groups on the Depression Anxiety and Stress scale (DASS-21) (Lovibond, 1995): Depression: F(3,1533) = 95.51** Anxiety: F(3,1534) = 49.87** Stress: F(3,1525) = 67.62** Recovered patients (via SF-36) also had higher Q-LES-Q scores: F(3,858) = 57.30** lower clinician-rated HONOS scores**:** F(3,1524) = 10.60** No hypothesis stated. (+)	(2 year period) RCI: 1.96**. Effect sizes were not calculated (+)
	Pukrop et al., 2003	NR		α= .79 -0.83 (+)	NR			CV: MSQoL: 0.61- 0.84** PANSS: NS Hamilton Depression Rating Scale (HAM-D) (Hamilton, 1960): -0.17 to -0.53* No hypothesis made. (+) DV: Between three groups of patients: Admission: F= 38.4-230.2*** Discharge: F= 12.4-132.9***	Responsiveness admission and discharge (minimum time of 3 weeks) Depressed patients: t= 3.51-9.39*** (2 NS results) Schizophrenic patients: t= 2.17-4.19*** (4 domains) No effect sizes.

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
								(+)	(+)
	Nishiyama et al., 2009	NR		MMSE<24 α= 0.3-0.91 MMSE <25 α= 0.26-0.88 MMSE <28 α= 0.63-0.9. (-)	NR			CV: Convergent validity was assessed by the percentage of items correlating ≥ 0.40 with their hypothesized domain, using multi-trait scaling analysis. Correlation rates ranged from 20% to 100% across domains and MMSE groups. Convergent validity declines with increasing cognitive impairment. (+) DV: Percentage of items correlating more strongly with their own scale than with others. MMSE < 24: 14%–100% MMSE < 25: 29%–100%. MMSE < 28: 75%–100%. (+)	
	Su et al., 2014	NR	CFA CFI: 0.918 Incremental fit index (IFI): 0.919 TLI: 0.885 RMSEA: 0.145 SRMR: 0.073 COSMIN guidelines will accept CFI and SRMR scores. (+)		NR	ICC= .87 (0.48-0.81) (2 weeks) (+)	Measurement error SEM= 10.03% smallest real difference (SRD) = 27.8% (acceptable) (+)	CV: WHOQOL-BREF: traits and methods correlated (CFI = 0.982, IFI=0.982 TLI 0.970 RMSEA = 0.068, SRMR = 0.050) (+) DV: Between groups: Taiwan general population: M= 524.4 +/- 144.10 Patients: M= 637.51 (p <0.01)	

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
								only *P* reported) (+)	
	Tunis et al., 1999	NR	CTT EFA: Factor loadings: 0.46-0.70 Variance explained: 56% (+)	α= 0.9 (0.76-0.91) (+)	NR			CV: mental health component and: BPRS: -.31*** CGI: -.15*** MADRs: -.55*** Physical component: NS (+) DV: Age: NS mental health component *r*=-0.19*** physical health Between general population: M= 60.9 - 84.2 Schizophrenic patients means: M= 41-82.1***, except bodily pain. (+)	Responsiveness (6 weeks) Mean change= 1.31- 14.23 NS to *p* <0.0001. No effect size calculated. (+)
13 WHOQOL-BREF	Su et al., 2014	NR	CFI: 0.967 IFI: 0.967 TLI: 0.900 RMSEA: 0.216 SRMR: 0.037 (+)		NR	ICC= 0.81-0.88 (2 weeks) (+)	Measurement error: SEM total: 5.55% SRD total score: 15.4% (+)	CV: SF-36: traits and methods correlated (CFI = 0.982, RMSEA = 0.068, SRMR = 0.050) (+) DV- Taiwan general population WHOQOL-BREF: M= 52.82 8.29 Patient population.: M= 57.35 *(p* <0.01) (only *P* reported) (+)	NR

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
	Oliveira et al., 2016	NR	CTT PCA:Factor loadings: 0.55-0.9 Varience explained: 59.51% CFA: X2=180.262*** CFI: 0.955 GFI: 0.914 TLI: 0.942 RMSEA: 0.050 SRMR: 0.055 (+)	α= 0.68-0.85 (-)	NR			CV: All domains predicted QOL significantly: General QoL adjusted: 0.51* (+)	NR
	Norholm & Bech, 2006	NR		Loevinger's homogeneity coefficients (0.34–0.47 in patients) Cronbach's alpha was not calculated. (?)	NR			DV: Schizophrenic population total QoL: M= 25.6 General population total QoL: M= 6.8 Schizophrenic population gender differences: p <0.05 Gender differences between QoL and MDI scores: male: M= 22 Female: M= 30 (p <0.01) (only *P* reported) (+)	NR
	Chan et al., 2003	NR			NR			CV: BPRS: -0.1 to -0.23* (finance not significant) (+) GAS: Psychological domain: 0.14* Environment domain: 0.15* Total negative life events: -0.42 to -0.23*	NR

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
								Subjective well-being sociodemographic and clinical variable: F= 3.914- 8.725**** (+) DV: Between populations: Life satisfaction: F=7.74*** Environment: F=3.53* all other domains were NS. (-). Education: Life satisfaction: -0.25**** Environment: -0.15* All other domains don't correlate (-)	
	Van de Willige et al., 2005	NR			NR			CV: EQ-5D: 0.55-0.58*** PANSS: -0.54*** Auditory Hallucination Rating Scale (AHRS) (Haddock et al., 1999) negative contents: -0.31** distress: -0.54*** Social functioning measured by Groningen Social Disabilities Schedule (GSDS) Wiersma et al., 1990): -0.57*** (+)	Significant changes in treatment group: t = 2.11*. No effect size (18 months) (+)
14 MSQOL	Pukrop et al., 2003	NR	NR	α= 0.83- 0.87, no evidence of unidimensionality provided. (?)	NR	NR	NR	CV: Sf-36: 0.61- 0.84** There were no significant correlations with negative and positive symptoms (PANSS positive and negative symptoms).	Admission and discharge (minimum time of 3 weeks) Depressed patients: t = 3.79-16.60***

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
								HAM-D: -0.18 to -0.51* (role physical NS) (+) DV: Comparison between patient groups: Admission: F= 40.8-364.5*** Discharge: F= 5-46.8*** (+)	Schizophrenic patients: t = 0.98-7.63*** (3 domains NS) no effect size. (+)
15 EQ-5D	Pikanen et al., 2011	(+)	PCA KMO: 0.701 Bartletts test of sphericity: χ 2 166, df 10,*** EFA: factors loadings: 0.52-0.708 Variance explained: 41.3% (-)	α= 0.63 (0.530-0.63) (-)	NR	NR	NR	CV- Q-LES-Q: ρ =0.445*** (+) DV: GAF: Lower functioning: Median score 0.64*** Higher functioning: Median score 0.89*** Diagnosis: NS (+)	
	Van de Willige et al., 2005				NR	NR	NR	CV- WHOQOL: 0.55-0.58*** PANSS factors: -0.31 to -0.47** AHRS negative contents and visual analogue scale (VAS) score: -0.46*** Negative contents and time trade off (TTO): NS Distress and TTO and VAS score: -0.25 to -0.50* Social functioning measured by GSDS: -0.29 to -0.41* (+)	(18 months): t = 2.37* No effect size. (+)
16 MHQOL	Ebadi & Rezaiye, 2025	(+)	CTT EFA and CFA	α= 0.876 (+)	Translation (WHO): N= 10	ICC= 0.97 (two-week interval)	SEM= 0.14 (+)	NR	NR

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
			59.45% variance explained with one factor removed. GFI (0.90), CFI (0.97), IFI (0.97), NFI (0.95), PNFI (0.64), RMSEA (0.14), CMIN/DF (3.5), and RMR (0.049). Factor loadings (0.39-0.948) scree plot used (+)		Loop analysis: N= 30 (r <0.3). (?)	(+)			
17 SF-12	Soysal Gunduz et al., 2021	NR	CTT EFA/PCA KMO: 0.83 Bartletts: X2 =938.30**** Factor 1 eigenvalue of 5.58 representing 46.57% of variance. Factor 2 eigenvalue of 1.57 representing 13.14% of variance. 59.7% cumulative variance (+)	Physical health component summary α= 0.80 (0.32-0.73) mental health component summary α= 0.88 (0.60-0.78) (+)	NR	NR	NR	CV: NHP health profile: -0.03 to -0.58 (no p reported) (?) DV: Surgical group: M= 46- 80.5 Psychiatric group: M= 41.7- 65, rheumatology group: M= 47.9-88.9 (p<0.0001) (only p reported) (+)	NR
18 SQLS	Kaneda et al., 2002	NR	CTT PCA: 55.8% variance accounted for and factor loadings greater than 0.5 (three domains), some factors failed to load (?)	α = 0.73-0.93 (?)	Forward and back translation: Pilot tested: N= 5 (?)	NR	NR	CV- SF36: -0.680*** and -0.375** No significant correlation with WHOQOL and Drug Induced Extrapyramidal Symptoms Scale (DIEPSS) (Peljto et al., 2017) Three out of four hypothesis confirmed (75%) (+)	NR

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
19 ReQoL	Keetharuth et al., 2018	(+) OutPatients	CFA RMSEA: ≤0.08 CFI: >0.95 IRT: *r* = 0.98 Expected A Posteriori (EAP): 0.95. CFA: 6 factor model EFA: 2 factor solution: 0.8. IRT analysis: *r* =0.98 MISFIT: *p* <0.05 DIF (age gender, ethnicity, gender) (+)	REQOL 10: α =0.92 REQOL20: α=0.96 (+)	Translatability assessment was conducted and DIF calculated (?)	ICC: 2-week interval ReQoL 10: ICC= 0.85 ReQoL 20: ICC= 0.87	NR	CV: Short Warwick-Edinburgh Mental Health Wellbeing Scale (SWEMWBS) (Stewart-Brown et al., 2009) and Clinical Outcomes in Routine Evaluation (CORE-10) (Barkham et al., 2013): >0.8** REQOL 10 and REQOL 20: 0.98** (+) DV: Between general population and patient population (hypothesis made). REQOL 10: SES= 0.93 REQOL 20: SES= 1.05 (+)	Effect size: small <0.2 to moderate <0.5 change over time. (no *p* reported) (+)
20 SQLS-R4	Kuo et al. 2007	NR		α= 0.84-0.92 for both domains (two domains) (+)	NR	ICC= 0.844 for both domains (two weeks)	SEM= 86% met criteria for changes smaller than 2 SEM (+)	CV: SF-36: -0.25 to -0.65* Negative correlations due to different polarities of the scores on the two instruments. expected correlations. Psychosocial domain: -0.67 (+)	Two-week period: Responsiveness coefficient: (RC) 0.55 and 0.03 improved and 0.27 and -0.09 in deteriorated (?)
	Kuo et al., 2009	NR	EFA (three domains) KMO: 0.87 Bartlett: 1763.12*** PCA varimax rotation: 68.1% variance explained factor loadings all above 0.3 (+)		NR			CV: PANSS and GAF: NS T-LQOLP: psychosocial domain: -0.46** and -0.22* Vitality domain: -0.54** and -0.18* (+)	

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
	Martin & Allen, 2007	NR	CFA: Two factor correlated CFI (0.95) TLI (0.98) RMSEA (0.11) SRMR (0.94) (-)	α = 0.83-0.96, (+)	NR			CV: HADS A: 0.89** HADS D: 0.70** (+)	
	Su et al., 2017	NR	CFA (three domains) CFI (0.967), TLI (0.964), RMSEA (0.068), SRMR (0.072), 3-dimensional construct all factor loadings significant factor loadings for 33 item model 0.47-0.76% variance explained (+)		NR			CV: WHOQOL BREF:-0.26 and -0.69** (+)	
21 QOL-BD	Michalak et al., 2010	(+)	Latent structure factor analysis. Factor loadings >0.3 KMO: 0.92 Bartlett's test of sphericity: v2 = 9075.85*** 12-factor extraction explained 71% of the shared variance. (+)	α= 0.79-0.93 (+)	NR	Internal reliability test-retest 0.46-0.77 (7 to 10 days) (-)	NR	CV- Expected correlations were found between QOL measures (SF36, QLESQ, SWL) and symptom severity measures (PANSS HAMD) HAM-D: -0.62 (b=)0.323, t=)2.04* Young Mania Rating Scale (YMRS) Young et al., 1978): 0.18 SF-36: 0.63 Q-LES-Q: 0.89 (no *p* reported) In line with hypothesis. (+)	NR

Instrument name	Reference	Content validity	Structural validity (rating)	Internal consistency (rating)	Cross-cultural validity (rating)	Reliability (rating)	Measurement error (rating)	Construct validity (rating) Convergent validity (*r*) Discriminant validity (rating)	Responsiveness (rating) (d)
22 Q-LES-Q-18	Ritsner et al., 2005	NR	Varimax rotation factor analysis Four factors were identified on the highest eigenvalues (>1.0): These factors accounted for 31.4%, 26.3%, 22.7% and 19.5% of the total variance. (+)	α= 0.96 (0.74-0.97) inpatient sample (+)	NR	ICC=0.91 (2 week) (outpatients) (not rated for inpatient)	NR	CV- LQOLP: 0.33- 0.51*** QLIS: 0.51-0.64*** Q-LES-Q: 0.55 to 0.64*** (+) DV: Between patients and healthy controls: F=42.0*** (+)	NR
23 Henrich's Quality of Life Scale	Simon-Abbadi et al., 1999	NR	CTT Factor analysis varimax rotation 53.5% variance explained. Negative factor loadings reported. (-)	α = (0.9) 0.6-0.9 (-)	NR	inter rater reliability, no time interval reported ICC- 0.72-0.79 (+)	Weighted Kappa= 0.1-0.7 (?)	CV: With subjective quality of life assessment: 0.3** GAF: 0.56*** (+)	NR

p<0.05* *p*<0.01**P<0.001*** P<0.0001****

Content Validity

Content validity is presented in Table 5. Information on the construct and origin was taken from the papers included in this review. Information on the target population and intended context has been gathered, highlighting which instruments were specifically developed for use in an adult inpatient psychiatric setting and which have been adapted for use in this context. Finally, a summary of content validation has been included, summarising the relevance, comprehensiveness and comprehensibility of the outcome measure as reviewed by patients and professionals.

Content validity was reported for eight OMI. The ReQoL reports content validity from patients in a community mental health team (CMHT), so the results may not be relevant to an inpatient population. This study was included because other measurement properties were tested with an inpatient population, and the review aims to highlight potential areas for the development of OMIs for this population. The SF-36, and EQ-5D were not explicitly developed for mental health or inpatient purposes; they were developed as generic instruments for measuring HRQoL and evaluating medical outcomes and healthcare interventions. Michalak et al. (2010) used appropriate methods to test relevance and comprehensibility of the QOLBD amongst staff and patients. Furthermore, they conducted a pilot study. This indicates that it was adequately assessed according to COSMIN, and it was rated “very good” against the risk of bias checklist.

Table 5

Content Validity

OMI	Reference	Clear construct (definition)	Origin of construct (theory or evidence base)	Target population for which the PROM was developed	Intended context of use	Content Validity Relevance, comprehensibility, comprehensiveness		
						Pilot testing	Patients	Professionals
1	Anderson et al., 2002	Objective QoL: based on functional access to resources	Response to requests for a briefer LQOLI.	NR	NR	NR	NR	NR
2	Russo et al., 1997	Personal characteristics, objective and subjective QoL. Multidimensional definition of QoL	National QoL studies and existing QoL measures	Mentally unwell patients	Inpatient and outpatient mental health		NR	
	Lehman, 1988					Comprehensibility: Interview format (45 minutes), pilot trials with target population. No description of method or design.	NR	Comprehensiveness: selection of items based upon national quality of life studies. Items were pulled from existing instruments. Comprehensibility: modified to increase
3	Goodwin & Madell, 2002	QoL measures of the degree of enjoyment and satisfaction experienced by subjects in areas of daily functioning.	Built on previous research of HRQoL, little research in QoL within psychiatry.	Medical and psychiatric disorders. Developed for use with depressed patients	NR	NR	NR	NR
	Bishop et al., 1999							
4	Ritsner et al., 2002 (10 items removed)	QoL enjoyment and satisfaction experienced by	NR	Patients with mental disorders	NR	NR	NR	NR

OMI	Reference	Clear construct (definition)	Origin of construct (theory or evidence base)	Target population for which the PROM was developed	Intended context of use	Content Validity Relevance, comprehensibility, comprehensiveness		
						Pilot testing	Patients	Professionals
		patients with mental disorders						
5	Pitkanen et al., 2011 Ritsner et al., 2002 (only 27 items considered) Kaiser et al., 1997 Gaite et al., 2000 Hansson et al., 1998	Defined by objective and subjective life circumstances. HRQoL, multidimensional.	Developed from LQOLI in the context of deinstitutionalisation.	Schizophrenic patients-mental illness related	Multiple psychiatric subpopulations. Mixed settings.	NR	NR	NR
6	Boyer et al., 2010	QoL in schizophrenia-multidimensional	Condensed version of S-QOL-41	Patients with schizophrenia	NR	NR	NR	NR
7	Auquier et al., 2003 Chou et al., 2010	Subjective QoL, multidimensional.	Somatic medicine and human sciences.	Patients with schizophrenia	Inpatient and outpatient mental health	Comprehensibility: Item generation, 90 items: pilot tested with target population N= 40 Relevance: Producing 83 question format. A second survey further reduced the items excluded irrelevant items.	Relevance: Videotaped interviews with Schizophrenic in and outpatients: N=20 Trained interviewer and interview guide used. Relevance: Cognitive debriefing N=27, asked to point to important domains	Comprehensiveness: Involvement of a steering committee who confirmed the content was meaningful.

OMI	Reference	Clear construct (definition)	Origin of construct (theory or evidence base)	Target population for which the PROM was developed	Intended context of use	Content Validity Relevance, comprehensibility, comprehensiveness		
						Pilot testing	Patients	Professionals
8	Angstman et al., 2009	Subjective QoL: important needs, goals, and wishes, and fulfilment.	Developed from LQOLI	Social phobia, generalised anxiety disorder, and depression	A range of settings	NR	NR	NR
9	Franz et al., 2012	Subjective QoL emphasising context.	Identifies that QoL needs a better theoretical foundation.	Schizophrenia	Hospital: day, in and outpatient	Comprehensibility: Pilot version developed and tested. Consideration to comprehensiveness and difficulty of questions. Final version tested.	Relevance: Open-ended interviews with schizophrenic patients from different care settings to identify relevant items	Comprehensiveness: Experts in the area paraphrased each item and excluded redundant items.
	Franz et al., 2012							
10	Katsavdakis et al., 1999 (study 2)	Considering experiences that the service user themselves finds important and the negative impact of mental health (subjective QoL)	Developed to address both the concerns of the consumer movement and the necessity of providing objective outcomes - reference to biopsychosocial model in differing context	Psychiatric patients	Inpatient and outpatient	Relevance: Pilot tests reduced initial 90 items that had been generated from the FACE domains to 62. The final version was then pilot tested.	Comprehensiveness and relevance: Focus groups inpatient and outpatient. To feedback on comprehensibility and removing irrelevant items and altering content reduced to 55 items.	Comprehensibility: Understanding from research to develop a scale that is more comprehensible.
11	Nair et al., 1996	QoL in chronically mentally ill (subjective)	Based on the material and social world together with subjective experiences as contributing to the individual patient's perception of their QoL	Chronically mentally ill.	Rehabilitation long stay psychiatric inpatients	Relevance: Pilot test final form: N=50 patients, reduced to exclude religion.	NR	Relevance and comprehensibility: Medical and nursing staff had a group discussion and provided recommendations. Following these 8

OMI	Reference	Clear construct (definition)	Origin of construct (theory or evidence base)	Target population for which the PROM was developed	Intended context of use	Content Validity Relevance, comprehensibility, comprehensiveness		
						Pilot testing	Patients	Professionals
								life domains were selected.
12	Newnham et al., 2007 (only subscales related to mental health)	QoL- objective and subjective as defined by the world health organisation (WHO)	Originated from health insurance and development in the article to apply to psychiatric population- The health insurance experiment (Brook et al., 1983).	Generic instrument for medical outcome measurement General adult population.	Various healthcare setting, in and outpatient.	NR	NR	NR
	Pukrop et al., 2003							
	Nishiyama et al., 2009							
	Su et al., 2014							
	Tunis et al., 1999							
13	Su et al., 2014	QoL as defined by WHO. Objective, subjective and global level of functioning.	QoL is the result of a patient's continuous efforts to adapt to both internal (illness related) and external circumstances.	Adults with disease or impairment and healthy adults	Across different contexts	NR	NR	NR
	Oliveira et al., 2016							
	Norholm & Bech, 2006							
	Chan et al., 2003							
	Van de Willige et al., 2005							
14	Pukrop et al., 2003	QOL as a multidimensional measure	Created from the combination of items from several questionnaires (LQLP, NHP, Psychological general well-being schedule,	Healthy controls, patients with depression and patients with schizophrenia	Designed for a psychiatric setting	NR	NR	NR

OMI	Reference	Clear construct (definition)	Origin of construct (theory or evidence base)	Target population for which the PROM was developed	Intended context of use	Content Validity Relevance, comprehensibility, comprehensiveness		
						Pilot testing	Patients	Professionals
			questionnaire for everyday living, questionnaire for the state of health, satisfaction questionnaire, SF-36, sickness impact profile).					
15	Pikanen et al., 2011 Van de Willige et al., 2005	A preference-based measure of HRQoL (subjective QoL)	Developed by the EuroQoL group.	Measure of QoL in various diseases. A generic QoL measure, that has been used in a range of population surveys, including the English national health survey.	To be used in the evaluation of different healthcare interventions	NR	NR	NR
16	Ebadi & Rezaiye, 2025	Based on WHO definition of QoL.	Origin of scale has stemmed from impact of QoL after psychiatric disorder.	Mental health	NR	Pilot test, Loop analysis: N= 30 patients	Relevance and comprehensibility: Feedback: N=10 psychiatric patients	Comprehensibility: Feedback on wording and scoring: : N=10 professionals Relevance: N=10 professionals provided ratings on relevance.
17	Soysal Gunduz et al., 2021	HRQoL- health status of a patient	HRQoL scales were initially developed for	Physical, mental health	Across multiple contexts	NR	NR	NR

OMI	Reference	Clear construct (definition)	Origin of construct (theory or evidence base)	Target population for which the PROM was developed	Intended context of use	Content Validity Relevance, comprehensibility, comprehensiveness		
						Pilot testing	Patients	Professionals
		was considered beyond their illness and objective criteria.	research purposes. Derived from SF36	and general population.				
18	Kaneda et al., 2002	QoL- developed in response to lengthier scales and to be self-report	Brief self-report QoL measure specific to people with schizophrenia, Wilkinson et al. (2000) developed a new scale, the SQLS.	Schizophrenia	Clinical trials	NR	NR	NR
19	Keetharuth et al., 2018	Based on aspects of QoL that matter to mental health service Users (subjective QoL)	Based on the following framework: Connectedness, Hope, Identity, Meaning and Empowerment (CHIME) and O'Connell (2014) framework.	Mental health	Multiple contexts including secondary care (in and outpatient) primary care, increasing access to psychological therapies (IAPTs), general population, CAMHS.	Relevance: N=19 outpatients' interview to generate 7 themes.	Comprehensiveness, comprehensibility and relevance: Qualitative interviews with adult from CMHT and young people added 12 new items and eliminated 39 items and rephrasing.	Comprehensiveness: Systematic review of literature: 1597 items were generated and reduced to 88 items. Scientific group considered all items and REQOL was developed. Feedback sought: N= 33 clinicians.
20	Kuo et al. 2007 Kuo et al., 2009 Martin & Allen, 2007	Subjective QoL, based on WHO definition.	Derived from Wilkinson et al. (2000) SQLS	Schizophrenic patients	NR	NR	NR	NR

OMI	Reference	Clear construct (definition)	Origin of construct (theory or evidence base)	Target population for which the PROM was developed	Intended context of use	Content Validity Relevance, comprehensibility, comprehensiveness		
						Pilot testing	Patients	Professionals
	Su et al., 2017							
21	Michalak et al., 2010	QoL as defined by WHO (Subjective experience)	Developed from the need for additional forms of assessment outside of objective measures to measure treatment outcome of recovery. Developed as disorder specific scales can potentially be more sensitive to change.	Bipolar	Clinical trials of pharmacological or psychosocial treatment interventions, longitudinal monitoring, or routine clinical care	Comprehensibility: Pilot questionnaire: N= 225 participants. Refinements made before final study.	Relevance and comprehensiveness: Interviews: N=35 in and outpatient. Tape recorded and to saturation.	Comprehensiveness: N= 5 caregivers N= 12 researchers participated in interviews. Data triangulation with literature.
22	Ritsner et al., 2005	HRQoL	Derived from QLESQ, to reduce the items to a core subset.	BD, depression, PMDD, and acute coronary syndrome, psychoses, PTSD, postsurgical states, schizophrenia	Used in QoL outcome studies of mentally disabled patients	NR	NR	NR
23	Simon-Abbadi et al., 1999	HRQoL- no consensus definition reported in paper.	Biopsychosocial perspective For use with schizophrenic deficit syndrome, and as a measure of psychosocial functioning	Schizophrenia specific	Psychopharmacologic treatments, predominantly in outpatients	NR	NR	NR

Quality of the Evidence (QOE)

The GRADE approach (Guyatt et al., 2011) was used to assess the overall QOE for each measurement property across the included studies, as presented in Table 6.

The MHQOL, S-QOL-41, SF-36, SQLS-R4, and ReQoL were initially identified as covering the most significant number of measurement properties; however, using the GRADE framework showed limitations in the QOE. As such, these tools may not represent the strongest OMIs despite their broader coverage of measurement properties.

Several OMIs demonstrated relatively comprehensive assessment across measurement properties. For instance, the Sf-36 shows a “moderate” to “high” quality of evidence across five measurement properties, indicating a strong measure for this population. However, QOE for responsiveness was “low”, and there was no reported content validity or cross-cultural validity.

The absence of sufficiently rated content validity, as per COSMIN, was a key limitation across most OMIs. It is recommended that instruments report content validity as a priority, and the outcome measure should not be rated if content validity is not reported (Mokkink et al., 2024). However, results for all measurement properties have been evaluated to highlight the strengths and weaknesses of each available outcome measure, and the information can be used to inform future research.

OMI 7, the SQOL-41, has covered most measurement properties except measurement error. The lower QOE ratings for structural validity, responsiveness, and reliability were impacted by sample size. Furthermore, the reliability results were rated as “inadequate” due to the risk of bias. Overall, this scale shows promising evidence for use in an adult psychiatric inpatient setting. However, this scale was designed as a Schizophrenic specific scale and might not be useful across multiple mental health diagnoses seen in an adult inpatient

population. Furthermore, the study was conducted with a mixed sample of patients, which is reflected in the QOE score.

OMI 16, the MHQOL, has assessed six out of eight measurement properties, with a "very low" QOE rating for cross-cultural validity, assessed as low due to the lack of group comparisons. However, this scale has been translated to a high standard and then psychometrically tested. The scale is rated positively across content validity, structural validity, internal consistency, reliability and measurement error. The QOE rating was reduced due to the availability of only one adequately rated study to compare and a low sample size. This scale shows promising evidence for use in an adult inpatient setting. Further research on construct validity, particularly with larger sample sizes, would support the strength of this scale.

The results indicate that no OMI validated in an adult inpatient psychiatric setting covers all measurement properties. The ReQoL shows promising scores across most measurement properties. However, there is an absence of cultural validity. Translational assessments were conducted, but there was not enough information to score using COSMIN standards. Measurement error was not reported in the evaluated study. This study was conducted with mixed populations, and only a small group of inpatients were included in different phases. Content validity was based on feedback from a CMHT sample, not an inpatient sample (Table 5), questioning its validity in an inpatient setting. However, it was important to include this study to highlight further areas for development for an inpatient psychiatric population.

Other outcome measures show some good evidence of measurement properties, but without adequate content validity conducted in an inpatient setting, it is difficult to know how meaningful these OMIs would be for adult psychiatric inpatients. In summary, none of the

OMIs reviewed cover all the required measurement properties set out by COSMIN (Mokkink et al., 2024).

Table 6

A Summary of the Quality of Evidence (QOE)

Measurement property		OMI																						
		1	2	3†††	4	5‡‡‡	6	7§§§	8	9	10	11	12 ****	13	14	15	16	17	18	19	20	21	22	23 ††††
Content Validity	Rating		+					+		+	+	+				+	+			+ ‡‡‡‡		+		
	QOE		VL					M		M	VL	VL				VL	L			VL		M		
Structural validity	Rating		+	?	+	-	+	+					+	+		-	+	+	?	+	+/-	+	+	-
	QOE		M	VL	H	VL	M	L					M	M		H	H	M	VL	M	VL	VL	M	VL
Internal consistency	rating		+/-	?	+	-/?	+	+		-		-	+/-	-/?	?	-	+	+	?	+	+	+	+	-
	QOE		M	M	H	M	M	M		M		M	L	L	H	H	M	M	M	M	M	M	M	L
Cross-cultural validity	Rating					?		?									?		?					
	QOE					VL		VL									VL		VL					
Reliability	rating		-			+/?	+	+		+**			+	+			+			+	+	-	+ §§§§	+
	QOE		VL			M	L	VL		VL			M	M			L			L	L	VL	VL	VL
Measurement Error	Rating												+	+			+				+			?
	QOE												M	M			L				VL			VL
Criterion validity	Rating																							
	QOE																							
Construct Validity	rating	+	+		+	+	+	+		+	+	+	+	+/-	+	+		?/+	+	+	+	+	+	+
	QOE	VL	M		M	M	M	M		M	M	VL	M	L	M	M		VL	M	M	M	L	L	VL
Responsiveness	Rating	-	+				+	+	+		+	-	+	+	+	+				+	?			
	QOE	VL	M				VL	L	VL		L	VL	L	VL	L	L				M	VL			

†††††one study has 10 items removed and can't be pooled

‡‡‡‡‡ one study only contained 27 items and so was not included in the summary of results

§§§ SQOL41 and SQOL41 Chinese version

**** combined score includes multiple language versions, one study excluded from summary as it only includes 14 items

†††† one reference a different language version, but same items. All items remain the same, but the number of domains varies across studies

‡‡‡‡ studies that use outpatients for these domains

§§§§ studies that use outpatients for these domains

Reliable Change Index (RCI)

RCIs have been calculated for each study that reported the means and standard deviations (SD) for a patient population, along with Cronbach's Alpha, and are presented in Table 7. Some studies did not report the total QoL score and instead reported domain scores; in these cases, the RCI for each domain was calculated.

Table 7

Reliable Change Index Calculations

Instrument name	Citation	Sample N	Domain	Mean (SD)	Cronbach's Alpha	RCI (CI)
OMI 3 Q-LES-Q	Ritsner et al., 2002	199 (inpatient)	Health	45.4 (10.8)	0.91	9.98 (8.32-9.98)
			Work	30.8 (10.9)	0.9	9.55 (8.77-10.52)
			Leisure time activities	21.1 (6.9)	0.84	7.65 (6.65-8.87)
			Social relationships	37.6 (10.4)	0.86	10.79 (9.33-12.45)
			Life satisfaction	3.5 (1.1)	-	-
			Perceived QOL index	3.5 (0.8)	0.94	0.54 (0.68-0.68)
			Subjective feelings	52.1 (13.2)	0.93	9.68 (8.8-11)
			Household duties	30.1 (14.5)	0.9	12.71 (11.39-14.02)
			General activities	47.6 (11.5)	0.95	7.13 (6.2-8.06)
			Satisfaction with medication	3.7 (1.1)	-	-
OMI 5 LQoLP	Gaite et al., 2000	404 (in and outpatients)	Life Satisfaction Score average score	4.67 (0.76)	0.87	0.76 (1 to 1)
			Global well being	4.37 (1.34)	0.83	1.34 (1-1)
OMI 6 SQLS	Boyer et al., 2010	517 (in and out patients)	Total score	56.4 (18.8)	0.88	18.05 (17.28-19.2)
OMI 7 SQOL-41	Auquier et al., 2003	207 (in and out)	Index score	55 (18.3)	0.94	12.42 (11.54-13.58)
OMI 9 QLIS	Franz et al., 2012	135 (in and outpatient)	Social contacts	4.47 (2.24)	0.63	3.78 (3.37 -5.06)
			Appreciation by others	6.37 (2.13)	0.70	3.23 (3.04-3.04)
			Relationship to family	6.15 (2.96)	0.71	4.42 (4.48-4.48)
			Appraisal of pharmacotherapy	5.97 (2.11)	0.71	3.15(2.99-2.99)
			Appraisal of psychopathological symptoms	5.97 (2.11)	0.82	2.82 (2.35-3.53)
			Cognitive functioning	5.49 (2.31)	0.79	2.93 (2.54-3.81
			Abilities to manage daily living	6.42 (2.43)	0.74	3.43 (2.93-2.24)
			Appraisal of accommodation/housing	6.14 (2.17)	0.73	3.13 (2.88-2.88)
			Financial situation	4.72 (2.65)	0.77	3.52 (2.66-3.99)
			Leading a normal life	4.56 (2.6)	0.70	3.95 (3.04-4.55)
			Confidence	5.74 (2.28)	0.74	3.22 (2.83-4.24)
			Global life satisfaction	4.64 (2.75)	0.88	2.64 (1.92-2.88)

OMI 17 SF-12	Soysal Gundaz et al., 2021	55 (inpatient)	Psychosocial	41.9 (20.4)	0.93	14.96 (12.47-18.33)
			Motivation/energy	48.1 (18.2)	0.73	26.21 (21.6-31.69)
			Symptoms/side effects	32.5 (17.6)	0.80	21.82 (18.59-27.27)
OMI 18 SQLS	Kaneda et al., 2002	55 (inpatient)	Psychosocial	41.9 (20.4)	0.93	14.96 (12.47-18.33)
			Motivation/energy	48.1 (18.2)	0.73	26.21 (21.6-31.69)
			Symptoms and side-effects	32.5 (17.6)	0.80	21.82 (18.59-27.27)
OMI 19 ReQoL	Keetharuth et al., 2018	4037 (mixed pop)	ReQoL 10	21.99 (10.26)	0.92	8.31 (7.84-8.62)
			ReQoL 20	21.63 (9.97)	0.96	5.53 (5.54-5.54)
OMI 20 SQLS-R4	Kuo et al. 2007	83 (in and outpatients)	Psychosocial mean change	(12.92)	0.92	10.13 (8.62-11.76)
			Vitality mean change	(12.37)	0.84	13.71 (12.2-16.63)
	Martin & Allen, 2007	100 (mixed populations)	Psychosocial feelings	38.27 (23.29)	0.96	12.91 (11.09-14.97)
			Cognition and vitality	40.18 (18.17)	0.82	21.37 (18.82-24.7)
			Total score	38.91 (20.99)	0.96	31.37 (18.82-24.7)
OMI 21 QOL-BD	Michalak et al., 2010	225 (In and outpatients)	Physical	11.77 (4.01)	0.79	5.09 (5.08-5.08)
			Sleep	11.42 (4.12)	0.83	4.71 (4.57-5.71)
			Mood	12.98 (3.98)	0.90	3.49 (3,51-3.51)
			Cognition	12.78 (4.22)	0.91	3.51 (3.33-4.16)
			Leisure	13.11 (4.16)	0.91	3.46 (3.33-4.16)
			Social	14.31 (4.00)	0.88	3.84 (3.84-3.84)
			Spirituality	13.05 (4.07)	0.93	2.98 (2.93-2.93)
			Finances	12.67 (4.64)	0.88	4.46 (3.84-4.8)
			Household	12.96 (4.11)	0.91	3.42 (3.33-4.16)
			Self-esteem	13.82 (3.70)	0.88	3.55 (2.88-3.84)
			Independence	15.78 (3.23)	0.81	3.9 (3.62-4.83)
			Identity	13.68 (4.14)	0.90	3.63 (3.51-4.38)
			Work	15.20 (3.52)	0.89	4.14 (3.53-4.7)
			Education	13.92 (4.96)	0.95	3.07 (3.1-3.1)
OMI 22 Q-LES-Q-18	Ritsner et al., 2005	339 (in and outpatients)	General QOL index	3.4 (0.8)	0.96	0.44 (0.55-0.55)
OMI 23 Henrich's Quality of Life Scale	Simon-Abbadi et al., 1999	60 (inpatients)	Mean global score HQLS	41.3 (16.6)	0.9	14.55 (12.27-17.53)

Discussion

Previous reviews have assessed the characteristics and dimensions of available QoL OMIs used in mental health settings (Van Krugten et al., 2021). This review evaluated the measurement properties of QoL OMIs used in inpatient psychiatric settings. Twenty-three outcome measures were identified. The results revealed that none of the outcome measures assessed covered all nine properties described by COSMIN.

Several studies reported construct validity but did not predefine a hypothesis for testing, and directional correlations were implied from the literature. Many studies employed methods unsuitable for measurement properties analyses and reported only mean scores, failing to calculate correlations or effect sizes. Some studies followed an appropriate translation process and validated outcome measures using a different language from the original development. However, most studies failed to test this through group comparisons or qualitative methods. Several studies did not use an adequate sample size for factor analysis to establish sufficient structural validity. When reliability was assessed, studies did not adequately report test conditions. Similarly, interventions used during sensitivity to change analysis were often poorly described, leaving the reader to assume that the conditions were appropriate; in some cases, the effect size was not reported. Numerous studies did not report measurement error, neglecting to highlight any potential biases that could have been overlooked. Many studies reported “very-good” and sufficient internal consistency, which demonstrates how well the items on the scale measure the same construct. This suggests that many scales may be accurately measuring what they were intended to measure.

The findings align with previous reviews, including a review of the SQLS and SQoLS-R4 (Zuniga Le-Bert et al., 2024), which highlighted similar results to the current review regarding structural validity. As noted in the current review, the low sample size and different

results reported across studies impact the overall GRADE results. Another systematic review, focusing on the EQ-5D and SF-36 in patients with schizophrenia (Papaioannou et al., 2011), reported similar findings to the current review, including convergent validity (Anquier et al., 2003) and responsiveness (Van de Villige et al., 2005). Furthermore, Azaiez et al. (2018) conducted a systematic review of HRQOL measures for patients with schizophrenia, reporting similar findings for the S-QOL-18 (Boyer et al., 2010), QLIS (Franz et al., 2012, 2013), and Q-LES-Q 18 (Ritsner et al., 2005). The consistency in reporting across these reviews underscores the reliability of the findings presented in the current study.

Among the outcome measures evaluated in this review, only eight reported content validity, and only three reported moderate QOE. This indicates that, among the studies that reported content validity, the majority did not provide sufficient information for analysis. COSMIN (Mokkink et al., 2024) recommends using qualitative methods to evaluate outcome measures at various stages, including development and pilot testing. Furthermore, both patients and professionals should assess comprehensiveness, relevance, and comprehensibility of OMIs. A scale should demonstrate content validity to ensure it is grounded in appropriate theory and evidence base, along with outcomes that are meaningful to its users.

Among the strongest measures reported, the MHQOL assessed six of nine measurement properties and was tested in an inpatient sample (Ebadi & Rezaiye, 2025). The original scale was developed by Van Krugten et al. (2022) on a subsample of mental health care users, not specifically inpatients, suggesting it could be helpful for a wide range of mental health care contexts and diagnoses. Further research should be conducted with a larger sample of inpatients to evaluate discriminant validity between different populations, such as inpatient and outpatient populations, and to assess convergent validity with other widely used QoL measures. Furthermore, responsiveness should be tested to assess sensitivity to change.

The SQOL-41 demonstrated strong performance across six of nine measurement properties, including moderate QOE for content validity (Auquier et al., 2003; Chou et al., 2010). This measure was evaluated in both inpatient and outpatient settings. However, it was developed specifically for a schizophrenic population and may not be generalisable to all adult inpatients.

The ReQoL evaluated six of nine measurement properties, with moderate QOE across multiple measurement properties. Keetharuth et al. (2018) developed the ReQoL as a measure for use across various contexts and mental health populations. However, the lack of content validity in an inpatient sample raises concerns about the measure's use within this setting and the relevance of the items to adult psychiatric inpatients. Further research could establish content validity with an inpatient sample.

Strengths and Limitations

This review was conducted following the COSMIN guidelines. The interpretation of these guidelines often relies on the reviewer’s expertise and judgment, making it complex and challenging to follow (Mckenna & Heaney, 2021). However, these standards are detailed and support the appraisal of outcome measures (Kwok et al., 2021).

Due to the complexity of the guidelines, Mokkink et al. (2024) recommends that an expert in the area conduct the review and that two reviewers review each measurement property. Furthermore, Cochrane review guidelines (Higgins et al., 2024) recommend that each stage of the review (screening, extraction, and quality assessment) be conducted by two reviewers. Due to time and resources, only a portion of the full text screening and quality assessment was conducted by a second reviewer. Furthermore, the first reviewer was a trainee psychologist working under the supervision of expertly qualified reviewers. The first reviewer assessed each measurement property, and consensus discussions with the review team were held whenever dilemmas arose to minimise the overall risk of bias. Transparency

of reporting was maintained, further minimising the risk of bias, and the review was pre-registered.

Forward citation searching was not noted in the pre-registered protocol. Although not originally planned, this step enhanced the comprehensiveness of the review and enabled the inclusion of literature that reflects the current direction of the evidence base.

Minor updates were made to the inclusion and exclusion criteria before conducting the formal search strategy. This involved clarifying that only studies reporting and evaluating measurement properties, such as validation studies, should be included. Future studies could examine interventional studies and extract data on the responsiveness of scales. This was unnecessary at this stage of the review process, as the focus was on understanding a breadth of measurement properties, not solely responsiveness.

One of the objectives in the pre-registered protocol was to compare the results to the O'Connell et al. (2014) framework. This was later deemed unnecessary to meet the aims of the review, as it is not specific to an inpatient setting.

Some studies included in the review contained mixed samples, which have the potential to introduce bias; this was considered when rating the QOE. A subgroup analysis (Ding et al., 2022) could be used to understand the results for an inpatient population further. However, it was deemed unnecessary to meet the aims of the current review. Including a mixed population allowed for a breadth and depth of literature coverage, including aspects of construct validity that might not have been covered otherwise. This approach enabled the presentation of knowledge gaps across a range of OMIs that have the potential to be beneficial in an inpatient setting.

COSMIN recommends that if there is insufficient or no content validity, reviewing the other measurement properties becomes redundant. This is because it may not measure the intended construct for the intended population. However, as part of this review, all

measurement properties were evaluated to gain a deeper understanding of the current evidence base. Interpretations can be made regarding OMIs whilst acknowledging content validity as a limitation. It is hoped that, by highlighting the strengths and weaknesses of each scale, future research can build on these areas.

The SF-36 reported three Cronbach's Alpha scores across different studies. However, due to the inconsistent ratings and lack of content validity, these results were not pooled together for meta-analysis. Although the LQoLP reports multiple internal consistency results across studies, only two studies reported the same version of the LQoLP with the same number of items. Therefore, results were not pooled for meta-analysis.

Many studies included in the review did not report study design, leaving the reader to establish this for themselves.

In the search strategy, the Health of the Nation Outcome (HONOS) (Wing et al., 1996) was included, as it was initially considered a QoL outcome measure. However, upon further reading, it was established that it is a measure of mental health and social functioning. For this reason, studies which reported the measurement properties of the HONOS were excluded. This is a potential limitation of the search strategy, which yields many redundant results; future replications might consider excluding HONOS from the search strategy.

Implications and future research

This review highlights the range of available OMIs with some measurement properties tested in an inpatient population. However, there is a need for further research to explore content validity in inpatient samples, particularly for the ReQoL, a general mental health quality of life measure with strong reliability, commonly used in practice within inpatient settings of the National Health Service (NHS). A specific inpatient version of the measure could be developed based on content validity and subsequently retested across various indices. Furthermore, content validity studies for QoL OMIs should be carried out in

inpatient settings. Additional research should be conducted using the MHQOL to gain insights into construct validity and responsiveness for inpatient populations.

RCI scores were calculated for several outcome measures and are presented in the results section. This has implications for understanding change over time using the identified outcome measures. RCIs determine whether observed changes in scores are genuine or due to measurement error; this can help guide treatment decisions and inform the use of outcomes in research to understand treatment effectiveness (Jacobson & Truax, 1991).

Conclusion

Given that no outcome measures were identified that encompassed all measurement properties, clinicians should approach the identified outcome measures with caution. Most measures demonstrated good internal consistency and construct validity, indicating some reliability and validity in this setting. Until a measure has been thoroughly tested in an inpatient setting, it can be inferred that there is currently no recommended measure for clinical practice.

References

Addington, D., Addington, J., & Maticka-tyndale, E. (1993). Assessing depression in schizophrenia: The Calgary Depression Scale. *British Journal of Psychiatry*, *163*(S22), 39–44. https://doi.org/10.1192/s0007125000292581

Amir-Behghadami, M., & Janati, A. (2020). Population, intervention, comparison, outcomes and study (picos) design as a framework to formulate eligibility criteria in systematic reviews. *Emergency Medicine Journal*, *37*(6), 387–387. https://doi.org/10.1136/emermed-2020-209567

Anderson, A. L., McNeil, D. C., & Reddon, J. R. (2002). Evaluation of Lehman's brief quality of life interview in assessing outcome in psychiatric rehabilitation in people with severe and persistent mental disorder. *Social Work in Mental Health*, *1*(2), 43–59. https://doi.org/10.1300/j200v01n02_04

Angstman, S., Schuldberg, D., Harris, K. J., Cochran, B., & Peterson, P. (2009). Use of the quality of life inventory for measuring quality of life changes in an inpatient psychiatric population. *Psychological Reports*, *104*(3), 1007–1014. https://doi.org/10.2466/pr0.104.3.1007-1014

Auquier, P., Simeoni, M. C., Sapin, C., Reine, G., Aghababian, V., Cramer, J., & Lançon, C. (2003). Development and validation of a patient-based health-related quality of life questionnaire in schizophrenia: The S-Qol. *Schizophrenia Research*, *63*(1–2), 137–149. https://doi.org/10.1016/s0920-9964(02)00355-9

Azaiez, C., Millier, A., Lançon, C., Clay, E., Auquier, P., Llorca, P. M., & Toumi, M. (2018). Health related quality of life in patients having schizophrenia negative symptoms - a

systematic review. *Journal of market access & health policy*, *6*(1), 1517573. https://doi.org/10.1080/20016689.2018.1517573

Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., Mellor-Clark, J., Richards, D., Unsworth, G., & Evans, C. (2013). The CORE-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling & Psychotherapy Research, 13*(1), 3–13. https://doi.org/10.1080/14733145.2012.729069

Bigelow, D. A., McFarland, B. H., & Olson, M. M. (1991). Quality of life of community mental health program clients: Validating a measure. *Community Mental Health Journal, 27*(1), 43–55. https://doi.org/10.1007/BF00752714

Bishop, S. L., Walling, D. P., Dott, S. G., Folkes, C. C., & Bucy, J. (1999). Refining quality of life: Validating a multidimensional factor measure in the severe mentally ill. *Quality of Life Research*, *8*(1/2), 151–160. https://doi.org/10.1023/a:1026489331009

Blome, C., & Augustin, M. (2015). Measuring change in quality of life: bias in prospective and retrospective evaluation. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, *18*(1), 110–115. https://doi.org/10.1016/j.jval.2014.10.007

Boyer, L., Simeoni, M.-C., Loundou, A., D'Amato, T., Reine, G., Lancon, C., & Auquier, P. (2010). The development of the S-Qol 18: A shortened quality of life questionnaire for patients with schizophrenia. *Schizophrenia Research*, *121*(1–3), 241–250. https://doi.org/10.1016/j.schres.2010.05.019

Brook, R. H., Ware, J. E., Jr, Rogers, W. H., Keeler, E. B., Davies, A. R., Donald, C. A., Goldberg, G. A., Lohr, K. N., Masthay, P. C., & Newhouse, J. P. (1983). Does free care improve adults' health? Results from a randomized controlled trial. *The New England Journal of medicine, 309*(23), 1426–1434. https://doi.org/10.1056/NEJM198312083092305

Burks, H. B., des Bordes, J. K. A., Chadha, R., Holmes, H. M., & Rianon, N. J. (2021). Quality of life assessment in older adults with dementia: A systematic review. *Dementia and Geriatric Cognitive Disorders*, *50*(2), 103–110. https://doi.org/10.1159/000515317

Chan, G. W., Ungvari, G. S., Shek, D. T., & Leung, J. J. (2003). Hospital and community-based care for patients with chronic schizophrenia in Hong Kong. *Social Psychiatry and Psychiatric Epidemiology*, *38*(4), 196–203. https://doi.org/10.1007/s00127-003-0616-5

Chou, C.-Y., Ma, M.-C., Yang, T.-T., & Chen-Sea, M.-J. (2010). Psychometric validation of the S-Qol Chinese (Taiwan) version for patients with schizophrenia. *Quality of Life Research*, *20*(5), 763–767. https://doi.org/10.1007/s11136-010-9809-z

De Jong, K., Conijn, J. M., Gallagher, R. A. V., Reshetnikova, A. S., Heij, M., & Lutz, M. C. (2021). Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: A multilevel meta-analysis. *Clinical Psychology Review*, *85*, 102002. https://doi.org/10.1016/j.cpr.2021.102002

Ding, Q., Wan, S., & Dowling, T. C. (2022). Research and scholarly methods: Subgroup analysis. *JACCP: Journal of the American College of Clinical Pharmacy, 5*(5), 551–558. https://doi.org/10.1002/jac5.1611

Ebadi, A., & Rezaiye, M. (2025). Translation and psychometric properties of the Persian version of the mental health quality of life (mhqol) questionnaire. *BMC Psychiatry*, *25*(1). https://doi.org/10.1186/s12888-025-06522-y

Eiser, C., & Jenney, M. (2007). Measuring quality of life. *Archives of Disease in Childhood*, *92*(4), 348–350. https://doi.org/10.1136/adc.2005.086405

Elsman, E. B. M., Mokkink, L. B., Terwee, C. B., Beaton, D., Gagnier, J. J., Tricco, A. C., Baba, A., Butcher, N. J., Smith, M., Hofstetter, C., Lee Aiyegbusi, O., Berardi, A., Farmer, J., Haywood, K. L., Krause, K. R., Markham, S., Mayo-Wilson, E., Mehdipour, A., Ricketts, J., Szatmari, P., … Offringa, M. (2024). Guideline for reporting systematic reviews of outcome measurement instruments (OMIs): PRISMA-COSMIN for OMIs 2024. *Journal of Clinical Epidemiology*, *173*, 111422. https://doi.org/10.1016/j.jclinepi.2024.111422

EndNote Team (2013). EndNote (version 2025). Philadelphia, PA: Clarivate.

Fantino, B., & Moore, N. (2009). The self-reported Montgomery-Åsberg Depression Rating Scale is a useful evaluative tool in major depressive disorder. *BMC Psychiatry, 9,* Article 26. https://doi.org/10.1186/1471-244X-9-26

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental State Examination. *PsycTESTS Dataset*. https://doi.org/10.1037/t07757-000

Franz, M., Fritz, M., & Meyer, T. (2012). Discriminant and convergent validity of a subjective quality-of-life instrument aimed at high content validity for schizophrenic persons. *Quality of Life Research*, *22*(5), 1113–1122. https://doi.org/10.1007/s11136-012-0221-8

Franz, M., Fritz, M., Gallhofer, B., & Meyer, T. (2012). QLiS – development of a schizophrenia-specific quality-of-life scale. *Health and Quality of Life Outcomes*, *10*(1). https://doi.org/10.1186/1477-7525-10-61

Gaite, L., Vázquez-Barquero, J. L., Arrizabalaga, A. A., Schene, A. H., Welcher, B., Thornicroft, G., Ruggeri, M., Vázquez-Bourgon, E., Retuerto, M. P., & Leese, M. (2000). Quality of life in schizophrenia: Development, reliability and internal consistency of the Lancashire Quality of Life Profile – European version. *British Journal of Psychiatry*, *177*(S39). https://doi.org/10.1192/bjp.177.39.s49

Ginty, A. T. (2020). Psychometric Properties. *Encyclopedia of Behavioral Medicine*, 1770–1771. https://doi.org/10.1007/978-3-030-39903-0_480

Goodwin, A. M., & Madell, D. L. (2002). Measuring the quality of life of people with severe and enduring mental health problems in different settings. *Journal of Mental Health*, *11*(3), 305–317. https://doi.org/10.1080/09638230020023688

Guy, W. (1976). Clinical global impressions scale. *PsycTESTS Dataset*. https://doi.org/10.1037/t48216-000

Guyatt, G. H., Oxman, A. D., Schünemann, H. J., Tugwell, P., & Knottnerus, A. (2011). GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *Journal of clinical epidemiology*, *64*(4), 380–382. https://doi.org/10.1016/j.jclinepi.2010.09.011

Haddock, G., McCarron, J., Tarrier, N., & Faraghier, E. B. (1999). Scales to measure dimensions of hallucinations and delusions: The psychotic symptom rating

scales (PSYRATS). *Psychological Medicine*, *29*(4), 879–889. https://doi.org/10.1017/s0033291799008661

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry*, *23*(1), 56–62. https://doi.org/10.1136/jnnp.23.1.56

Hansson, L., Svensson, B., & Björkman, T. (1998). Quality of life of the mentally ill. reliability of the Swedish version of the Lancashire Quality of Life Profile. *European Psychiatry*, *13*(5), 231–234. https://doi.org/10.1016/s0924-9338(98)80010-2

Happell, B. (2008). The value of routine outcome measurement for consumers of mental health services: master or servant?. *The International Journal of Social Psychiatry, 54*(4), 317–327. https://doi.org/10.1177/0020764008090285

Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Page, M., & Welch, V. (2024, August). *Cochrane Handbook for Systematic Reviews of Interventions*. Cochrane Handbook for Systematic Reviews of Interventions | Cochrane Training. https://training.cochrane.org/handbook

Hunt, S. (1991). Quality of life and technology assessment. monograph of the Council on Health Care Technology of the Institute of Medicine. *Health Policy*, *17*(1), 90–91. https://doi.org/10.1016/0168-8510(91)90122-e

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19. https://doi.org/10.1037//0022-006x.59.1.12

Kagawa-Singer, M., Padilla, G. V., & Ashing-Giwa, K. (2010). Health-related quality of life and culture. *Seminars in Oncology Nursing*, *26*(1), 59–67. https://doi.org/10.1016/j.soncn.2009.11.008

Kaiser, W., Priebe, S., Barr, W., Hoffmann, K., Isermann, M., Röder-Wanner, U.-U., & Huxley, P. (1997). Profiles of subjective quality of life in schizophrenic in- and out-patient samples. *Psychiatry Research*, *66*(2–3), 153–166. https://doi.org/10.1016/s0165-1781(96)02958-7

Kaneda, Y., Imakura, A., Fujii, A., & Ohmori, T. (2002). Schizophrenia quality of Life Scale--Japanese version. *PsycTESTS Dataset*. https://doi.org/10.1037/t39265-000

Katsavdakis, K. A., Clifford, P. I., Evans, R. B., Graham, P., Allen, J. G., Sargent, J., Lyle, J., & Frager, D. C. (1999). The How Are You? scale: a quality-of-life outcomes measure for routine practice. *Bulletin of the Menninger Clinic*, *63*(3), 366–387.

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261–276. https://doi.org/10.1093/schbul/13.2.261

Keetharuth, A. D., Brazier, J., Connell, J., Bjorner, J. B., Carlton, J., Taylor Buck, E., Ricketts, T., McKendrick, K., Browne, J., Croudace, T., & Barkham, M. (2018). Recovering quality of life (reqol): A new generic self-reported outcome measure for use with people experiencing mental health difficulties. *The British Journal of Psychiatry*, *212*(1), 42–49. https://doi.org/10.1192/bjp.2017.10

Kuo, P.-J., Chen-Sea, M.-J., Lu, R.-B., Chung, M.-S., Kuo, C.-C., Huang, W.-C., & Ma, H.-I. (2007). Validation of the Chinese version of the schizophrenia quality of Life Scale

Revision 4 (SQLS-R4) in Taiwanese patients with schizophrenia. *Quality of Life Research*, *16*(9), 1533–1538. https://doi.org/10.1007/s11136-007-9262-9

Kuo, P. J., Ma, H. I., Kuo, C. C., Huang, W. C., & Chung, M. S. (2009). Factor analysis of the schizophrenia quality of Life Scale Revision 4 (SQLS-R4) Chinese version and related factors. *International Journal of Psychiatry in Clinical Practice*, 1–7. https://doi.org/10.1080/13651500902919376

Kwok, E. Y., Rosenbaum, P., Thomas-Stonell, N., & Cunningham, B. J. (2021). Strengths and challenges of the Cosmin tools in outcome measures appraisal: A case example for speech–language therapy. *International Journal of Language & Communication Disorders*, *56*(2), 313–329. https://doi.org/10.1111/1460-6984.12603

Lehman, A. F. (1988). A quality of life interview for the chronically mentally ill. *Evaluation and Program Planning*, *11*(1), 51–62. https://doi.org/10.1016/0149-7189(88)90033-x

Lovibond, S. H., & Lovibond, P. F. (1995). Depression anxiety stress scales. *PsycTESTS Dataset*. https://doi.org/10.1037/t01004-000

Martin, C. R., & Allan, R. (2007). Factor structure of the schizophrenia quality of Life Scale Revision 4 (SQLS-R4). *Psychology, Health & Medicine*, *12*(2), 126–134. https://doi.org/10.1080/13548500500407383

McKenna, S. P., & Heaney, A. (2021). Setting and maintaining standards for patient-reported outcome measures: Can we rely on the cosmin checklists? *Journal of Medical Economics*, *24*(1), 502–511. https://doi.org/10.1080/13696998.2021.1907092

Mezzich, J. E., Cohen, N. L., Ruiperez, M. A., Banzato, C. E., & Zapata-Vega, M. I. (2011). The Multicultural Quality of Life Index: presentation and validation. *Journal of*

Evaluation in Clinical Practice, *17*(2), 357–364. https://doi.org/10.1111/j.1365-2753.2010.01609.x

Michalak, E. E., & Murray, G. (2010). Development of the qol.bd: A disorder-specific scale to assess quality of life in bipolar disorder. *Bipolar Disorders*, *12*(7), 727–740. https://doi.org/10.1111/j.1399-5618.2010.00865.x

Mokkink, L. B., Elsman, E. B. M., & Terwee, C. B. (2024). COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *33*(11), 2929–2939. https://doi.org/10.1007/s11136-024-03761-6

Murray, G. C., McKenzie, K., Kidd, G. R., Psychologist, S. L., & Sinclair, B. (1998). The Five accomplishments: A framework for obtaining customer feedback in a Health Service Community Learning Disability Team. *British Journal of Learning Disabilities*, *26*(3), 94–99. https://doi.org/10.1111/j.1468-3156.1998.tb00058.x

Nair, V., Tiwari, S., Wee, M., Leong, S. F., Sarjit, K., Thavamani, & Liew, C. S. (1996). Quality of life survey among long-stay mentally ill patients: patient and staff perspectives. *Singapore Medical Journal*, *37*(5), 512–516.

Newnham, E. A., Harwood, K. E., & Page, A. C. (2007). Evaluating the clinical significance of responses by psychiatric inpatients to the mental health subscales of the SF-36. *Journal of Affective Disorders*, *98*(1–2), 91–97. https://doi.org/10.1016/j.jad.2006.07.001

Nishiyama, T., & Ozaki, N. (2009). Measurement limit of quality-of-life questionnaires in psychiatric settings. *Quality of Life Research*, *19*(1), 25–30. https://doi.org/10.1007/s11136-009-9556-1

Nørholm, V., & Bech, P. (2006). Quality of life in schizophrenic patients: Association with depressive symptoms. *Nordic Journal of Psychiatry*, *60*(1), 32–37. https://doi.org/10.1080/08039480500504966

O'Connell, J., O'Cathain, A., & Brazier, J. (2014). Measuring quality of life in mental health: are we asking the right questions? *Social Science & Medicine (1982)*, *120*, 12–20. https://doi.org/10.1016/j.socscimed.2014.08.026

Oliveira, S. E. H., Carvalho, H., & Esteves, F. (2016). Toward an understanding of the quality of life construct: Validity and reliability of the WHOQOL-Bref in a psychiatric sample. *Psychiatry Research*, *244*, 37–44. https://doi.org/10.1016/j.psychres.2016.07.007

Overall, J. E., & Gorham, D. R. (1988). The Brief Psychiatric Rating Scale (BPRS): Recent developments in ascertainment and scaling. *Psychopharmacology Bulletin, 24*(1), 97–99.

Papaioannou, D., Brazier, J., & Parry, G. (2011). How valid and responsive are generic health status measures, such as EQ-5D and SF-36, in schizophrenia? A systematic review. *Value in Health*, *14*(6), 907–920. https://doi.org/10.1016/j.jval.2011.04.006

Peljto, A., Zamurovic, L., Milovancevic, M. P., Aleksic, B., Tosevski, D. L., & Inada, T. (2017). Drug-induced Extrapyramidal Symptoms Scale (DIEPSS) Serbian Language

version: Inter-rater and Test-retest Reliability. *Scientific Reports*, *7*(1), 8105. https://doi.org/10.1038/s41598-017-08706-3

Pitkänen, A., Välimäki, M., Endicott, J., Katajisto, J., Luukkaala, T., Koivunen, M., Kuosmanen, L., & Hätönen, H. (2011). Assessing quality of life in patients with schizophrenia in an acute psychiatric setting: Reliability, validity and feasibility of the EQ-5D and the Q-les-Q. *Nordic Journal of Psychiatry*, *66*(1), 19–25. https://doi.org/10.3109/08039488.2011.593099

Pukrop, R., Schlaak, V., Möller-Leimkühler, A. M., Albus, M., Czernik, A., Klosterkötter, J., & Möller, H.-J. (2003). Reliability and validity of quality of life assessed by the short-form 36 and the modular system for quality of life in patients with schizophrenia and patients with depression. *Psychiatry Research*, *119*(1–2), 63–79. https://doi.org/10.1016/s0165-1781(03)00110-0

Ritsner, M., Kurs, R., Gibel, A., Ratner, Y., & Endicott, J. (2005). Validity of an abbreviated quality of life enjoyment and satisfaction questionnaire (Q-les-Q-18) for schizophrenia, schizoaffective,and mood disorder patients. *Quality of Life Research*, *14*(7), 1693–1703. https://doi.org/10.1007/s11136-005-2816-9

Ritsner, M., Kurs, R., Kostizky, H., Ponizovsky, A., & Modai, I. (2002). Subjective quality of life in severely mentally ill patients: a comparison of two instruments. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 11*(6), 553–561. https://doi.org/10.1023/a:1016323009671

Rodgers, M., Sowden, A., Petticrew, M., Arai, L., Roberts, H., Britten, N., & Popay, J. (2009). Testing methodological guidance on the conduct of narrative synthesis in

systematic reviews. *Evaluation*, *15*(1), 49–73. https://doi.org/10.1177/1356389008097871

Rosenberg, M. (1965). Rosenberg self-Esteem Scale. *PsycTESTS Dataset*. https://doi.org/10.1037/t01038-000

Russo, J., Roy-Byrne, P., Reeder, D., Alexander, M., Dwyer-O'Connor, E. D., Dagadakis, C., Ries, R., & Patrick, D. (1997). Longitudinal assessment of quality of life in acute psychiatric inpatients: Reliability and validity. *The Journal of Nervous & Amp;Amp Mental Disease*, *185*(3), 166–175. https://doi.org/10.1097/00005053-199703000-00006

Simon-Abbadi, S., Guelfi, J. D., & Ginestet, D. (1999). Psychometric qualities of the French version of the Heinrichs Quality of Life Rating Scale. *European Psychiatry*, *14*(7), 386–391. https://doi.org/10.1016/s0924-9338(99)00231-x

Soysal Gunduz, Mutlu, S., Aslan Basli, A., G, C., Akgul, Yilmaz, E., & Aydemir. (2021). Validation of the Turkish form of Short Form-12 Health Survey Version 2 (SF 12V2). *Archives of Rheumatology*, *36*(2), 280–286. https://doi.org/10.46497/archrheumatol.2021.8458

Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J., & Weich, S. (2009). Short Warwick–Edinburgh mental Well–being Scale. *PsycTESTS Dataset*. https://doi.org/10.1037/t80221-000

Strathdee, G. (2024). *Rapid Review into data on mental health inpatient settings: Final report and recommendations*. GOV.UK. https://www.gov.uk/government/publications/rapid-review-into-data-on-mental-health-inpatient-settings-final-report-and-

recommendations/rapid-review-into-data-on-mental-health-inpatient-settings-final-report-and-recommendations#:~:text=review%20the%20data%20that%20is,collection%2C%20particularly%20for%20frontline%20staff

Su, C.-T., Ng, H.-S., Yang, A.-L., & Lin, C.-Y. (2014). Psychometric Evaluation of the short form 36 health survey (SF-36) and the World Health Organization Quality of life scale brief version (WHOQOL-BREF) for patients with schizophrenia. *Psychological Assessment*, *26*(3), 980–989. https://doi.org/10.1037/a0036764

Su, C.-T., Yang, A.-L., & Lin, C.-Y. (2017). The construct of the schizophrenia quality of life scale revision 4 for the population of Taiwan. *Occupational Therapy International*, *2017*, 1–11. https://doi.org/10.1155/2017/5328101

Swan, K., Speyer, R., Scharitzer, M., Farneti, D., Brown, T., Woisard, V., & Cordier, R. (2023). Measuring what matters in healthcare: A practical guide to psychometric principles and instrument development. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1225850

Teoli, D., & Bhardwaj, A. (2023). Quality Of Life. In *StatPearls*. StatPearls Publishing.

Tracy, D. K., Gadelrab, R., Rahim, A., Pendlebury, G., Reza, H., Bhattacharya, R., Bachlani, A., Worlley, K., Rigby, D., Scott, J., & Dave, S. (2022). Digital Literacy in contemporary mental healthcare: Electronic patient records, outcome measurements and social media. *BJPsych Advances*, *30*(1), 36–43. https://doi.org/10.1192/bja.2022.74

Tunis, S. L., Croghan, T. W., Heilman, D. K., Johnstone, B. M., & Obenchain, R. L. (1999). Reliability, validity, and application of the Medical Outcomes Study 36-item short-

form health survey (SF-36) in schizophrenic patients treated with olanzapine versus haloperidol. *Medical Care*, *37*(7), 678–691. https://doi.org/10.1097/00005650-199907000-00008

Van de Willige, G., Wiersma, D., Nienhuis, F. J., & Jenner, J. A. (2005). Changes in quality of life in chronic psychiatric patients: A comparison between EuroQol (EQ-5D) and WHOQOL. *Quality of Life Research*, *14*(2), 441–451. https://doi.org/10.1007/s11136-004-0689-y

Van Krugten, F. C. W., Feskens, K., Busschbach, J. J. V., Hakkaart-van Roijen, L., & Brouwer, W. B. F. (2021). Instruments to assess quality of life in people with mental health problems: a systematic review and dimension analysis of generic, domain- and disease-specific instruments. *Health and Quality of Life Outcomes*, *19*(1), 249. https://doi.org/10.1186/s12955-021-01883-w

Van Krugten, F. C. W., Busschbach, J. J. V., Versteegh, M. M., Hakkaart-van Roijen, L., & Brouwer, W. B. F. (2022). The Mental Health Quality of Life Questionnaire (MHQoL): development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation. 31*(2), 633–643. https://doi.org/10.1007/s11136-021-02935-w

WHOQOL Group (1994). Development of the WHOQOL: Rationale and Current Status. *International Journal of Mental Health*, *23*(3), 24–56. https://doi.org/10.1080/00207411.1994.11449286

Wing, J. K., Curtis, R. H., & Beevor, A. S. (1996). Health of the Nation Outcome Scales. *PsycTESTS Dataset*. https://doi.org/10.1037/t28370-000

Young, R. C., Biggs, J. T., Ziegler, V. E., & Meyer, D. A. (1978). A rating scale for mania: reliability, validity and sensitivity. *The British Journal of Psychiatry: The Journal of Mental Science*, *133*, 429–435. https://doi.org/10.1192/bjp.133.5.429

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Zúñiga Le-Bert, M., Wiessner, M., Wehr, S., Weigel, L., & Leucht, S. (2024). Schizophrenia quality of life scale and schizophrenia quality of Life Scale Revision 4: A systematic review of measurement properties. *Schizophrenia Bulletin*. https://doi.org/10.1093/schbul/sbae119

Appendix

Appendix A

Cosmin Definitions of Domains and Measurement Properties

REDACTED

Appendix B

Completed COSMIN Systematic Review Checklist and Page Numbers

REDACTED

Appendix C

Full Search Strategy

Number	Search Term
1	"psy* hospital*"
2	"psy* unit*
3	"psy* inpatient*"
4	"psy* ward*"
6	"psy* facility*"
6	"psy* institution*"
7	"mental* ward*"
8	mental* hospital*"
9	"mental* unit*"
10	"mental* inpatient*"
11	"mental* facility*"
12	"mental* institution*"
13	"psychiatric care"
14	"psychiatric services"
15	"psychiatric treatment"
16	"psychiatric rehabilitation"
17	"mental health rehabilitation"
18	"psychiatric admission*"
19	"psychiatric hospitalization"
20	"psychiatric readmission*"
21	"psychiatric discharge*"
22	"Tier 4 mental health"
23	"Tier 4 service*"
24	"highly specialized mental health"
25	"specialist mental health inpatient*"

26	"tertiary mental health care"
27	"tertiary psychiatric care"
28	"specialist psychiatric unit*"
29	"high dependency mental health unit*"
30	"long-stay psychiatric hospital*"
31	"chronic psychiatric inpatient care"
32	"specialist inpatient mental health service*"
33	"hospital-based mental health care"
34	Psychometric*
35	“factor analysis”
36	Valid*
37	Reliab*
38	“responsiveness”
39	“internal consistency”
40	Measure*
41	instrument
42	scale
43	outcome
44	Index
45	PROM
46	“quality of life”
47	SQOL
48	QOL
49	ReQol
50	“EuroQol-5
51	q-les-q
52	EQ-5D
53	EuroQol-5 Dimension

54	MHQoL
55	SQLS
56	BQOL
57	QLDS
58	SF-36
59	Short Form Health Survey-36
60	HoNOS
61	SF-12
62	SF-6D
63	WHOQOL
64	EUROHIS-QOL
65	AQoL
66	“15 dimensional”
67	QLI
68	QOLS
69	CDC HRQOL-4
70	PWI-A
71	“Personal well-being index-adult”
72	QOL5

Scopus Search

(TITLE-ABS-KEY (psychometric* OR "factor analysis" OR valid* OR reliab* OR responsiveness OR "internal consistency" OR measure* OR outcome* OR instrument OR scale OR index OR prom)
AND
TITLE-ABS-KEY ("quality of life" OR qol OR sqol OR reqol OR "EuroQOL-5" OR q-les-q OR eq-5d OR mhqol OR sqls OR bqol OR qlds OR sf-36 OR "short form health survey- 36" OR honos OR sf-12 OR sf-6d OR wholqol OR eurohis-qol OR aqol OR "15 dimensional" OR qli OR qols OR "CDC HRQOL-4" OR pwi-a OR "personal well-being index-adult" OR qol5)
AND
TITLE-ABS-KEY ("psy* hospital*" OR "psy* unit*" OR "psy* inpatient*" OR "psy* ward*" OR "psy* facility*" OR "psy* institution*" OR "mental* ward*" OR "mental* hospital*" OR "mental* unit*" OR "mental* inpatient*" OR "mental* facility*" OR "mental* institution*" OR "psychiatric care" OR "psychiatric services" OR "psychiatric treatment" OR "psychiatric rehabilitation" OR "mental health rehabilitation" OR "psychiatric admission*" OR "psychiatric hospitalization" OR "psychiatric readmission*" OR "psychiatric discharge*" OR "Tier 4 mental health" OR "Tier 4 service*" OR "highly specialized mental health" OR "specialist mental health inpatient*" OR "tertiary mental health care" OR "tertiary psychiatric care" OR "specialist psychiatric unit*" OR "high dependency mental health unit*" OR "long-stay psychiatric hospital*" OR "chronic psychiatric inpatient care" OR "specialist inpatient mental health service*" OR "hospital-based mental health care"))
AND
(LIMIT-TO (LANGUAGE , "English"))

PsycInfo Search

1. (psychometric* or "factor analysis" or valid* or reliab* or responsiveness or "internal consistency" or measure* or outcome* or instrument or scale or index or prom).ab,ti.

2. treatment outcomes/ or psychotherapeutic outcomes/ or patient reported outcome measures/patient reported outcome measures/ or "treatment process and outcome measures"/ or self-report/ or treatment outcomes/rating scales/ or likert scales/
3. foreign language translation/ or cross cultural test adaptation/ or cross cultural validity/
   exp Test Validity/ or exp Psychometrics/ or exp Test Reliability/

4. factor structure/ or factor analysis/

5. exp Construct Validity/ or exp Confirmatory Factor Analysis/

6. 1 or 2 or 3 or 4 or 5

7. ("quality of life" or qol or sqol or reqol or "EuroQOL" or q-les-q or eq-5d or mhqol or sqls or bqol or qlds or sf-36 or "short form health survey-36" or honos or "health of nation outcome scale" or sf-12 or sf-6d or whoqol or eurohis-qol or aqol or "15-dimensional" or qli or qols or "CDC HRQOL-4" or pwi-a or "personal well-being index-adult" or qol5).ab,mf,ti.

8. "quality of life"/ or "quality of life measures"/
9. 7 or 8

10. ("psy* hospital*" or "psy* unit*" or "psy* inpatient*" or "psy* ward*" or "psy* facility*" or "psy* institution*" or "mental* ward*" or "mental* hospital*" or "mental* unit*" or "mental* inpatient*" or "mental* facility*" or "mental* institution*" or "psychiatric care" or "psychiatric services" or "psychiatric treatment" or "psychiatric rehabilitation" or "mental health rehabilitation" or "psychiatric admission*" or "psychiatric hospitalization" or "psychiatric readmission*" or "psychiatric discharge*" or "Tier-4 mental health" or "Tier-4 service*" or "highly specialized mental health" or "specialist mental health inpatient*" or "tertiary mental health care" or "tertiary psychiatric care" or "specialist psychiatric unit*" or "high dependency mental health unit*" or "long-stay psychiatric hospital*" or "chronic psychiatric inpatient care" or "specialist inpatient mental health service*" or "hospital-based mental health care").ab,mf,ti.

11. psychiatric hospitalization/ or psychiatric hospital admission/ or psychiatric hospital discharge/ or psychiatric hospital readmission/ or mental health commitment/ psychiatric units/ or psychiatric hospitals/
12. 10 or 11
13. 6 and 9 and 12

Medline Search

1. psychometric* or "factor analysis" or valid* or reliab* or responsiveness or "internal consistency" or measure* or outcome* or instrument or scale or index or prom).ab,ti.

2. Thesaurus search: Psychometrics/

3. 1 or 2

4. "quality of life" or qol or sqol or reqol or "EuroQOL" or q-les-q or eq-5d or mhqol or sqls or bqol or qlds or sf-36 or "short form health survey-36" or honos or "health of nation outcome scale" or sf-12 or sf-6d or whoqol or eurohis-qol or aqol or "15-dimensional" or qli or qols or "CDC HRQOL-4" or pwi-a or "personal well-being index-adult" or qol5).ab,ti.

5. Thesaurus search: "quality of life"/ or psychological well-being/

6. 4 or 5

7. ("psy* hospital*" or "psy* unit*" or "psy* inpatient*" or "psy* ward*" or "psy* facility*" or "psy* institution*" or "mental* ward*" or "mental* hospital*" or "mental* unit*" or "mental* inpatient*" or "mental* facility*" or "mental* institution*" or "psychiatric care" or "psychiatric services" or "psychiatric treatment" or "psychiatric rehabilitation" or "mental health rehabilitation" or "psychiatric admission*" or "psychiatric hospitalization" or "psychiatric readmission*" or

"psychiatric discharge*" or "Tier-4 mental health" or "Tier-4 service*" or "highly specialized mental health" or "specialist mental health inpatient*" or "tertiary mental health care" or "tertiary psychiatric care" or "specialist psychiatric unit*" or "high dependency mental health unit*" or "long-stay psychiatric hospital*" or "chronic psychiatric inpatient care" or "specialist inpatient mental health service*" or "hospital-based mental health care").ab,ti.

8. Hospitals, Psychiatric/

9. 7 or 8

10. 3 and 6 and 9

Cinahl Search

"psychiatric inpatient*" OR "mental health inpatient*" OR "psychiatric unit*" OR "mental health facility" OR "psychiatric care" OR "mental health services" OR "specialist inpatient mental health" OR “psychiatric hospital” OR “psychiatric unit” OR “psychiatric inpatient” OR “psychiatric ward” OR “psychiatric facility” OR “psychiatric institution” OR “mental health ward” OR “mental health hospital” OR “mental health unit” OR “mental health inpatient” OR “mental health facility” OR “mental health institution” OR “psychiatric care” OR “psychiatric services” OR “psychiatric treatment” OR “psychiatric rehabilitation” OR “mental health rehabilitation” OR “psychiatric admission” OR “psychiatric hospitilization” OR “psychiatric readmission” OR “psychiatric discharge” OR “tier 4 mental health” OR “highly specialised mental health” OR “specialist mental health inpatient” OR “tertiary mental health care” OR “tertiary psychiatric care” OR “specialist psychiatric unit” OR “high dependency mental health unit” OR “long stay psychiatric hospital” OR “chronic psychiatric inpatient care” OR “specialist inpatient mental health service” OR “hospital based mental health care”

AND

"recovering quality of life" OR "reqol" OR "reqol-10" OR "reqol-20" OR qol OR sqol OR reqol OR eurqol OR q-les-q OR eq-5d OR mhqol OR sqls OR bqol OR qlds OR sf-36 OR “short form health survey-36” OR honos OR “health of nation outcome scale” OR sf-12 or sf-6d OR whoqol OR eurohis-qol OR aqol OR 15-dimensional OR qli OR qols OR CDC hrqol-4 OR pwi-a OR “personal wellbeing index adult” OR qol5 OR “health related quality of life”

AND

psychometrics OR validity OR reliability OR “psychometric properties” OR “factor analysis” OR responsiveness OR “internal consistency” OR “outcome measures” OR “outcome assessment” OR “outcome tool” OR “outcome instrument” OR “outcome scale” OR “likert scale” OR PROM OR “patient reported outcome measure”

CINAHL headings
Quality of life
Psychometrics
Psychiatric inpatients

Appendix D

COSMIN Risk of Bias Assessment

REDACTED

Appendix E

COSMIN Criteria for Good Measurement Properties

REDACTED

Appendix F

Grade Approach

Grade factor	Downgrading	Definition
Risk of bias	0	Multiple studies of at least adequate quality OR one study of very good quality
	-1	Only one study of adequate quality OR multiple studies of doubtful quality
	-2	Only one study of doubtful quality OR multiple studies of inadequate quality
	-3	Only one study of inadequate quality
Imprecision (not for content validity, structural validity, and cross-cultural validity/measurement invariance)	0	Total sample size of all studies >100
	-1	Total sample size of all studies 50-100
	-2	Total sample size of all studies <50
Inconsistency	0	Results are consistent OR results are summarised and rated per subset of studies, and subsequently graded
	-1	Overall rating based on the majority of consistent results
Indirectness	0	One population (population of interest)
	-1	Mixed populations*
	-2	Sample is not target population*

*additional rating were included to reflect mixed populations

Appendix G

Decisions Log

Content Validity	If the study reports on content validity and gives an indication of relevance, comprehensibility and comprehensiveness as reported by patients or professionals. This was scored as a +. If this was not present, it was left blank. Although COSMIN requires content validity before progressing, it is important to include all studies anyway to show gaps in the research in an inpatient setting were content validity should be gathered.
Structural validity	If the sample size was not sufficient this was downgraded. COSMIN guidelines were strictly followed.
Internal consistency	If the scale reported Cronbach's alpha and was a continuous scale with indication of unidimensionality this was scored as VG. If there is evidence of unidimensionality found from other studies in the same review, this is enough for internal consistency to be score + with the appropriate alpha score.
Cross-cultural validity	Cosmin requires that articles report on MGCFA or other methods of comparing groups. It also should be tested on a sufficient sample size. If the study just reported the translation process, although this is sufficient, it is not sufficient with COSMIN and so this was marked as a ? because although not incorrect, there is information about relationships between groups missing.
Reliability	If description of stability was not made or consistency of test environment this was marked down as 'A' depending on the overall quality of the study, if the study was poor, it cannot be assumed that test conditions are the same. If ICC was reported and about <.07 it was scored as a + if only the subdomains scores were reported and some fell below this then the overall scale was marked as a -, this is because we were looking for total QOL rather than individual subscales.
Measurement error	Similarly, there needs to be a statement of consistent test conditions. If measurement error was good, this was score as a + even when COSMIN looks for specific report. This is because some articles are older and use different statistical methods.
Construct validity	If no hypothesis was made, but it was clear what was anticipated even without hypothesis the score was dropped to 'A' for methods and design. Furthermore, for discriminant validity if no attempt was made at correlating results or showing effect size or change, this was marked as 'D'. For convergent validity, if no correlations were made it was marked as 'D' for analysis. Scoring was marked as a + if most results showed expected relationships and were score ? if it was unclear what was expected and what the result meant.
Responsiveness	If no hypothesis was made, but it was clear what was anticipated even without hypothesis the score was dropped to 'A' for methods and design.
Overall ratings	If the majority of studies score + or – and only some ? the overall score was summarised as the majority. Scales with some items missing were not included in the summary score. Scales of different languages were summarised together if they were comparable items. One study was removed after contacting the author to confirm inpatient setting.

Part Two: Empirical Study

Implementing the DIALOG Outcome Measure and DIALOG+, A Solution-Focused Intervention, in a Psychiatric Rehabilitation Inpatient Setting: A Mixed Methods Randomised Feasibility and Acceptability Study.

Word Count: 7970

Abstract

This study aimed to determine the feasibility and acceptability of implementing DIALOG and DIALOG+ in a psychiatric rehabilitation inpatient setting.

Method

A sequential explanatory mixed-methods design was employed. Firstly, a randomised between-groups comparison design was used to collect data on the implementation of DIALOG and DIALOG+ over five months. Secondly, patient participants completed surveys, while staff participants engaged in focus groups and interviews about their experiences using DIALOG and DIALOG+.

Results

Results indicated that the recruitment rate failed to meet the established criteria for progression, and engagement rates were low, suggesting that the intervention was not presented regularly to patient participants. Thematic analysis identified four themes that support the feasibility and acceptability outcomes: affective attitude towards feasibility objectives; affective attitude towards the intervention; alignment with current practice; and the inpatient setting.

Discussion

Some aspects of the trial met the criteria for progression, while others did not. The study had some limitations, including sample size, linked to low recruitment rate. It is recommended that further feasibility trials be conducted to investigate these areas, particularly in terms of recruitment and engagement.

Practitioner Points

- Regular staff training, coaching, and supervision should be used when implementing a new intervention in this setting.

- A lead should be based within the team who can organise and communicate regularly with the team system to ensure DIALOG+ can be effectively integrated into the current practice.

Keywords

Feasibility, acceptability, clinical trial, inpatient psychiatric rehabilitation, DIALOG, DIALOG+, mixed methods.

Introduction

Outcome measures have routinely been used in mental health settings to monitor changes over time. When employed effectively, they can enhance patient outcomes and the quality of care (Philpot et al., 2018). They can influence clinical decision-making, such as readiness for discharge and referrals to subsequent services (Boyce & Brown, 2013; Roe et al., 2022). In inpatient settings, quality of life (QoL) outcome measures are a priority, particularly with the deinstitutionalisation of inpatient practices and emphasis on patient-centred care (McCarron et al., 2019). In mental health settings, patient-reported outcome measures (PROMs) are recommended for measuring QoL (Shepherd et al., 2014). However, published data on the use of PROMs in psychiatric inpatient settings is limited (Boyce & Browne, 2013; Churruca et al., 2021; Gelkopf et al., 2021).

Given the importance of embedding meaningful outcomes into mental health care, DIALOG was developed as a PROM of subjective quality of life (SQoL) and treatment satisfaction (Priebe et al., 2012). To further structure patient-clinician conversations, the DIALOG+ intervention was developed (Priebe et al., 2013). It is grounded in solution-focused and person-centred theory. It emphasises the patient's strengths and resources, which have the potential to reduce the workload of healthcare professionals and lead to efficient practices (Priebe et al., 2017). Moreover, it has been demonstrated to be an effective intervention when used in community mental health teams (CMHTs), with people experiencing psychosis, and in various countries, showing positive effect sizes (Fichtenbauer et al., 2019; Omer et al., 2016; Priebe et al., 2013). However, to our knowledge, it has not yet been tested in a psychiatric rehabilitation inpatient setting.

Inpatient psychiatric rehabilitation care aims to facilitate recovery and integration into the community, and individuals may develop goals related to housing, employment, living independently, and relationships (Booker et al., 2020). Although DIALOG+ has not

yet been tested in this setting, it could offer benefits that align with its goals, such as enhancement of self-management, improved patient-clinician communication and contributions to personalised care plans, improved coproduction, and enhancing QoL by focusing on aspects of life that are important to the patient (Lundmark et al., 2024; Priebe et al., 2017; VanLoggerenberg et al., 2023).

There are recognised barriers to implementing outcome measures and new interventions in inpatient settings, such as the lack of available appropriate technology, clinician attitudes, concerns that measures are depersonalising, and that data may not be in the best interest of the service user (Sharples et al., 2016). Furthermore, Millar et al. (2020) highlighted that patients experiencing psychosis may find it challenging to engage. Other barriers may include higher discharge rates (Berry et al., 2016), low staff levels and resources, concerns about risk, inadequate staff training (Ball et al., 2022), and challenging patient-clinician dynamics and safety issues on the ward (Millar et al., 2020). It was found that having a predictable time for the intervention acted as an implementation facilitator.

DIALOG and DIALOG+ have demonstrated benefits and have been successfully implemented in CMHTs in the UK. This study aimed to implement them in a psychiatric rehabilitation setting; to our knowledge, no previous studies have been conducted in this context. A feasibility study is conducted when there is little to no prior research on the proposed subject matter, as is the case in the current study, asking the question, “Can it be done?” (Blatch-Jones et al., 2018). This provides an opportunity to refine methods and interventions before proceeding with further randomised controlled trials (RCTs) (Pearson et al., 2020).

Research Aims

The current study aimed to implement DIALOG and DIALOG+ in a psychiatric rehabilitation inpatient setting to understand its feasibility and acceptability. This study

aligned with the local National Health Service (NHS) Foundation Trust's plans to implement DIALOG within its inpatient settings.

Feasibility Objectives

To assess the feasibility of the intervention, criteria were established reporting on seven aspects of feasibility, including eligibility, recruitment, engagement, completion, dropout, follow-up, and affective attitude towards the research and intervention (Bugge et al., 2013; Patterson et al., 2020). Data on staff and patient participants were collected for each feasibility criterion. Affective attitudes were explored through interviews and focus groups with staff, as well as surveys with patients, which asked questions related to their thoughts on the feasibility criteria.

Acceptability Objectives

A healthcare intervention is considered appropriate or acceptable based on the cognitive, emotional, and experiential responses of those delivering or receiving it (Sekhon et al., 2017). Jacobsen et al. (2020) propose qualitative methods to understand staff and patient feedback related to acceptability. Therefore, acceptability objectives were achieved through qualitative semi-structured interviews, focus groups, and questionnaires. To assess acceptability, the theoretical acceptability framework was used, encompassing attitudes towards research and the intervention, ethical considerations, intervention coherence, cost, and perceived effectiveness (Sekhon et al., 2017).

Secondary Aim

Feasibility trials can indicate effectiveness before implementing an RCT (Blatch-Jones et al., 2018). To assess whether DIALOG+ improved outcomes for patients, a comparison of pre- and post-intervention scores was suggested. To achieve this, DIALOG scores and demographic details were collected to evaluate the potential generalisability of these measures.

Method

Ethical approval was obtained from an NHS research ethics committee (REC 24/SC/0209), and the Health and Research Authority approved the study (9th July 2024; Appendix A). The study protocol was registered in an available repository clinicaltrials.gov (Reference number: 187799). Participant data is stored securely and anonymously. A detailed data management plan (DMP) has been uploaded to DMPOnline (ref: 140413; Appendix B). A data-sharing agreement between the NHS trust and the University of Sheffield was implemented (Appendix C). Informed consent, confidentiality, and the right to withdraw guided all aspects of the research design as described throughout.

Patient and Public Involvement (PPI)

Feedback from staff and patients in participating wards was gathered to inform the design of study materials, including posters, consent forms, information sheets, and the DIALOG and DIALOG+ booklets, resulting in minor changes to presentation before final ethical approval. The patient engagement team was also involved as part of the broader PROM implementation program within the trust.

Design

The Journal Article Reporting Standards (JARS) for mixed methods studies (2018) and the CONSORT checklist for randomised pilot and feasibility trials (Eldridge et al., 2016; Appendix D) were followed to ensure accurate reporting. A sequential explanatory mixed-methods design (Ivankova et al., 2006) was employed. The purpose of the mixed-methods design was to triangulate data by combining quantitative and qualitative components of the trial (Burt, 2015). Findings were integrated through narrative reporting using a contiguous approach (Fetters et al., 2013). Feasibility data were collected during the quantitative phase of the study, while qualitative methods were used to gain insights from participants on feasibility and acceptability outcomes.

Quantitative Method

Design

A randomised between-groups comparison design was employed (Simkus, 2024). This design adheres to the pre-planned implementation of DIALOG in this setting and aligns with the previous application of DIALOG+ in community settings (Priebe et al., 2015). Data were collected longitudinally (Caruana et al., 2015) over a period of five months. The independent variable was group membership, either in the comparison (DIALOG) or treatment (DIALOG+) group. The dependent variable was the scores on the DIALOG measure after the intervention period.

Feasibility Criteria

To meet the study's aims, feasibility criteria (Table 1) were developed based on the requirements outlined by Patterson et al. (2018) and Bugge et al. (2013). These criteria were used to inform data collection and analysis. Some of the feasibility criteria were met through qualitative methods, including affective attitude and willingness of clinicians to recruit participants.

Table 1

Feasibility Criteria

Feasibility Criteria	Description and Data Sources
Eligibility and Descriptive statistics	This is the number of eligible participants out of the total number of patients. Age, gender, ethnicity, history of mental illness or previous inpatient admissions, drug-use, diagnosis, baseline severity on DIALOG, length of stay.
Recruitment rate/ consent rate	The number of people who agreed and consented to participate in the study out of the total number approached. Furthermore, the willingness of clinicians to recruit participants, investigated qualitatively.
Staff participation rate	Number of staff members who attended DIALOG+ training and implemented the intervention.
Average number of sessions.	Average number of times the intervention was successfully implemented and average number of multi-disciplinary team (MDT) meetings to check treatment fidelity.
Patient Participation rate	Out of those who consented and agreed to engage in DIALOG and DIALOG+ how many engaged in the intervention after consenting. (Total sample size / who accepted)
Completion rate	The proportion of those who consented and engaged in DIALOG and DIALOG+ and who completed the intervention.
Dropout rate	Proportion of participants who initially engaged with the intervention and then chose to drop out. The total number of patients enrolled minus the number who completed the intervention divided by the total number of enrolled participants.
Follow up rate	The number of participants willing to engage in follow-up out of the total number of participants.
Affective Attitude	Participant's (staff and patient) attitudes towards feasibility objects that were met or not met during the trial. Understood through qualitative methods.

Participants

The study was conducted with inpatients within an NHS Trust, including three psychiatric rehabilitation inpatient wards (Table 2). Patient eligibility criteria are outlined in Table 3 and were derived from previous studies (Priebe et al., 2013; Wong et al., 2022). Staff eligibility criteria are outlined in Table 4.

Table 2

Table of Sample Size

Ward	Average length of stay	Number of beds
Male Bungalow 1	336 days	8
Male Bungalow 1A	336 days	14
Female bungalow 2	336 days	8

Table 3

Patient Inclusion and Exclusion Criteria

Inclusion	Exclusion
• Aged between 18-65 • Capacity to give informed consent as determined by the MDT and through following informed consent processes. • Patients admitted under the Mental Health Act (MHA) were eligible to participate if they could consent, as determined by the MDT • Admitted to the inpatient ward • Can complete DIALOG and DIALOG+	• No capacity to consent • Insufficient command of the English Language

Table 4

Staff Participants Inclusion Criteria

Inclusion	Exclusion
• Had received PROMS training from the trust to use DIALOG • Worked across one of the participating wards. • Any patient facing role	• Had not received DIALOG training • Agency staff • Non-patient facing role

Sample Size

A power calculation is not appropriate for a feasibility trial (Jacobson et al., 2020; Sims & Lewis, 2012). Instead, sample size was determined by the study's duration and the availability of participants within the specified time frame. Several authors propose a sample size ranging from 12 to 36 (Browne, 1995; Julious, 2005; Varese et al., 2021; Billingham et al., 2013), and a recruitment rate between 42.04% and 65% would meet feasibility requirements (Berry et al., 2016; Millar et al., 2020; Young et al., 2019).

Measurement Instruments

DIALOG (Appendix E) consists of 11 questions, eight related to SQoL and three to treatment satisfaction. Patients can rate their satisfaction on a 7-point Likert scale for each of the eight life domains and three treatment aspects, resulting in scores for both SQoL and treatment satisfaction. The DIALOG scale was found to be reliable and valid as an outcome measure in community mental health settings, with good internal consistency and convergent validity (Priebe et al., 2002; 2007; 2012).

DIALOG+ (Appendix F) utilises the DIALOG outcome measure and includes additional solution-focused questions for clinicians to ask patients, employing a “four-step approach” (understanding, looking forward, considering options, and agreeing on actions) (Priebe et al., 2020).

DIALOG+ is a web-based app, but due to data security issues and a lack of funding for IT support, we were unable to use it. Instead, a paper booklet was created with permission from the original developers (Priebe et al., 2020; Appendix G) to explore the viability for future app integration. However, using paper versions has limitations, such as being cumbersome and potentially disrupting conversations.

Procedure

Posters (Appendix H) were distributed amongst the team. Eligible staff received information sheets in person (Appendix I) and time to ask questions, and the researcher received informed written consent (Appendix J). Participating staff were trained in the principles of Good Clinical Practice (GCP) in a Trust-led training session before commencing recruitment. Once trained, staff could receive consent from patient participants. Patients received information sheets (Appendix K) and were given time to ask questions. Written consent (Appendix L) was received by trained staff in person. Recruitment occurred throughout the trial's duration for staff and patients who met the eligibility criteria. Participants could withdraw from the trial at any time before the data analysis.

Consenting staff received DIALOG+ training provided by the primary researcher and supervisor using resources developed by the East London Foundation Trust (ELFT) (Priebe et al., 2020). Participating wards were then randomly allocated by the primary researcher to two groups: a treatment group (DIALOG+) and a comparison group (DIALOG), using a computer-generated randomisation schedule called random.org (2025), which produces random integers.

Data Collection

Once trained, participating staff began collecting data between 27th November 2024 and 30th April 2025. Krasny-Pacini and Evans (2018) recommend a minimum of ten sessions for intervention in a comparison group study. Therefore, DIALOG and DIALOG+ were implemented over five months, with the opportunity to complete them across ten sessions.

To fulfil feasibility and acceptability criteria, staff were prompted to make a note in the clinical note system if participation in the trial had been offered to patients and the reasons for declining it. Staff were asked to complete either the DIALOG or the DIALOG+

with consenting patients, depending on their group assignment, and record outcomes in the respective booklets. DIALOG data collection took between 3 and 15 minutes (M = 5.89), and DIALOG+ took between 25 and 45 minutes (M = 33).

Data were collected fortnightly in the respective booklets, with an option for staff to record whether patients were offered to participate and declined. The scores, key discussion areas, and progress were intended to feed into the fortnightly MDT meetings.

DIALOG and DIALOG+ data were collected and stored in booklets, which were kept securely following the DMP. DIALOG and demographic data were extracted at the end of the trial, also following the DMP. Staff were asked to input data into patient records. To help staff manage any risks associated with the study, a risk management plan was developed (Appendix M).

Proposed Analysis

The primary analysis of feasibility was conducted using the feasibility criteria. Each feasibility domain was assessed by collecting count data and calculating the corresponding rates, expressed as percentages. Some domains are evaluated through qualitative methods, including affective attitude and the willingness of clinicians to recruit participants.

Progression Criteria

Young et al. (2019) proposed a set of progression criteria to determine whether feasibility has been achieved. These criteria have been adapted to be relevant to the current study (Table 5). Criteria that are “stop” indicate that they do not meet the requirements for progression to RCT, whereas those rated “go” do. Criteria that are neither “stop” nor “go” are deemed “consider,” indicating they require further evaluation and adaptation (Avery et al., 2017). Progression criteria for drop-out rate were not specified in the original protocol. For analysis, it was assumed to align with the loss-to-follow-up criterion used. It is essential to consider both drop-out and loss-to-follow-up to understand at which stage participants

stop engaging. Results will be given a red, amber, green (RAG) rating (Mellor et al., 2023) to identify areas of improvement visually.

Table 5

Table of Progression Criteria

Criteria	*Stop*	*Go*	*Current study*
Eligibility	20% of patients eligible	More than 50% of patients eligible	Majority of patients likely to be eligible to participate in DIALOG and DIALOG+. Understood by eligibility criteria and patient demographics.
Recruitment	Less than 25% of eligible patients recruited	more than 50% of eligible patients recruited.	Recruitment may be difficult in the inpatient psychiatric wards. Understood by analysing recruitment rates.
Completion	less than 30% adherence to DIALOG and DIALOG+ sessions	more than 70% adherence to DIALOG and DIALOG+ sessions	Understood through completion rates of DIALOG and DIALOG+
Participation	Less than 25%	More than 50%	Understood by the number of those who consent to engage and then engaged in the intervention after.
Dropout	More than 40% loss to drop-out	Less than 20% loss to drop out	Understood by those who initially engaged with the intervention and then chose to drop out.
Loss to follow-up	more than 40% loss to follow up	less than 20% loss to follow up	Understood by considering participation in follow-up patient surveys.

Secondary analysis

The proposed analysis was a linear regression model (Arkes, 2023) to assess whether the implementation of DIALOG+ as an intervention, in addition to using the DIALOG outcome measure, leads to improved outcomes. However, insufficient data were collected to enable any meaningful comparisons.

Qualitative Method

Qualitative methods offer participants a voice in research (Sutton, 2015) and provide insight into participants' subjective experiences (Clarke & Braun, 2013). This research is grounded in a critical realist perspective (Bhaskar, 2013), facilitating an understanding of the

real barriers and facilitators encountered, while acknowledging the subjectivity inherent in social constructions that shape how they're perceived (Braun & Clarke, 2022b; Fryer, 2022).

Acceptability Framework

To meet the study's aims, an acceptability framework (Sekhon et al., 2017; Table 6) was utilised to guide data collection and analysis. A critical realist position in this study meant that the researcher employed the pre-established feasibility and acceptability criteria to steer analysis, recognising that there was a clear focus of the study and that the findings would likely align with these criteria. However, this stance also acknowledges that these experiences occur within a social context, including the researcher's bias, which is further reflected upon through reflexivity.

Table 6

Acceptability Framework

Acceptability Criteria	Description
Affective attitude	How the staff and patient participants feel about the intervention.
Perceived effort of implementation	Information from staff and patient participants on the response effort of implementing and participating in the study intervention.
Ethicality (current practice)	Gathering information about how the intervention aligns with the current inpatient context and values.
Intervention coherence	Gathering feedback from staff and patient participants on their understanding and experience with the intervention.
Cost	Analysing the cost and benefits of implementing the intervention (including staff training and time).
Perceived effectiveness	Gathering feedback from staff and participants on the intervention's effectiveness and perceived success.

Data collection

Focus groups and supplementary semi-structured interviews were completed with staff participants between 3rd and 21st of April, 2025. An interview guide (Appendix N) was developed based on previous research (Omer et al., 2016; Priebe et al., 2017).

Focus groups enable the collective construction of meaning (Kook, 2019; Kitzinger, 1995), which may represent the daily interactions of staff within inpatient team settings, providing insight into group dynamics that could impact the feasibility of the study. The interaction can stimulate more memories, debate and disclosure (Millward et al., 1995). A hybrid method combining focus groups and interviews was used to gather both breadth and depth of data, encompassing group dynamics and individual perspectives (Morgan, 1997). It was hoped that using a small number of interviews would help to validate the focus groups' findings and triangulate data, and avoid power imbalances by interviewing managers separately.

A printed survey with open-ended questions, based on Matanov et al. (2021), and a theory-informed questionnaire (Appendix O), developed by Sekhon et al. (2022), were offered to patient participants in both groups following the completion of the intervention or upon drop-out. Participants completed these with a staff member who knew them well. With pilot research, surveys are used to identify issues, particularly when there is little previous research (Palinkas, 2014). Surveys have the benefit of adding context and offering additional explanation, particularly important in a feasibility trial.

Sample characteristics

In focus groups, it is generally expected that the sample is heterogeneous; in this case, however, the staff worked together in similar roles. Strategies to mitigate staff feeling uncomfortable expressing honesty were implemented, such as asking prompt questions e.g. "Does anyone disagree?" (Breakwell et al., 2020).

Participants

Purposeful sampling (Palinkas, 2013) was used to invite eligible staff (Table 7) to focus groups and senior leadership team members to interviews.

Table 7

Staff Participants' Eligibility Criteria

Inclusion	Exclusion
Staff who work on one of the participating wards and have consented to participate,	Agency staff
Professionals who are involved in leadership roles within the service and consent to participate in the study.	Staff who have not participated in the trial and do not consent to participate in feedback for research.

The recommended sample size for focus groups varies based on study aims, and some researchers advise between three and eight participants (Braun & Clarke, 2022a; Breakwell et al., 2020). Smaller groups may lead to less diverse viewpoints and potential issues with engagement (Morgan, 1997). We aimed for four to five participants per group for sufficiently diverse views and discussion. We aimed to have 12 participants in total to achieve an adequate depth and breadth of perspectives from a range of professionals (Guest et al., 2006).

Procedure

Focus groups and interviews were conducted by the primary researcher in person in the MDT room, located within the inpatient service. Interviews took place online via Microsoft Teams. An encrypted recording device was used to record focus groups and interviews, and an external transcription agency was used for verbatim transcription. The data was anonymised by removing participants' names and locations from the transcripts.

Interviews lasted between 16 and 42 minutes (M = 29). Focus groups lasted between 69 and 77 minutes (M = 73). Confidentiality was explained at the beginning of the session. Participants had a professional relationship with the moderator, who was involved in

training, coaching, and data collection throughout the trial's duration. The moderator's role was to design the conversation and notice when participants began to lose track (Breakwell et al., 2020), and to direct the discussion using the interview guide, which consisted of open questions.

Reflexivity

To decrease bias, researchers are advised to actively engage in reflexivity and evaluate their subjectivity (Olmos-Vega et al., 2023; Watt, 2007). The primary researcher was a trainee clinical psychologist. As a trainee, they might feel pressured to find the "correct" answer. Additionally, the primary researcher previously worked as a Behaviour Analyst in residential care homes. This involved collaboratively developing positive behaviour support plans (PBS) with staff and clients, including sharing knowledge about behaviour strategies and implementing changes. These experiences meant the researcher had encountered barriers and facilitators when applying behaviour change strategies in similar settings, which could lead to biases such as assumptions about what barriers and facilitators may be present in the current study, based on prior experiences, potentially influencing data interpretation. A short reflexivity log was maintained to help recognise and reduce these biases; selected entries are included in Appendix P.

Analysis

Transcribed and anonymised interviews and focus groups were analysed using reflexive thematic analysis (RTA) (Braun & Clarke, 2006). RTA was chosen as a method that acknowledges the researcher's bias and reflexivity (Creswell & Poth, 2016). Furthermore, it is theoretically flexible (Clarke & Braun, 2014), allowing for the incorporation of the acceptability framework. Alternative methods of analysis were considered, including framework, content, or template analysis. However, framework analysis is typically used with large sample sizes to quantify patterns and is employed in policymaking (Goldsmith, 2021);

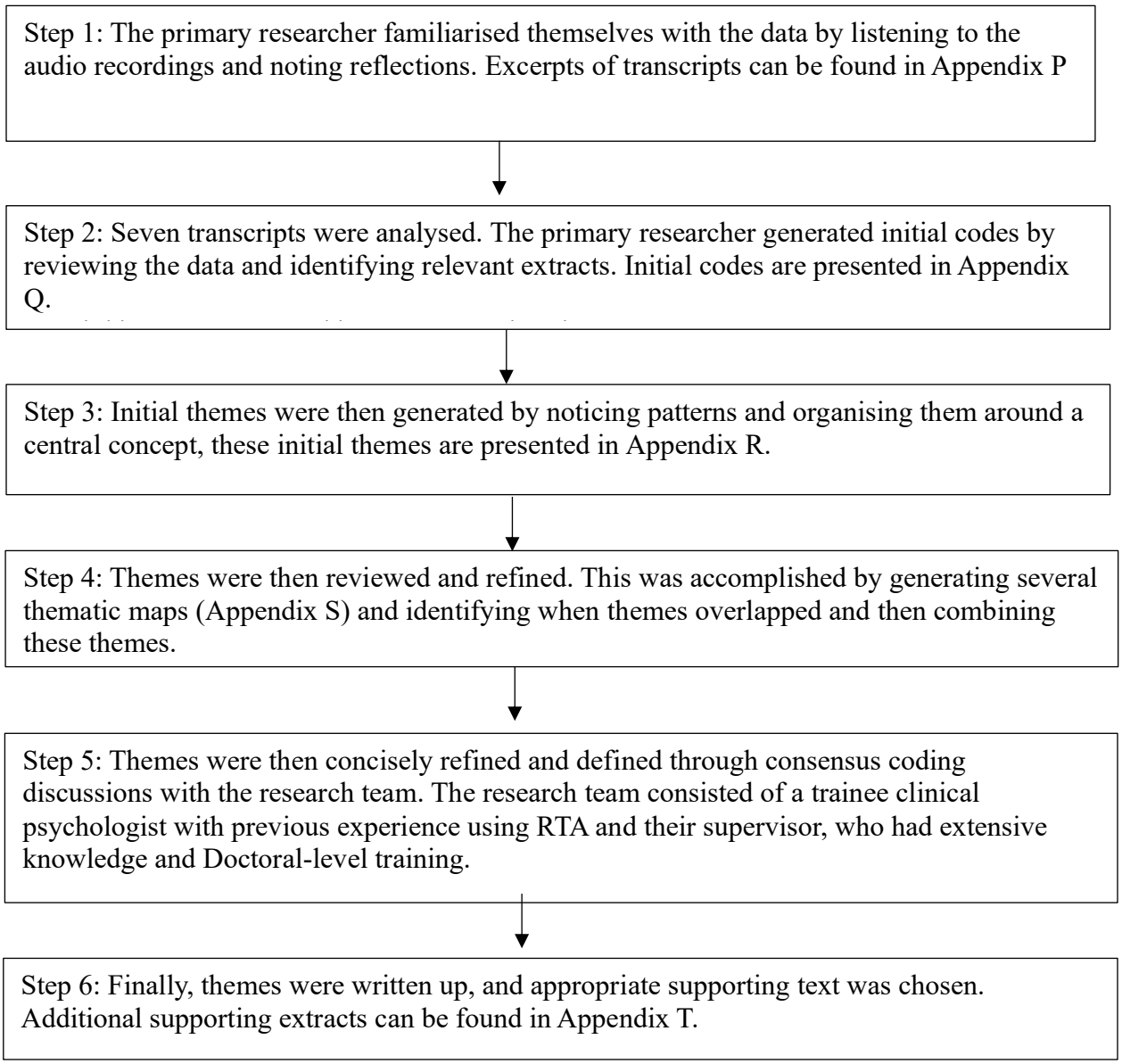
other methods may not provide the same analytical flexibility as RTA, which utilises both inductive and deductive reasoning (Neuendorf, 2018). Furthermore, RTA draws both explicit and implicit ideas from the texts, relying on the researcher's interpretation (Guest et al., 2012). The collaborative effort of researchers and participants in co-constructing meaning proves especially beneficial for exploratory studies and aligns most closely with the study's aims (Braun & Clarke, 2019).

The analysis adhered to Braun and Clarke's six-phase framework (2006), as illustrated in Figure 1. To assist with the coding and mapping process, NVivo 15 (2024) software was utilised. We did not aim to reach saturation as it is typically not relevant when employing RTA, as the researcher actively constructs themes; consequently, it does not align with the values and assumptions of RTA (Braun & Clarke, 2024). Consensus checking was used to establish agreement through discussions with the research team regarding the themes, thereby reducing bias (Hill & Knox, 2021).

Unfortunately, not enough survey data was collected to analyse thematically, and so this was presented descriptively to ensure the service users' perspective was still considered (Bingham, 2020).

Figure 1

Six-Phase Framework of RTA



Results

Quantitative Results

Staff and patient characteristics are presented in Table 8. There was a total of eight patient participants and 17 staff participants. Ethnicity and diagnosis data were redacted due to low numbers to maintain patient and staff anonymity.

Table 8

Patient and Staff Characteristics

Demographic	DIALOG	DIALOG+
Patient Sample size (N)	5	3
Mean age (SD)	32.6 (8.35)	38 (13.22)
% male	60%	100%
Staff sample size (N)	11	6
Staff job role	Support worker, Nurse, Deputy ward manager	Support worker, clinical associate psychologist, Nurse

The feasibility outcomes are presented in Table 9, alongside the progression criteria, and are RAG-rated. Results are reported as percentages; however, due to the small sample size, percentages may appear inflated, so numerators are also reported.

The study met the requirements for progression in several areas, including patient eligibility (83.78%), recruitment rate for staff (68%), and participation rates for patients (90.47%) and staff (58.82%). However, the engagement rate was low (30%), highlighting that patients weren't offered the intervention regularly. Furthermore, patient participation rates were higher than those of the staff, suggesting a potential issue with staff engagement in the intervention. This was explored in more detail qualitatively.

The total average number of sessions completed over the five-month intervention period was low (1.87); the goal was to complete 10 sessions, and no outcomes were brought

to the MDT. There were several areas identified that indicated that adjustments are needed to progress with the trial, including patient recruitment rate (27.58%), engagement rate (30%), completion rate 62.5% and dropout rate (25%). Results indicated that the DIALOG+ group's recruitment rate (17.64%) and the total patient (37.5%) and staff (52.94%) follow-up rates did not meet the criteria for progression.

Table 9

Feasibility Results

Progression Criteria	Participant	DIALOG	DIALOG+	Total	Descriptive information
Eligibility rate Number of eligible participants Stop: <20% of patients eligible Go: >50% of patients eligible	Staff	Eligible for training: 10 Attended training: 6 Training in DIALOG later: 5	Eligible for training: 10 Attended training: 7	Eligible for training: 20 Later: 25	Reasons for not attending training: Sickness No work capacity Carers leave Non-working day Administration error
	Patient	Total patients: 19 Eligible patients: 14 Percent eligible: 73.68% Criteria met: GO	Total patients: 18 Eligible patients: 17 Percent eligible: 94.44% Criteria met: GO	Total patients:37 Total eligible:31 Total percent eligible: 83.78% Criteria met: GO	Some participants were not asked due to staff team changes.
Recruitment/ consent rate The percentage of people who agreed and consented to participate in the study out of the total number approached. Stop: <25% of eligible patients recruited Go: >50% of eligible patients recruited Recruitment occurred between 27th November 2024 to 30th April, 2025	Total eligible staff sample/ who accepts	Approached: 15 Agreed:11 Consent rate: 73% Criteria met: GO	Approached: 10 Agreed: 6 Consent rate: 60% Criteria met: GO	Total recruitment rate: 68% Criteria met: GO	
	Total eligible patient sample size/ who accepts	Number asked to participate: 14 Consented: 5 Consent rate: 35% Criteria met: Consider	Number asked to participate: 17 Consented: 3 Consent rate: 17.64% Criteria met: Stop	Total asked to participate: 29 Total consented: 8 Consent rate: 25.81% Criteria met: Consider	
Participation rate Stop (<25% recruited) Go (> 50% recruited) Out of those who consented and agreed to engage in DIALOG and	Number of staff who attended training, consented to participate and then implemented intervention	Number of staff consented: 11 No. who implemented DIALOG: 6 Participation rate: 54.54% Criteria met: GO	Number of staff consented: 6 No. who implemented DIALOG+: 4 Participation rate: 66.66% Criteria met: GO	Total staff consented: 17 Total staff participation: 10 Total participation rate: 58.82% Criteria met: GO	

DIALOG+ how many engaged in the intervention after consenting.	Patients	Number of potential opportunities: 40 Number of times asked: 15 Number of times engaged: 11 Engagement rate: 15/40= 37.5% Criteria met: Consider Participation rate: 11/15= 73.33% Criteria met: GO	Number of potential opportunities: 30 Number of times asked: 6 Number of times engaged: 4 Engagement rate: 6/30 20% Criteria met: STOP Participation rate: 4/6 66.66% Criteria met: GO	Total number of potential opportunities: 70 Total number of times asked: 21 Total number of times engaged: 19 Total engagement rate: 30% Criteria met: Consider Total participation rate: 90.47% Criteria met: GO	Reasons for not engaging: XMAS holiday, on leave, too tired Staff feedback that they sometimes didn't ask patients to engage if they felt they wouldn't and therefore data collection opportunities were missed.
Completion rate: proportion of those who consented and engaged in DIALOG and DIALOG+ who completed the intervention. Stop: <30% adherence Go: > 70% adherence.	Patients	Number who participated: 5 Number who finished: 3 Completion rate: 60% Criteria met: Consider	Number who participated:3 Number who finished: 2 Completion rate: 66.66% Criteria met: Consider	Total number who participated: 8 Total number who finished: 5 Total completion rate: 62.5% Criteria met: Consider	
Average number of sessions	Patients	Total number of sessions: 11 Number of participants: 5 Mean session: 2.2 Number of MDTs: 0	Total number of sessions: 4 Number of participants: 3 Mean:1.33 Number of MDTs: 0	Total number of sessions: 15 Total number of participants: 8 Total mean: 1.87 Total MDT: 0	
Dropout rate: Proportion of participants who initially engaged with the intervention and then chose to drop out. Stop: >40% dropped out Go: < 20% dropped out	Staff	No. of staff: 11 No dropped out: 3 Dropout rate: 27.27% Criteria met: Consider	No. of staff: 6 No dropped out: 2 Dropout rate: 33.33% Criteria met: Consider	Total: 17 Total drop out: 5 Dropout rate: 29.41% Criteria met: Consider	No reasons given.
	Patients	No. of patients: 5 No dropped out: 1 (a further one lost capacity) Dropout rate: 20% Criteria met: Consider	No. of patients: 3 No dropped out: 1 Dropout rate: 33.33% Criteria met: Consider	Total: 8 Total drop out: 2 Dropout rate: 25% Criteria met: Consider	No longer had capacity to participate, decided they didn't want to do it anymore, no reason given.

Follow-up rate: The number of participants willing to engage in follow-up surveys related to the study Stop: >40% loss to follow-up Go: <20% loss to follow-up	Staff	No. participants: 11 No. engaged in follow-up: 5 Follow-up rate: 45.45% Criteria: STOP	No. participant: 6 No. engaged in follow-up: 4 Follow-up rate: 66.66% Criteria: Consider	Total:17 Total follow-up: 9 Follow-up rate: 52.94% Criteria: STOP	Some staff who did not attend follow-up gave drop-out as a reason and not participating in the intervention. There were some from the DIALOG group who did not attend due to scheduling.
Follow-up for patient participants occurred throughout the duration of the trial from 27th November, 2024, to 30th April, 2025. Staff follow-up occurred in April 2025.	Patients	No. participant: 5 (one lost capacity) No. engaged in follow-up: 2 Follow-up rate: 40% Criteria: STOP	No. participant: 3 No. engaged in follow-up: 1 Follow-up rate: 33.33% Criteria: STOP	Total: 8 Total follow-up: 3 Follow-up rate: 37.5% Criteria: STOP	Patient reasons: no longer had capacity, discharged, no reason given.

Initial plans to use regression analysis to understand efficacy were not feasible due to the limited data available. Results (Table 10) show the DIALOG+ group had only one participant who repeated the measure over time, and so differences between groups can't be compared. DIALOG+ was offered on average twice per participant, and DIALOG three times per participant, throughout the trial, indicating an issue with engagement in offering the intervention. No DIALOG scores were brought to MDT, further examined through RTA.

Table 10

DIALOG and DIALOG+ Scores

	DIALOG	DIALOG+
Number of participants who completed the intervention	3	2
Average number of DIALOG/DIALOG+ completed	2.2	1.33
Average number offered	3	2
Average number of MDTS	0	0
First DIALOG scores	65.33	44
Final DIALOG score	67	52 (Only one participant had repeated measures)

Patient Qualitative Results

Patient participants were asked to provide feedback on their experiences of the trial and intervention. Unfortunately, not enough participants responded to the feedback questionnaires to draw any conclusions. Therefore, the results are presented descriptively in Table 11.

Three participants completed a questionnaire (Appendix U), some participants reported difficulty answering the question about acceptability, suggesting some refinements to the questionnaire may be required. Two participants from the DIALOG group reported

either liking the scale or having no opinion about it. They felt confident using it, as it didn't interfere with their priorities, was acceptable, and ethical. However, one participant felt it didn't improve their clinical outcome, while the other participant thought it did. One participant from the DIALOG+ group reported that the scale required a lot of effort; however, they felt confident in completing it, found it acceptable, and believed it was clear how it improved their clinical outcome.

These participants also completed a survey. There were contrasting views between groups about the value of the scale. The participant from the DIALOG+ group found the scale more useful than those in the DIALOG group. All participants found the questions essential and relevant. The participant in the DIALOG+ group felt that the intervention helped lead to joint working. In contrast, those in the DIALOG group didn't notice any impact on their relationship with staff. The participant from the DIALOG+ group felt the scale was more thorough, offering structure and clear focus. They highlighted the importance of their results being fed into MDT meetings.

Table 11

Patient Participant Survey Feedback Results

Description	Quote
Reflections about DIALOG scale: There were some contrasting views about the scale with service users in the DIALOG group finding the scale less useful than those in the DIALOG+ group.	DIALOG: P 1: "Not very (useful). I answered a few questions every now and again, didn't really have any benefit. I didn't understand what it was doing for me." DIALOG+ P3: "Gave you a clue where you were on things and any progress. Sometimes staff would use it to give me reassurance in specific areas."
Relevance: All participants found the questions important and relevant.	P1: "The questions were relevant and important" P2: "Questions on accommodation, mental and physical health are important to me" P3 (DIALOG+): "Covered areas that were important to me: such as physical health. Helped me think of specific goals in specific areas. Some areas less important due to being in hospital."
Ease of use: All participants found the scale and intervention easy.	P1: "Easy" P2: "Just did it with staff, it was easy… questions were fine." P3: Very clear- rating, set a goal, come out with some objective's facilitators read it out and filled it out which was helpful and made it easier to think
Relationship with staff: Participants had different opinions on the impact on relationship with staff. Some felt they already have a sufficient relationship, and others felt it led to more joint working.	P1: "No (it didn't impact relationship with staff) because I already had a good relationship with staff" P2: "I preferred doing it with staff than on my own, because I could ask them about it." DIALOG+ P3: Felt it impacted positively with the person delivering it- gave us both a clearer joint understanding
Other Outcome measures: The participant from the DIALOG+ group felt that it was more thorough than other outcome measures	P1: "No, I don't use any other outcome measure" P2: "used GASS for medication, that's similar." P3: Hard to say- used measures before but a while ago. This felt more thorough. I liked the process as it gave me a better focus and helped me to organise what was in my head
DIALOG+ improvements: The participant felt that it was important for results to be fed into MDT	P3: A clearer system for feeding into and out of the MDT. Sometimes communication of plans not clear. Feel the feedback survey needs completing by the person who has done the measures. Consistency across the measures and having the same facilitator felt important.
DIALOG+ organisation: The participant felt that DIALOG+ offered structure and a clear focus, further explaining it was clear and consistent.	P3: very satisfied with certain staff responses- when staff used clear consistent responses, and I understand plans/ objectives I was happy need to understand plan and rationale P3: I liked the process as it gave me a better focus and helped me to organise what was in my head

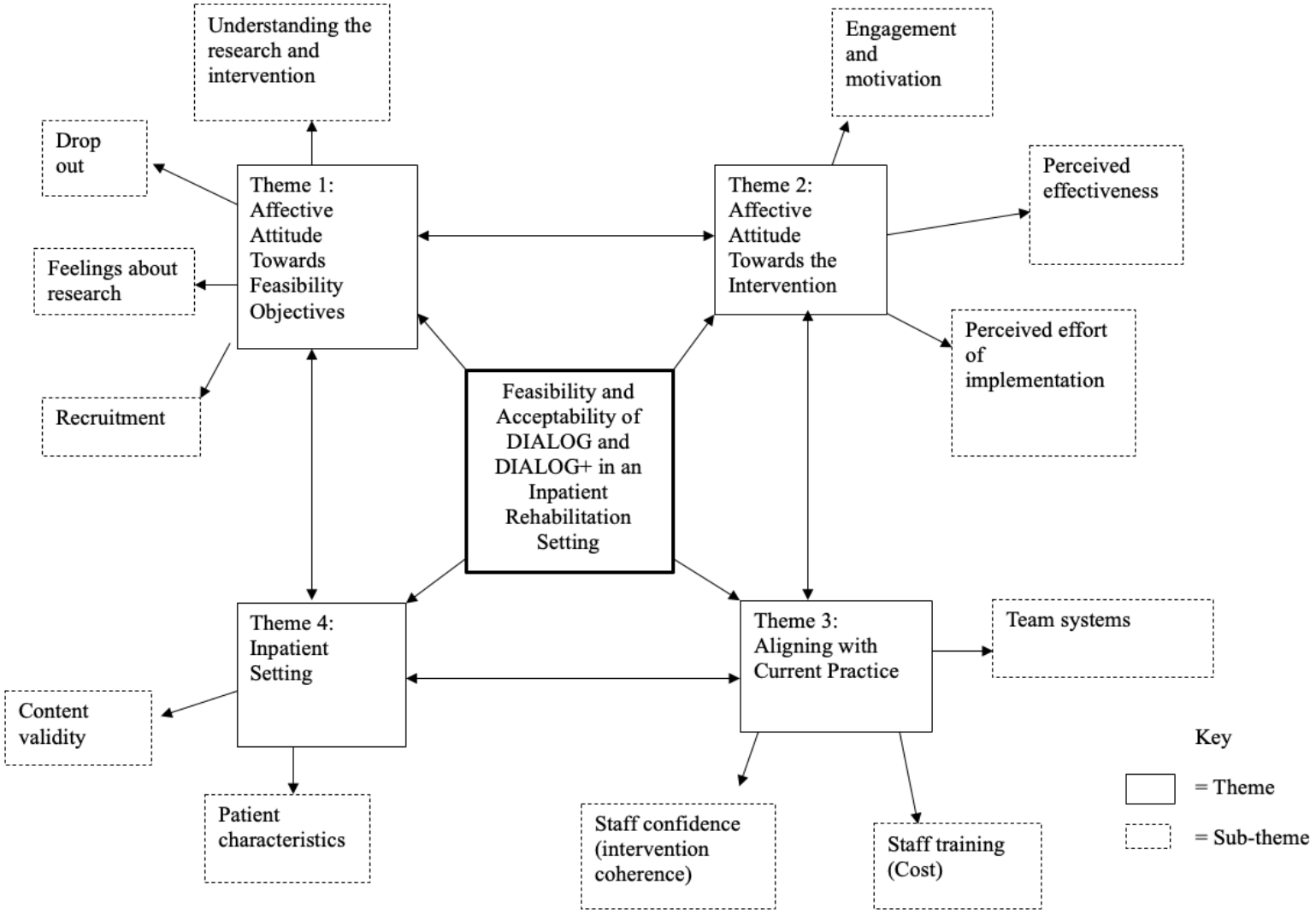
Staff Qualitative Results

Thirteen staff participants provided feedback through either focus groups or interviews. Two focus groups, each consisting of four participants, comprised either nurses, clinical associate psychologists or support workers. Those interviewed included the ward manager, two consultant psychiatrists, a Trainee Advanced Practitioner, and a support worker who was unable to attend the focus group. Some participants were involved in data collection, while others held leadership roles in the teams. Demographic details, such as age and ethnicity, were not included to preserve participant anonymity.

To meet the study's aims, the identified themes align with the feasibility and acceptability criteria. Themes and connections are illustrated in a thematic map (Figure 1). Most participants contributed to all themes, highlighting the shared and overlapping nature of their experiences. Participants are labelled (P), and focus groups are labelled (F). Supplementary illustrative quotes supporting each theme can be found in Appendix R.

Figure 1

Thematic Map



Theme 1: Affective Attitude: Feasibility Objectives

This theme explores thoughts and feelings about research, including confidentiality, drop-out, eligibility, recruitment, follow-up, and future studies. Staff discussions revealed their personal experiences with research, shedding light on the lack of data collection.

Subtheme: Feelings about Research

Participants had contrasting views on the value and purpose of research, with some appreciating its integration into practice and some believing it to be too formal.

P1: "I think it's valuable to have service users involved in and to be participating in studies."

P4: "I felt that it was too formal … because of it being a research study … that maybe posed a barrier."

Subtheme: Understanding the Research Process

Several participants recommended using posters and leaflets, gathering "feedback from patients" (F2P1), offering a "PowerPoint presentation" (P2), and focusing on key domains in DIALOG+.

P1 "I think leaflets would be helpful to be left out in the offices … people can just pick it up and have a look at it."

F1P2: "If they pick three areas that they want to focus on for the entire time … only asking three questions each time."

These suggestions for improving the research were already embedded as part of the trial, which may reflect some of the challenges experienced in inpatient settings regarding sharing understanding and communication around projects. This suggests that having a dedicated research role, with some participants suggesting "a lead to keep an eye on it" (P1), and more frequent and comprehensive training, may be necessary to ensure staff have a sufficient understanding of the protocol.

Subtheme: Recruitment

Discussions highlighted that the team have a strong understanding of the client group.

F2P4 "Straight away I knew who'd do it."

However, it remains unclear whether the assumptions about patient participation may have influenced recruitment attempts.

Some participants highlighted effective methods of recruitment, such as considering the timing of recruitment, adding the outcome measure as part of the admission pack to set expectations and allowing service users time to read the information. In contrast, some participants struggled with recruitment, finding it challenging to describe research processes and the benefits of the study.

F1P4 "Add it to the admission pack, … this is part of your rehabilitation"

F2 P4: "I'd say "oh just have a read of this first, I'll come back to you, if you agree to do it we'll go from there""

F2P1 "I found it very clunky explaining to people why we are doing it … the purpose of it and where the data will be used."

Subtheme: Participation and Drop-out

Service user drop-out was relatively low in the study; however, staff provided some comments on drop-out, indicating reasons including when "they've been discharged" (P1), they have a desire to move on. Some found the intervention "too much" (F2P2). However, quantitative results show that staff drop-out rates exceeded patient drop-out rates.

Some staff in the focus groups had worked within the service for over 30 years, many witnessing the deinstitutionalisation of inpatient services; these experiences may have influenced clinical decision-making. Participants described using their clinical judgement when deciding who would engage with the intervention and when. However, it was unclear

whether this clinical judgement may have acted as a barrier to trialling a new intervention, potentially influenced by their previous experiences within the service.

F2P2 "when they are on the ward, we don't need this to say look at this every two week and … we know, we just know with us patients."

F1P3: "People that will engage will … but people who won't engage are not gonna."

However, it was unclear if clinical judgement acted as a barrier to trialling a new intervention based on previous experiences as a staff member.

Theme 2: Affective Attitude: The intervention

Subtheme: Perceived Effectiveness

There was an ongoing theme that service users did not value the intervention and had other competing priorities, as well as a disconnect between using a formal intervention and its value in day-to-day life. There were different opinions on the value of the intervention. Some think that this would be the last priority, and others think it would align with current priorities.

F1P3: "They know what's important and what isn't so that (DIALOG+) wouldn't be nowhere near as important."

P3: "You'd have to make it a matter of life and death really."

F2P1: "A lot of valuable information there that we wouldn't have got if we hadn't done the DIALOG."

Discussions highlighted that the appearance of the booklet impacted how the intervention was perceived and acted as a barrier to engagement, and some felt that having the DIALOG App would have changed service users' perception of the intervention.

P2 "If it's informal and it's not written … you know into a booklet."

F2P1 "They would be much happier to sit and do it on a phone or an App rather than writing it out."

Subtheme: Perceived Effort of Implementation

Most participants highlighted the effort required to implement the intervention as a barrier. This included the ease of use, perceived effort, time, and duration of the intervention.

F1 P4: "Forty-five minutes is a long time for them."

F2P1 "It was quite resource heavy for both myself and the person I was doing it with."

Other barriers included clients changing their minds on the day. However, this was not reflected in the quantitative data collection.

P2 "We're going to do it that day ... and then come the day he either would not want to do it or he'd want to do something else."

F1P4 ""oh I don't want to do that today" even though you've done your first fifteen minutes"

Subtheme: Engagement and Motivation

Many participants highlighted engagement and motivation as key barriers to implementing new interventions. Throughout the discussion, themes of how staff members felt while working began to emerge, including feelings of burnout, being burdened, guilty, helpless, and at times, unmotivated.

F2P2 "it's not fair on the staff as well, we get burnt out, everybody gets burnt out so when you've got stressors on a ward as well, that's probably a barrier for this."

At times, a sense of despondency seemed to prevail among some staff, potentially stemming from past experiences with limited success in engaging service users in day-to-day activities. It was evident that motivation posed a challenge for service users in all areas of care.

F1P2: "(SIGH), erm (SIGH) like, like we've said it's not about them fitting in the setting, it's getting the clients to engage and do it."

F1P3 "'what's point, there's no', it's like (SIGH) … they're just not gonna do it and then we get despondent."

F1P4 "You've got a patient that's got an emergency appointment at the dentist ... and then won't go to it."

Engagement was linked to how tasks are presented. Participants reflected that "quick wins" (P4) are needed to maintain engagement, and without immediate results, the intervention might be viewed as ineffective. In contrast, other participants suggested successful ways of presenting the intervention, such as having a natural conversation.

P2 "They are not the quick, immediate results that they were hoping for."

F2P2 "It's more like a natural interaction rather than there being a piece of paper and sort of filling a form out together."

Some staff implied that the benefits of DIALOG+ may not be perceived as enough incentive. In contrast, some participants recognised its intrinsic motivation, but it would require an effortful response to contact rewards within the intervention.

F1P1: "There ought to be … a reward for the person that's done it as well because … it's not rewarding for them"

P4: "To see the scores on the graph, certainly for DIALOG+. You need to have done a few scales, which would be about 3 months' worth of work."

Some participants in leadership positions who were interviewed viewed the intervention as part of person-centred care and could see the long-term benefits.

P3: "Because this is person-centred … people are already motivated and willing to engage in rehab … additional tool to help them focus their goals … to achieving them."

Discussions highlighted that staff recognised the importance of consistency when implementing a new intervention. However, results have shown that it was not consistently implemented.

P3: "Yeah, I think consistency is key. If you keep … working at it."

Some participants offered solutions and felt hopeful that the intervention could be achieved. This included distributing the workload and breaking it down "into fifteen-minute slots" (F1P4).

P1: "It's too much really for one person to do all of them with one person, but then also it's easier for one person to do them all, cos it's where you left off."

Theme 3: Aligning with Current Practice

Subtheme: Team Systems

Many participants reflected on different solutions relating to the team systems, including involving the named nurse, using "one-to-ones", integrating with care plans, having protected time, planning the intervention with rotas and staffing, using "mini teams", using handovers, and involving the ward manager. Most participants highlighted the importance of bringing results to the MDT, although results show that DIALOG was not integrated into the MDT throughout the trial.

F2P1 "Protected time to do it otherwise its hard to get the time, especially support workers when you've got all the other jobs to do."

P 1: "Yeah, it's not been embedded in MDT meetings and I think that would really help."

Many participants reflected on the barriers in current practice, including an already overcrowded and burdened MDT meeting, staff shortages, and last-minute rota changes. Highlighting a strained system with multiple new changes happening simultaneously.

P 1 "It's a difficult time 'cause. We're trying to change lots at the moment and I think it's been quite confusing for a lot of people."

F2P1 "I would lose motivation to do it with people if we weren't seeing it properly used, so the idea of I think for it to be reviewed in MDT's … but that never happened because we haven't got enough time in MDT's"

Furthermore, staff emphasised that questions asked in DIALOG were already being asked in practice, which might lead to questionnaire fatigue among staff and service users. In contrast, some participants felt that DIALOG reflected conversations that were already happening, "but we're not capturing it" (P4), offering a way to improve communication.

P2 "That fatigue … you know, every day that they want to go out there, they've got to do the risk assessments, and you know what I mean? … I think the sort of chats that are in the DIALOG are things… You have in passing."

Subtheme: Staff Training (cost)

The feedback from staff indicated that further training, ongoing coaching, supervision, opportunities for practice and feedback, "bite-sized teaching" (P5), and "regular audits" (P1) and "someone to go through it together" (F1P1) are essential for successfully implementing a new intervention and might contribute to increased staff confidence.

F1P2 "If everybody…was trained in it, then any given day…somebody can pick it up and run with it, but if there's only a small number around, it won't get done."

P4 "With the staff turnover, we need to ensure that there's probably a robust system for training."

Although different suggestions were made for training, the general theme was that staff felt more training, coaching, and supervision would be necessary to help them feel confident in presenting and implementing the intervention.

Subtheme: Staff Confidence (intervention coherence)

Emerging themes related to staff ability and confidence in implementing the intervention were identified, with several participants mentioning confidence as a barrier to presenting the intervention. Some participants suggested that professional training or qualifications were necessary to implement the intervention. There was a sense that support workers were not traditionally expected to implement clinical interventions, which may reflect underlying team dynamics in how interventions are designed and delivered, where knowledge and responsibility flow from qualified professionals downward.

P2: "Not got the sort of qualifications and the sort of the training … So maybe psychology OT might be the better people delivering."

F1 P4: "Some of our staff might be … intimidated doing that or not feel confident enough."

Theme 4: Inpatient Setting

Unique aspects of a rehabilitation setting were highlighted, including the stage of recovery and severity of illness, with some participants reflecting that service users might be more likely to participate in the intervention towards the end of their recovery journey. However, at that point, they can be overwhelmed by the demands of multiple professionals. Furthermore, some participants highlighted barriers such as patient leave and incidents on the ward as unique aspects of this setting.

P1 "Particularly rehab, … service users have lots to work on. It's a longer admission."

P5 "I think it's specifically helpful for rehab as well … it helps build relationships"

P4 "Implement it at the right time in someone's care … We almost jump on that because we're so enthusiastic that someone's doing well and then suddenly … the person's then got too many demands and too many things going on."

Subtheme: Patient Characteristics

Most discussions highlighted patient characteristics such as fluctuating mental health, a history of services, and types of diagnoses (psychosis, paranoid thoughts, autism), and individual differences among service users may influence their likelihood to participate.

P1 "*The type of patient that you would find in a rehab setting would probably pose difficulties with engaging in studies more than most."*

P3 "Schizophrenia is the biggest group of patients … of which some of them have negative symptoms, which means they're not as motivated to carry out the basic activities of daily life, never mind doing DIALOG+."

Subtheme: Content Validity

There were several discussions about the content of the scale and its application in an inpatient setting. Participants highlighted job situations, medication, and family as areas that were difficult to discuss with inpatients. Accommodation was difficult to rate because patients had both home and hospital, and they were confused about which to rate. It felt clear that further scale content validity was required in an inpatient setting. Some participants questioned whether the scale would be culturally relevant to certain service users in the setting.

F2P3 "Like this job one … they've never had a job cos they're just not well enough to work."

P2 "Personal safety … paranoid patients … thinks someone is going to kill his family and … thinks his family are still out to get him … could be very triggering."

F2P1: "People were saying what's the difference really between 'fairly', 'very', and 'totally'."

Discussion

The study aimed to understand the feasibility and acceptability of implementing DIALOG and DIALOG+ in a psychiatric rehabilitation inpatient setting. Findings suggest that patient eligibility, staff recruitment, and staff and patient participation rates met criteria for progression, demonstrating the feasibility of these aspects of the study. However, the patient recruitment, engagement, completion and drop-out rates did not fulfil the requirements for the progression criteria, indicating that these elements of the study are not feasible without revisions. The follow-up rate for staff and patient participants, as well as the recruitment rate for the DIALOG+ group, showed that the 'stop' requirements were met, suggesting potential challenges to feasibility within this setting.

Four themes related to feasibility and acceptability were identified from focus groups and interviews with staff, including*: 'affective attitude towards feasibility objectives; affective attitude towards the intervention; alignment with current practice; and the inpatient setting',* providing an understanding of acceptability in this setting. The main barrier to feasibility appeared to be recruitment, which links to all the qualitative themes. Patient feedback indicated promising interest regarding acceptability, with responding patients recognising the benefits of DIALOG+ and finding the DIALOG scale easy to complete.

Findings suggested that *'affective attitudes'* impacted recruitment and motivation to engage. For instance, staff reported that they found engagement and motivation particularly challenging in their day-to-day roles. This issue was not unique to DIALOG and DIALOG+ but was observed across all demands presented to the patient population. Staff perceived the intervention as demanding and time-consuming, noting that the interaction between patient characteristics, motivation, and engagement acted as a barrier to participation. They discussed relying on their clinical judgment when recruiting participants for the trial and when presenting the intervention to them; consequently, service users were given fewer

opportunities to participate in the trial. Previous research has highlighted that clinician judgment can be biased, depending on the assumptions and expectations held by the clinician (Thirsk et al., 2022). It is therefore essential for future studies to understand the potential impact of this by employing more objective approaches, such as documenting the reasons for approaching patients or not and exploring patient versus staff perspectives.

Results indicated that staff may not have fully *"understood the research and processes"* involved. Learning theory indicates that there are many ways in which the learner acquires knowledge and skills (Gandhi, 2023). One example is social learning theory (Bandura, 1977), which emphasises the importance of observing interventions being implemented and modelling the skills involved, such as through coaching.

The findings suggest that '*aligning with current practice'* through ongoing training, coaching, supervision, and audits are necessary when implementing a new intervention in this setting to ensure staff feel confident and equipped. This is in line with previous reviews that highlighted understanding of the intervention as a contributing factor influencing motivation to engage (Raphael et al., 2021), and that staff skills, abilities, confidence, understanding, awareness, role identity, and attitude all play a role in the success of implementation (Geerligs et al., 2018). Findings suggest that having a lead based at the research site could help bridge the gap between research and implementation. Additionally, staff felt that dedicated time was necessary to implement the intervention. This is supported by previous research, which highlights a lack of dedicated therapy time as a barrier to implementing psychological interventions (Ince et al., 2015).

DIALOG+ is designed to be implemented via an app. However, due to limitations within the NHS and concerns about data security, we were unable to use the app. Staff highlighted that the booklet and the element of written work contributed to the '*perceived effort'* of the intervention, and previous research has found that a lack of appropriate

technology can pose a barrier to implementing outcome measures in mental health services (Sharples et al., 2016).

Bronfenbrenner's ecological systems theory (Bronfenbrenner, 2005) proposes that individuals are shaped by interconnected environmental systems, ranging from the immediate context to broader societal influences. This framework can be applied to how both staff and service users behave within an inpatient setting. A system influenced by NHS governance, media narratives, funding structures, job roles, management, and MDT. This suggests that when implementing a new intervention in this context, it is important to consider the broader ecological system and how the various layers may facilitate or hinder implementation. Furthermore, this may include individual difficulties experienced with engagement, which may reflect broader systemic pressures.

Considering this theory, the findings revealed specific insights related to the '*inpatient setting*' in which the research was conducted and the team systems involved. For example, although staff and patients valued the integration of DIALOG and DIALOG+ within MDT, this did not occur during the trial. It was reported that MDT meetings can feel overwhelming due to insufficient time to cover all the necessary topics. Utilising existing systems to support the implementation of DIALOG and DIALOG+ could increase the likelihood of success. In this context, elements such as designated nurses, individualised one-on-one sessions, and mini teams represent unique aspects of the team system that, if engaged in the intervention, could help shape and enhance its success. These may help the intervention to become more embedded within other processes, such as MDTs. These insights are further supported by a previous systematic review, which highlighted the significance of the system, environmental context, culture, and communication processes in influencing the successful implementation of new interventions in hospital settings (Geerligs et al., 2018).

The *'inpatient setting'* was highlighted as a potential barrier to feasibility. Previous research has highlighted that mental health challenges, such as experiencing psychosis, may complicate the implementation of a new intervention, but they should not act as a barrier (Millier et al., 2020). Conversely, a systematic review conducted by Ince et al. (2015) emphasised that service users who are overly medicated or symptomatic are contributing factors that hinder the successful implementation of psychological interventions for schizophrenia. Staff reported that patient mental health and the inpatient setting serve as barriers to the implementation of DIALOG+ in this context. However, it is worth noting that there was also some positive feedback from patients about the intervention, so this should be explored on a case-by-case basis.

Content validity is considered an essential aspect of developing outcome measurement tools (Mokkink et al., 2024). However, DIALOG has not been tested for content validity in an inpatient setting. This became clear in the study's results, as questions about accommodation were confusing when patients had home leave and weren't sure which accommodation to rate. Staff also reported that some of the questions seemed redundant and duplicated work that is already being conducted. Questions about family and job situation might seem controversial to some patients. However, patient participants who provided feedback felt that these questions were important to them. Although the content validity of the scale is unclear, the study highlights a gap in the literature, and further studies validating the scale in an inpatient setting are warranted.

In summary, more rigorous staff training, coaching, and supervision are needed. Furthermore, all team systems must be aligned and in agreement with the intervention to ensure effective communication and engagement. Finally, methods of presenting tasks to service users should be considered, as the results indicated that staff and patients may be fatigued from repeated demands with insufficient perceived benefits.

The secondary exploratory aims of the study were not met. However, the aim of the study was not to determine the efficacy of the intervention in this setting; the intention was to gather an indication of effectiveness (Thabane et al., 2010). Although a linear regression would not have been feasible during this trial, the author wished to highlight the methods that would be recommended for an RCT.

Strengths and Limitations

The sample, although diverse in some respects, was based on a single inpatient setting in South Yorkshire. Therefore, findings from the study may not be generalisable to other inpatient settings in different locations. Limited qualitative feedback was received from patient participants, which may reflect the need to use alternative feedback formats, such as group discussions or interviews. These methods may have provided a richer and deeper insight into participants' experiences; however, surveys are often used when there is little or no previous research about the topic (Palinkas, 2014). Those who completed the feedback found it challenging to understand the question about the acceptability of the DIALOG. While PPI was used for the study materials, including surveys in the booklets, increased and targeted involvement in adapting these questions may benefit future feasibility studies.

Focus groups were used with support staff and nurses who participated in the trial, while interviews were held with the senior leadership team (SLT). Using a focus group for the SLT could have allowed for the observation of potential team dynamics. However, due to the challenges in availability of the SLT, interviews were chosen for practical reasons.

A progression criterion was used based on Young et al. (2019). However, literature suggests that establishing progression can be complicated, variable, and somewhat subjective (Avery et al., 2017); this poses a limitation to such criteria. Furthermore, drop-out rates were not predefined in the study protocol, as there was an overlap between follow-up rates and drop-out. Predefined criteria from the literature were used to minimise bias.

The author adopted a critical realist position, recognising that the criteria for feasibility and acceptability would inform the themes identified. Consequently, the themes aligned with both the research question and these criteria. This may have resulted in some more nuanced findings within the data being missed, potentially introducing bias to the results (Byrne, 2021). However, the subthemes emphasised more nuanced data, providing additional insights generated from the analysis.

The study was time-limited due to the predetermined hand-in date associated with a Doctoral project. However, this may have limited the study's progression. It was clear from the beginning of the trial that recruitment was difficult; having more time would have allowed the research team to reformulate its approach, including additional training days and more staff recruitment and coaching.

The study used comprehensive and predefined feasibility, acceptability and progression criteria, minimising bias regarding feasibility outcomes. Furthermore, multiple staff members received training to implement DIALOG and DIALOG+, as well as GCP training, which offered benefits to the ward, including staff development opportunities and readiness for research projects. Literature suggests that services that engage in research have better outcomes (Boaz et al., 2015).

The mixed-methods design enabled an overall breadth and depth of the trial to be explored (Aschbrenner et al., 2022). Results were analysed using thematic analysis, which is theoretically flexible (Clarke & Braun, 2014) and can be easily translated for a broad audience. Furthermore, perspectives from both staff and patients were sought, enhancing the breadth of information gathered. Moreover, the feasibility and acceptability criteria interact, with both elements contributing to the study's results.

Conclusions

If the trial's feasibility criteria were mostly ‘stop’, no further research should be conducted (Thabane et al., 2010). This study identified areas that met the progression criteria. With some modifications based on the feedback collected in this study, DIALOG and DIALOG+ may be feasible in this setting. It is recommended that further feasibility trials be conducted, incorporating the recommended adaptations, before undertaking RCTs to test efficacy.

References

Arkes, J. (2023). *Regression analysis: A practical introduction* (2nd ed.). Routledge. https://doi.org/10.4324/9781003285007

Aschbrenner, K. A., Kruse, G., Gallo, J. J., & Plano Clark, V. L. (2022). Applying mixed methods to pilot feasibility studies to inform intervention trials. *Pilot and Feasibility Studies*, *8*(1). https://doi.org/10.1186/s40814-022-01178-x

Avery, K. N., Williamson, P. R., Gamble, C., O'Connell Francischetto, E., Metcalfe, C., Davidson, P., Williams, H., & Blazeby, J. M. (2017). Informing efficient randomised controlled trials: Exploration of challenges in developing progression criteria for Internal Pilot Studies. *BMJ Open*, *7*(2). https://doi.org/10.1136/bmjopen-2016-013537

Ball, H., Yung, A., & Bucci, S. (2022). Staff perspectives on the barriers and facilitators to exercise implementation in inpatient mental health services: A qualitative study. *Mental Health and Physical Activity*, *22*(10), 0452. https://doi.org/10.1016/j.mhpa.2022.100452

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295x.84.2.191

Berry, K., Haddock, G., Kellett, S., Roberts, C., Drake, R., & Barrowclough, C. (2016). Feasibility of a ward-based psychological intervention to improve staff and patient relationships in psychiatric rehabilitation settings. *British Journal of Clinical Psychology*, *55*(3), 236-252. https://doi.org/10.1111/bjc.12082

Bhaskar, R. (2013). *A realist theory of science*. Routledge.

Billingham, S. A., Whitehead, A. L., & Julious, S. A. (2013). An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the

United Kingdom Clinical Research Network Database. *BMC Medical Research Methodology*, *13*(1). https://doi.org/10.1186/1471-2288-13-104

Bingham, A. (2020, October 1). *Structuring a qualitative findings section*. Andrea J. Bingham, Ph.D. https://www.andreajbingham.com/resources-tips-and-tricks/structuring-a-qualitative-findings-section

Blatch-Jones, A. J., Pek, W., Kirkpatrick, E., & Ashton-Key, M. (2018). Role of feasibility and pilot studies in Randomised Controlled Trials: A cross-sectional study. *BMJ Open*, *8*(9). https://doi.org/10.1136/bmjopen-2018-022233

Boaz, A., Hanney, S., Jones, T., & Soper, B. (2015). Does the engagement of clinicians and organisations in research improve healthcare performance: A three-stage review. *BMJ Open*, *5*(12). https://doi.org/10.1136/bmjopen-2015-009415

Booker, C., Rahman-Ali, F., & Paget, S. (2020). *Standards for inpatient mental health rehabilitation*. https://www.rcpsych.ac.uk/docs/default-source/improving-care/ccqi/quality-networks/rehabilitation-wards-aims-rehab/aims-rehab-4th-edition-inpatient-standards-publishable-document.pdf?sfvrsn=da6249fc_2

Boyce, M. B., & Browne, J. P. (2013). Does providing feedback on patient-reported outcomes to healthcare professionals result in better outcomes for patients? A systematic review. *Quality of Life Research*, *22*(9), 2265–2278. https://doi.org/10.1007/s11136-013-0390-0

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2022a). *Successful qualitative research: A practical guide for beginners*. SAGE.

Braun, V., & Clarke, V. (2022b). Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a *knowing* researcher. *International Journal of Transgender Health*, *24*(1), 1–6. https://doi.org/10.1080/26895269.2022.2129597

Braun, V., & Clarke, V. (2024). A critical review of the reporting of reflexive thematic analysis in health promotion international. *Health Promotion International*, *39*(3). https://doi.org/10.1093/heapro/daae049

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. https://doi.org/10.1080/2159676x.2019.1628806

Breakwell, G. M., Barnett, J., & Wright, D. B. (2020). Research methods in psychology. SAGE 3rd ed.

Bronfenbrenner, U. (Ed.). (2005). *Making human beings human: Bioecological perspectives on human development.* Sage Publications Ltd.

Browne R. H. (1995). On the use of a pilot sample for sample size determination. *Statistics in medicine*, *14*(17), 1933–1940. https://doi.org/10.1002/sim.4780141709

Bugge, C., Williams, B., Hagen, S., Logan, J., Glazener, C., Pringle, S., & Sinclair, L. (2013). A process for Decision-making after Pilot and feasibility Trials (ADePT): development following a feasibility study of a complex intervention for pelvic organ prolapse. *Trials*, *14*(1), 1-13. https://doi.org/10.1186/1745-6215-14-353

Burt, J. (2015). Following the mixed methods trail: Some travel advice. *British Journal of General Practice*, *65*(634), 264–265. https://doi.org/10.3399/bjgp15x685045

Caruana, E. J., Roman, M., Hernández-Sánchez, J., & Solli, P. (2015). Longitudinal studies. *Journal of Thoracic Disease*, *7*(11), E537–E540. https://doi.org/10.3978/j.issn.2072-1439.2015.10.63

Churruca K., Pomare C., Ellis L. A., Long J. C., Henderson S. B., Murphy L. E. D., et al. (2021). Patient-reported outcome measures (PROMs): A review of generic and condition-specific measures and a discussion of trends and issues. *Health Expect. 24* (4), 1015–1024. https://doi.org/10.1111/hex.13254

Clarke, V., & Braun, V. (2013). Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The psychologist*, *26*(2), 120-123.

Clarke, V., & Braun, V. (2014). Thematic analysis. *Encyclopedia of Critical Psychology*, 1947–1952. https://doi.org/10.1007/978-1-4614-5583-7_311

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage Publications.

Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., & Lancaster, G. A. (2016). Consort 2010 statement: Extension to randomised pilot and feasibility trials. *BMJ*, i5239. https://doi.org/10.1136/bmj.i5239

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs-principles and practices. *Health Services Research*, *48*(6 Pt 2), 2134–2156. https://doi.org/10.1111/1475-6773.12117

Fichtenbauer, I., Priebe, S., & Schrank, B. (2019). Die deutsche version von Dialog+ Bei Patientinnen mit psychose – eine pilotstudie. *Psychiatrische Praxis*, *46*(07), 376–380. https://doi.org/10.1055/a-0961-3328

Fryer, T. (2022). A critical realist approach to thematic analysis: Producing causal explanations. *Journal of Critical Realism*, *21*(4), 365–384. https://doi.org/10.1080/14767430.2022.2076776

Gandhi, M. H. (2023, July 17). *Learning theories*. StatPearls [Internet]. https://www.ncbi.nlm.nih.gov/books/NBK562189/

Geerligs, L., Rankin, N. M., Shepherd, H. L., & Butow, P. (2018). Hospital-based interventions: A systematic review of staff-reported barriers and facilitators to Implementation Processes. *Implementation Science*, *13*(1). https://doi.org/10.1186/s13012-018-0726-9

Gelkopf, M., Mazor, Y., & Roe, D. (2021). A systematic review of patient-reported outcome measurement (PROM) and provider assessment in Mental Health: Goals, implementation, setting, measurement characteristics and barriers. *International Journal for Quality in Health Care*, *34*(1), ii13–ii27. https://doi.org/10.1093/intqhc/mzz133

Goldsmith, L. J. (2021). Using Framework Analysis in Applied Qualitative Research. *The Qualitative Report*, *26*(6), 2061-2076. https://doi.org/10.46743/2160-3715/2021.5011

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? *Field Methods*, *18*(1), 59–82. https://doi.org/10.1177/1525822x05279903

Guest, G., MacQueen, K., & Namey, E. (2012). *Applied Thematic Analysis*. https://doi.org/10.4135/9781483384436

Hill, C. E., & Knox, S. (2021). Conceptual Foundations of consensual qualitative research. *Essentials of Consensual Qualitative Research.*, 3–10. https://doi.org/10.1037/0000215-001

Ince, P., Haddock, G., & Tai, S. (2015). A systematic review of the implementation of recommended psychological interventions for schizophrenia: Rates, barriers, and improvement strategies. *Psychology and Psychotherapy: Theory, Research and Practice*, *89*(3), 324–350. https://doi.org/10.1111/papt.12084

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods*, *18*(1), 3–20. https://doi.org/10.1177/1525822x05282260

Jacobsen, P., Peters, E., Robinson, E. J., & Chadwick, P. (2020). Mindfulness-based crisis interventions (MBCI) for psychosis within acute inpatient psychiatric settings; a feasibility randomised controlled trial. *BMC Psychiatry*, *20*, 1-12. https://doi.org/10.1186/s12888-020-02608-x

Julious, S. A. (2005). Sample size of 12 per group rule of thumb for a pilot study. *Pharmaceutical Statistics*, *4*(4), 287–291. https://doi.org/10.1002/pst.185

Kitzinger J. (1995). Qualitative research. Introducing focus groups. *BMJ (Clinical research ed.)*, *311*(7000), 299–302. https://doi.org/10.1136/bmj.311.7000.299

Kook, R., Harel-Shalev, A., & Yuval, F. (2019). Focus groups and the collective construction of meaning: Listening to minority women. *Women's Studies International Forum*, *72*, 87–94. https://doi.org/10.1016/j.wsif.2018.12.004

Krasny-Pacini, A., & Evans, J. (2018). Single-case experimental designs to assess intervention effectiveness in rehabilitation: A practical guide. *Annals of Physical and Rehabilitation Medicine, 61*(3), 164-179. https://doi.org/10.1016/j.rehab.2017.12.002

Lundmark, M., Allerby, K., Gremyr, A., & Andersson, A.-C. (2024). *Patients' Experience of Patient-Reported Outcomes, Continuous Feedback, and a Solution-Focused Approach (Using Dialog+) in Psychosis Care in Sweden*. https://doi.org/10.21203/rs.3.rs-4579887/v1

Matanov, A., McNamee, P., Akther, S., Barber, N., & Bird, V. (2021). Acceptability of a technology-supported and solution-focused intervention (dialog+) for Chronic Depression: Views of service users and clinicians. *BMC Psychiatry*, *21*(1). https://doi.org/10.1186/s12888-021-03256-5

McCarron, M., Lombard-Vance, R., Murphy, E., May, P., Webb, N., Sheaf, G., McCallion, P., Stancliffe, R., Normand, C., Smith, V., & O'Donovan, M. A. (2019). Effect of deinstitutionalisation on quality of life for adults with intellectual disabilities: a systematic review. *BMJ Open*, *9*(4), e025735. https://doi.org/10.1136/bmjopen-2018-025735

Mellor, K., Albury, C., Dutton, S. J., Eldridge, S., & Hopewell, S. (2023). Recommendations for progression criteria during external randomised pilot trial design, conduct, analysis and reporting. *Pilot and Feasibility Studies*, *9*(1). https://doi.org/10.1186/s40814-023-01291-5

Millar, A., Tip, L., Lennon, R., Macinnes, M., Michalska, B., Lawrie, S. M., & Schwannauer, M. (2020). The introduction of mindfulness groups to a psychiatric rehabilitation in-patient setting: A feasibility study. *BMC Psychiatry*, *20*(1). https://doi.org/10.1186/s12888-020-02725-7

Millward, L. J. (1995). Focus groups. *Research methods in psychology*, *2*, 274-292.

Mokkink, L. B., Elsman, E. B. M., & Terwee, C. B. (2024). COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *33*(11), 2929–2939. https://doi.org/10.1007/s11136-024-03761-6

Morgan, D. L. (1997). *Focus groups as qualitative research*. SAGE Publications, Inc., https://doi.org/10.4135/9781412984287

Neuendorf, K. A. (2018). Content analysis and thematic analysis. *Advanced Research Methods for Applied Psychology*, 211–223. https://doi.org/10.4324/9781315517971-21

NVivo (Version 15) (2024). *Lumivero.* www.lumivero.com

Olmos-Vega, F. M., Stalmeijer, R. E., Varpio, L., & Kahlke, R. (2023). A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. *Medical Teacher*, *45*(3), 241-251. https://doi.org/10.1080/0142159X.2022.2057287

Omer, S., Golden, E., & Priebe, S. (2016). Exploring the mechanisms of a patient-centred assessment with a solution focused approach (dialog+) in the community treatment of patients with psychosis: A process evaluation within a cluster-randomised controlled trial. *PLOS ONE*, *11*(2). https://doi.org/10.1371/journal.pone.0148415

Palinkas, L. A. (2014). Qualitative and mixed methods in mental health services and Implementation Research. *Journal of Clinical Child & Adolescent Psychology*, *43*(6), 851–861. https://doi.org/10.1080/15374416.2014.910791

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2013). Purposeful sampling for qualitative data collection and analysis in Mixed Method Implementation Research. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Patterson, J. M., Fay, M., Exley, C., McColl, E., Breckons, M., & Deary, V. (2018). Feasibility and acceptability of combining cognitive behavioural therapy techniques with swallowing therapy in head and neck cancer dysphagia. *BMC Cancer*, *18*, 1-11. https://doi.org/10.1186/s12885-017-3892-2

Pearson, N., Naylor, P.-J., Ashe, M. C., Fernandez, M., Yoong, S. L., & Wolfenden, L. (2020). Guidance for conducting feasibility and pilot studies for implementation trials. *Pilot and Feasibility Studies*, *6*(1). https://doi.org/10.1186/s40814-020-00634-w

Philpot, L. M., Barnes, S. A., Brown, R. M., Austin, J. A., James, C. S., Stanford, R. H., & Ebbert, J. O. (2018). Barriers and Benefits to the Use of Patient-Reported Outcome Measures in Routine Clinical Care: A Qualitative Study. *American Journal of Medical Quality: The Official Journal of the American College of Medical Quality*, *33*(4), 359–364. https://doi.org/10.1177/1062860617745986

Priebe, S., Golden, E., Katevas, K., Healey, P., & McCabe, R. (2020). Dialog+ manual. https://www.elft.nhs.uk/sites/default/files/DIALOG%20Manual%20(2021).pdf

Priebe, S., Golden, E., Kingdon, D., Omer, S., Walsh, S., Katevas, K., McCrone, P., Eldridge, S. & McCabe, R. (2017). Effective patient–clinician interaction to improve treatment outcomes for patients with psychosis: a mixed-methods design. *Programme Grants for Applied Research*, *5*(6). https://doi.org/10.3310/pgfar05060

Priebe, S., Golden, E., McCabe, R., & Reininghaus, U. (2012). Patient-reported outcome data generated in a clinical intervention in Community Mental Health Care - psychometric properties. *BMC Psychiatry*, *12*(1). https://doi.org/10.1186/1471-244x-12-113

Priebe, S., Huxley, P., Knight, S., & Evans, S. (2002). Manchester Short Assessment of Quality of Life. *International Journal of Social Psychiatry*. https://psycnet.apa.org/doi/10.1037/t10432-000

Priebe, S., Kelley, L., Golden, E., McCrone, P., Kingdon, D., Rutterford, C., & McCabe, R. (2013). Effectiveness of Structured Patient-clinician communication with a solution focused approach (Dialog+) in community treatment of patients with psychosis – a cluster randomised controlled trial. *BMC Psychiatry*, *13*(1). https://doi.org/10.1186/1471-244x-13-173

Priebe, S., Kelley, L., Omer, S., Golden, E., Walsh, S., Khanom, H., Kingdon, D., Rutterford, C., McCrone, P., & McCabe, R. (2015). The effectiveness of a patient-centred assessment with a solution-focused approach (dialog+) for patients with psychosis: A pragmatic cluster-randomised controlled trial in community care. *Psychotherapy and Psychosomatics*, *84*(5), 304–313. https://doi.org/10.1159/000430991

Priebe, S., McCabe, R., Bullenkamp, J., Hansson, L., Lauber, C., Martinez-Leal, R., Rössler, W., Salize, H., Svensson, B., Torres-Gonzales, F., Van Den Brink, R., Wiersma, D., &

Wright, D. J. (2007). Structured patient–clinician communication and 1-year outcome in Community Mental Healthcare. *British Journal of Psychiatry*, *191*(5), 420–426. https://doi.org/10.1192/bjp.bp.107.036939

Random.org - true random number service. (2025). https://www.random.org/

Raphael, J., Price, O., Hartley, S., Haddock, G., Bucci, S., & Berry, K. (2021). Overcoming barriers to implementing ward-based psychosocial interventions in acute inpatient mental health settings: A meta-synthesis. *International Journal of Nursing Studies*, *115*(10), 3870. https://doi.org/10.1016/j.ijnurstu.2021.103870

Roe, D., Slade, M., & Jones, N. (2022). The utility of patient-reported outcome measures in Mental Health. *World Psychiatry, 21*(1), 56–57. https://doi.org/10.1002/wps.20924

Sekhon, M., Cartwright, M., & Francis, J. J. (2022). Development of a theory-informed questionnaire to assess the acceptability of healthcare interventions. *BMC Health Services Research*, *22*(1). https://doi.org/10.1186/s12913-022-07577-3

Sekhon, M., Cartwright, M., & Francis, J. J. (2017). Acceptability of healthcare interventions: An overview of reviews and development of a theoretical framework. *BMC Health Services Research*, *17*(1). https://doi.org/10.1186/s12913-017-2031-8

Sharples, E., Qin, C., Goveas, V., Gondek, D., Deighton, J., Wolpert, M., & EdbrookeChilds, J. (2016). A qualitative exploration of attitudes towards the use of outcome measures in child and Adolescent Mental Health Services. *Clinical Child Psychology and Psychiatry*, *22*(2), 219–228. https://doi.org/10.1177/1359104516652929

Shepherd, N., Meehan, T., & Humphries, S. (2014). Supporting recovery: Challenges for inhome psychiatric support workers. *Mental Health Review Journal*, *19*(2), 73–83. https://doi.org/10.1108/mhrj-07-2013-0027

Sim, J., & Lewis, M. (2012). The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. *Journal of Clinical Epidemiology*, *65*(3), 301–308. https://doi.org/10.1016/j.jclinepi.2011.07.011

Simkus, J. (2024). Between-subjects design: Overview & examples. *Simply Psychology.* https://www.simplypsychology.org/between-subjects-design.html

Sutton, J., & Austin, Z. (2015). Qualitative Research: Data Collection, Analysis, and Management. *The Canadian Journal of Hospital Pharmacy*, *68*(3), 226–231. https://doi.org/10.4212/cjhp.v68i3.1456

Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., Robson, R., Thabane, M., Giangregorio, L., & Goldsmith, C. H. (2010). A tutorial on pilot studies: The what, why and how. *BMC Medical Research Methodology*, *10*(1). https://doi.org/10.1186/1471-2288-10-1

The Journal Article Reporting Standards (2025). American Psychological Association. https://apastyle.apa.org/jars

Thirsk, L. M., Panchuk, J. T., Stahlke, S., & Hagtvedt, R. (2022). Cognitive and implicit biases in nurses' judgment and decision-making: A scoping review. *International Journal of Nursing Studies*, *133*(10), 4284. https://doi.org/10.1016/j.ijnurstu.2022.104284

Van Loggerenberg, F., Akena, D., Alinaitwe, R., Birabwa-Oketcho, H., Méndez, C. A. C., Gómez-Restrepo, C., Kulenović, A. D., Selak, N., Kiseljaković, M., Musisi, S., Nakasujja, N., Sewankambo, N. K., & Priebe, S. (2023). Feasibility and outcomes of using DIALOG+ in primary care to improve quality of life and mental distress of patients with long-term physical conditions: an exploratory non-controlled study in Bosnia and Herzegovina, Colombia and Uganda. *BMC Primary Care*, *24*(1), 241. https://doi.org/10.1186/s12875-023-02197-0

Varese, F., Sellwood, W., Aseem, S., Awenat, Y., Bird, L., Bhutani, G., Carter, L., Davies, L., Davis, C., Horne, G., Keane, D., Logie, R., Malkin, D., Potter, F., van den Berg, D., Zia, S., & Bentall, R. (2020). Eye movement desensitization and reprocessing therapy for psychosis (emdrp): Protocol of a feasibility randomized controlled trial with Early Intervention Service Users. *Early Intervention in Psychiatry*, *15*(5), 1224–1233. https://doi.org/10.1111/eip.13071

Watt, D. (2007). On Becoming a Qualitative Researcher: The Value of Reflexivity. *The Qualitative Report*, *12*(1), 82-101. https://doi.org/10.46743/2160-3715/2007.1645

Wong, S. A., Duffy, L. A., Layfield, S. D., Long, M., Rodriguez-Villa, F., Gelda, S. E., Gelwan, E. M., Sugarman, D. E., Busch, A., Eisen, J., Ressler, K. J., & Yip, A. G. (2022). Using electronic patient-reported measures to characterize symptoms and improvement in inpatient psychiatric units. *Psychiatry Research*, *317*(11), 4839.https://doi.org/10.1016/j.psychres.2022.114839

Young, H. M., Goodliffe, S., Madhani, M., Phelps, K., Regen, E., Locke, A., ... & Conroy, S. (2019). Co-producing progression criteria for feasibility studies: a partnership between patient contributors, clinicians, and researchers. *International Journal of*

Environmental Research and Public Health, *16*(19), 3756.

https://doi.org/10.3390/ijerph16193756

Appendix

Appendix A

Ethical Approval Documents

Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

Dr Claire Bone
Senior Clinical Psychologist
SHSC and University of Sheffield
Burbage ward Michael Carlisle Centre
Osborne Road
Sheffield
s11 9Bf

Email: approvals@hra.nhs.uk
HCRW.approvals@wales.nhs.uk

09 July 2024

Dear Dr Bone

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**Implementing the DIALOG Outcome Measure and a Solution-Focused Intervention (DIALOG+) in a Psychiatric Rehabilitation Inpatient Setting: A Mixed Methods Feasibility Study.**
IRAS project ID:	**338942**
Protocol number:	**187799**
REC reference:	**24/SC/0209**
Sponsor	**University of Sheffield**

I am pleased to confirm that **HRA and Health and Care Research Wales (HCRW) Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

Please now work with participating NHS organisations to confirm capacity and capability, in line with the instructions provided in the "Information to support study set up" section towards the end of this letter.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report

(including this letter) have been sent to the coordinating centre of each participating nation. The relevant national coordinating function/s will contact you as appropriate.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?
HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?

The standard conditions document "*After Ethical Review – guidance for sponsors and investigators*", issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **338942**. Please quote this on all correspondence.

Yours sincerely,

List of Documents

The final document set assessed and approved by HRA and HCRW Approval is listed below

Document	*Version*	*Date*
Copies of materials calling attention of potential participants to the research [DIALOG+ poster]	2	24 May 2024
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [certificate of insurance]		24 May 2024
Interview schedules or topic guides for participants [intervention DIALOG+ booklet]	2	16 May 2024
Interview schedules or topic guides for participants [comparison group DIALOG booklet]	2	16 May 2024
Interview schedules or topic guides for participants [Post-intervention survey questions for patient participants]	2	16 May 2024
Interview schedules or topic guides for participants [Interview schedule for Staff participants]	2	16 May 2024
IRAS Application Form [IRAS_Form_04072024]		04 July 2024
Letter from sponsor [Letter from sponsor]		19 May 2024
Letters of invitation to participant [Study poster]	2	16 May 2024
Organisation Information Document [OID from sponsor]	2	04 June 2024
Other [IRAS amendments table]	1	26 June 2024
Participant consent form [consent form with initials]	3	26 June 2024
Participant consent form [Participant Information Sheet version 3]	1	03 July 2024
Participant information sheet (PIS) [Participant Information Sheet quick version 3 read V2.0]	3	25 June 2024
Participant information sheet (PIS) [Participant Information Sheet quick version 3read STAFF]	3	25 June 2024
Participant information sheet (PIS) [participant information sheet staff version 3]	3	25 June 2024
Participant information sheet (PIS) [Participant Information Sheet version 3]	3	25 June 2024
Protocol	4	16 May 2024
Referee's report or other scientific critique report [scientific approval]		19 May 2024
Schedule of Events or SoECAT [SoECAT]	1	04 June 2024
Schedule of Events or SoECAT [SoECAT]	2	21 June 2024
Summary CV for Chief Investigator (CI) [Summary CV Sinead Mclernon]		
Summary CV for Chief Investigator (CI) [Summary CV Claire Bone]		
Summary CV for student [Student cv]		
Summary CV for supervisor (student research) [Supervisor cv]		29 May 2024
Summary of any applicable exclusions to sponsor insurance (non-NHS sponsors only) [liability certificate]		24 May 2024
Summary, synopsis or diagram (flowchart) of protocol in non technical language [Flow chart in sinmple language]		16 May 2024
Validated questionnaire [DIALOG scale]		

IRAS project ID	338942

Information to support study set up

The below provides all parties with information to support the arranging and confirming of capacity and capability with participating NHS organisations in England and Wales. This is intended to be an accurate reflection of the study at the time of issue of this letter.

Types of participating NHS organisation	Expectations related to confirmation of capacity and capability	Agreement to be used	Funding arrangements	Oversight expectations	HR Good Practice Resource Pack expectations
Research activities and procedures as per the protocol and other study documents will take place at participating NHS organisations.	Research activities should not commence at participating NHS organisations in England or Wales prior to their formal confirmation of capacity and capability to deliver the study in accordance with the contracting expectations detailed.	An Organisation Information Document has been submitted and the sponsor is not requesting and does not expect any other agreement to be used with participating NHS organisations of this type.	Study funding arrangements are detailed in the Organisation Information Document	A Principal Investigator should be appointed at participating NHS organisations.	Where an external individual will be conducting any of the research activities that will be undertaken at this site type then they would be expected to hold a Letter of Access. This should be issued be on the basis of a Research Passport (if university employed) or an NHS to NHS confirmation of pre-engagement checks letter (if NHS employed). These should confirm Occupational Health Clearance in the case of direct access.These should confirm standard DBS checks

Other information to aid study set-up and delivery

This details any other information that may be helpful to sponsors and participating NHS organisations in England and Wales in study set-up.
The applicant has indicated that they intend to apply for inclusion on the NIHR CRN Portfolio.

Health Research Authority

South Central - Berkshire Research Ethics Committee
Health Research Authority
2 Redman Place
Stratford
E20 1JQ

Please note: This is the favourable opinion of the REC only and does not allow you to start your study at NHS sites in England until you receive HRA Approval

09 July 2024

Dr Claire Bone
Senior Clinical Psychologist
SHSC and University of Sheffield
Burbage ward Michael Carlisle Centre
Osborne Road
Sheffield
s11 9Bf

Dear Dr Bone

Study title:	**Implementing the DIALOG Outcome Measure and a Solution-Focused Intervention (DIALOG+) in a Psychiatric Rehabilitation Inpatient Setting: A Mixed Methods Feasibility Study.**
REC reference:	**24/SC/0209**
Protocol number:	**187799**
IRAS project ID:	**338942**

Thank you for responding to the Research Ethics Committee's (REC) request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair, together with Committee members Dr Deborah Scholey Dr Michael Yardley.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation

as revised, subject to the conditions specified below.

Good practice principles and responsibilities

The UK Policy Framework for Health and Social Care Research sets out principles of good practice in the management and conduct of health and social care research. It also outlines the responsibilities of individuals and organisations, including those related to the four elements of research transparency:

1. registering research studies
2. reporting results
3. informing participants
4. sharing study data and tissue

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of the study.

Confirmation of Capacity and Capability (in England, Northern Ireland and Wales) or NHS management permission (in Scotland) should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA and HCRW Approval (England and Wales)/ NHS permission for research is available in the Integrated Research Application System.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations

Registration of Clinical Trials

All research should be registered in a publicly accessible database and we expect all researchers, research sponsors and others to meet this fundamental best practice standard.

It is a condition of the REC favourable opinion that **all clinical trials are registered** on a public registry before the first participant is recruited and no later than six weeks after. For this purpose, 'clinical trials' are defined as:

- clinical trial of an investigational medicinal product
- clinical investigation or other study of a medical device
- combined trial of an investigational medicinal product and an investigational medical device
- other clinical trial to study a novel intervention or randomised clinical trial to compare interventions in clinical practice.

A 'public registry' means any registry on the WHO list of primary registries or the ICMJE list of registries provided the registry facilitates public access to information about the UK trial.

Failure to register a clinical trial is a breach of these approval conditions, unless a deferral has been agreed by the HRA (for more information on registration and requesting a deferral see: Research registration and research project identifiers).

Where a deferral is agreed we expect the sponsor to publish a minimal record on a publicly accessible registry. When the deferral period ends, the sponsor should publish the full record on the same registry, to fulfil the condition of the REC favourable opinion.

If you have not already included registration details in your IRAS application form you should notify the REC of the registration details as soon as possible.

Where the study is registered on ClinicalTrials.gov, please inform deferrals@hra.nhs.uk and the Research Ethics Committee (REC) which issued the final ethical opinion so that our records can be updated.

Publication of Your Research Summary

We will publish your research summary for the above study on the research summaries section of our website, together with your contact details, no earlier than three months from the date of this favourable opinion letter. Where a deferral is agreed, a minimum research summary will still be published in the research summaries database. At the end of the deferral period, we will publish the full research summary.

Should you wish to provide a substitute contact point, make a request to defer, or require further information, please visit: Research summaries - Health Research Authority (hra.nhs.uk)

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

After ethical review: Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study, including early termination of the study
- Final report
- Reporting results

The latest guidance on these topics can be found at Managing your approval - Health Research Authority (hra.nhs.uk)

Ethical review of research sites

NHS/HSC sites

The favourable opinion applies to all NHS/HSC sites taking part in the study, subject to confirmation of Capacity and Capability (in England, Northern Ireland and Wales) or management permission (in Scotland) being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Non-NHS/HSC sites

I am pleased to confirm that the favourable opinion applies to any non-NHS/HSC sites listed in the application, subject to site management permission being obtained prior to the start of the study at the site.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Copies of materials calling attention of potential participants to the research [DIALOG+ poster]	2	24 May 2024
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [certificate of insurance]		24 May 2024
Interview schedules or topic guides for participants [intervention DIALOG+ booklet]	2	16 May 2024
Interview schedules or topic guides for participants [comparison group DIALOG booklet]	2	16 May 2024
Interview schedules or topic guides for participants [Post-intervention survey questions for patient participants]	2	16 May 2024
Interview schedules or topic guides for participants [Interview schedule for Staff participants]	2	16 May 2024
IRAS Application Form [IRAS_Form_04072024]		04 July 2024
Letter from sponsor [Letter from sponsor]		19 May 2024
Letters of invitation to participant [Study poster]	2	16 May 2024
Other [IRAS amendments table]	1	26 June 2024
Participant consent form [Consent to participate]	3	26 June 2024
Participant information sheet (PIS) [Staff participant information sheet]	3	25 June 2024
Participant information sheet (PIS) [Staff participant information sheet, quick read]	3	25 June 2024
Participant information sheet (PIS) [patient participant information sheet, quick read]	3	25 June 2024
Participant information sheet (PIS) [patient participant information sheet]	3	25 June 2024
Protocol	4	16 May 2024
Referee's report or other scientific critique report [scientific approval]		19 May 2024
Summary CV for Chief Investigator (CI) [Summary CV Sinead Mclernon]		
Summary CV for Chief Investigator (CI) [Summary CV Claire Bone]		
Summary CV for student [Student cv]		

Summary CV for supervisor (student research) [Supervisor cv]		29 May 2024
Summary of any applicable exclusions to sponsor insurance (non-NHS sponsors only) [liability certificate]		24 May 2024
Summary, synopsis or diagram (flowchart) of protocol in non technical language [Flow chart in sinmple language]		16 May 2024
Validated questionnaire [DIALOG scale]		

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website: Quality assurance - Health Research Authority (hra.nhs.uk)

HRA Learning

We are pleased to welcome researchers and research staff to our HRA Learning Events and online learning opportunities– see details at: Learning - Health Research Authority (hra.nhs.uk)

IRAS project ID: 338942 Please quote this number on all correspondence

With the Committee's best wishes for the success of this project.

Appendix B

Data Management Plan (DMP)

Plan Overview
A Data Management Plan created using DMPonline

Title: Implementing the DIALOG Outcome Measure and a Solution-Focused Intervention (DIALOG+) in a Psychiatric rehabilitation Inpatient Setting: A Mixed Methods Feasibility Study.

Creator: Sinead Mclernon

Affiliation: The University of Sheffield

Template: Postgraduate Research DMP (The University of Sheffield)

Project abstract:
A randomised feasibility trial will be conducted, using a mixed-methods between-subject longitudinal design to study the feasibility and acceptability of implementing the DIALOG+ intervention in an NHS psychiatric inpatient setting. The DIALOG+ solution-focused intervention will be introduced alongside the DIALOG measure in the treatment phase. Scores on the DIALOG will be examined pre-and post-intervention, and qualitative interviews and focus groups will be conducted with staff to generate themes relating to acceptability and feasibility. A survey will also be conducted with patients following completion of the intervention or after dropout with consent.

ID: 140413

Start date: 01-05-2024

End date: 31-07-2025

Last modified: 06-04-2024
Implementing the DIALOG Outcome Measure and a Solution-Focused Intervention (DIALOG+) in a Psychiatric rehabilitation Inpatient Setting: A Mixed Methods Feasibility Study.
Defining your data

What digital data (and physical data if applicable) will you collect or create during the project?
How will the data be collected or created, and over what time period?
What formats will your digital data be in? (E.g. .doc, .txt, .jpeg)
Approximately how much digital data (in GB, MB, etc) will be generated during the project?
Are you using pre-existing datasets? Give details if possible, including conditions of use.
What Type of digital data?
DIALOG outcome measures will be collected for patients who consent to participate. This will be collected and stored on the patient record. These data can then be extracted and

anonymised for research purposes. Anonymised data will be collected from patient records and stored on the University of Sheffield's secure servers as an Excel file. These data will be collected over approximately 5 months.
Audio recordings will be taken of focus groups and interviews involving staff members and stored on the University of Sheffield secure server, X:Drive. Staff participating will be made aware of the recordings and prompted to keep confidentiality during focus groups and interviews. Qualitative data will be collected over 3 months. Audio recordings will be made using an encrypted recording device provided by the University of Sheffield. This data will then be taken to the transcribers to transcribe, according to the University of Sheffield transcribing guidelines. This will then be emailed back to me using my university email address and a separate email containing the password. Further MS word documents (doc) will be created to analyse and code the data. This will be saved on the X:Drive.
Anonymised written survey data will be collected from patients who agree to participate and scanned and stored on the University of Sheffield secure server as a doc.

How will the data be collected?
A data-sharing agreement should be made between (redacted trust name) and The University of Sheffield.
Anonymised data collected from patient records will be exported and stored on the University of Sheffield's secure servers as an Excel file when it is ready to be analysed. Participants will be assigned a unique code/number which can be used to track their questionnaire data, and these numbers will be documented in a password-protected file, stored on the University of Sheffield secure servers, on the X:Drive.
How much data and in what format?
Approximately 2 GB of data will be stored in the University's secure server. Data will be stored as an excel file on the X: Drive in the secure servers at the University of Sheffield.
Audio recordings will be saved as jpeg.
Analysed transcriptions will be saved as a doc.
Extracted patient participant data will be saved as excel.
survey data will be save as a doc.

Looking after data during your research

Where will you store digital data during the project to ensure it is secure and backed up regularly? (E.g. University research storage, or University Google drive)
How will you name and organise your data files? (An example filename can help to illustrate this)
If you collect or create physical data, where will you store these securely?
How will you make data easier to understand and use? (E.g. include file structure and methodology in a README file)
Will you use extra security precautions for any of your digital or physical data? (E.g. for sensitive and/or personal data)

Anonymised data collected from patient records will be exported and stored on University of Sheffield's secure servers. Participants will be assigned a unique code/number which can be used to track their questionnaire data, and these numbers will be documented in a password-protected file, stored on the University of Sheffield secure servers.
Consent forms will be scanned and stored securely on the University of Sheffield secure servers and once scanned the paper copy will be stored in a locked cabinet behind a locked

door within the redacted inpatient wards. Electronic copies can be deleted after 5 years as per GDPR compliance (Clarke et al., 2019). Paper consent forms can be destroyed at the end of the study, my supervisor and myself will be responsible for destroying the paper consent forms.
Participant's unique codes and consent forms will be saved in separate encrypted folders on the X:Drive on the University of Sheffield secure folders.
A README file will be created to keep track of the location and naming of files.

Storing data after your research

Which parts of your data will be stored on a long-term basis after the end of the project?
Where will the data be stored after the project? (E.g. University of Sheffield repository ORDA, or a subject-specific repository)
How long will the data be stored for? (E.g. standard TUoS retention period of minimum 10 years after the project)
Who will place the data in a repository or other long-term storage? (E.g. you, or your supervisor)
If you plan to use long-term data storage other than a repository, who will be responsible for the data?

The audio recordings will be saved for 10 months after the study completion and then destroyed. The data used for analysis, transcriptions and associated versions of this which include coding and analysis (MS document (.doc)) will be saved on a long-term basis (10 years) on the University of Sheffield repository ORDA. Myself or my supervisor will place this data in ORDA. All transcriptions will be checked for secondary identifiers.
Archiving of data will be included in the participant information sheet and consent form.

Sharing data after your research

How will you make data available outside of the research group after the project? (E.g. openly available through a repository, or on request through your department)
Will you make all of your data available, or are there reasons you can't do this? (E.g. personal data, commercial or legal restrictions, very large datasets)
If there are reasons you can't share all of your data, how might you make as much of it available as possible? (E.g. anonymisation, participant consent, sharing analysed data only)
How will you make your data as widely accessible as possible? (E.g. include a data availability statement in publications, ensure published data has a DOI)
What licence will you apply to your data to say how it can be reused and shared? (E.g. one of the Creative Commons licences)

Data may be made openly available through the data repository at the University of Sheffield depending on the data sharing agreement with the (redacted trust name)
Participants will be asked to consent to data sharing:

"I understand and agree that other authorised researchers may use my data in publications, reports, web pages and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form."
Data sharing will be available depending on consent.
All data will be fully anonymised and all identifiable information will be removed before researchers gain access.
Both myself and supervisor will be able to give access to data if needed.
The trial will be written for publication and shared with (redacted trust name). Any reports will be secure and stored in (redacted trust name) or the University of Sheffield secure drive.
This study aims to generate evidence to improve the quality of care.
Data sharing will be available depending on consent. Data may be made openly available through the data repository at the University of Sheffield depending on the data sharing agreement with the redacted.

Putting your plan into practice

Who is responsible for making sure your data management plan is followed? (E.g. you with the support of your supervisor)
How often will your data management plan be reviewed and updated? (E.g. yearly and if the project changes)
Are there any actions you need to take in order to put your data management plan into practice? (E.g. requesting University research storage via your supervisor.)

My supervisor, Dr Claire Bone, and I will be responsible for ensuring the data management plan is followed, along with support from my supervisor.
The plan will be reviewed for any changes or updates to the project, as well as at least yearly.

Appendix C

NHS Data Sharing Agreement

Data Sharing Agreement

Implementing the DIALOG Outcome Measure and a Solution-Focused Intervention (DIALOG+) in an Acute Psychiatric Inpatient Setting: A Mixed Methods Feasibility Study.
This data-sharing agreement is made

1. **Between:**
Redacted Trust
And
The University of Sheffield

2. **Purpose and objectives of the information sharing**

Be clear and concise about the reasons for data sharing. Give as detailed a description as possible. You should set out what objective you are hoping to achieve by sharing personal data between organisations. You can number each purpose separately.

This study is being conducted by Sinead McLernon (Clinical Psychologist in Training) as part of the qualification towards becoming a Doctor of Clinical Psychology at the University of Sheffield. She is being supervised by Dr Claire Bone and Professor Jaime Delgadillo, who are also based at the University of Sheffield. The research is being carried out in collaboration with the NHS, specifically the Redacted NHS Foundation Trust. The purpose of this data collection is to inform the research question of whether DIALOG and DIALOG+ are acceptable and feasible in a psychiatric inpatient setting.

3 **Controllers**
List here all organisations which are controllers and for which purposes, as part of this agreement
Redacted

4. **Processor**s
List here all organisations acting as processors and sub-processors as part of the agreement.
Redacted

5. **Data items to be processed**

List here all the data items you plan to process.

Descriptive statistics including age, gender, ethnicity, history of mental illness or previous inpatient admissions, drug-use, diagnosis, baseline severity on DIALOG, length of stay.

DIALOG outcome measures will be collected for patients who consent to participate. This will be collected and stored on the patient record. These data can then be extracted and anonymised for research purposes. An Excel file will be generated with predetermined headings as a template, participants will be given an ID number and data will be extracted onto this Excel document and stored within the trust secure server initially until data is ready to be analysed. Anonymised data will then be password protected prior to transferring to the research team, to be stored on the University of Sheffield's secure server. These data will be collected over approximately 5 months.

Audio recordings will be taken of focus groups and interviews involving staff members and stored on the University of Sheffield secure server. Staff participating will be made aware of the recordings and prompted to keep confidentiality during focus groups and interviews. Qualitative data will be collected over 3 months.

Anonymised Survey data will be collected from patients who agree to participate and scanned and stored on the University of Sheffield secure server as a doc. Paper copies will be stored in a locked filing cabinet behind a locked door on the ward.

6. **Article 6 Condition: Personal data**

State which Article 6 condition, (legal basis) is met. One from the list below must apply whenever you process personal data.

Participants will be informed of article 6 on a participant leaflet:

"As we will be collecting some data that is defined in the legislation as more sensitive (i.e. information about your ethnic origin and health), we also need to let you know that we are applying the following condition in law: that the use of your data is 'necessary for scientific or historical research purposes'."

The data collected aims to further understand the use of DIALOG and DIALOG+ in an inpatient acute psychiatric setting. The benefits of using DIALOG and DIALOG+ have already been assessed in community settings but not yet within inpatient settings. The NHS trust aims to trial the use of DIALOG across inpatient and community settings. It is therefore important that we understand whether DIALOG and DIALOG+ are beneficial and acceptable to use for patients and staff in these settings.

7. Article 9 condition: Special categories of personal data

N/A

8. Individual rights and preferences

Explain how individual rights and preferences will be managed by the parties to this agreement.

For each below, say how the right will be managed or why it is not applicable.

The right to be informed: Informed consent will be sought at the beginning of intervention and participants will have the opportunity to leave the trial at any point up until the data is being analysed. Capacity to consent will be informed by questions with the person and as an MDT, to check they understand and can weigh up the pros and cons of what they're consenting to.

The right of access: Participants will have access to previous DIALOG scores if needed.

The right to rectification: Rectification will be possible at the time of the weekly measure or shortly after. However, participants won't be able to rectify the previous week's score as this is an ongoing dynamic data collection.

The right to erasure: Confidentiality and record keeping will be discussed with participants and it will be documented on the patient record. Participants can withdraw their data from the study up until the date that it is combined for analysis, at which point it will no longer be identifiable; this date will be made clear in consent forms.

The right to restrict processing: Participants have the right to withdraw from the trial at any time up until data has been extracted and combined for analysis.

The right to portability: Participants can have copies of their DIALOG outcome measures. Participants maintain the right to request their DIALOG data and provide it to another controller (service) if that is their wish.

The right to object: Participants have the right to withdraw from the study, there are channels to discuss any concerns with clinicians on the ward and any concerns can also be forwarded to the research team or escalated within the trust if needed.

Rights in relation to automated decision-making profiling: The DIALOG app will automate some things such as summing up total scores, however, automated decision making will not be conducted as part of this study. The app sums scores but no automated decision is made based on the scores.

Please state here how you will manage any complaints raised regarding the proposed data sharing:

Participants can raise complaints with clinicians on the ward who can put them in touch with the chief investigator at the University via phone call. There will be a point of contact for the study on each ward.

Does the National Data Opt-out apply to proposed purposes for data sharing?

No, individual consent will be taken from all participating patients.

9. Compliance with duty of confidentiality or right to privacy

Please state here how you will be satisfying the duty of confidentiality. This must be in addition to how you have already explained meeting data protection requirements to process personal data.

A data-sharing agreement should be made between (redacted trust name) and The University of Sheffield. Anonymised data collected from patient records will be exported and stored on the University of Sheffield's secure servers.

Participants will be assigned a unique code/number which can be used to track their questionnaire data, and these numbers will be documented in a password-protected file, stored on the University of Sheffield secure servers.

As per trust policy for any clinical information obtained whilst patients are under the care of redacted their responses will be kept confidential and will not be shared with anyone who is not directly involved in their care.

Please provide an explanation if you are relying on statutory gateway. Describe which you are using and confirm whether it sets aside the common law duty of confidentiality. Anonymised data is being shared with the University of Sheffield's chief research investigator on the project so the data can be analysed for scientific research purposes. This will only happen where the patient has given their express informed consent for this to happen.

Is there any interference with Human Rights Article 8?

NO

If yes, describe and explain why it is necessary to interfere with human rights and why it is proportionate to do so.

10. Transparency

Describe here how you will communicate with the public. For example, update privacy notice, patient information leaflets, posters, information on websites.
Consent forms, posters and patient leaflets will be used and created with patients' involvement.

Participants will be asked to consent to data sharing:
"I understand and agree that other authorised researchers may use my data in publications, reports, web pages and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form."

"I understand and agree that other authorised researchers will have access to this data only if they preserve the confidentiality of the information as requested in this form."
Data sharing will depend on informed consent. Participants will be assigned a unique code/number alongside their initials whilst data collection takes place on the wards. Participants' initials will be kept alongside the random participant ID so that they can withdraw data up until it is transferred to the research team. Prior to data transfer, the initials will be removed to preserve confidentiality. The lead researcher will not have access to any clinically identifiable information and does not work on the wards, therefore it is only clinicians who are involved in the participants' care who could identify them based on demographic data, and they would have access to this information regardless as part of routine care.

11. How will the data sharing be carried out?

Describe the mechanism by which you will share data. Give an explanation of why this is secure and which organisation is responsible for ensuring security.

Data sharing will be carried out using a secure NHS.net email account. The excel file containing data collected will be stored securely in within the redacted servers until access is required for data analysis. The excel file will be password protected prior to be transferred and the password only provided at an agreed time for the transfer to take place. The files will be uploaded immediately to the secure university server and the email deleted (and recycle bin emptied). All of this will be confirmed live at the agreed time.

Describe how you will share any outputs or analysis and explain why this is secure, necessary and proportionate.

The trial findings will be written up for publication and shared with redacted. This study aims to generate evidence to improve the quality of care.

Data sharing will be available depending on consent. Data may be made openly available through the data repository at the University of Sheffield depending on the data sharing agreement with the redacted

Describe the frequency you expect to share data including the security precautions you will use proportionate to the level of frequency.

Data will be shared from redacted to the university research team once at the end of Phase one and then again at the end of Phase two. Data sharing security as described above will be observed at each phase.

Paper signed consent forms will be stored in a locked cabinet in the nurses office and scanned onto the clinical notes record. These forms will have patient names and date that it was signed, which will allow them to be located on insight. Participants in each phase of the study will be identified by date of consent.

Explain whether any information is being transferred outside the EU and, if so, relevant safeguards (this is to ensure compliance with Article 45 of the GDPR).

No

12. Accuracy of the data being shared

Describe the processes or procedures that you will follow to make sure that the data you hold and share is accurate. Explain how you will share any updates with all recipients of the data.

Regular supervision will occur within the research team and when needed there will be opportunities to consult with senior academics at the University of Sheffield. Clinicians on the ward will be offered training specifically related to carrying out the research project as well as completing Essentials of Good Clinical Practice training prior to joining the project. Updates to the data governance teams within (redacted trust name) will be offered.

13. **Rectification of data that has been shared**

Describe the procedures that are in place or will be put in place, for the parties to this agreement to rectify:

inaccurate data that has been shared

data that has been identified as inaccurate after sharing

This is separate to the individual's right to rectification.

DIALOG scores will be saved to the patient record and can be rectified shortly after completion if the patient request this. However as weekly scores are designed to provide a snapshot of how the person is feeling at that moment, it will not be possible to change these scores retrospectively (e.g. after a day has passed). It is unlikely that there will be any need to change these scores, as this will be an in-person conversation between clinician and patient at the time the scores have been taken.

The excel file can be amended up until the point in which the data is being analysed, retrospectively identifying participant information through the clinician on the ward. This won't be possible after the data has been amalgamated.

A consultation with the data governance team will be made before data extraction.

14. Retention and disposal requirements for the information to be shared

Consent forms will be scanned and stored securely on the University of Sheffield secure servers paper copies of consent forms will be kept in a locked cabinet in the nurses offices on the ward and scanned onto the clinical notes system. Electronic copies can be deleted after 5 years as per GDPR compliance (Clarke et al., 2019).

Anonymised data collected from patient records will be exported and stored on University of Sheffield's secure servers. Participants will be assigned a unique code/number which can be used to track their questionnaire data, and these numbers will be documented in a password-protected file, stored on the University of Sheffield secure servers for up to 10 years.

15. Breach management

Outline the process for how any breach of data security or confidentiality will be managed by relevant parties.

In the event of a data security breach, the research and clinical leadership team will meet to take immediate action to prevent any further breaches and reduce the potential for harm as a result of existing breaches. In line with the duty of candour, the clinical team will meet to support any affected participants, to explain the circumstances and potential impact of the breach, and the actions taken to rectify it. The participants will be reminded of how to make a complaint should they wish to do so. A follow-up meeting will be arranged shortly after between the clinical team and the university for a post-incident review and to put any further

precautions in place. An incident form will be completed at the time the breach is noted for all participants who have been affected and a note made on their clinical record, including details of any conversations had with the participant and the follow-up support and actions offered.

16. State what each organisation will be responsible for by the terms of this agreement For example, organisation A will deal with complaints, organisation B will handle risk.

The University of Sheffield will act as the Data Controller for this study. This means that the University is responsible for looking after your information and using it properly. The redacted will be responsible for keeping paper consent forms and data safe on the premises. There will be an allocated clinician responsible for this on the ward.
(redacted trust name) will initially manage any complaints about the study via the field supervisor, and may involve the university if necessary or at the participant's request. Any risk that is identified as part of collecting the DIALOG measure will be managed by the participant's clinical team (redacted trust name) as part of routine practice.
Any risks that arise in terms of data breaches will be managed by both organisations as outlined above.

17. Contacts – Information Governance (IG) and Caldicott Guardian
List here the IG contacts for each organisation.
The University of Sheffield

18. Start date of the agreement

19. Review of the agreement
Agreement will be reviewed if any changes are made to the project and will be conducted by redacted

20. Review period
Three monthly (or at each data collection phase).

21. Variation
State here if the any party can vary the terms of this agreement. Describe how this will be done.

22. Ending the agreement
State how a party can end their participation in the agreement, and how data will be managed by the exiting party.

23. End date
State the date the agreement ends. The project end date is anticipated to be May 2025

Appendix D

CONSORT Checklist

REDACTED

Appendix E

DIALOG Data Collection Booklet



DIALOG



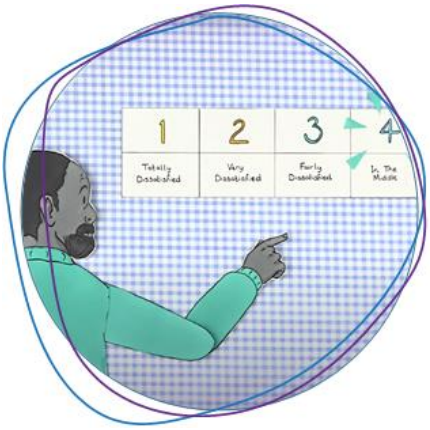
ID: _____

Credit: Stefan Priebe,
East London
Foundation Trust

NHS
East London
NHS Foundation Trust

Table of Contents

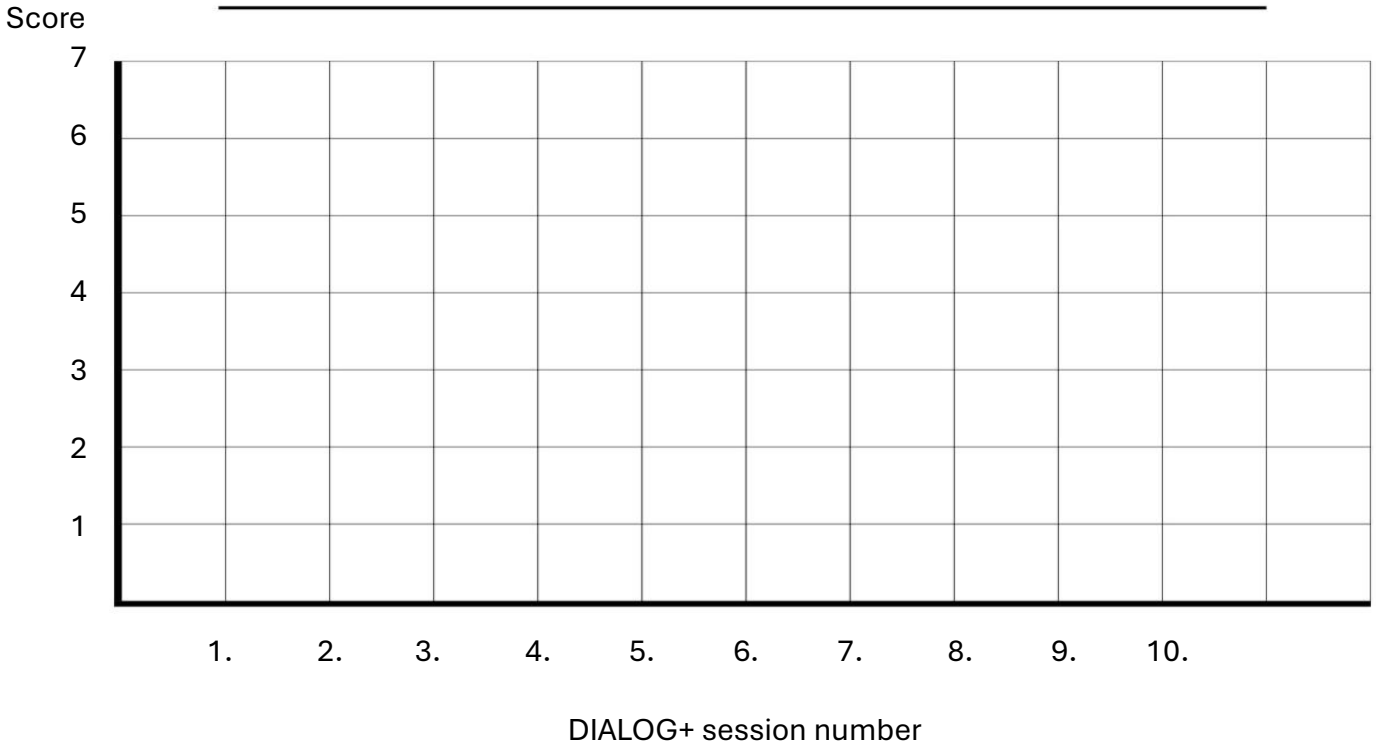
DIALOG SCALE 1 ..... ***3***
DIALOG SCALE 2 ..... ***4***
DIALOG SCALE 3 ..... ***5***
DIALOG SCALE 4 ..... ***6***
DIALOG SCALE 5 ..... ***7***
DIALOG SCALE 6 ..... ***8***
DIALOG SCALE 7 ..... ***9***
DIALOG SCALE 8 ..... ***10***
DIALOG SCALE 9 ..... ***11***
DIALOG SCALE 10 ..... ***12***
Plot progress. ***14***
Ongoing Feedback ***15***
Feedback Survey ***17***



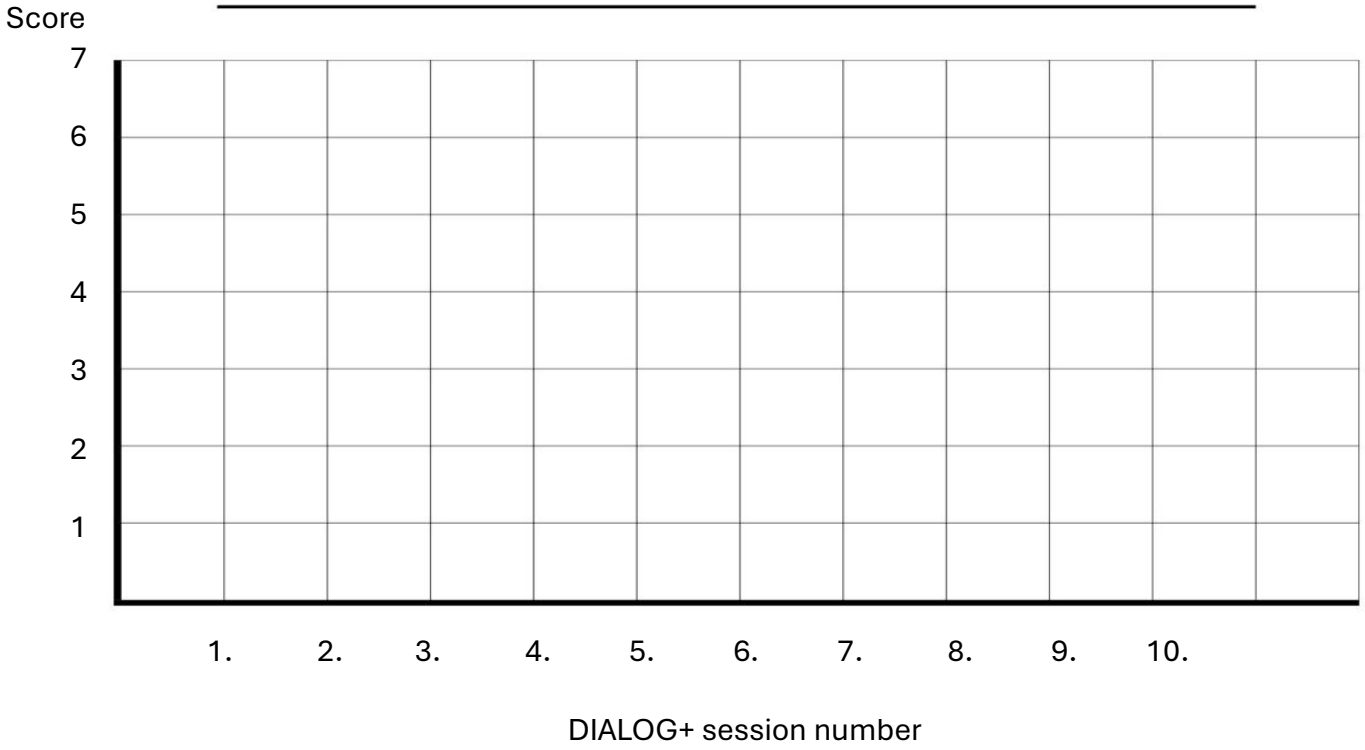
DIALOG SCALE 1

REDACTED OUTCOME MEASURE

Plot progress.
Quality of Life



Patient Experience



Ongoing Feedback

Please use this section to make ongoing notes of client feedback.

Appendix F

DIALOG+ Data Collection Booklet



DIALOG+



ID: _____

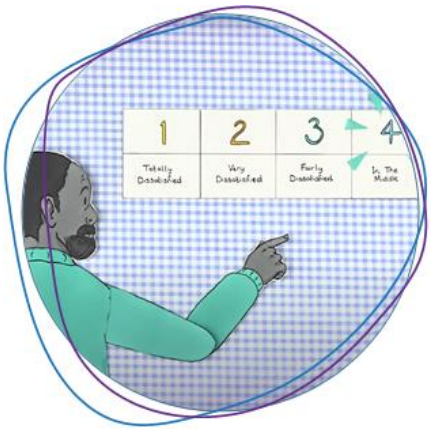
Credit: Stefan Priebe,

East London
Foundation Trust

NHS
East London
NHS Foundation Trust

Table of Contents

DIALOG SCALE 1 ..... ***4***
DIALOG SCALE 2 ..... ***6***
DIALOG SCALE 3 ..... ***8***
DIALOG SCALE 4 ..... ***10***
DIALOG SCALE 5 ..... ***12***
DIALOG SCALE 6 ..... ***14***
DIALOG SCALE 7 ..... ***16***
DIALOG SCALE 8 ..... ***18***
DIALOG SCALE 9 ..... ***20***
DIALOG SCALE 10 ..... ***22***
DIALOG SCALE 11 ..... ***24***
DIALOG SCALE 12 ..... ***26***
Plot progress. ***28***
Ongoing Feedback ***31***
Post-Intervention Feedback Survey ***33***
Further Question Assistance ***36***



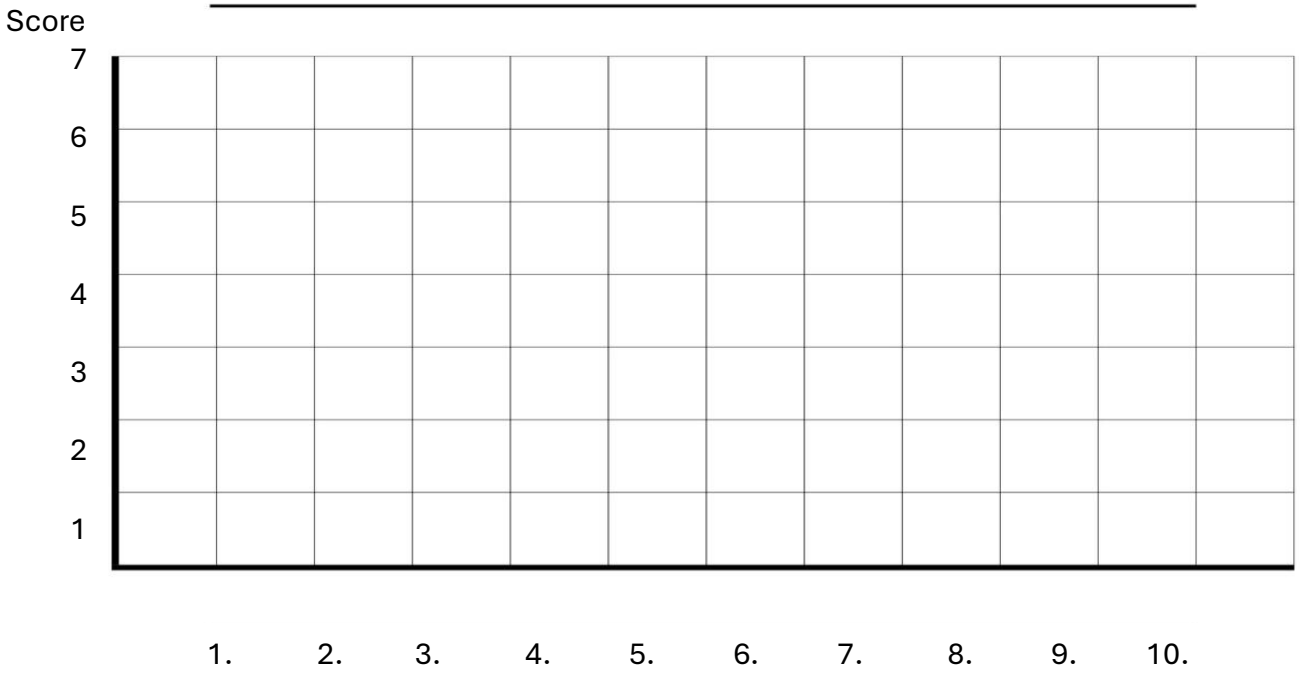
Date:
Completed by (staff member initials):
Duration to complete:
Refused to participate and why:

REDACTED OUTCOME MEASURE

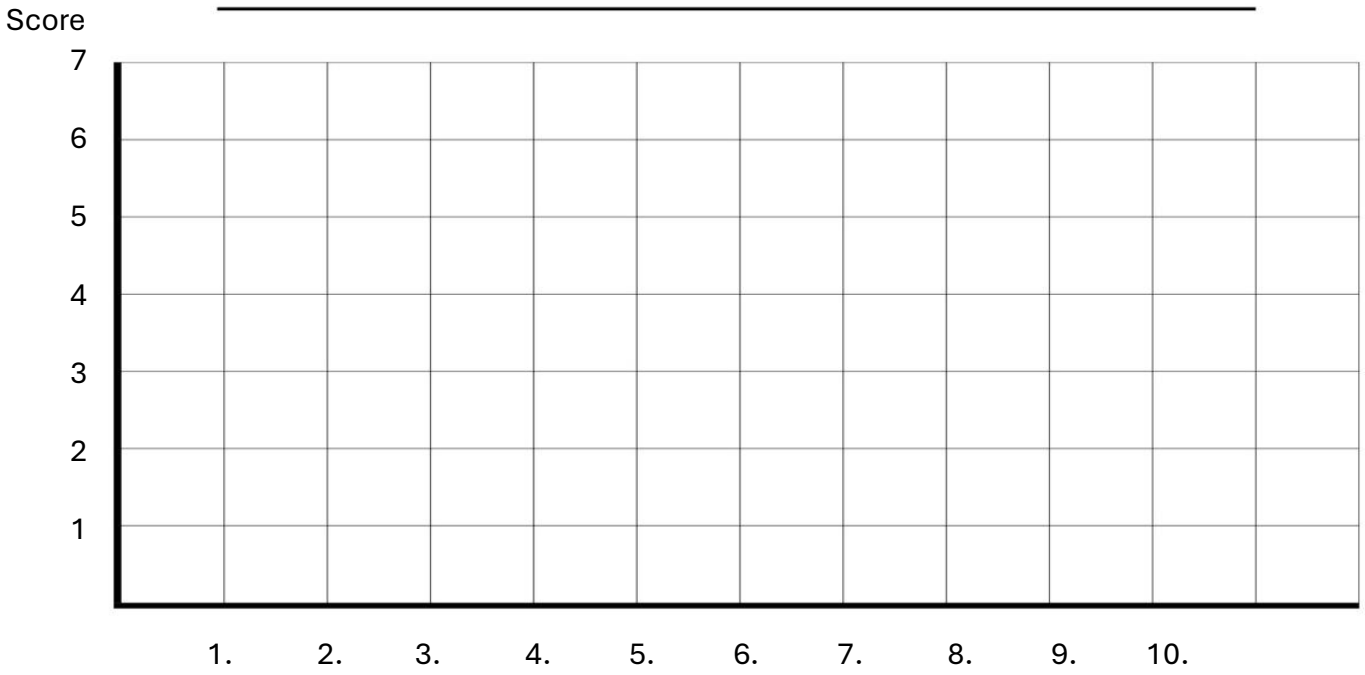
Pick up to three items from the scale you would like to discuss further.

REDACTED FOUR STEP APPROACH

Plot progress.
Quality of Life

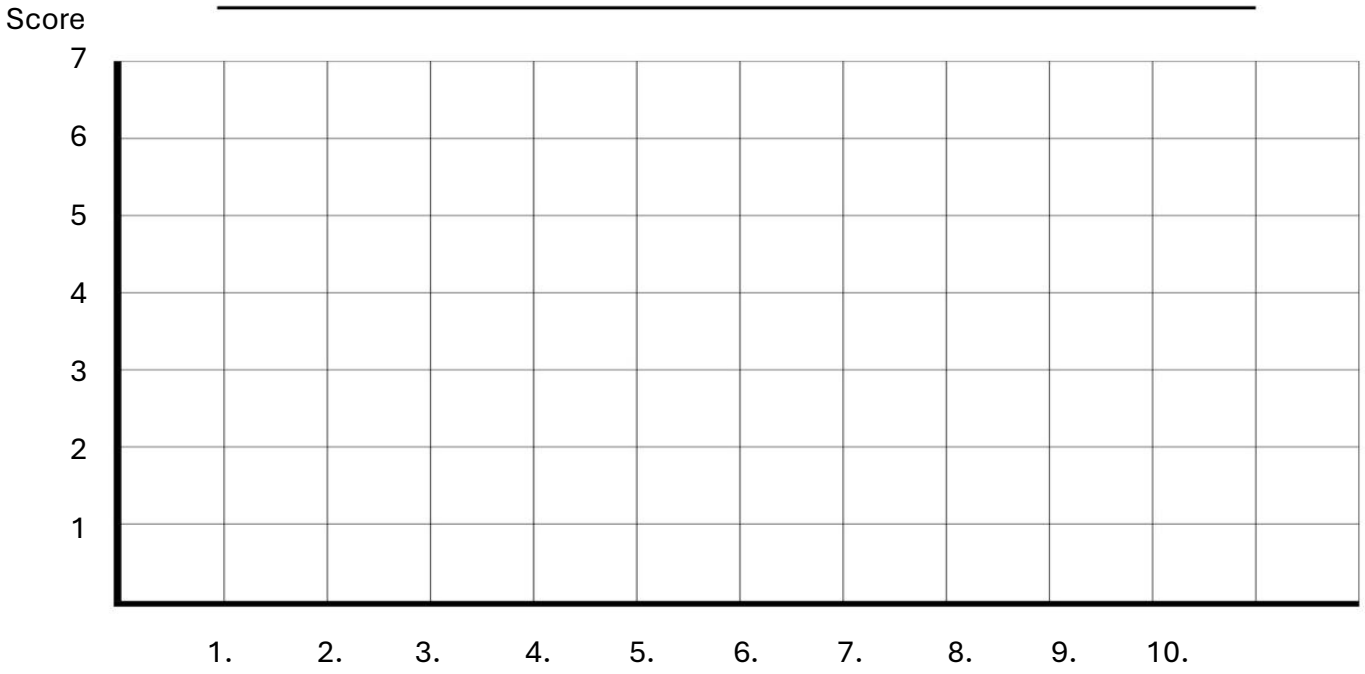


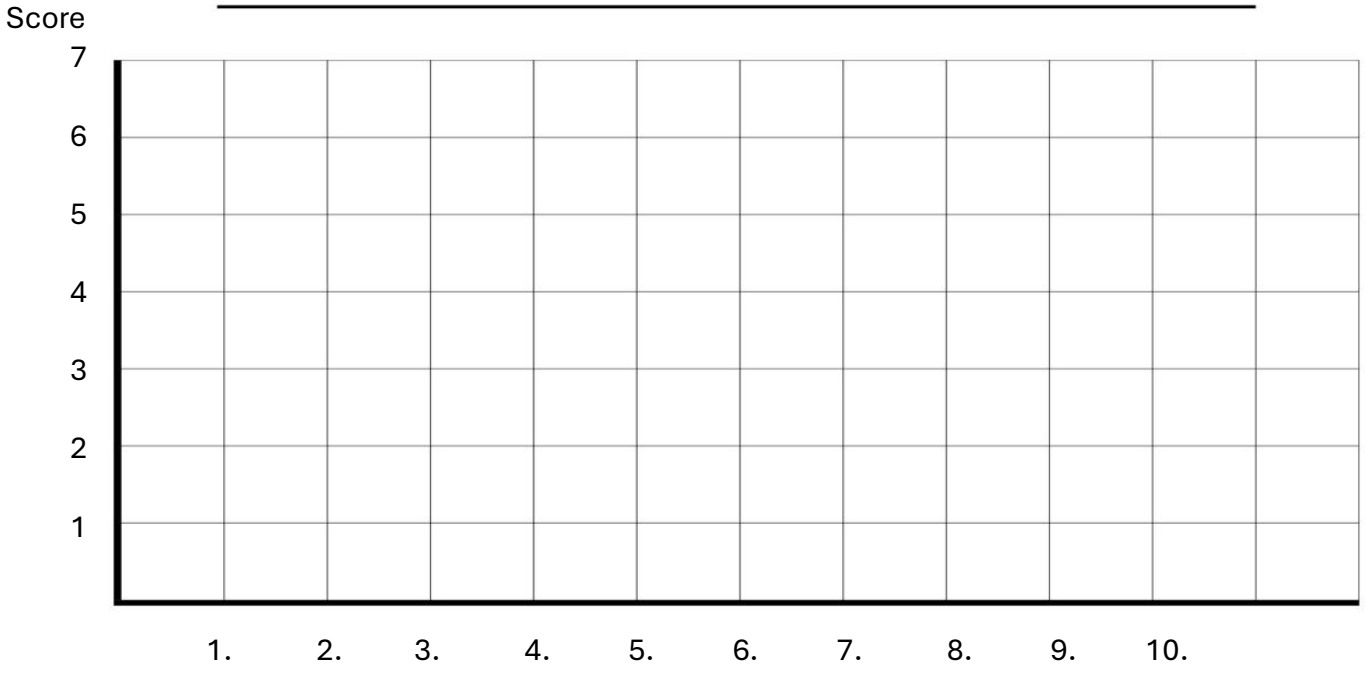
Pati



Areas of Interest

Patient Experience





Ongoing Feedback

Please use this section to make ongoing notes of client feedback.

Further Question Assistance

1. **Understanding**

In this section explore:

a. **Explore** reasons for dissatisfaction/wishes for additional help e.g. "What makes you dissatisfied with….?"
b. **What works** and what are the positive aspects within the situation e.g. "What makes you rate your medication 3 rather than 1?"

2. **Looking forward**

In this section explore:

a. **The best-case scenario** for the person in the given domain. E.g. "If your satisfaction with your accommodation was 7 (totally satisfied) how would you know?"
b. **Small Changes** that would be a sign of improvement for the patient e.g. "What would need to happen for you to move from 3 (fairly dissatisfied) to 4 (in the middle) with your accommodation?"

3. **Exploring options**

In this section explore:

a. **What the patient can do** to make the desired changes happen
b. **What the clinician can do** and what services they can offer to help with that?
c. **What others can do** and how should they be involved in the process?

4. **Agreeing on action guidelines**

This section includes:

a. **A summary** of what was discussed.
b. **Documentation** of what the patient and clinician have agreed on.

Appendix G

Permission to use Paper version of DIALOG+

REDACTED

Appendix H

Research Poster

IRAS 338942. 24/5/2024
Version 2



DIALOG+

Research study

What is DIALOG+

- DIALOG+ is an intervention that involves having a conversation with your clinician about different aspects of life that might impact your quality of life.
- This conversation would involve rating how you are feeling on a scale across several life areas such as mental health, physical health, accommodation, friendships etc.
- Following this you would then think of small changes that could be made together to help improve that score gradually over time.
- This might involve considering medication and reviewing your current situation.

What is involved in research?

- **DIALOG is an outcome measure that will be implemented in this NHS trust, DIALOG+ is an extension of this outcome measure.**
- If patients are interested in participating, they will be asked to consent, there are patient information sheets available.
- Feedback from staff and patients will be gathered from consenting participants as part of the study.
- DIALOG will be collected bi-weekly and DIALOG+ will be collected alongside this when implemented for those in the treatment group, those in the comparison group will continue with biweekly DIALOG.

Who are the researchers?

- Sinead McLernon is a trainee clinical psychologist and will be writing up the results as part of the completion of her training as a Clinical psychologist.
- Dr Claire Bone is supervising Sinead McLernon as a qualified Clinical psychologist.
- Further details can be found on information sheets located in the nurse's office.



Appendix I

Staff Participant Information Sheets

IRAS 338942. 16/5/2024
Version 2

Participant Information Sheet (quick read) STAFF

You are being invited to take part in a research project on a feasibility trial of DIALOG+.



What is the purpose of this study?

- The NHS trust has planned to introduce DIALOG, which measures patient experience and patient outcomes.
- DIALOG+ asks additional questions which are solution-focused, and your answers can form a goal to work towards.
- The study aims to identify how feasible and acceptable it is to implement DIALOG+ in an inpatient setting, this might include identifying potential barriers and facilitators to implementation.

Why am I being asked to take part?

- We would like you to take part as you are currently a staff member on a psychiatric ward.
- This is a great opportunity to be involved in providing feedback on outcome measures and a ward-based intervention.
- It is also an opportunity to learn more about outcome measures and develop skills in implementing DIALOG+ a solution-focused intervention.

Do I have to take part?

- Taking part in this research is entirely voluntary.
- You may also discontinue your participation at any time, without giving a reason why.
- You may withdraw your data without giving a reason why up until April 2025, which is when we will begin analysing the data.

What will happen if I take part? What will I have to do?

- You will be asked to attend training on DIALOG+ this will be facilitated by the research team.
- You will meet with patients biweekly to complete DIALOG+ questions relating to the DIALOG measure and pick topics to bring to MDT meeting.
- At the end of the intervention phase, clinicians and staff on the ward participating in the DIALOG+ intervention will be asked to take part in an interview or focus group online to reflect on their experiences of the intervention.

Will my taking part in the project be kept confidential?

- All the information that you provide will be kept strictly confidential, and the research team will only have access to anonymised data.

What will happen to the data collected, and the results of the research project?

- The results of this research will form part of a Clinical Psychology Doctoral thesis.
- We aim to publish the results in an academic journal. As stated above you will not be personally identified in any reports or publications.

Contact Details

Lead Researcher Sinead McLernon

Supervisor Dr Claire Bone

Contactable on the ward.

STAFF

Implementing the DIALOG Outcome Measure and a Solution-Focused Intervention (DIALOG+) in a Psychiatric Rehabilitation Inpatient Setting: A Mixed Methods Feasibility Study.

You are invited to participate in a research project sponsored by the University of Sheffield. The research is a feasibility trial of DIALOG+ (feasibility in this case asks the question: is it possible to do DIALOG+ in this setting?).

DIALOG+ is a solution-focused and person-centred intervention (this means that questions and rating scales will be used to make progress towards the patient's goals). This leaflet will provide you with all the information about the trial and you can have the opportunity to opt out of the intervention. If anything is unclear or you would like more information, please ask us; our contact details are at the end of this information sheet.

What is the purpose of this study?

The trust has planned to introduce DIALOG into psychiatric inpatient wards in Sheffield. DIALOG is an outcome measure based on quality of life and patient experience. DIALOG captures areas of a patient's life that they might want to focus on during their recovery. These areas can then be discussed in the multidisciplinary team (MDT) meeting. DIALOG+ builds on this outcome measure, asking follow-up questions related to the area of the patient's life the patient wants to focus on. The patient can then set goals related to this and bring this to the MDT every two weeks to discuss. DIALOG+ has been used in community settings and we would like to understand if it can be used in an inpatient setting.

Why am I being asked to take part?

We would like you to take part as you are currently a staff member in a psychiatric inpatient ward that has been identified as appropriate for this research project. The intervention being studied is used by staff and linked to clinician-patient interaction.

We will be looking at how feasible this intervention is for NHS inpatient settings and whether it is appropriate in your work context. The overall aim is for the intervention to increase the quality of life of patients on the ward. Part of the aims of the study is to identify how feasible it is to implement DIALOG+ in an inpatient setting, this might include identifying what makes it easier or difficult to implement this intervention in this setting.

Do I have to take part?

Taking part in this research is entirely voluntary. If you do not wish to take part, there will be no negative consequences. You may also discontinue your participation at any time, without giving a reason why. It is up to you to decide whether to take part. If you decide to take part, you will be asked to sign a consent form before you participate. If you decide not to take part in the study, your decision will have no impact on any appraisal processes or similar.

What will happen if I take part? What will I have to do?

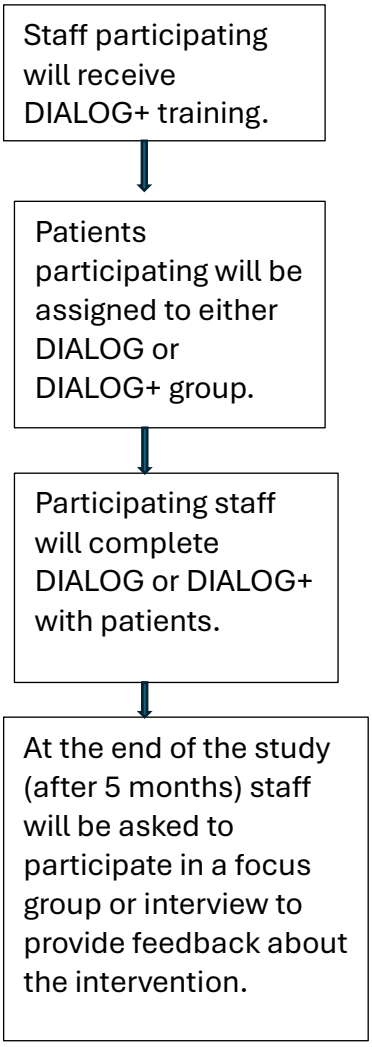
You will be asked to attend training on DIALOG+ this will be facilitated by the research team (details below). After completing the training, you will be required to gather DIALOG information from patients. The trial will run for 5 months. You will meet with patients for a meeting every two weeks to complete DIALOG+ questions relating to the DIALOG measure and pick topics to bring to MDT meetings. DIALOG+ will be completed up to 10 times during the trial and will take around 30 minutes each time to complete.

DIALOG+ asks about many aspects of the patient's life such as friendships, job situation, health, medication, and how satisfied they are with the professionals that they work

with. They will then be asked to think about best-case scenarios or smallest improvements in these areas, how to get there, and agree on actions for MDT.

At the end of the intervention phase, clinicians and staff on the ward participating in the DIALOG+ intervention will be asked to participate in either a focus group or an interview to reflect on their experiences of the intervention. It is important to gather this information to help inform the service going forward, as well as meeting the aims of the project.

Summary of the study



What are the possible disadvantages and risks of taking part?

The interviews and focus groups could be time-consuming.

What are the possible benefits of taking part?

Staff will have the opportunity to provide feedback on a ward-based intervention and on outcome measures which are being implemented across the wards.

Staff will have the opportunity to develop skills in using DIALOG and DIALOG+ a solution-focused intervention.

DIALOG+ is an effective intervention improving patient outcomes and reducing clinician workload, this is an opportunity to be part of researching these outcomes in an inpatient setting.

It is hoped that this work will help us to see if this type of intervention method can be used in inpatient settings in (redacted trust name) and will provide us with enough data to do a preliminary randomised control trial.

Will my taking part in the project be kept confidential?

All the information that you provide will be kept strictly confidential and will only be accessible to members of the research team. The only exception for this would be if information arose that caused the research team concerns for the safety of you or others. In these circumstances, we would have a duty of care to pass the information on. You will not be identified in any reports or publications unless you have given your explicit consent for this. If you agree to us sharing information you provide with other researchers (e.g. by making it available in a data archive) then your personal details will not be included unless you explicitly request this.

What is the legal basis for processing my personal data?

According to data protection legislation, we are required to inform you that the legal basis we are applying in order to process your personal data is that 'processing is necessary for the performance of a task carried out in the public interest ' (article 6 (1)(e)).

As we will be collecting some data that is defined in the legislation as more sensitive (i.e. information about your ethnic origin and health), we also need to let you know that we are applying the following condition in law: that the use of your data is 'necessary for scientific or historical research purposes'.

What will happen to the data collected, and the results of the research project?

Your data will be stored securely at the University of Sheffield, accessible only to members of the research team. Data will be anonymised as you will be assigned a pseudonym. Your data will be stored on a password-protected computer at the University of Sheffield. The data collected within the study will be stored on a University of Sheffield server. Once we have finished the study, we will keep some of the data so we can check the results. All anonymised information will be retained for 10 years.

The results of this research will form part of a Clinical Psychology Doctoral thesis. We also aim to publish the results in an academic journal. As stated above you will not be personally identified in any reports or publications.

How will we use information about you?

We will need to use information from you for this research project. This information will be transcripts from focus groups. People will use this information to do the research. People who do not need to know who you are will not be able to see your name or contact details. We will keep all information about you safe and secure. Once we have finished the study, we will keep some of the data so we can check the results or for future research. We will write our reports in a way that no-one can work out that you took part in the study.

Sensitive information will be removed from transcripts. All anonymised information will be retained for 10 years.

What are your choices about how your information is used?

You can stop being part of the study at any time, without giving a reason, but we will keep information about you that we already have.

We need to manage your records in specific ways for the research to be reliable. This means that we won't be able to let you see or change the data we hold about you.

Where can you find out more about how your information is used?

You can find out more about how we use your information:

- At www.hra.nhs.uk/information-about-patients/
- By asking one of the research teams (details at the end of this form).
- www.hra.nhs.uk/patientdataandresearch
- https://www.sheffield.ac.uk/govern/data-protection

Who is organising and funding the research?

This study is being conducted by Sinead Mclernon (Clinical Psychologist in Training), as part of the qualification towards becoming a Doctor of Clinical Psychology at the University of Sheffield. Sinead is being supervised by Dr Claire Bone, who is also based at the University of Sheffield. The research is being carried out in collaboration with the NHS, redacted trust name

The University of Sheffield will act as the Data Controller for this study. This means that the University is responsible for looking after your information and using it properly.

Who has ethically reviewed the project?

The Research Ethics Committee will review this project, protecting the rights, safety, dignity and wellbeing of research participants.

Insurance

The study is sponsored by The University of Sheffield who will provide insurance for the project. This means that The University has in place insurance against liabilities for which it may be legally liable, and this cover includes any such liabilities arising out of the study.

What if something goes wrong and I wish to complain about the research?

If you wish to make a complaint about this project in the first instance you should contact the lead researcher or their supervisor. If you do not feel satisfied that your complaint has been dealt with appropriately you can contact the Head of the Psychology Department, who can be contacted at the following address: Department of Psychology, University of Sheffield, Cathedral Court, 1 Vicar Lane, Sheffield, S1 2LT.

If your complaint relates to how your personal data has been handled, additional information about how to raise a concern can be found in the University's Privacy Notice: https://www.sheffield.ac.uk/govern/data-protection/privacy

Information can be found on the patient liaison service (PALS) website:

https://www.sth.nhs.uk/patients/patient-experience/feedback/patient-advice-and-liaison-service-pals

This is an independent source of information on studies and also a route to raise complaints about treatment in the NHS

Contact Details

Researcher
Sinead Mclernon

Supervisor

Dr Claire Bone **Thank you very much for taking time to read about the project.**

Appendix J

Staff Participant Consent Forms

IRAS 338942 3/7/2024
Version 1

NHS Sheffield Health and Social Care NHS Foundation Trust

The University Of Sheffield.

Implementing the DIALOG Outcome Measure and a Solution-Focused Intervention (DIALOG+) in an Acute Psychiatric Inpatient Setting: A Mixed Methods Feasibility Study. (STAFF)

Consent Form

Please initial the appropriate boxes	**Yes**	**No**
Taking Part in the Project		
I have read and understood the project information sheet, or the project has been fully explained to me. (If your answer is No to this question, please do not proceed with this consent form until you are fully aware of what your participation in the project will mean.)		
I have been given the opportunity to ask questions about the project.		
I understand that my participation is voluntary and that I can withdraw from the study at any time. I do not have to give any reasons for why I no longer want to take part and there will be no adverse consequences if I choose to withdraw from the project.		
How my information will be used during and after the project		
Any data accessed or shared outside the project will be non-identifiable. I understand my personal details such as name, age, gender, diagnoses etc will not be revealed to people outside the project.		

I understand and agree that other authorised researchers may use my data in publications, reports, web pages and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form.		
I understand and agree that other authorised researchers will have access to this data only if they preserve the confidentiality of the information as requested in this form.		
So that the information you provide can be used legally by the researchers		
I agree to assign the copyright I hold in any materials generated as part of this project to The University of Sheffield.		
How will my consent form be filed?		
You will receive a copy of the consent form and a copy will be held in the site file.		

Appendix K

Patient Participant Information Sheets

IRAS 338942. 25/6/2024
Version 3

Participant Information Sheet (PATIENT PARTICIPANTS)

You are invited to participate in a research project on a feasibility trial of DIALOG+.



The University Of Sheffield.

What is the purpose of this study?

- The NHS trust has planned to introduce DIALOG, which measures patient experience and patient outcomes.
- DIALOG+ asks additional questions and your answers can form a goal to work towards.

Why am I being asked to take part?

- We would like you to take part as you are currently an inpatient in a psychiatric ward.

Do I have to take part?

- Taking part in this research is entirely voluntary.
- You may also discontinue your participation at any time, without giving a reason why.
- You can stop being part of the study at any time, without giving a reason, but we will keep information about you that we already have.

What will happen if I take part? What will I have to do?

- You will be assigned to a group that either receives DIALOG or DIALOG+.

- You will complete the DIALOG measure every two weeks and those in the DIALOG+ group will have a meeting with a clinician before your multidisciplinary team (MDT) meeting to complete DIALOG+.
- You will be asked about many aspects of your life such as friendships, job situation, health, medication and how satisfied you are with the professionals that you work with.
- The study will last for 5 months and DIALOG+ will be completed up to 10 times. DIALOG+ will take around 30 minutes to complete.

Will my taking part in the project be kept confidential?

- All the information that you provide will be kept strictly confidential, and the research team will only have access to anonymised data.

What will happen to the data collected, and the results of the research project?

- The results of this research will form part of a Clinical Psychology Doctoral thesis.
- We aim to publish the results in an academic journal. As stated above you will not be personally identified in any reports or publications.

The use of personal data

- In this research study we will use information from you. We will only use information that we need for the research study or future research. We will let very few people know your name or contact details, and only if they really need it for this study.
- Everyone involved in this study will keep your data safe and secure. We will also follow all privacy rules.
- At the end of the study, we will save some of the data in case we need to check it. We will make sure no-one can work out who you are from the reports we write.
- The information pack tells you more about this.

Contact Details

Researcher: (redacted)

Participant information sheet (patient)

Implementing the DIALOG Outcome Measure and a Solution-Focused Intervention (DIALOG+) in a Psychiatric Rehabilitation Inpatient Setting: A Mixed Methods Feasibility Study.

You are invited to participate in a research project sponsored by the University of Sheffield. The research is a feasibility trial of DIALOG+ (feasibility in this case asks the question: is it possible to do DIALOG+ in this setting?).

DIALOG+ is a solution-focused and person-centred intervention (this means that questions and rating scales will be used to make progress towards your goals). This leaflet will provide you with all the information about the trial and you can have the opportunity to opt out of the intervention. If anything is unclear or you would like more information, please ask us; our contact details are at the end of this information sheet.

What is the purpose of this study?

The redacted trust has planned to introduce DIALOG into psychiatric inpatient wards in Sheffield. DIALOG is an outcome measure based on quality of life and patient experience. DIALOG is a measure to capture areas of a patient's life that they might want to focus on during their recovery. These areas can then be discussed in the multidisciplinary team (MDT) meeting (these are the meetings when all the professionals involved in your care meet to

discuss your care). DIALOG+ builds on this outcome measure, asking follow-up questions related to the area of the patient's life the patient wants to focus on. The patient can then set goals related to this and bring this to the MDT every two weeks to discuss.

DIALOG+ has been used in community settings and we would like to understand if it can be used in an inpatient setting.

Why am I being asked to take part?

We would like you to take part as you are currently an inpatient in a psychiatric inpatient ward. The study will compare two groups of patients, those who receive DIALOG and those who receive DIALOG+.

If you are in the DIALOG group, we would like you to complete DIALOG every two weeks. It will take around 15 minutes to complete. The study will last 5 months, DIALOG will be completed up to 10 times throughout the study.

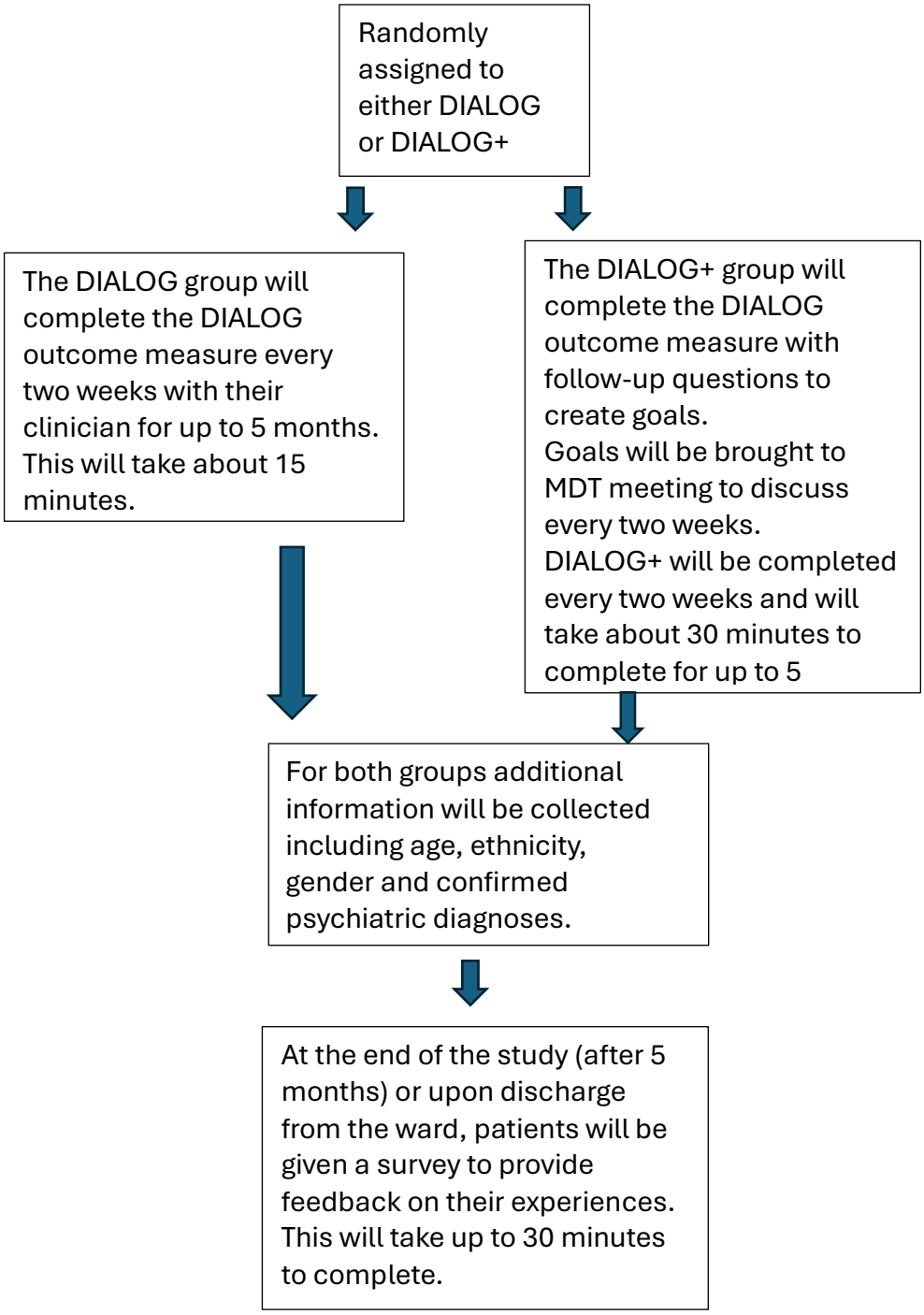
If you are in the DIALOG+ group, we would like you to meet with a clinician to complete the DIALOG+ questions and pick topics to bring to your MDT meeting. You will be asked about many aspects of your life such as friendships, job situation, health, medication and how satisfied you are with professionals that you work with. You will then be asked to think about your best-case situation or smallest improvements in these areas and how to get there. Together with your clinician you will agree on actions for the MDT.

DIALOG+ will be completed every two weeks and will take around 30 minutes to complete. The study is planned to last 5 months, DIALOG+ will be completed every two weeks for the duration of the study. This means DIALOG+ will be completed up to 10 times throughout the study.

We are also collecting additional information from you including your age, ethnicity, gender, and confirmed psychiatric diagnoses. We will be using this information to aid in the analyses of data.

You will be asked to complete a survey of your experiences of completing DIALOG and DIALOG+. The survey will be offered at the end of the study or if you are discharged from the ward, it will be offered before your discharge date.

Summary of the trial



Do I have to take part?

Taking part in this research is entirely voluntary. If you do not wish to take part, there will be no negative consequences. You may also stop your participation at any time, without giving a reason why. If you wish to withdraw, please contact redacted (details are at the end of this form). It is up to you to decide whether to take part. If you decide you do not wish to take part, you can say no.

What will happen if I take part? What will I have to do?

At the beginning of the study, you will be assigned to either a DIALOG group or a DIALOG+ group. The decision will be made by randomly allocating the wards to either group through the process of flipping a coin.

You will complete the DIALOG measure every two weeks and if you are in the DIALOG+ group you will have a meeting with a clinician before your MDT meeting to complete DIALOG+ as described above.

You will be asked to complete a survey of your experiences of completing DIALOG and DIALOG+.

What are the possible disadvantages and risks of taking part?

The questionnaires ask about topics that we hope are useful to you and that we discuss as part of your regular care. If any of the questions do cause you distress, you can discuss this with the clinical ward team. If you have any questions about the research project contact a member of the research team using the details below.

The post-study survey may take some time to complete.

What are the possible benefits of taking part?

It is hoped that the conversations we have during these DIALOG+ meetings will support positive outcomes. The learning we gather from this study will drive change and improvement for other people who need inpatient admissions in the future.

You will have the opportunity to provide feedback on the intervention which may have an impact on future research and use of DIALOG+.

Will my taking part in the project be kept confidential?

All the information that you provide will be kept confidential, and the research team will only have access to anonymised data, such as your scores on the DIALOG measure and demographic information such as age, gender, ethnicity, and diagnosis. You will not be identified in any reports or publications. If you agree to us sharing information you provide with other researchers (e.g. by making it available in a data archive) then your personal details will not be included.

How will we use information about you?

We will need to use information from you for this research project.

This information will include a unique participant number which will be identifiable by staff who work directly with you on the ward. People will use this information to do the research or to check your records to make sure that the research is being done properly.

People who do not need to know who you are will not be able to see your name or contact details. Your data will have a code number instead.

We will keep all information about you safe and secure.

Once we have finished the study, we will keep some of the data so we can check the results or for future research. We will write our reports in a way that no-one can work out that you took part in the study.

What are your choices about how your information is used?

You can stop being part of the study at any time, without giving a reason, but we will keep information about you that we already have.

We need to manage your records in specific ways for the research to be reliable. This means that we won't be able to let you see or change the data we hold about you.

Where can you find out more about how your information is used?

You can find out more about how we use your information:

- At www.hra.nhs.uk/information-about-patients/
- By asking one of the research teams (details at the end of this form).
- www.hra.nhs.uk/patientdataandresearch
- https://www.sheffield.ac.uk/govern/data-protection

What will happen to the data collected, and the results of the research project?

Your data will be stored securely at the University of Sheffield, accessible only to members of the research team. Data will be anonymised as you will be assigned a unique participant number/ pseudonym. Your data will be stored on a password-protected computer at the University of Sheffield. The data collected within the study will be stored on a University of Sheffield server. Once we have finished the study, we will keep some of the data so we can check the results. All anonymised information will be retained for 10 years.

The results of this research will form part of a Clinical Psychology Doctoral thesis. We also aim to publish the results in an academic journal. As stated above you will not be personally identified in any reports or publications.

What is the legal basis for processing my personal data?

According to data protection legislation, we are required to inform you that the legal basis we are applying to process your personal data is that 'processing is necessary for the performance of a task carried out in the public interest' (article 6 (1)(e)).

As we will be collecting some data that is defined in the legislation as more sensitive (i.e. information about your ethnic origin and health), we also need to let you know that we are applying the following condition in law: that the use of your data is 'necessary for scientific or historical research purposes'.

Who is organising and funding the research?

This study is being conducted by Sinead Mclernon (Clinical Psychologist in Training), as part of the qualification towards becoming a Doctor of Clinical Psychology at the University of Sheffield. Sinead is being supervised by Dr Claire Bone, who is also based at the University of Sheffield. The research is being carried out in collaboration with the NHS, specifically and redacted.

The University of Sheffield will act as the Data Controller for this study. This means that the University is responsible for looking after your information and using it properly.

Who has ethically reviewed the project?

The Research Ethics Committee will review this project, protecting the rights, safety, dignity and wellbeing of research participants.

Insurance

The study is sponsored by The University of Sheffield who will provide insurance for the project. This means that The University has in place insurance against liabilities for which it may be legally liable, and this cover includes any such liabilities arising out of the study.

What if something goes wrong and I wish to complain about the research?

If you wish to make a complaint about this project in the first instance you should contact the lead researcher or their supervisor. If you do not feel satisfied that your complaint has been dealt with appropriately you can contact the Head of the Psychology Department, who can be contacted at the following address: Department of Psychology, University of Sheffield, Cathedral Court, 1 Vicar Lane, Sheffield, S1 2LT.

If your complaint relates to how your personal data has been handled, additional information about how to raise a concern can be found in the University's Privacy Notice: https://www.sheffield.ac.uk/govern/data-protection/privacy

Information can be found on the patient liaison service (PALS) website. This is an independent source of information on studies and a route to raise complaints about treatment in the NHS.

https://www.sth.nhs.uk/patients/patient-experience/feedback/patient-advice-and-liaison-service-pals

Contact Details

Researcher

Sinead Mclernon

Department of Psychology, University of Sheffield,

Cathedral Court, 1 Vicar Lane,

Sheffield, S1 2LT.

Telephone: Please leave a message with research officer Amrit Sinha on 0114226650 and Sinead will return your call.

Supervisor

Dr Claire Bone

Thank you very much for taking time to read about the project

Appendix L

Patient Participant Consent Forms

IRAS 338942. 26/6/2024
Version 3

Sheffield Health and Social Care NHS Foundation Trust

The University Of Sheffield.

Implementing the DIALOG Outcome Measure and a Solution-Focused Intervention (DIALOG+) in an Acute Psychiatric Inpatient Setting: A Mixed Methods Feasibility Study. (PATIENT)

Consent Form

Please initial the appropriate boxes	**Yes**	**No**
Taking Part in the Project		
I have read and understood the project information sheet, or the project has been fully explained to me. (If your answer is No to this question, please do not proceed with this consent form until you are fully aware of what your participation in the project will mean.)		
I have been given the opportunity to ask questions about the project.		
I agree to take part in the project. I understand that taking part in the project will involve me completing either the DIALOG outcome measure or DIALOG+ intervention for 10 sessions over a period of 5 months.		
I understand that my participation is voluntary and that I can withdraw from the study at any time. I do not have to give any reasons for why I no longer want to take part and there will be no adverse consequences if I choose to withdraw from the project.		
How my information will be used during and after the project		

Any data accessed or shared outside the project will be non-identifiable. I understand my personal details such as name, age, gender, diagnoses etc will not be revealed to people outside the project.		
I understand and agree that other authorised researchers may use my data in publications, reports, web pages and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form.		
I understand and agree that other authorised researchers will have access to this data only if they preserve the confidentiality of the information as requested in this form.		
So that the information you provide can be used legally by the researchers:		
I agree to assign the copyright I hold in any materials generated as part of this project to The University of Sheffield.		
I understand that relevant sections of my medical notes and data collected during the study, may be looked at by individuals from The University of Sheffield, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.		
How will my consent forms be filed?		
A copy of your consent form will be given to you, a copy will be kept in the site file and a copy will be stored on your medical record.		

Name of Participant: Signature: Date:

Name of person taking consent: Signature: Date:

Appendix M

Risk Management Plan

Risk management plan

Research team

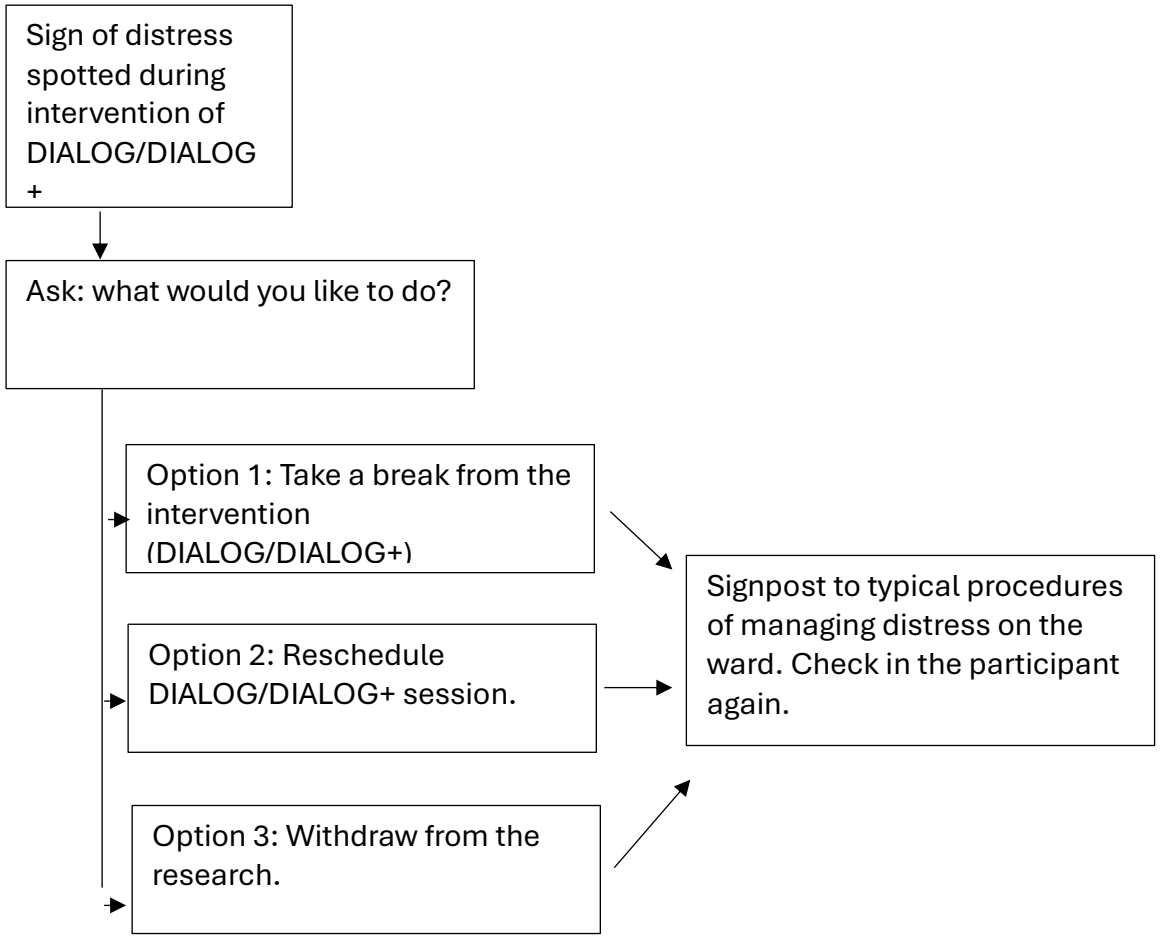
- **Sinead Mclernon** Trainee Clinical Psychologist at the University of Sheffield. Email:
- **Dr Claire Bone,** Project Supervisor, and Clinical Psychologist. Email: redacted

Potential for harm/distress to the participants

The intervention used will be implemented by trained clinicians on the ward who will have expert knowledge in managing distress. DIALOG and DIALOG+ are questionnaire-based interventions which will discuss sensitive topics such as mental health, physical health, employment, accommodation, lifestyle, relationships, friendships, safety, and medication. All these topics have the potential to evoke emotions and with that, signs of distress. The intervention will be implemented by experienced ward staff who will have training in managing distress on the ward.

Participants will also be offered the opportunity to feedback on their experience of the intervention in the form of a survey. This may be a valuable opportunity for the participants to give feedback but may also evoke feelings of distress depending on their life experiences.

Ensure appropriate protection and well-being of the participants.



Support for staff implementing DIALOG/DIALOG+

Staff may find it challenging to implement DIALOG and DIALOG+ on top of their typical duties. To ensure staff wellbeing staff will be offered training on implementing DIALOG and DIALOG+ which aims to offer an extra tool for clinicians to use during sessions. Staff can consult with the project researchers for more information on the project and for debriefs as needed. Staff will be provided with the researcher's contact details for any queries and support implementing the intervention.

Appendix N

Interview Guide

1. **Understanding & Experience with DIALOG/DIALOG+ (Affective Attitude, Coherence)**

(Tell me about) What was your experience of using DIALOG+?

- How did you find incorporating this method into your usual style?
- How did you introduce the new structure to your clients?
- How do you think your clients responded to this new style?
- What was the impact of structuring the communication between you and your clients in this way?
- Was it useful for you to see how your clients rate their satisfaction?
- What effect did using DIALOG+ during meetings with your clients have compared with meetings before DIALOG+?
    - How did you feel about the intervention?
    - Did you feel you understood the purpose of the intervention?
    - Did you feel confident implementing the intervention?
- Did you feel the intervention fit well within an inpatient setting?

2. **Integration into Routine Practice (Perceived Effectiveness, Ethicality)**

How effective do you feel the intervention was?

Did you feel the intervention fit well within an inpatient setting?

What was the impact of introducing a DIALOG+ booklet into your routine meetings?

Was DIALOG+ compatible with your routine meetings/way of working?

What effect did using DIALOG+ during meetings with your clients have compared with meetings before DIALOG+?

- Have you noticed any changes in your clients; behaviour and communication in your meetings since using DIALOG+?
- Did this method change how information was elicited from you clients?

2. Practicality & Feasibility (Perceived Effort, Implementation)

- How did you find the length of sessions? /Was the intervention easy/convenient to implement?

- What do you think of the frequency of times you were asked to use this with your clients?
- How did it feel to collect DIALOG data biweekly?
- What helped when implementing DIALOG+?

6. Recruitment & Consent (Feasibility)

How did you feel recruiting patients to participate in the study?

How did you feel engaging patients in the trial?

7. Barriers & Solutions (Feasibility)

- What were the barriers to successfully implementing DIALOG+?
- What helped when implementing DIALOG+?

8. Participation & Dropout (Engagement)

- How do you think your clients responded to this new style?
- Were there any patients that DIALOG did not work well with?
- Do you feel the process varied significantly between individual clients?

why did you drop out of using the DIALOG+ intervention?

9. Follow-Up & Long-Term Engagement (Follow-Up Rate, Retention)

- How easy or difficult was it to keep patients engaged in follow-up sessions?

10. Future Research & Scaling Up

Would you continue using the DIALOG+ procedure with your clients?

Appendix O

Patient Feedback Surveys

Post-Intervention Feedback Survey (DIALOG+ group)

1. How useful did you find using DIALOG and in what ways?
2. To what extent did the DIALOG scale monitor outcomes that were important to you?
3. How easy did you find it to use the DIALOG+ intervention?
4. How satisfied were you with your treatment whilst using DIALOG+?
5. How much did you like going through the process of DIALOG+ (including completing the DIALOG scale and 4-step approach)?
6. Did receiving DIALOG+ affect the relationship with your clinician?
7. How easy was it to use the booklet during the sessions?
8. Did you prefer DIALOG+ to treatment you received previously and if so, why?
9. Question for those who dropped out of the intervention but consented to study: why did you drop out of using the DIALOG+ intervention?
10. Are there any suggestions for how to improve DIALOG?

Did you like or dislike DIALOG+? (circle)

Strongly Dislike	Dislike	No Opinion	Like	Strongly like
1	2	3	4	5

How much effort did it take to engage with DIALOG+?

No effort at all	A little effort	No Opinion	A lot of effort	Huge Effort
1	2	3	4	5

There are moral or ethical consequences to using DIALOG+ (Do you agree or disagree?)

Strongly Disagree	Disagree	No Opinion	Agree	Strongly Agree
1	2	3	4	5

DIALOG+ has improved your clinical outcome?

Strongly Disagree	Disagree	No Opinion	Agree	Strongly Agree
1	2	3	4	5

It is clear to me how DIALOG+ will help improve clinical outcomes.

Strongly Disagree	Disagree	No Opinion	Agree	Strongly Agree
1	2	3	4	5

How confident did you feel about engaging with DIALOG+?

Very unconfident	unconfident	No Opinion	Confident	Very Confident
1	2	3	4	5

DIALOG+ interfered with my other priorities.

Strongly Disagree	Disagree	No Opinion	Agree	Strongly Agree
1	2	3	4	5

How acceptable was DIALOG+ to you?

Completely unacceptable	Unacceptable	No Opinion	Acceptable	Completely Acceptable
1	2	3	4	5

Feedback Survey (DIALOG group)

1. How useful did you find using DIALOG?
2. To what extent did the DIALOG monitor outcomes that were important to you?
3. How easy did you find completing DIALOG?
4. Did DIALOG improve the relationship with your clinician?
5. How easy was it using the booklet during sessions?
6. Did you prefer DIALOG to other outcome measures and why?

Did you like or dislike DIALOG? (circle)

Strongly Dislike	Dislike	No Opinion	Like	Strongly like
1	2	3	4	5

How much effort did it take to engage with DIALOG?

No effort at all	A little effort	No Opinion	A lot of effort	Huge Effort
1	2	3	4	5

There are moral or ethical consequences to using DIALOG (Do you agree or disagree?)

Strongly Disagree	Disagree	No Opinion	Agree	Strongly Agree
1	2	3	4	5

DIALOG has improved your clinical outcome?

Strongly Disagree	Disagree	No Opinion	Agree	Strongly Agree
1	2	3	4	5

It is clear to me how DIALOG will help improve clinical outcomes.

Strongly Disagree	Disagree	No Opinion	Agree	Strongly Agree
1	2	3	4	5

How confident did you feel about engaging with DIALOG?

Very unconfident	unconfident	No Opinion	Confident	Very Confident
1	2	3	4	5

DIALOG interfered with my other priorities.

Strongly Disagree	Disagree	No Opinion	Agree	Strongly Agree
1	2	3	4	5

How acceptable was DIALOG to you?

Completely unacceptable	Unacceptable	No Opinion	Acceptable	Completely Acceptable
1	2	3	4	5

Appendix P

Reflexivity Log Excerpts

These reflections are based on the primary researcher's notes from listening to the audio tapes. To notice when feelings related to previous experiences of working with staff teams in a residential setting arose and to monitor any potential biases. Only relevant quotes are included.

"Participants worked in supporting roles or nursing roles, and some participants hadn't collected any data. Given my experience working in residential and inpatient settings, I had certain presumptions about the barriers and facilitators to collecting data that might arise. From my past experiences, poor leadership and accountability have been significant barriers to implementing new interventions. Note: remain open to hearing new barriers and facilitators"

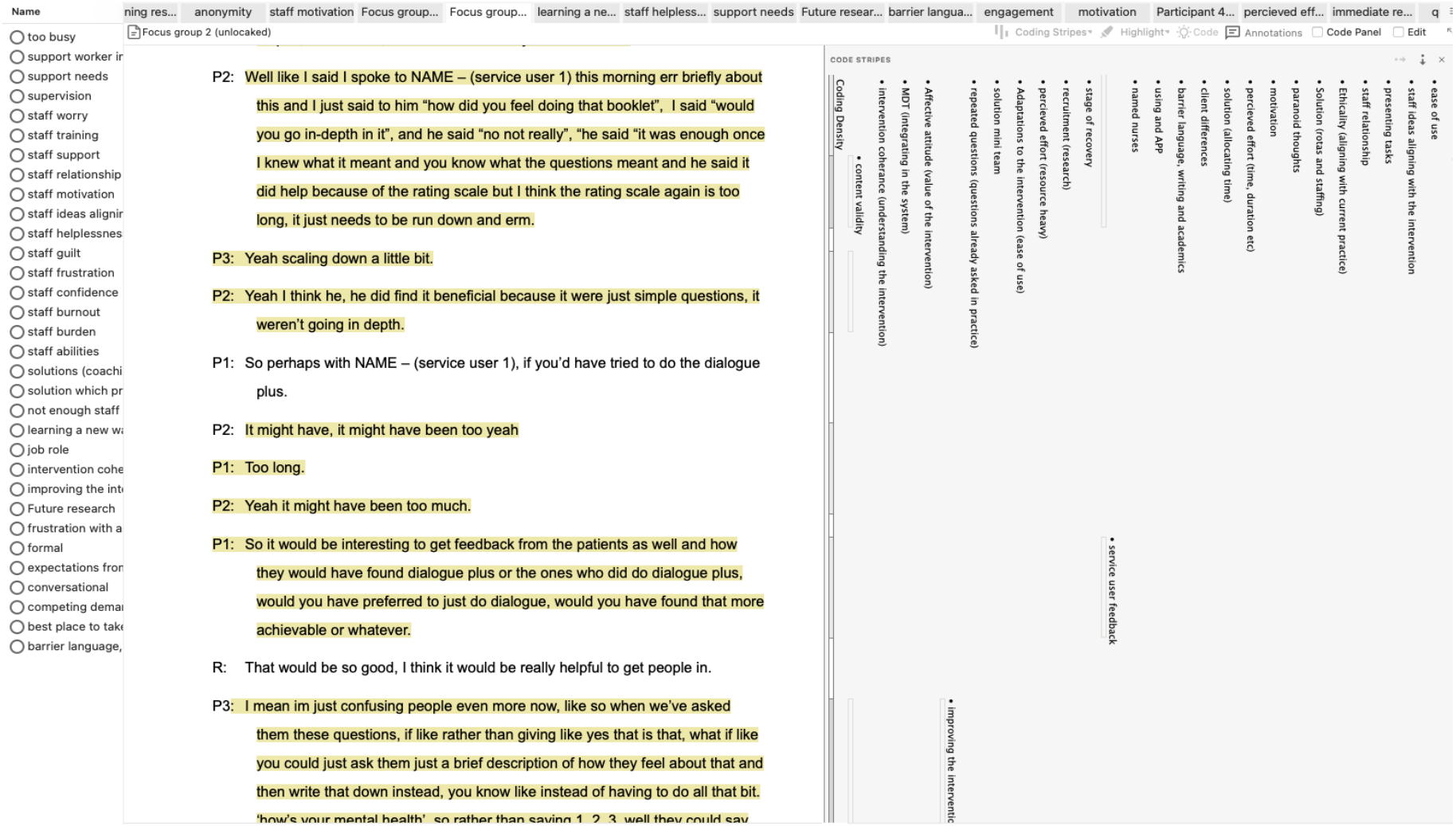
"I noticed that discussions focused on institutional processes as a barrier to progress, such as the need for service users to have more freedom within an institutional setting. I thought about how systems had changed to make the services less institutional."

"One staff member was a clinical associate psychologist, and I had higher expectations for them, and I wondered if this was part of my bias based on my role as a trainee clinical psychologist."

"I noticed that the participant frequently spoke about confidence, as well as their perception of their job role and how that influenced the implementation of the intervention. I felt surprised to hear that the participant felt that the intervention didn't align with their job role, and I wondered if this was linked to confidence."

Appendix Q

An Example of Transcript Excerpt



Appendix R

Exploratory Code Output

Initially, six themes were identified; upon further review, some themes were combined after consensus checking with the research team. Preparing the staff team involved aligning with current practices, and the perceived effort and effectiveness were combined to form the theme of the staff's affective attitude towards the intervention.

Name	Codes	References
Focus group 1	59	331
Focus group 2	80	372
Participant 1	39	93
Participant 2	44	113
Participant 3	22	46
Participant 4	50	122
Participant 5	18	33

Initially themes and subthemes	Code	Sources	References
Initial Theme: Preparing the staff team Subthemes: 1. Understanding research and the intervention 2. Drop out 3. Feelings about research 4. recruitment	Barrier language, writing and academics	2	11
	Best place to take action	1	1
	Competing demands	1	1
	Conversational	2	9
	Expectations from staff	1	1
	Follow-up	3	4
	Formal	1	1
	Frustration with administration	1	2
	Future research	4	6
	Improving the Intervention	1	13

	Intervention Coherence (understanding the intervention)	4	37
	Job role	3	12
	Learning a new way of working	2	2
	Not enough staff trained in the intervention	3	8
	Solution- which professional?	3	13
	Solution-coaching	4	9
	Staff abilities	1	1
	Staff burden	1	1
	Staff burnout	2	3
	Staff confidence	3	18
	Staff frustration	2	5
	Staff guilt	2	2
	Staff helplessness	1	6
	Staff ideas aligning with ther intervention	2	4
	Staff motivation	2	12
	Staff relationship	2	8
	Staff support	1	2
	Staff training	4	9
	Staff worry	1	1
	supervision	1	4
	Support needs	1	1
	Support worker instinct	1	1
	Too busy	2	2
Affective attitude towards research Subthemes: 1. Staff training 2. Staff confidence 3. Staff relationship	Feelings about research	3	6
	Anonymity	1	2
	Drop out	4	7
	Eligibility	2	3
	Explaining research	2	2
	Posters and leaflets	1	3
	Recruitment (research)	4	20
	Service user feedback	2	7
Aligning with current practice (team systems)	As part of admission	4	6
	Audit	1	2
	Barrier (communication)	1	1

	Barrier (staffing, rotas and organising the diary)	5	16
	Care plans	2	7
	aligning with current practice	6	28
	Having a lead	3	10
	MDT (integrating in thee system)	7	46
	Named nurses	4	18
	New admissions	1	1
	One to ones	4	14
	Ongoing service development	3	4
	Overcrowded MDTs	2	7
	Protected time	2	6
	Repeated questions (questions already asked in practice)	3	21
	RIO (new note system)	4	8
	Solution (allocating time)	5	20
	Solution (rotas and staffing)	3	12
	Solution (handover)	3	6
	Solution (mini team)	2	9
	Strained systems (too many changes)	1	1
	Team dynamics	1	1
	Team systems and pathways	5	11
	Ward manager	1	1
Inpatient setting Subtheme: 1. Content validity 2. Patient characteristics	Activity planner	2	2
	Age of service user	1	1
	Autism	2	9
	Barrier (patient leave)	2	4
	Barrier (mental health)	6	20
	Client characteristics	2	4
	Client differences	3	23
	Client group	2	9
	Client history impacting participation	2	3
	Client preferences	2	6
	Closed off	1	2

	Consent and capactiy	1	2
	Content validity	3	32
	Culturally relevant	1	3
	Expectations of a rehab setting	7	18
	Fluctuating mental health	4	6
	General barriers with client group	1	6
	Incidents on the wards	1	1
	Individual differences	2	9
	Inpatient setting	1	1
	Lack of insight	1	1
	Length of stay	1	1
	Paranoid thoughts	2	8
	Patient characteristics impacting participation	2	7
	Private	1	1
	Psychosis	2	2
	Recovery	4	7
	Scale content validity	1	1
	Section	1	1
	Severity of the illness	1	7
	Stage of recovery	4	30
Perceived Effectiveness Subthemes: 1. Engagement and motivation 2. Affective attitude about the intervention	Perceived value of the intervention	7	49
	Barrier (appearance of the booklet)	2	7
	Consistency	5	12
	Engagement	7	27
	Enthusiasm	1	1
	Immediate results	1	1
	Meaningful	1	1
	More than just an outcome measure	1	1
	Motivation	6	46
	Motivation solutions	2	6
	Motivation food	2	3
	Patient buy-in	2	7
	Perceived effectiveness (other priorities)	2	12

	Perceived effectiveness (benefits of DIALOG+)	2	7
	Person centred are	2	4
	Quick wins	1	2
	Relevance	1	4
	Strength of DIALOG+	2	9
	Using an APP	1	11
	Willing to participate	1	1
Perceived Effort	Adaptations to the intervention	2	13
	Demand heavy	3	5
	Ease of use	1	7
	Effort (too many questions)	1	3
	Goal based	3	4
	Improving communication	2	4
	Living day to day	1	1
	Patient changing their mind on the day	3	5
	Perceived effort	3	10
	Perceived effort (resource heavy)	1	10
	Perceived effort (time and duration)	4	28
	Practicality and feasibility (planning with clients)	4	7
	Presenting tasks	6	39
	Question fatigue	1	3
	routine	2	6
	Time frame	2	6

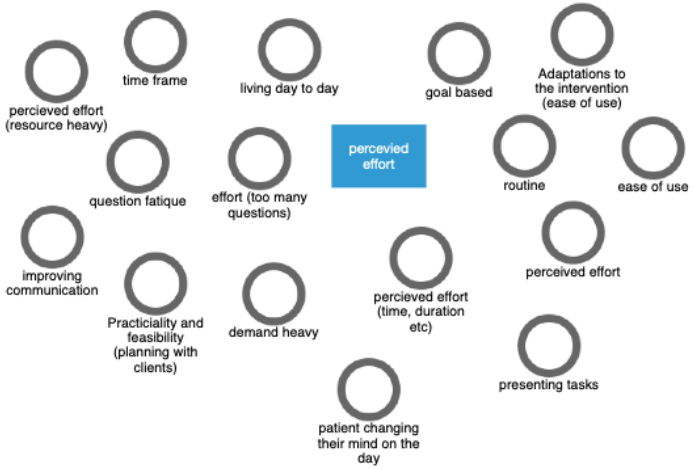
Appendix S

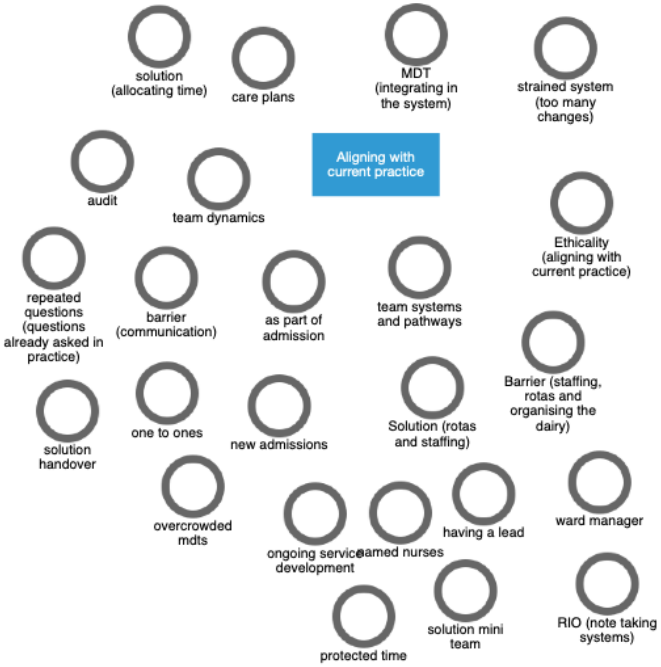
Example Thematic Map Revisions

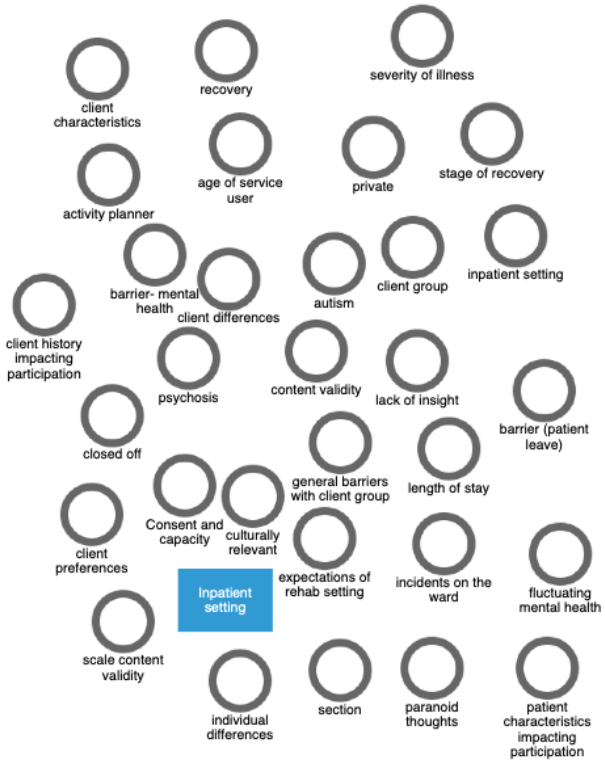
Different visualisation tools were used to rearrange codes, organise them, and create various maps. Included are some examples to illustrate the process. NVivo provides a visual tool to demonstrate the impact of each code.

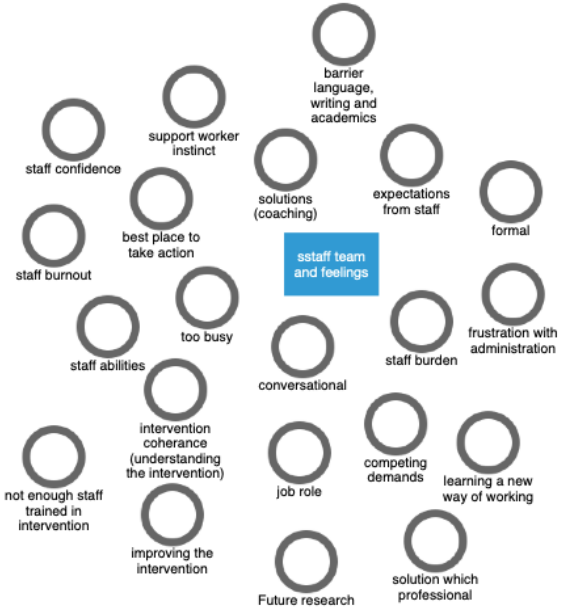


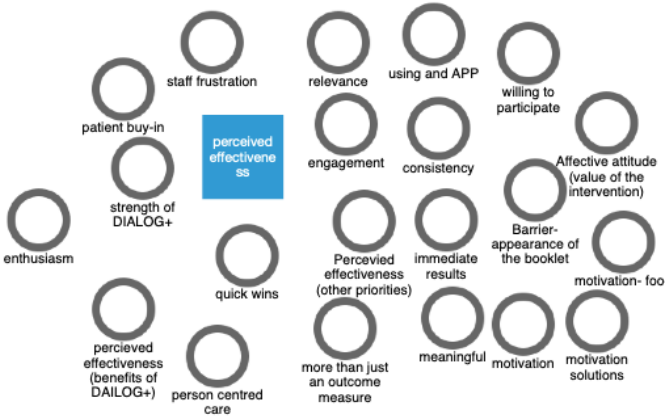

drop out
eligibility
explaining research
anonymity
recruitment (research)
follow up
Affective attitude thoughts and feelings about research
service user feedback
Affective attitude (feelings about research outcomes)
posters and leaflets

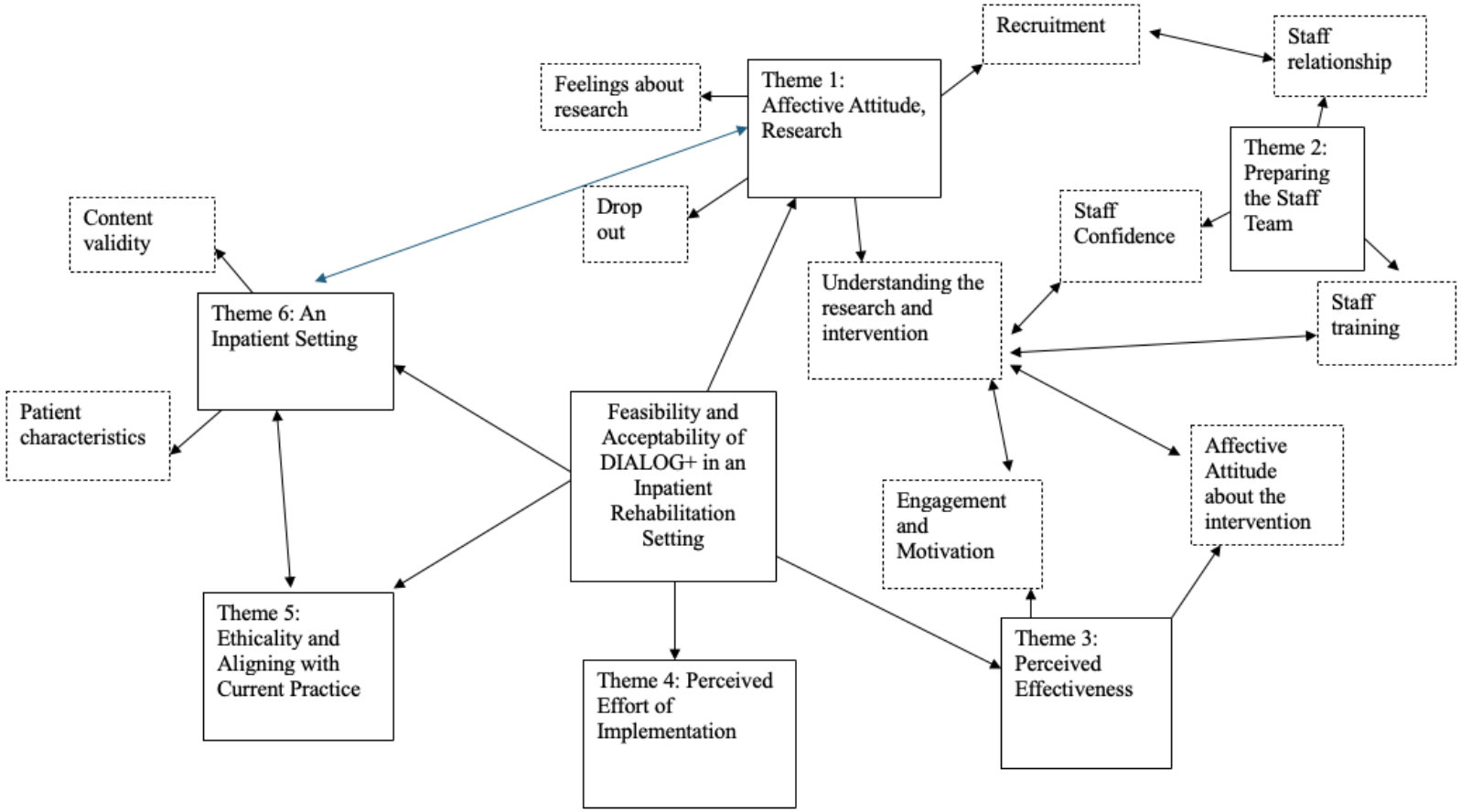

percieved effort (resource heavy)
time frame
living day to day
goal based
Adaptations to the intervention (ease of use)
question fatique
effort (too many questions)
percevied effort
routine
ease of use
improving communication
Practicality and feasibility (planning with clients)
demand heavy
percieved effort (time, duration etc)
perceived effort
presenting tasks
patient changing their mind on the day

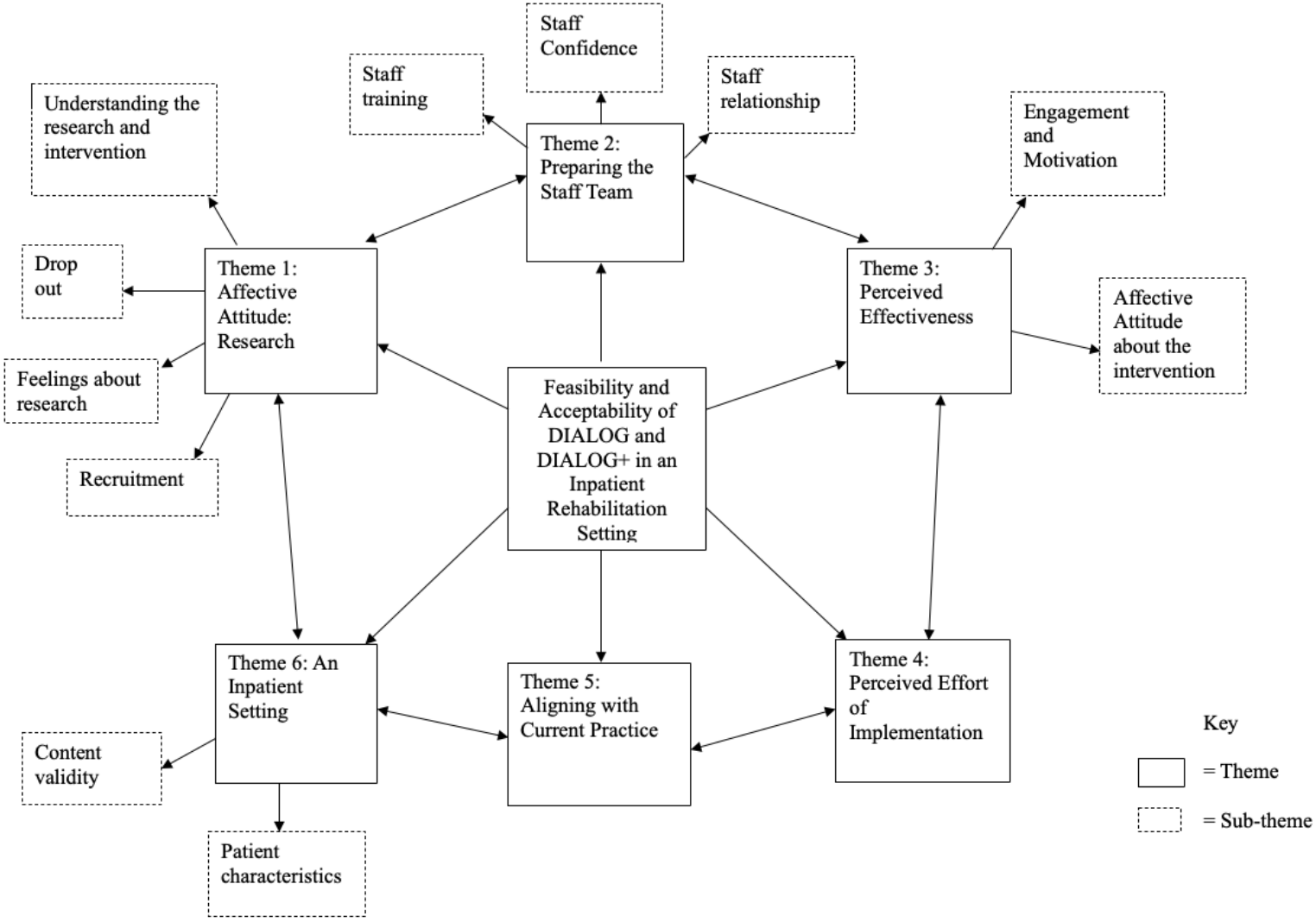

solution (allocating time)
care plans
MDT (integrating in the system)
strained system (too many changes)
audit
team dynamics
Aligning with current practice
Ethicality (aligning with current practice)
repeated questions (questions already asked in practice)
barrier (communication)
as part of admission
team systems and pathways
Barrier (staffing, rotas and organising the dairy)
solution handover
one to ones
new admissions
Solution (rotas and staffing)
overcrowded mdts
ongoing service development
seamed nurses
having a lead
ward manager
protected time
solution mini team
RIO (note taking systems)


client characteristics
recovery
severity of illness
activity planner
age of service user
private
stage of recovery
barrier- mental health
client differences
autism
client group
inpatient setting
client history impacting participation
psychosis
content validity
lack of insight
barrier (patient leave)
closed off
general barriers with client group
length of stay
Consent and capacity
culturally relevant
client preferences
Inpatient setting
expectations of rehab setting
incidents on the ward
fluctuating mental health
scale content validity
individual differences
section
paranoid thoughts
patient characteristics impacting participation


barrier language, writing and academics
staff confidence
support worker instinct
solutions (coaching)
expectations from staff
formal
best place to take action
staff burnout
sstaff team and feelings
staff abilities
too busy
conversational
staff burden
frustration with administration
intervention coherance (understanding the intervention)
competing demands
learning a new way of working
not enough staff trained in intervention
job role
improving the intervention
Future research
solution which professional


staff frustration
relevance
using and APP
willing to participate
patient buy-in
perceived effectiveness
engagement
consistency
Affective attitude (value of the intervention)
strength of DIALOG+
enthusiasm
quick wins
Percevied effectiveness (other priorities)
immediate results
Barrier- appearance of the booklet
motivation- foo
percieved effectiveness (benefits of DAILOG+)
person centred care
more than just an outcome measure
meaningful
motivation
motivation solutions


Recruitment
Staff relationship
Feelings about research
Theme 1: Affective Attitude, Research
Theme 2: Preparing the Staff Team
Content validity
Drop out
Staff Confidence
Understanding the research and intervention
Staff training
Theme 6: An Inpatient Setting
Patient characteristics
Feasibility and Acceptability of DIALOG+ in an Inpatient Rehabilitation Setting
Affective Attitude about the intervention
Engagement and Motivation
Theme 5: Ethicality and Aligning with Current Practice
Theme 4: Perceived Effort of Implementation
Theme 3: Perceived Effectiveness


Staff Confidence
Staff training
Staff relationship
Understanding the research and intervention
Theme 2: Preparing the Staff Team
Engagement and Motivation
Drop out
Theme 1: Affective Attitude: Research
Theme 3: Perceived Effectiveness
Affective Attitude about the intervention
Feelings about research
Feasibility and Acceptability of DIALOG and DIALOG+ in an Inpatient Rehabilitation Setting
Recruitment
Theme 6: An Inpatient Setting
Theme 5: Aligning with Current Practice
Theme 4: Perceived Effort of Implementation
Content validity
Patient characteristics
Key
= Theme
= Sub-theme

Appendix T

Additional Extracts for Theme Illustration

Theme	Subtheme	Additional Extracts
Affective attitude towards research Participant contributions: P1 P2 P3 P4 F2, P1,2,3,4 F1, P1,2,3,4	Understanding the research and intervention	*F2 P1: So it would be interesting to get feedback from the patients as well and how they would have found dialogue plus or the ones who did do dialogue plus, would you have preferred to just do dialogue, would you have found that more achievable or whatever.* *P 1: I don't know how much information you've got, but I think anything too busy. Or too much might lose interest, but have you got any sort of posters and or leaflets, things like that?* *P3: We have used ReQoL previously. We haven't used dialogue. Or dialogue plus so it's little bit unfamiliar territory in that way clinically.* *F1P2: Some of our clients are very poor education, have left school or not even to school, they've been kids but have been to school like but not very often, they've been kicked out of school and stuff like that and then gone into other areas and ended up in mental health and stuff.* *F2P1: It is confusing isn't it cos when I were looking at it, how am I supposed to say this to NAME – (service user) when I don't know myself, but it took me a bit of time just to get it., that's why I separated it cos it were easier.*
	Drop out	*F2P3: But maybe if you had like a, got like a couple of service users in like two or three or four, not too many at a time and sit with them as a group, I think you'd probably get a lot more out of them* *F1P3: Yes, yeah, people that will engage will be, we will follow up and we'll engage in a follow up but people who won't engage are not gonna.* *F1P2: And if they've dropped out it would be really, really difficult to get them back engaged again.*

		P2 If he'd be more willing to sort of participate in that, I think he could have seen the good in it. But I think he kind of signed up to it and then changed his mind but didn't want to say as much.
	Feelings about research	*F2P1: other people like NAME – (service user 2) was very keen to be part of research because it helps other people and it helps like future patients and blah, blah.* *P1: So you're always likely to sort of come across those issues? Although I do think it's valuable to have service users involved in and and to be participating in studies.* *P4: Because it was for research, it felt much more formal than I think it would if it was actually implemented regularly within our service as part of our offer of care.* *F2P4: So because we had the consent forms and we had to explain that they could drop out at any point. I felt that the people that I'd consented were almost daunted by that.* *F1P1:Yeah and also when I was explaining it was part of a research project, he got quite paranoid and suspicious about why we were doing it.*
	Recruitment	*F2 P4: Yeah do you know what I mean, so like some of them I knew when I asked them they went "no", do you know what I mean cos I just knew them, we just know them do you know what I mean, take time to get to know them but like we were saying like some of them like when NAME – (service user 2) were here, you know he'd do that do you know what I mean.* *F2 P3: you know it might be a good time, we've got new admissions come in, see if they would be interested in doing it.* *P2: I think you'd have the ones that wanted. It would go for it and the ones that might benefit from it, but don't really see it that way. I think, you know, they'd just be straight up no's.*

		P2: One of the guys I asked about it was straight line. No no. And I almost felt bad asking him because. Like. Knowing him as I do, I don't think he likes to reflect on his own mental health. *P4: But in terms of recruiting? I felt that it was too formal in summary, because of it being a research study and I felt that may be posed a barrier Immediately. I also recruited.*
Affective attitude towards the intervention Participant contributions: P1 P2 P3 P4 P5 F2, P1,2,3,4 F1, P1,2,3,4	Engagement and motivation	*F1P3: I used to love doing stuff like that. I'm not as enthusiastic these days, I'm coming towards end of my working life but when I were younger I'd take up on anything like that, you know, I'd plan anything like that, you know I used to be essence of care coordinator and everything for my unit but now I just, I just want a quiet carry on.* *F1P1: No there is some enthusiastic younger ones, there are, course there are. P2: There are but there's a lot in middle ground.* *F2P1: I would lose motivation to do it if I was to thinking it was just a tick box exercise and no one's actually picking this up and being like right where now,* *F1P3: Why am I doing it do you mind if I just go on ground leave or have a smoke. So you know they're more interested in going and having a cigarette than sitting down with that.realise until they've started it but when it comes to the end of it there's nothing for them, so that could be something to think about as well.* *F1P2: Yeah ReQoL and all other things that we do with them, so it's probably like sometimes I think (SIGH) sometimes I think 'oh another one, another one', do you know what I mean so, and it just depend how they are, it's how you catch them and how they're feeling at the time basically,* *F2P2: Because it is its consistency. P1: So it's really about thinking about the right person to do with the person and prioritising some time for them to do it, like a day every two weeks to do the dialogue plus with certain patients.*
	Perceived effectiveness	*F2P1: I also think people drop out if they cant see that its them fitting them so like I think with NAME – (service user 2) when we were talking about genuinely one of the examples was bins, cos*

		then I went back to him the next week and we reviewed the accommodation one and I was like oh it had been actioned as a result of this and something had been done, he was then happy to do the next weeks because there is a point to this and I can see changes and things happening, whereas I guess if they're not getting that feedback or they're not seeing anything shifting its just another tick box exercise so they will lose motivation or whatever that I imagine. *P4 Because to see the scores on the graph, certainly for dialogue plus.You need to have done a few scales, which would be about 3 months worth of work.So I wonder if it felt.That they was investing time into it, but not necessarily seeing outcomes, particularly if we weren't then discussing the scales at MDT.* *F1 P1: And that's what NAME (service user) felt as well but I just couldn't offer him anything at the end of it because there wasn't nothing there.* *P2: Not speaking from my own experience of it, but maybe some sort of false sense of why if I do this, things are going to get better and then it not being they are not the quick immediate results that they were hoping for.* *P5: Should hopefully help improve their understanding of of what the team are looking for and working towards and you know, hopefully I guess.*
	Perceived effort of implementation	*P4: It felt like a huge task. As someone involved in implementing the study, it felt huge to keep both service user and staff enthusiasm up for the period of time.* *F1P2: rather than two weeks because if they've done one and then they're doing another one in two weeks, its like (SIGH), whereas if you do one and then they've got time to even achieve the things they need to achieve, do you know what I mean.* *P2: There's a lot of repetition for these guys, you know, and there's a lot of how are you questions and you know, how's this? How's that? Where's your mood? You know, if you are you experiencing anything?*

		P3: keep requesting that every so often in different ways. You know, approach it in different ways until you you get that collaboration. Sometimes no matter what you try, you don't. And that's when you know you exhausted it all and you leave it for some time and then try again after a few weeks or yeah. *F2 P1: And I do think like there's something about the practicality of it isn't there, like its too much really for one person to do all of them with one person but then also its easier for one person to do them all cos its where you left off.* *F2P2 It is confusing isn't it cos when I were looking at it, how am I supposed to say this to NAME – (service user) when I don't know myself, but it took me a bit of time just to get it., that's why I separated it cos it were easier.* *F2 P1: I mean they could make it much easier couldn't they, they could have one question page on the App, you just tick and then it goes to the next one. P2: Yeah one question at a time, yeah.*
Aligning with current practice Participant contributions: P1 P2 P3 P4 P5 F2, P1,2,3,4 F1, P1,2,3,4	Team systems	*P 1: So when you've not got consistent people in Mdt's, it's hard to start to embed things. But I think it's something that we should continue to try and embed.* *P3 Every two weeks, they're supposed to have once a week one to one sessions. So it can easily be achieved as part of that, we will just need the word managers and other tasks to buy into that.* *P 1: Because if you and (NAME staff) weren't sort of pushing and planning, it might be something that falls. Aside, so I think if there was a lead on it, making sure that the meetings were planned in with the staff that were trained. And keeping an eye on that, I think that would really help it.* *P4: Or care planning around it. There's a lot about the framework that would support implementation of dialogue, plus to be able to move forward with the things that are coming from the study to be able to actually conduct the work.* *F2P2: Yeah, again it goes down to your mini teams as well and probably meetings with them and you know getting together. Sometimes MDT, its you're putting that much on it, that's why its taking*

		a long time but things like this are relevant just to be on the ward base and just for your mini teams and to take that lead if you will, end this and the scoring on it maybe and bits what need to be improved. *P2: Easy to integrate, easy to sort of bring in. I mean we, you know, we're already sort of completing daily progress notes on people and you know, we will include interactions that we had, you know signs where they, you know said something positive or said something that would worry us* *P5: Yeah. So I mean, as we did sort of like another quality improvement project with the national Mental Health Service. And I think part of our quality work in terms of sort of culture of care. Because I think the MDT isn't the problem.* *F2P2: MDT's is too busy I think to be.. P1: Its too busy and its P2: And it can go on and on for hours.* *F1 P4: Unless we did like we said, introduce it on admission saying something that we have to complete while you're an inpatient and every two weeks or whatever..*
	Staff training (cost)	*P4: New starters and ensuring that they could seek support was needed, and that also the right people are supervising them, I suppose as well.* *P1: I think obviously like we said before, with with more staff training on all disciplines* *F1P2: I got a little bit of help from NAME – (staff) as well, NAME – (staff) went through it with me, it were just the scoring weren't quite sure on it really to be fair,* *P5: So I think it's been helpful just having that one person as the anchor for that and hopefully that information will get disseminated through the bite sized teaching. Discussions.*

		F1 P2: If everybody, because NAME- (male staff) was trained in it, then any given day, a good day regardless of who is around, somebody can pick it up and run with it, but if there's only a small number around, it won't get done will it. *P2: Just getting staff more aware of it, maybe a little bit more. Staff, anything that I've not covered that you thought that you wanted to provide feedback on and we've not said that.* *F2P2: The scoring I found difficult but doing the questions erm after a bit you kind of got into it and it were good*
	Staff confidence (intervention coherence)	*F1P4: Cos as well some of our staff might be a little bit intimidated doing that or not feel confident enough to do that with a client because we're all individuals as well aren't we, I know like we're professionals and we're support workers and.. People aren't confident and like even coming to work with this little forum here, people would come, especially with you taping it, some people would be totally intimidated and not be able to speak up and you know, so that as well.* *F2P1: I thought he were gonna bite my head off (LAUGH). I was like "alright NAME – (service user) I'm sure you've been asked these questions before", "what are you asking me that for, what do you mean, I wouldn't be in here if I hadn't got mental health problems, of course mental health is bad", he's saying (LAUGH), and then I mentioned physical health, is there a physical health.* *P4: And we found that there were staff that were potentially under confident and then we didn't have the resources to support them. And I wonder if we'd have had more resource in terms of time allocated to the participants, whether or not we would have had better outcomes with the. Staff feeling confident to conduct the measure independently.* *P2: guess a lot of the service users see support workers as sort of the nearest to them. They've not got the sort of qualifications and the sort of the training and you know, so they tend to sort of reflect towards like as peers more. So maybe psychology OT might be the better people delivering.*

Inpatient setting Participant contributions: P1 P2 P3 P4 P5 F2, P1,2,3,4 F1, P1,2,3,4	Inpatient setting	*F1 P4: So he's not, so all our clients aren't all just in here, they're all on different sections and different sections mean different things.* *P1: It will. I suppose it because we're very goal based in terms of rehabilitation. It's something that can open up those discussions on sort of a person's wants and wishes for the future. So I think it fits really well into sort of the recovery model of rehab* *P5: suppose there's some people who don't want to work with us. In general. But normally they tend not to be the type of clients we work with anyway. I I I don't think that being an impatient should be a a barrier towards. I think it's helpful for us 'cause we can sort of monitor change and wanted to progress. I think it's specifically helpful for for rehab as well, just in terms of that sort of, I guess it helps build that relationships with with patient* *F2P2 I'm still concerned that the fact there's one short on our ward because one of service users is not well, so again its making sure that's everything's, you know safe on your ward to do that and have that time so you're not worried, that's probably a thing to think about cos its not fair on the staff as well, we get burnt out, everybody gets burnt out so when you've got stressors on a ward as well, that's probably a barrier for this.* *P3: Well, as a rehab service and given the luxury of time and resource that we have, although it's fast dwindling given the NHS pressures in general. Yesterday I heard that they might be even more cuts coming through to our service.*
	Patient characteristics	*F2P2: I think NAME – (service user) is quite a well person and I do think like for some other patients who aren't as well as NAME – (service user) they would really, really struggle with that, I just think its too much.* *P4: I think one of the things was that people were at different stages of their recovery, so some people were prepping for discharge, so they were taking extended periods of leave. We had a couple of service users discharged during the study.*

		F2P3: There's also negative effects isn't there, erm I'm trying to word it a bit better but I can't cos my head won't work. Erm negative symptoms isn't it which is poor motivation and whatever in't it, poorly motivated. We are getting more people that are on the autistic spectrum, so for them we… *P3: Are experiencing some negative symptoms of schizophrenia because that is. A major diagnosis we work with on the ward, so schizophrenia is by far the biggest group of patients that that we have. And of which some of them have negative symptoms, which means they're not as motivated. To carry out the basic activities of daily life, never mind doing a dialogue plus so that can be another barrier.* *F1 P4: Yeah and what you have to take into consideration as well, so like an admission pack has to be done within 72 hours. Is it ever done in 72 hours, no because our patients cant, can't do all like the admission pack and all the work that needs to be done on admission in 72 hours, so and it can be still going on weeks after cant it, you know certain topics, certain items of the admission the pack still might not be done and have got to be done and then you've got the powers that be saying "this was not done" but the because…*
	Content validity	*F2 P2: Cos like you say number 1 and number 2 is like totally satisfied and then very. Its same thing. P4: It is yeah. P2: It is to us so.. P4: So 1, 2 in middle and then 3.* *F2: Its either good for some people and it's not for others but again we have erm a lot of different cultures here as well and that's why I said a different, a completely different working scale, the same but and easy learning one or an easy read one maybe or erm so we've got, if were all doing this as like a research it might have to be for different people as well and different cultures erm because that wouldn't have done for some people that we've got.* *P1: There are some quite difficult questions for service users and I know that if they don't get actually if they if somebody just gives a service user the outcome measure* *P4: We noted during the training that there were a couple of questions that weren't necessarily fitting for our setting. Maybe one about the job situation and certainly the one about*

		accommodation and we struggled to be able to think about what our blanket application of that would be and for someone. *F2P3: I mean I'm just confusing people even more now, like so when we've asked them these questions, if like rather than giving like yes that is that, what if like you could just ask them just a brief description of how they feel about that and then write that down instead, you know like instead of having to do all that bit. 'how's your mental health', so rather than saying 1, 2, 3, well they could say yeah 1, 2, 3 or 4 and then 'are you okay to explain on that' and then you know do that, you know "my mental health's like this, blah de blah". I don't know.*

Appendix U

Patient Questionnaire Results

Question	DIALOG	DIALOG+
Did you like or dislike DIALOG?	No opinion= 1 Strongly like= 1	Like= 1
How much effort did it take to engage with DIALOG?	No effort at all = 2	A lot of effort= 1
Do you think there are ethical consequences to using DIALOG?	Strongly disagree= 1 Disagree =1	No opinion =1
DIALOG has improved your clinical outcome?	Disagree= 1 Agree= 1	Agree= 1
It is clear to me how DIALOG will help improve clinical outcomes	Strongly disagree= 1 Agree= 1	Agree= 1
How confident did you feel about engaging with DIALOG?	Confident =2	Confident= 1
DIALOG interfered with my other priorities	Strongly disagree= 2	No opinion= 1
How acceptable was DIALOG to you?	Acceptable= 1 Completely acceptable= 1	Completely acceptable = 1