Contents

Original Papers

Identification of Predictors of Mood Disorder Misdiagnosis and Subsequent Help-Seeking Behavior in Individuals With Depressive Symptoms: Gradient-Boosted Tree Machine Learning Approach (e50738)
Jiri Benacek, Nimotalai Lawal, Tommy Ong, Jakub Tomasik, Nayra Martin-Key, Erin Funnell, Giles Barton-Owen, Tony Olmert, Dan Cowell, Sabine Bahn. ................................................................. 2

Understanding Public Perceptions of Virtual Reality Psychological Therapy Using the Attitudes Towards Virtual Reality Therapy (AVRT) Scale: Mixed Methods Development Study (e48537)
Aislinn Gomez Bergin, Aoife Allison, Cassie Hazell. ................................................................. 15

Research Letter

The Frequency of Design Studies Targeting People With Psychotic Symptoms and Features in Mental Health Care Innovation: Secondary Analysis of a Systematic Review (e54202)
Lars Veldmeijer, Gijs Terlouw, Jim Van Os, Job Van ’t Veer, Nynke Boonstra. ................................................................. 27
Identification of Predictors of Mood Disorder Misdiagnosis and Subsequent Help-Seeking Behavior in Individuals With Depressive Symptoms: Gradient-Boosted Tree Machine Learning Approach

Jiri Benacek¹, PgDip; Nimotalai Lawal¹, MEng; Tommy Ong¹, MEng; Jakub Tomasik¹, PhD; Nayra A Martin-Key¹, PhD; Erin L Funnell¹,², BSc; Giles Barton-Owen², MA; Tony Olmert¹, MD; Dan Cowell², BSc; Sabine Bahn¹,², MD, PhD

¹Department of Chemical Engineering and Biotechnology, Cambridge Centre for Neuropsychiatric Research, University of Cambridge, Cambridge, United Kingdom
²Psyomics Ltd, Cambridge, United Kingdom

Corresponding Author:
Sabine Bahn, MD, PhD
Department of Chemical Engineering and Biotechnology
Cambridge Centre for Neuropsychiatric Research
University of Cambridge
Philippa Fawcett Drive
Cambridge, CB3 0AS
United Kingdom
Phone: 44 1223334151
Email: sb209@cam.ac.uk

Abstract

Background: Misdiagnosis and delayed help-seeking cause significant burden for individuals with mood disorders such as major depressive disorder and bipolar disorder. Misdiagnosis can lead to inappropriate treatment, while delayed help-seeking can result in more severe symptoms, functional impairment, and poor treatment response. Such challenges are common in individuals with major depressive disorder and bipolar disorder due to the overlap of symptoms with other mental and physical health conditions, as well as, stigma and insufficient understanding of these disorders.

Objective: In this study, we aimed to identify factors that may contribute to mood disorder misdiagnosis and delayed help-seeking.

Methods: Participants with current depressive symptoms were recruited online and data were collected using an extensive digital mental health questionnaire, with the World Health Organization World Mental Health Composite International Diagnostic Interview delivered via telephone. A series of predictive gradient-boosted tree algorithms were trained and validated to identify the most important predictors of misdiagnosis and subsequent help-seeking in misdiagnosed individuals.

Results: The analysis included data from 924 symptomatic individuals for predicting misdiagnosis and from a subset of 379 misdiagnosed participants who provided follow-up information when predicting help-seeking. Models achieved good predictive power, with area under the receiver operating characteristic curve of 0.75 and 0.71 for misdiagnosis and help-seeking, respectively. The most predictive features with respect to misdiagnosis were high severity of depressed mood, instability of self-image, the involvement of a psychiatrist in diagnosing depression, higher age at depression diagnosis, and reckless spending. Regarding help-seeking behavior, the strongest predictors included shorter time elapsed since last speaking to a general practitioner about mental health, sleep problems disrupting daily tasks, taking antidepressant medication, and being diagnosed with depression at younger ages.

Conclusions: This study provides a novel, machine learning–based approach to understand the interplay of factors that may contribute to the misdiagnosis and subsequent help-seeking in patients experiencing low mood. The present findings can inform the development of targeted interventions to improve early detection and appropriate treatment of individuals with mood disorders.

(JMIR Ment Health 2024;11:e50738) doi:10.2196/50738
KEYWORDS
misdiagnosis; help-seeking; gradient-boosted trees; machine learning; depression; bipolar disorder; diagnose; diagnosis; mood; mental health; mental disorder; mental disorders; depression; depressive; predict; predictive; prediction; depressed; algorithm; algorithms

Introduction

Mood disorders are debilitating psychiatric conditions that negatively affect a person’s emotional state. They result in impaired ability to function and complete daily tasks, and an increased risk of self-harm and suicide [1]. Two of the most common mood disorders are major depressive disorder (MDD) and bipolar disorder (BD), which affect approximately 3.4% and 0.5% of the global population, respectively, at any given time [2]. Beyond the impact on the affected individuals, there are also economic and social consequences such as lost productivity, increased health care costs, and costs incurred by unpaid carers. In the United Kingdom alone, the economic burden of managing MDD and BD is estimated at £7.5 billion (US $9.55 billion) and £5.2 billion (US $6.62 billion), respectively [3], with a significant portion of this burden attributed to underdiagnosis and high rates of misdiagnosis of mood disorders.

Although misdiagnosis is prevalent in all areas of medicine, the heterogeneous nature of mental illness and lack of objective diagnosis make it more common for mental health conditions [4]. The diagnosis of mental health disorders is currently based on assessing patient symptom profiles using diagnostic manuals such as the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5) [5] or the International Statistical Classification of Diseases and Related Health Problems, 11th Revision (ICD-11) [6]. As such, diagnosis relies heavily on symptom reporting and patients who do not recognize and thus do not report their symptoms or present with complex symptoms are more likely to be misdiagnosed [7]. For example, issues with symptom reporting are considered a major cause of BD misdiagnosis [8], with many patients with BD only seeking medical help during depressive episodes [9], which makes mania more difficult to identify. Consequently, as many as 78% of mood disorder diagnoses are missed in primary care [10], including approximately 40% of patients with BD who are initially misdiagnosed with MDD [11]. This, in turn, leads to incorrect treatment of BD with antidepressants which have lower efficacy than mood stabilizers in alleviating bipolar symptoms and have been associated with prolonged episodes of mania and accelerated cycling between manic and depressive states [12,13]. Understanding factors that lead to misdiagnosis could guide the development of more effective means for early identification and intervention in individuals at high risk.

An additional barrier to receiving a correct diagnosis and necessary care is the reluctance of affected individuals to speak to medical professionals about their mental health. The European Study of the Epidemiology of Mental Disorders carried out across 6 countries found that only 25.4% of respondents spoke to a medical professional about their mental health problems [14]. Likewise, active engagement with mental health services is consistently low, with almost 75% of patients experiencing a mental illness in England receiving no treatment [15]. One of the reasons for the low rates of help-seeking in individuals experiencing mental health symptoms is concerns of potential public and self-internalized stigma. Consequently, individuals struggling with their mental health often turn to coping mechanisms such as social withdrawal, secrecy, and label avoidance [16,17] rather than seeking help [18]. Therefore, it is imperative to recognize barriers to help-seeking in mental health to facilitate early and accurate diagnosis in un and misdiagnosed individuals.

Although previous studies have investigated factors contributing to the misdiagnosis, poor help-seeking behavior, and barriers to receiving a diagnosis, only a few have used machine learning methods to do so [19]. The use of machine learning in mental health research has increased in recent years, with many studies focusing on detection and diagnosis, treatment and support, public health, and research and clinical administration [19]. While not without limitations, the use of machine learning can offer data-driven insights into complex relationships between high-dimensional data [20,21]. Although other, mostly qualitative investigations have identified the predictors of help-seeking and misdiagnosis by considering factors individually, this study aims to take a more holistic approach. By developing machine learning models based on extensive self-reported patient data, we aim to identify and quantify interdependent predictive factors for the misdiagnosis of mental health disorders, specifically mood disorders, and help-seeking behavior in individuals who may have been misdiagnosed. Identifying such predictive factors could aid in avoiding preventable misdiagnosis, encourage help-seeking, and improve outcomes in patients presenting with depressive symptoms.

Methods

Data Acquisition

Overview

The data used in this report were collected as part of the Delta Study—a study aiming to facilitate a more accurate and earlier diagnosis of BD and MDD; carried out in the United Kingdom by the Cambridge Centre for Neuropsychiatric Research between 2018 and 2020 [22,23]. The study consisted of an adaptive digital questionnaire, the Composite International Diagnostic Interview (CIDI) [24], and 2 follow-up questionnaires at 6 and 12 months. The stages of the Delta Study are summarized in Figure 1. Participants were recruited nonrandomly through email, the Cambridge Centre for Neuropsychiatric Research (CCNR) website, and paid Facebook advertisements. The eligibility criteria included at least mild depressive symptoms, indicated by a score of ≥5 on the Patient Health Questionnaire-9 (PHQ-9) [25] at the time of recruitment, aged between 18 and 45 years, and residency in the United Kingdom. Participants who indicated current suicidal ideation or intent, were pregnant, or breastfeeding, were excluded.
Adaptive Digital Questionnaire

In total, 3232 participants completed the adaptive digital questionnaire available on the Delta Study digital platform. The questionnaire consisted of 635 questions, divided into six sections: (1) demographic information and personal history; (2) manic and hypomanic symptoms; (3) depressive symptoms; (4) personality profiling; (5) treatment, medication, substance use, and family psychiatric history; and (6) other psychiatric conditions. As the questionnaire was adaptive to answers given by participants, the maximum number of questions an individual could answer was 382, with an average of 284. Within the questionnaire, participants reported their baseline diagnosis, and their current well-being (within the previous 14 days) was quantified using the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) [26].

Composite International Diagnostic Interview

Participants who completed the web-based mental health questionnaire were invited to complete the CIDI version 3.0 via telephone. The CIDI is a structured diagnostic interview for mental disorders created by the World Health Organization based on the International Classification of Diseases and Related Health Problems, 10th Revision (ICD-10). It was developed primarily for epidemiological studies and has been extensively validated, demonstrating high diagnostic reliability [27]. In this study, only sections pertaining to mood disorder diagnoses were applied, that is, the demographics, depression, and mania modules. Interviewers were trained by CIDI-certified instructors prior to conducting the interviews. In total, 924 participants completed the CIDI and received one of the following diagnoses in their results report: BDI, BDII, subthreshold BD, MDD with subthreshold BD, MDD, or no mood disorder diagnosis (referred to as “low mood”).

Follow-Up Questionnaires

Participants who completed the digital questionnaire were invited to fill out 2 follow-up questionnaires, 6 and 12 months after receiving their results report. The follow-up questionnaires aimed to determine the effects of participation in the Delta Study on participants’ quality of life and record subsequent changes in diagnosis and treatment. A total of 2064 participants
completed at least 1 of the follow-up questionnaires, with 1780 respondents at 6 months and 1542 respondents at 12 months.

Outcomes

Overview

For the purposes of this study, 2 dependent variables were defined.

Misdiagnosis

For participants who completed the CIDI, the mood disorder diagnosis reported at baseline was compared to the diagnosis obtained from the CIDI, including patients with no mood disorder diagnosis at baseline who should have been diagnosed. CIDI diagnosis was used as the gold standard, and any mismatch with the baseline diagnosis was defined as misdiagnosis. This definition of misdiagnosis was consistent with previous studies investigating under- and misdiagnosis of mood disorders based on comparing patient-reported diagnoses to the outcomes of structured clinical interviews [28,29].

Help-Seeking Behavior

In the 6- and 12-month follow-up questionnaires, participants were asked: “Have you had an appointment with a GP or psychiatrist to talk about your mental health in the past 6 months?” A positive response to this question at either time point was defined as help-seeking. In order to examine help-seeking in misdiagnosed individuals, only those who were identified as misdiagnosed within outcome 1 were included in the analysis.

Analysis

Overview

Raw data processing and feature engineering were performed in R (version 3.6.3; R Core Team) [30]. Subsequent analyses and modeling were carried out using Python (version 3.9.7; Python Software Foundation) [31]. Main libraries used included Pandas (version 1.5.2; Pandas Development Team) [32] and NumPy version 1.23.5 [33] for data manipulation; scikit-learn version 1.0.2 [34], XGBoost (version 1.6.1; The XGBoost Contributors) [35], and SHAP version 0.41.0 [36] for modeling and interpretation; and Seaborn version 0.12.1 [37] and Matplotlib version 3.6.2 [38] for plotting.

Data Preparation

Prior to analysis, constant and duplicate variables were removed. Answers to questions examining the same symptom or construct were concatenated, and new features were created to represent these aggregated answers. Missing data were imputed where possible (for example, the answer to the question asking “Has anyone suggested you drink less?” was set to 0 for participants who had indicated they do not drink), and otherwise remained nonrandomly missing. Categorical variables were 1-hot encoded, that is, unique dummy variables were created where the presence of each category was denoted by “1,” and its absence was represented by “0.”

Modeling and Interpretation

This analysis aimed to develop predictive models to identify variables influencing (1) misdiagnosis and (2) help-seeking behavior in participants who were identified as potentially misdiagnosed. A decision tree–based machine learning algorithm Extreme Gradient Boosting (XGBoost) [35] was chosen to train the classification models due to being robust to outliers, agnostic to data distribution, having the ability to handle nonrandom missing data, offering good predictive power, and due to its allowing for good model interpretability. Repeated nested cross-validation (rNCV) was used for model training and evaluation to obtain accurate estimates of model performance in unseen data. rNCV relies on performing a k-fold cross-validation (CV) within each round of another CV. This allows for model-specific hyperparameter optimization in the inner loop, with the final model being trained using the best-performing set of parameters, and later evaluated in the outer loop of rNCV. For this analysis, a 4-fold stratified CV was used in both the inner and outer loops, where 3 of the folds acted as a training set and 1 as a test set. Tuned model parameters included the number of estimators (1 to 100), shrinkage rate (0.1 to 0.3) to prevent overfitting, and tree depth (1 or 2) to allow for first-order interactions between predictors. The training was repeated 100 times, generating a total of 400 models for each of the objectives. Generalized model performance was evaluated by calculating the area under the receiver operating characteristic curve (AUC). The classification cutoff was optimized for the Youden index [39] to balance the true positive and true negative rates and offset potential imbalances between classes. SHAP (Shapley additive explanations) analysis [36], which combines local interpretable model–agnostic explanations (LIME) [40] and Shapley sampling values [41] approaches, was used for model interpretation. Feature occurrence frequency was calculated as the percentage of the models that incorporated a given feature to generate predictions. Reported results represent mean and SD values across the rNCV models.

Ethical Considerations

The study protocol was approved by the University of Cambridge Human Biology Research Ethics Committee (HBREC 2017.11) and all enrolled participants signed a digital informed consent form.

Results

Misdiagnosis

The self-reported baseline diagnosis did not match the diagnosis assigned by CIDI for 471 (50.97%) of the 924 participants who completed the CIDI interview. These participants were therefore considered misdiagnosed. No between-group differences were observed in terms of age, sex, ethnicity, highest achieved education level, or relationship status between the correctly diagnosed and misdiagnosed groups (Table S1 in Multimedia Appendix 1). However, there were significant differences in employment status as well as well-being and PHQ-9 scores, with misdiagnosed individuals, on average, reporting lower well-being and more severe depressive symptoms.

On average, the models correctly classified 70% (SD 9%) of misdiagnosed participants and 71% (SD 9%) of correctly diagnosed participants, with a mean accuracy of 70% (SD 3%) and the out-of-fold AUC of 0.75 (SD 0.03; Figure 2 and Table...
S2 in Multimedia Appendix 1). Among the 1045 variables evaluated, the strongest predictors of misdiagnosis were more severe composite depressive symptoms and unstable self-image (Figure 3). Unstable self-image was measured by a 4-level Likert scale question “Is your image and sense of yourself and what you believe in unstable and constantly changing?” The next strongest predictor was the diagnosing clinician, with those who were undiagnosed at baseline or reported a diagnosis by a psychiatrist more likely to be misdiagnosed. The top 10 predictors also included variables related to age at diagnosis of BD and MDD, with late (≥35 years of age) diagnosis or no diagnosis at all, increasing the likelihood of being misdiagnosed (Figure S1 in Multimedia Appendix 1). Misdiagnosed participants were also more likely to recklessly spend money, experienced more frequent intense mood swings or mania in general, had higher weight gain during low mood episodes, and were more sexually active than usual at the time of data collection.

Figure 2. Out-of-fold model performance in predicting misdiagnosis. Green lines represent predictive performance on unseen out-of-fold data for each of the 400 final models. The thick blue line represents the average of all ROC curves. The grey area represents 1 SD. AUC: area under the receiver operating characteristic curve; ROC: receiver operating characteristic.
Model performance was largely driven by the top 5 predictors, with a steady decline in SHAP scores for subsequent variables. Of the top 10 predictors, 9 were selected in more than 75% (n=300) of the models, suggesting a relatively stable model composition. The exception was a variable related to “being more sexually active than usual,” which was selected in 71% (n=284) of the models. More detailed information on feature selection frequency is provided in Figure S3 in Multimedia Appendix 1.

**Help-Seeking Behavior**

Help-seeking behavior was investigated in 379 participants who were misdiagnosed at the baseline and who had completed at least 1 of the follow-up questionnaires. Of those, 229 (60.42%) participants sought an appointment with a medical professional during the follow-up period to discuss their mental health and were therefore defined as “help-seekers.” The help-seeker and non–help-seeker groups differed significantly in the highest achieved education level, relationship status, well-being, and severity of depressive symptoms (Table S3 in Multimedia Appendix 1). Participants more likely to seek help were on average less formally educated, more likely single, reported higher mean severity of symptoms, and worse overall well-being.

The model achieved an AUC of 0.71 (SD 0.04; Figure 4), with a sensitivity of 65% (SD 13%), specificity of 72% (SD 13%), and average accuracy of 67% (SD 4%; Table S4 in Multimedia Appendix 1). The strongest predictor was the shorter time since patients last spoke to a general practitioner (GP) about their mental health at baseline (Figure 5). It was followed by sleep problems disrupting daily tasks and taking prescribed antidepressants, both associated with increased help-seeking. Consistent with this, lower help-seeking was observed in participants who had never been prescribed antidepressants, namely selective serotonin reuptake inhibitors (SSRIs). Furthermore, there was a lower likelihood of help-seeking with higher age at both the first episode of low mood and diagnosis of depression, which was similarly predictive to not having been previously diagnosed with depression. Finally, impaired ability to work, lower well-being scores, feeling worthless, and lower self-rated mental health were associated with help-seeking behavior.
**Figure 4.** Out-of-fold model performance in predicting help-seeking. Green lines represent predictive performance on unseen out-of-fold data of each of the 400 final models. The thick blue line represents an average of all ROC curves. The grey area represents 1 SD. AUC: area under the receiver operating characteristic curve; ROC: receiver operating characteristic.

**Figure 5.** Results for top 10 variables in the help-seeking model in misdiagnosed individuals. Shown is feature SHAP importance (in brackets) and feature SHAP values (data points). SHAP values below 0 show directionality toward low help-seeking (ie, no appointment with GP or psychiatrist to discuss mental health), whereas values above 0 show directionality toward high help-seeking. GP: general practitioner; SHAP: Shapley additive explanations; SSRI: selective serotonin reuptake inhibitor.
The 3 variables, namely, time since last spoken to a GP, sleep problems disrupting daily tasks, and still taking prescribed antidepressants, were selected in nearly all models (Figure S4 in Multimedia Appendix 1), suggesting their high relevance for model predictions. Among other predictors, only age when diagnosed with depression was selected in more than 75% (n=300) of models, with the remaining features only selected in approximately 50% (n=200) of models, indicating their lower relevance.

**Discussion**

**Principal Findings**

This study aimed to develop machine learning models to explore factors potentially contributing to misdiagnosis and subsequent help-seeking in individuals experiencing low mood. For this purpose, we used data obtained through an extensive digital questionnaire concerning demographic, personality, and mental health data, as well as, the validated and standardized diagnostic interview, CIDI. Developed models achieved a fair level of predictive power, with AUCs of 0.75 and 0.71 for predicting misdiagnosis and help-seeking, respectively. Below, we discuss the main findings as well as the strengths and limitations of this analysis.

**Misdiagnosis**

The strongest predictor of misdiagnosis was the severity of depressed mood, with more severe depressive symptoms being associated with a greater risk of being misdiagnosed. This directionality was consistent with other top predictors of misdiagnosis, including unstable self-image, reckless spending, frequent intense mood swings, mania, weight gain during low mood, and being more sexually active than usual. Except for the instability of self-image, these predictors can be divided into either depression or mania or bipolar-related symptoms. Overall, the finding that individuals with more severe mental health symptoms are at a greater risk of being misdiagnosed is surprising, given the opposite could be expected as milder symptoms are harder to detect [42]. Several factors could contribute to this association, including the complexity of diagnosing mental health disorders [43], variability in symptom presentation [44,45], and the high degree of symptom overlap across different diagnoses [5]. A possible explanation for the increased risk of misdiagnosis among individuals with more severe symptoms is that they may present with prominent mood instability, such as that observed in patients with personality disorder, or rapidly cycling symptoms, making accurate diagnosis more challenging [9]. In addition, individuals with more severe symptoms often lack motivation to seek help, hence their symptoms may remain unrecognized for a longer time [46].

In the case of mood disorders, misdiagnosis of individuals with higher depressive symptom severity may result from the fact that patients with BD generally seek medical help during depressive episodes and often present with more severe depressive symptoms than patients with MDD, while underreporting manic phases [47,48]. In fact, less than a third of patients with BD report the presence of reckless behavior, excessive spending, and increased sexual interest or activity [49]. This contributes to approximately 40% of patients with BD receiving an incorrect initial diagnosis of unipolar depression [50]. Also, the association of frequent intense mood swings with mood disorder misdiagnoses may be related to incorrect treatment of depressive symptoms of BD with antidepressants, rather than mood stabilizer medication, which has the potential to induce mania and rapid cycling [51,52].

The second most predictive feature of misdiagnosis identified in this study was unstable self-image. Previous literature has shown that an unstable sense of self is associated with frequent changes in diagnosis, and often linked to complex and unstable personality characteristics [53]. The high ranking of self-image stability could, however, be a result of the high comorbidity rates between BD and other disorders featuring unstable self-image that were not evaluated by the diagnostic interview used in this study, such as borderline personality disorder [54]. This is especially important considering that such disorders may share a high number of similarities with BD, leading to frequent misdiagnoses [55,56]. The 2 additional symptoms that are ranked high in terms of predictive value for misdiagnosis in this analysis regard reckless spending and increased sexual activity, representing reckless or impulsive behavior, which are included in the diagnostic criteria of both BD and borderline personality disorder [17].

Finally, among the top predictors of misdiagnosis were 3 variables related to psychiatric history, including psychiatrist involvement in the diagnosis, age at depression diagnosis, and age at BD diagnosis. Interestingly, the models attributed a higher risk of misdiagnosis to individuals whose depression was diagnosed by a psychiatrist. This may be caused by the fact that patients at high risk of misdiagnosis, such as those with more complex symptom presentation or suspected comorbidities, are usually referred to secondary care, following the National Institute for Health and Care Excellence (NICE) guidelines [57]. However, this finding should be interpreted with caution, as diagnoses made by psychiatrists are generally more accurate than those derived from the CIDI. Also, participants who received a diagnosis of a mood disorder at an older age, or not at all, were more likely to be misdiagnosed. This finding is surprising, as previous literature suggests that the severity and impact of symptoms decline with age, with 86% of patients with BD diagnosed by the age of 25 [58]. However, it is possible that due to milder symptoms, patients who are older may remain undiagnosed for longer periods of time.

**Help-Seeking**

Analysis of participants with a mismatch between their self-reported formal diagnosis and the CIDI outcome revealed several predictors of help-seeking related to patients’ mental health history and symptoms.

The most predictive feature was time since last spoken to a GP at baseline, with patients who had visited their GP more recently being more likely to seek help. Interestingly, that was not the case for the time since last spoken to a psychiatrist, likely due to most participants not being under secondary care and the long waiting times for psychiatric assessment [59]. In line with previous literature [60], these findings indicate that help-seeking was also associated with more severe psychiatric symptoms and...
having a previous diagnosis of mood disorder. Similarly, participants seeking help reported lower well-being, feeling more worthless, and more functional impairment in carrying out daily tasks and at work caused by symptoms and sleep problems.

Interestingly, while there was not a significant overall age difference between the help-seekers and non–help-seekers, further analyses showed a lower tendency to seek help in individuals who were over 35 years old at initial diagnosis of depression (Figure S2 in Multimedia Appendix 1). The pattern of people who are younger being more likely to seek help is in line with the published literature [61]. Together with the finding that the initial diagnosis at older age was a strong predictor of misdiagnosis [62], this result indicates that patients who are most likely to be misdiagnosed are also the least likely to seek help. Thus, older patients may require more support to tackle potential barriers to help-seeking and receiving a diagnosis, such as stigma and inadequate mental health education [63].

The final set of predictors of help-seeking was related to previous medication. Higher help-seeking was observed in misdiagnosed individuals who were still taking previously prescribed antidepressants, in particular SSRIs, as opposed to misdiagnosed individuals who either had never been prescribed SSRIs or other antidepressant medication or had stopped taking it. The association of antidepressant treatment with help-seeking indicates that the prescribed medication may have been ineffective, as is often the case when attempting to treat depressive episodes of BD with antidepressant monotherapy [64]. Compared with the patients with MDD, the patients with BD respond worse to antidepressant medication, with short-term nonresponse rates of 51.3% in BD versus 31.6% in MDD [65]. This difference is even more pronounced in the long-term, where the loss of response to antidepressants is 3.4 times more frequent in patients with BD, while withdrawal relapse into depression is 4.7 times less frequent in BD compared to patients with MDD [65]. Moreover, individuals with unrecognized BD who are treated with antidepressants sometimes develop symptoms of mania, which in turn may motivate patients or their relatives to seek consultation with a specialist [66].

**Limitations**

The main limitation of this study is the reliance on CIDI as the gold standard for mood disorder diagnosis. Although the CIDI demonstrates good agreement with structured diagnostic interviews conducted by clinicians [67], future studies should consider either retrospective or longitudinal study designs, and ideally access medical records for more accurate diagnoses, including those beyond mood disorders. Additionally, the study participants were recruited online following strict inclusion criteria and were predominantly White, necessitating further research in traditionally underrepresented ethnic minorities and more representative patient cohorts. Another limitation is the exclusion of individuals with current suicidal ideation, a characteristic that could be an important indicator of misdiagnosis. Finally, the observed associations do not necessarily imply causality, which can only be evaluated through prospective causal inference study designs.

**Conclusions**

This analysis leveraged comprehensive patient data, a robust machine learning algorithm, and an extensive validation framework, to identify predictors of mood disorder misdiagnosis in individuals experiencing depressive symptoms, and subsequent help-seeking. The results highlight the increased risk for misdiagnosis associated with incomplete symptom profiles, more severe or harder to detect symptoms, and older age. Therefore, comprehensive symptom monitoring outside of depressive episodes, mental health screening at earlier ages, and clinician knowledge of the influence of advanced age on misdiagnosis risk are important considerations for early and accurate diagnosis of mood disorders. Moreover, prior engagement with mental health services, functional impairment in performing daily tasks, and younger age were associated with a higher likelihood of help-seeking. Together, these results add to the growing application of machine learning techniques in examining existing barriers to accessing mental health services [19], and may ultimately lead to the development of novel screening tools or procedures for a comprehensive mental health risk assessment in individuals presenting with mood-related symptoms.

**Acknowledgments**

We are most grateful to all participants of the Delta Study for their time and efforts. We are also grateful to all members of the Delta Study Service User Advisory Group. We also thank all those involved in designing and conducting the Delta Study for their valuable input including Jason D Cooper, Sung Yeon Sarah Han, Lynn P Farrag, Emily Bell, Lauren V Friend, Sharmelee Thiahulan, Pawel Eljasz, Mark Agius, Neill Hunt, and the CIDI interviewers. This study was funded by the Stanley Medical Research Institute (07R-1888) and Psyomics Ltd. Stanley Medical Research Institute was not involved in any part of the research. Psyomics Ltd was involved in the design and conduct of the study; recruitment, data collection, and management; and review or approval of the study.

**Authors’ Contributions**

SB and DC conceived the Delta Study, conceptualized, and supervised the development of the web-based mental health questionnaire. SB, DC, GBO, and T Olmert contributed to the design of the study. GBO and T Olmert collected the web-based mental health questionnaire data. GBO and JT processed the web-based mental health questionnaire data. JB, NL, and T Ong analyzed the data. SB, JT, NAMK, and ELF advised the analysis. JB and JT wrote the first draft of the study, with contributions from NL, T Ong, NAMK, ELF, and SB. All authors contributed to the final version of the study.
Conflicts of Interest

SB is a director of Psynova Neurotech Ltd and Psyomics Ltd. SB, ELF, and DC have financial interests in Psyomics Ltd. GBO had financial interests in Psyomics Ltd. SB, JT, and T Olmert have received payments from the University of Cambridge for licensing of data from the Delta Study. SB and JT may benefit financially from patents arising from the Delta Study. ELF is a consultant for Psyomics Ltd. All other authors declare no competing interests.

Multimedia Appendix 1
Demographics, model performance metrics, dependence plots, and feature selection frequencies for all objectives.

References
15. Davies SC. Annual report of the chief medical officer 2013, public mental health priorities: investing in the evidence. Department of Health. 2014. URL: https://assets.publishing.service.gov.uk/media/5a8034e0e5274a2e87d87a0/CMO_web_doc.pdf [accessed 2023-12-19]


59. Hidden waits force more than three quarters of mental health patients to seek help from emergency services. Royal College of Psychiatrists. 2022. URL: https://www.rcpsych.ac.uk/news-and-features/latest-news/detail/2022/10/10/hidden-waits-force-more-than-three-quarters-of-mental-health-patients-to-seek-help-from-emergency-services [accessed 2023-12-19]


Abbreviations

- AUC: area under the receiver operating characteristic curve
- BD: bipolar disorder
- CCNR: Cambridge Centre for Neuropsychiatric Research
- CIDI: Composite International Diagnostic Interview
CV: cross-validation
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
GP: general practitioner
ICD-10: International Classification of Diseases and Related Health Problems, 10th Revision
ICD-11: International Statistical Classification of Diseases and Related Health Problems, 11th Revision
LIME: local interpretable model-agnostic explanations
MDD: major depressive disorder
NICE: National Institute for Health and Care Excellence
PHQ-9: Patient Health Questionnaire-9
rNCV: repeated nested cross-validation
SHAP: Shapley additive explanations
SSRI: selective serotonin reuptake inhibitor
WEMWBS: Warwick-Edinburgh Mental Well-Being Scale
XGBoost: Extreme Gradient Boosting
Understanding Public Perceptions of Virtual Reality Psychological Therapy Using the Attitudes Towards Virtual Reality Therapy (AVRT) Scale: Mixed Methods Development Study

Abstract

Background: Virtual reality (VR) psychological therapy has the potential to increase access to evidence-based mental health interventions by automating their delivery while maintaining outcomes. However, it is unclear whether these more automated therapies are acceptable to potential users of mental health services.

Objective: The main aim of this study was to develop a new, validated questionnaire to measure public perceptions of VR therapy (VRTX) guided by a virtual coach. We also aimed to explore these perceptions in depth and test how aspects such as familiarity with VR and mental health are associated with these perceptions, using both quantitative and qualitative approaches.

Methods: We used a cross-sectional mixed methods design and conducted an exploratory factor analysis of a questionnaire that we developed, the Attitudes Towards Virtual Reality Therapy (AVRT) Scale, and a qualitative content analysis of the data collected through free-text responses during completion of the questionnaire.

Results: We received 295 responses and identified 4 factors within the AVRT Scale, including attitudes toward VRT, expectation of presence, preference for VRT, and cost-effectiveness. We found that being more familiar with VR was correlated with more positive attitudes toward VRT (factor 1), a higher expectation of presence (factor 2), a preference for VRT over face-to-face therapy (factor 3), and a belief that VRT is cost-effective (factor 4). Qualitative data supported the factors we identified and indicated that VRT is acceptable when delivered at home and guided by a virtual coach.

Conclusions: This study is the first to validate a scale to explore attitudes toward VRT guided by a virtual coach. Our findings indicate that people are willing to try VRT, particularly because it offers increased access and choice, and that as VR becomes ubiquitous, they will also have positive attitudes toward VRT. Future research should further validate the AVRT Scale.

(KEYWORDS: psychological interventions; digital; virtual reality; virtual agent; mental health; presence)
Introduction

Background
Virtual reality (VR) is an immersive environment where people can interact using either computer equipment, such as a screen and mouse, or VR-enabled headsets and controllers, where additional sensors can track the users’ actions in real time. The latter application provides people with a greater sense of presence, a term used to describe how closely a virtual environment is interpreted as real [1]. In recent years, VR has been used successfully in a range of health care settings to improve and increase access to treatment [2]. In particular, VR has been used in the delivery of psychological therapies for a range of mental health problems, with several decades of evidence demonstrating its clinical efficacy in the treatment of psychosis, depression, anxiety, and eating disorders [3-8].

VR therapies (VRTs) were initially developed to be used by therapists as an adjunct or tool in their delivery of therapy. However, the need for a real-world therapist to deliver VRTs presents a key challenge for their widespread implementation [9]. Researchers have shown that the automation of some therapeutic elements may overcome this barrier to meet the increase in demand for treatment globally [10,11]. Emerging evidence demonstrates that VRTs can be successfully delivered with little to no therapist involvement, with virtual coaches supporting people receiving therapy for fear of heights and agoraphobia in the context of psychosis [10,11]. Virtual coaches are also known as virtual agents [12]. These characters are not under human control and therefore offer automation of therapies, in which dialogue and responses are scripted instead of the formulation that is offered by real-world therapists.

There are financial and resource incentives for mental health services to offer more automated therapies [13]. Clinicians also appear to be in support of VRTs. For example, cognitive behavior therapists [14] and psychiatric health care staff [15] reported positive attitudes toward VRT; particularly when they were more familiar with VR. However, these studies do not consider how staff feel about VRT guided by a virtual coach and, notably, do not explore patient and public perceptions of VRT. VRT dropout rates have been used as a proxy measure of patient experience, and these figures show similar dropout rates to therapies delivered without VR [16]. However, dropout rates from research do not provide us with a clear picture of whether people will engage in therapies delivered using VR, including those guided by a virtual coach. A content analysis of social media posts by the public appears to suggest an interest in the application of VR in mental health care [17]. Staff and service users also have positive views toward their use in mental health inpatient facilities [18]. However, these studies still do not directly ask potential users of mental health services whether they would be willing to try VRT guided by a virtual coach or the factors that relate to this willingness.

Health care staff, when asked for their views regarding service users’ opinions of VRT, had concerns regarding patients’ willingness to accept their use as part of their mental health care package [19]. Furthermore, the literature has highlighted a lack of personalization as a barrier to engagement with digital mental health interventions [20]. It is unclear whether this indicates that automated VRTs can be sufficient when scripts are relevant to the experience of the individual. It is possible that the presence of a virtual coach may encourage more positive attitudes and a willingness to try VRTs. There is a need to understand service user and public perspectives on the use and delivery of VRTs and those guided by a virtual coach and how different factors may affect the uptake of such interventions.

Aim
The main aim of this study was to establish a new, validated questionnaire to measure the perceptions of VRT guided by a virtual coach. Second, we aimed to explore how these perceptions are associated with familiarity with VR and mental health, using both quantitative and qualitative approaches.

Methods

Study Design
This study used mixed methods with a cross-sectional design. Data were collected from a web-based questionnaire using Jisc software [21].

Participants
The participants were recruited via social media to complete the web-based questionnaire. We aimed to recruit a minimum of 200 participants in line with sample size recommendations for exploratory factor analysis (EFA) [22]. To be eligible to participate, persons were required to be a resident of the United Kingdom or Ireland and aged ≥18 years. A link to the web-based survey was included in all promotional materials.

Measures

Demographics
The participants were asked to provide basic demographic information, including their age and sex, as well as whether they were identified as having a mental health condition, had ever experienced therapy, or had supported anyone with a mental health condition. Furthermore, they were asked about their experience of VR (from never to ≥10 times) and their familiarity with VR, VRT, and mental health conditions.

The Attitudes Towards Virtual Reality Therapy Scale
The Attitudes Towards Virtual Reality Therapy (AVRT) Scale was developed by AMA and ADGB. Items were based on themes identified in previous literature that contribute to perceptions of VRT and digital mental health interventions [14,23-29]. Items surrounding the virtual coach drew on the literature related to therapeutic alliance and focused on trust, comfort, and need [30].

We designed 54 items all assessing different aspects of attitudes toward VRT, including 9 items related to attitudes toward VRT delivered by a virtual coach. Each item used a 7-point Likert scale where participants rated their agreement from “strongly agree” to “strongly disagree.” Strong agreement or disagreement with 16 of these items triggered a free-text question for participants to provide context using free-text responses. Responses were scored from 1 (strongly disagree) to 7 (strongly agree).
agree). A higher score indicates more positive perceptions of VRT. A total of 27 items were reverse-worded and therefore reverse-coded.

Furthermore, participants were invited to respond to 3 additional free-text questions asking what they would like to know more about, how they think their level of VR experience has influenced their perceptions, and what they think the best setting for VRT would be.

**Procedure**

Upon opening the questionnaire, participants were first shown the information sheet, followed by a consent statement. After consenting, participants were asked to enter a unique identification code so that their anonymized responses could be identified later. Participants were then asked to provide basic demographic information, followed by an explanatory paragraph (Multimedia Appendix 1) about VRT and the virtual coach, which was described as “a computer-generated avatar” that “guides the patients through the scenarios and offers advice and encouragement.” This was followed by items on experience with VR and mental health, the AVRT Scale, and the 3 free-text questions. After completing these questionnaires, participants were presented with a debrief statement.

**Ethics Approval**

Ethics approval was granted by the Division of Psychiatry and Applied Psychology Ethics Subcommittee of the University of Nottingham (Project ID 1534).

**Statistical Analysis**

Raw data were downloaded from Jisc [21] into SPSS Statistics software (version 25; IBM Corp) [31]. We removed responses from participants who did not meet the inclusion criteria, did not provide consent, or had missing data. Sample characteristics were summarized using descriptive statistics.

To validate our new questionnaire, we conducted an EFA using principal component analysis with a varimax rotation. We assessed the suitability of the data for factor analysis using Bartlett’s test of sphericity and the Kaiser-Meyer-Olkin measure of sampling adequacy (acceptable adequacy ≥0.6). All 54 items using Likert scales were included in the EFA. Items were first screened to check for multicollinearity and poor correlations with the other items. We operationalized this screening by assessing the determinant and searching for any interitem correlations of ≥0.7 or where most coefficients were nonsignificant or <0.4. Any items that failed this initial screening were removed, and the EFA was rerun. Factors were derived using eigenvalues ≥1, where the Kaiser criterion [32] was met, and in combination with the point of inflection on the scree plot, where they were not. For factors to be retained, they must comprise at least 3 items. Where items loaded on >1 factor, the item was assigned to the factor with which it made the most thematic sense.

The questionnaire included both positively and negatively worded items. Once the factors were established, we reverse-scored negatively worded items so that a higher score indicated a more favorable attitude. We then assessed the internal consistency of the final factor structure using Cronbach α, with an acceptable internal consistency of ≥0.7 [33]. Further items may be removed at this point, where the scale reliability can be substantially improved if the item is removed. The factor scores were computed using the mean and SD of the scale sum.

We conducted a series of Pearson r correlations to assess whether there was a relationship among the scale totals of the derived factors and lived experience of VR, VRT, and mental health problems.

**Qualitative Analysis**

All responses to the free-text response questions (ie, 16 free-text boxes triggered by extreme responses to survey questions and 3 additional free-text questions) were uploaded to NVivo (version 12 for Mac; QSR International). Qualitative content analysis [34] was used to quantify and summarize the qualitative data within the broader context of the AVRT Scale. All data were coded inductively by a qualitative researcher (ADGB), where several codes could be applied to a single response. The codes were collated by questions or items. The study team met to review and revise any discrepancies or discuss any questions. Findings were then summarized according to each question or item and were presented within the factors of the AVRT Scale.

**Results**

**Sample Characteristics**

We collected 295 responses to the survey. Our sample reflected a range of age groups. The majority were female, had used VR at least once, and had no personal or professional experience with mental health problems. However, most participants had supported a friend or family member with poor mental health (Table 1).
Table 1. Sample characteristics (N=295).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age group (years), n (%)</td>
<td></td>
</tr>
<tr>
<td>18-24</td>
<td>89 (30.2)</td>
</tr>
<tr>
<td>25-29</td>
<td>39 (13.2)</td>
</tr>
<tr>
<td>30-39</td>
<td>45 (15.3)</td>
</tr>
<tr>
<td>40-49</td>
<td>49 (16.6)</td>
</tr>
<tr>
<td>50-64</td>
<td>64 (21.7)</td>
</tr>
<tr>
<td>&gt;65</td>
<td>9 (3.1)</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>83 (28.1)</td>
</tr>
<tr>
<td>Female</td>
<td>209 (70.8)</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>3 (1)</td>
</tr>
<tr>
<td>Frequency of experiencing VR(^a), n (%)</td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>112 (38)</td>
</tr>
<tr>
<td>Once</td>
<td>48 (16.3)</td>
</tr>
<tr>
<td>&lt;5 times</td>
<td>88 (29.8)</td>
</tr>
<tr>
<td>5-9 times</td>
<td>20 (6.8)</td>
</tr>
<tr>
<td>≥10 times</td>
<td>27 (9.2)</td>
</tr>
<tr>
<td>Participants identifying as having a mental health condition, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>102 (34.6)</td>
</tr>
<tr>
<td>No</td>
<td>181 (61.4)</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>12 (4.1)</td>
</tr>
<tr>
<td>Participants with experience in therapy for a mental health condition, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>131 (44.4)</td>
</tr>
<tr>
<td>No</td>
<td>160 (54.2)</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>4 (1.4)</td>
</tr>
<tr>
<td>Participants who have supported a friend or family member or colleague with a mental health condition, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>253 (85.8)</td>
</tr>
<tr>
<td>No</td>
<td>39 (13.2)</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>3 (1)</td>
</tr>
<tr>
<td>Participants who have worked in a caring role for people with mental health conditions, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>108 (36.6)</td>
</tr>
<tr>
<td>No</td>
<td>186 (63.1)</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>1 (0.3)</td>
</tr>
<tr>
<td>Participants familiar with VR, mean (SD)</td>
<td>2.33 (1.02)</td>
</tr>
<tr>
<td>Participants familiar with VR therapy, mean (SD)</td>
<td>1.35 (0.75)</td>
</tr>
<tr>
<td>Participants familiar with mental health conditions, mean (SD)</td>
<td>3.65 (0.96)</td>
</tr>
</tbody>
</table>

\(^a\)VR: virtual reality.

Quantitative Analysis

Item Screening

The data were found to be appropriate for EFA (Kaiser-Meyer-Olkin=0.93; \(\chi^2_{1431}=10,205.5, \ P<.001\)). The determinant suggested that there was multicollinearity, and inspection of the correlation coefficients revealed 2 pairs of items that were highly correlated (items 7 and 8=0.83; items 36 and 37=0.87); therefore, we removed 1 item from each pair of correlations (items 8 and 36). Items 13, 32, and 33 were removed, as the majority of interitem correlations were

https://mental.jmir.org/2024/1/e48537

JMIR Ment Health 2024 | vol. 11 | e48537 | p.18
(page number not for citation purposes)
non-significant \( (P>0.05) \). Furthermore, we removed items 6, 19, 25, 26, 28, 31, and 47 as either all or all but one of the correlation coefficients was <0.4. In total, we removed 12 items and then reran the EFA on the remaining 42 items.

The determinant again indicated that multicollinearity was an issue. We identified 3 pairs of correlations with coefficients >0.7 (items 10 and 11=0.76; items 29 and 49=0.76; items 35 and 37=0.71); therefore, we removed 1 item from each pair (items 10, 29, and 35) and reran the EFA on the remaining 39 items.

**Exploratory Factor Analysis**

The Kaiser criterion was met \( (n=295; \text{average communalities } 0.64) \) [32]. Therefore, the factor structure was determined based on eigenvalues >1. The rotated factor solution suggested 7 factors, which explained 63.64% of the variance. However, 3 factors were not retained because they contained <3 items. The removal of these factors resulted in the removal of items 7, 23, 24, 27, 48, and 49. The resulting 33 items were entered into a final EFA. A 4-factor solution was suggested based on the eigenvalues and the scree plot, which explained 58.61% of the variance.

Factor 1 had 13 items that assessed respondents’ support for VRT, including 6 reverse-worded items (factor 1: attitude toward VRT). Factor 2 had 9 items. These items, including 7 reverse-worded items, assessed the extent to which the respondents expected VRT to be immersive (factor 2: the expectation of presence). Factor 3 had 7 items asking respondents to compare VRT to aspects of face-to-face therapies (factor 3: preference for VRT). Factor 4 had 4 items each assessing different aspects of the cost-effectiveness of VRT (factor 4: cost-effectiveness). Refer to Multimedia Appendix 2 for the final factor structure.

**Scale Reliability**

After reverse-scoring the reverse-worded items, we computed Cronbach \( \alpha \) values for each scale. All scales had strong internal consistency (all Cronbach \( \alpha \geq 0.82 \)). The scale reliabilities could not be improved by removing any of the items. A higher score on each of the subscales suggested a more favorable attitude (factor 1), increased perceived presence (factor 2), a preference for VRT over traditional therapies (factor 3), and agreement that VRT is cost-effective (factor 4). The desired direction for each subscale to demonstrate support for VRT was high for factors 1, 3, and 4 and low for factor 2.

**Relationship Between Scales and Lived Experience**

There was a significant relationship between the participants’ familiarity with VR and their scores on all the factors. Familiarity with VR was positively associated with a more favorable attitude toward VRT (factor 1), higher expectations of presence (factor 2), a preference over face-to-face therapy (factor 3), and a belief that VRT is cost-effective (factor 4). We also found significant positive correlations between factors 1, 2, and 3, but not factor 4, and familiarity with the VRT. There was no significant correlation between mental health familiarity and the scores for any factors. Multimedia Appendix 2 presents the correlation coefficients and associated significance scores.

**Qualitative Results**

**Qualitative Questions**

Table 2 presents the initial qualitative questions that all participants were asked.

| Type of data | Question | Response, n (%) |
|--------------|----------|----------------|}
| Qualitative question 1 | Which aspects of virtual reality therapy guided by a virtual coach would you like to know more about? | 226 (76.6) |
| Qualitative question 2 | How has your previous experience of virtual reality (minimal or extensive) influenced your perceptions of virtual reality therapy guided by a virtual coach? | 236 (80) |
| Qualitative question 3 | If you were offered virtual reality therapy guided by a virtual coach, where do you think the best place to do the therapy would be? | 245 (83) |

**Which Aspects of VRT Guided by a Virtual Coach Would You Like to Know More About?**

Of the 226 participants, 19 (8.4%) indicated that they did not want to know anything more. Those who provided reasons indicated that they did not want to try or did not know enough. Of them, 22 (9.7%) participants indicated that, as they did not know enough, they would like to find out more; 5 (2%) indicated that they would like to try out VRT; and 9 (3.9%) asked regarding its cost. In total, 13 (5.8%) participants were curious about the conditions that could be targeted with VRT, specifically regarding its use for anxiety disorders, depression, and emotion regulation.

Many participants asked how it could be tailored or personalized for them (29/226, 12.8%). This meant thinking about their position within the interaction, asking about safety or how much control they would have, and whether VRT might have a negative effect and how this would be monitored. Of 226 participants, 47 (20.8%) asked about the virtual coach, wanting to know how real it would be, how much of the language would be generic or responsive to them, and how they could build a relationship with a virtual coach. Of these, many wanted more information about whether there was a real therapist involved and how involved they would be (11/47, 23%), whether they would be able to meet them in person, whether they would deliver the therapy live or preprogram the coach, or whether the virtual coach would be completely artificially intelligent. The reality of the virtual coach and the VRT (16/226, 7.1%) was also an important question posed by participants, including asking whether it would be realistic enough and comparing it to “real” or face-to-face therapy.
Most additional responses indicated that participants would like to know more about the process (67/226, 29.6%). This included practical questions regarding the frequency of use, the length of sessions, and how it would be delivered (eg, in which location). Furthermore, many participants asked about the content (26/226, 11.5%), particularly not only the scenarios that could be represented but also other aspects including the appearance, the script, and how the content could link with face-to-face therapy. Of the 226 participants, 11 (4.9%) asked about the technical aspects including the development of the coach (eg, whether an algorithm or artificial intelligence was used) and what equipment would be used to deliver the VRT.

**How Has Your Previous Experience of VR (Minimal or Extensive) Influenced Your Perceptions of VRT Guided by a Virtual Coach?**

The largest group of those answering this question indicated that they had no previous experience (79/236, 33.5%). A few without experience were positive or curious (19/236, 8.1%) whereas others (8/236, 3.1%) expressed more negative perceptions about its effectiveness as a therapeutic tool, anticipating that it would not feel real or tailored enough to the individual. The second largest group (59/236, 25%) felt that their previous experience had helped them to be more positive and linked it to their own experience of mental health and how it could be used for treatment. Although several mentioned using VR for gaming, they felt that it was effective at producing a level of presence that would be conducive to therapy and help to invoke real emotions and responses. They felt it was easy to use, could potentially lower costs, make therapy more accessible, and even with negative experiences, such as motion sickness or technical difficulties, they still had a positive perception of VRT.

However, 34 (14.4%) of the 236 participants with a more negative perception reported nausea or dizziness, whereas others perceived VR as more suited to games. This included problems with the quality of their experience, feeling that the VRT had not offered enough presence. However, those with a negative experience comprised the smallest group (8/236, 3.1%). Finally, the third largest group felt that their previous experience would not influence how they felt about VRTs guided by a virtual coach (40/236, 16.9%). For some, their previous experiences could not inform their perception of VRTs as it had been for entertainment purposes or they had too little experience to be able to make a judgment (13/236, 5.5%).

**If You Were Offered VRT Guided by a Virtual Coach, Where Do You Think the Best Place to Do the Therapy Would Be?**

The largest group of respondents who identified a single location felt that it would be best delivered within the home (84/245, 34.3%), whereas the second largest group felt that it would be best delivered in a more professional location (43/245, 17.6%). Several felt that it could be offered in both settings (38/245, 15.5%), whereas others suggested that access could first be through a clinic (12/245, 4.9%), where they could access technical or therapeutic support, or from home (4/245, 1.6%), where they would feel more comfortable. When respondents highlighted delivery from home, they described it as being safe, comfortable, and familiar. They felt that they might feel susceptible or disorientated when coming out of a VRT session and that being at home would be preferable. More professional locations, such as physician surgeries or clinics, were also described as safe and familiar, although by fewer people. Professional or clinical settings were often viewed as a better location because of the presence of support. Other reasons included the level of cleanliness offered and that there would be fewer distractions.

Those without a preference identified elements of the location that were necessary to optimize the experience, including having a space to move, feeling safe and secure (eg, in an enclosed space), having privacy and quiet, and having few distractions. They also felt it would need to consider the condition being treated (including severity) and the individual’s preferences.

**Factors With Item Responses**

Table 3 presents those items where either strong agreement or disagreement elicited a qualitative response.
Table 3. Qualitative responses to items.

<table>
<thead>
<tr>
<th>Question</th>
<th>Strongly agree, n (%)</th>
<th>Strongly disagree, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor 1: Attitudes toward VRT(^a) (42 and 43)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 41: If the virtual coach encouraged me to do something between sessions, I would try to do it. (n=19)</td>
<td>18 (95)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Item 51: I would never be willing to try virtual reality therapy. (n=63)</td>
<td>3 (5)</td>
<td>60 (95)</td>
</tr>
<tr>
<td>Item 50: I would be willing to try virtual reality therapy if I had more information about it. (n=24)</td>
<td>21 (87)</td>
<td>3 (13)</td>
</tr>
<tr>
<td>Item 52: I would encourage the people I care about to try virtual reality therapy, if it was offered to them. (n=16)</td>
<td>14 (87)</td>
<td>2 (13)</td>
</tr>
<tr>
<td>Item 53: I would discourage the people I care about to try virtual reality therapy, if it was offered to them. (n=25)</td>
<td>1 (4)</td>
<td>24 (96)</td>
</tr>
<tr>
<td>Item 54: I cannot imagine virtual reality therapy being useful for someone with mental health problems. (n=24)</td>
<td>2 (8)</td>
<td>22 (92)</td>
</tr>
<tr>
<td>Item 42: I would feel comfortable interacting with the virtual coach. (n=14)</td>
<td>10 (71)</td>
<td>4 (29)</td>
</tr>
<tr>
<td><strong>Factor 2: Expectation of Presence (47, 49, 50, 52, and 40)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 43: I would find the characters in the virtual reality therapy unsettling. (n=6)</td>
<td>1 (17)</td>
<td>5 (83)</td>
</tr>
<tr>
<td>Item 45: I am skeptical about the effectiveness of virtual reality therapy. (n=13)</td>
<td>6 (46)</td>
<td>7 (54)</td>
</tr>
<tr>
<td>Item 44: I think that the virtual reality therapy would make me feel present enough to be effective. (n=7)</td>
<td>3 (43)</td>
<td>4 (57)</td>
</tr>
<tr>
<td><strong>Factor 3: Preference for VRT (54, 39, 53, and 41)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 39: I think virtual reality therapy would be better than face-to-face therapy. (n=29)</td>
<td>2 (7)</td>
<td>27 (93)</td>
</tr>
<tr>
<td>Item 40: I would trust a virtual coach the same amount as a real therapist. (n=20)</td>
<td>4 (20)</td>
<td>16 (80)</td>
</tr>
<tr>
<td><strong>Factor 4: Cost-effectiveness (46)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 46: I think virtual reality therapy will be worth the cost. (n=9)</td>
<td>6 (66)</td>
<td>3 (34)</td>
</tr>
<tr>
<td><strong>Nonfactor answers (45, 48, 44, and 51)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 47: I think I would be able to use the virtual reality equipment easily. (n=26)</td>
<td>24 (92)</td>
<td>2 (8)</td>
</tr>
<tr>
<td>Item 49: I think that virtual reality equipment could spread diseases. (n=19)</td>
<td>0 (0)</td>
<td>19 (100)</td>
</tr>
<tr>
<td>Item 48: If my skills with technology were poor, I would feel confident using virtual reality therapy if the health care professional accompanying me was trained to a high standard. (n=19)</td>
<td>17 (89)</td>
<td>2 (11)</td>
</tr>
</tbody>
</table>

\(^{a}\)VRT: virtual reality therapy.

**Factor 1: Attitude Toward VRT**

Individuals who scored highly on this factor had a positive attitude toward VRTs and VRTs delivered by a virtual coach, whereas those who scored low had a negative attitude. Within the items where strong agreement or disagreement elicited a text response (items 41, 51, 50, 52, 53, 54, and 42), those with positive attitudes highlighted the value of having a choice in mental health therapies. They emphasized the need to be willing to try different treatments to find the one that worked, reflecting on how more automated and digital options can help to increase access. Those with more negative attitudes indicated that it would be a type of therapy that they would not choose.

**Factor 2: Expectation of Presence**

Individuals who scored highly on this factor felt that VR would not be real, that is, low presence. Individuals who scored low felt that VR was immersive. Within the items where strong agreement or disagreement elicited a text response (items 43, 45, and 44), the respondents indicated several factors that affected their expectation that VR would be “real enough.” Previous experiences appeared to be linked to the expectations of presence. People who enjoyed their experiences had higher expectations of presence. Those with lower expectations felt that VRT would be too much like a game, whereas others indicated that experiencing cybersickness meant they had not felt present.

**Factor 3: Preference for VRT**

Individuals who scored highly on this factor showed a preference for VRT, whereas those who scored low showed a preference for face-to-face therapy. Within the items where strong agreement or disagreement elicited a text response (items 39 and 40), there was a strong sense that those who preferred face-to-face therapy would feel the loss of human interaction most and feel that a real person was needed to build a relationship and trust. Those who were more in favor of VRT...
and the virtual coach felt that it would be more convenient and potentially enable more disclosures related to their mental health.

**Factor 4: Cost-Effectiveness**

Individuals who scored high on this factor felt that VRT was cost-effective, whereas those who scored low did not. Within the item where strong agreement or disagreement elicited a text response (item 46), those who felt it was cost-effective highlighted the decreasing costs of equipment and the benefits this could bring to mental health services. For those who felt VR was still too expensive, there was also recognition of the difficulties that services might have in adopting VR.

### Discussion

#### Overview

This study aimed to develop a new instrument for assessing the public perception of VRT delivered by a virtual coach. We received 295 responses. We found that a 4-factor solution was the best fit for the AvRT Scale, with all subscales having excellent internal consistency. The 4 factors were (1) attitudes toward VRT, (2) expectation of presence, (3) preference for VRT, and (4) cost-effectiveness. We found that being more familiar with VR was correlated with more positive attitudes toward VRT (factor 1), a higher expectation of presence (factor 2), preference for VRT over face-to-face therapy (factor 3), and belief that VRT is cost-effective (factor 4). Familiarity with mental health was not associated with any factor. The qualitative data supported the quantitative findings, with many respondents stating that their previous experience with VR may have affected their perception of VRT. Respondents identified their homes and spaces that felt safe and quiet as the best locations for delivering VRT. The virtual coach was a salient concept throughout the qualitative responses, with participants wanting to better understand it and the relationships that could be facilitated.

#### Principal Findings

Previous literature has indicated a correlation between VR familiarity and more positive attitudes toward VRTs [14]; however, this is the first study to demonstrate this through a public survey. Although we do not know the direction of this association, the qualitative findings suggest that as the reach of VR headsets increases, VRTs will likely be viewed more positively. The perceptions of potential patients are important in determining the efficacy of VRT, as positive expectations of any psychological therapy are associated with better treatment outcomes [35,36]. Therefore, an increase in the popularity of VR kits may indirectly improve the efficacy of VRTs.

This study is the first to explore people's perceptions of VRTs guided by a virtual coach. Although most participants had no personal or professional experience of mental health therapy, many mentioned aspects relating to the virtual coach that draw parallels with “therapeutic alliance”; in psychotherapy, this denotes the importance of the relationship between the therapist and service user. In psychotherapy research, a strong therapeutic alliance is associated with better treatment outcomes [37]. The concept of therapeutic alliance has been studied more broadly in relation to VR-assisted therapies [30] and digital mental health [38]. The effects are similar but may be predictive of treatment outcomes to a lesser extent and more so predict engagement.

Understanding whether it is possible to foster a “therapeutic alliance” with a virtual coach and, if so, the nature of this relationship is something that the public is concerned with and therefore requires further investigation. Our findings provide initial insights into how therapeutic alliances may operate in VRT with a virtual coach. Respondents who showed a preference for VRTs indicated that the presence of a virtual coach would aid disclosure. This may be because of the anonymity that this form of communication offers [39]. Furthermore, it is notable that many were curious about the level of automation and formulation offered by the virtual coach. Qualitative findings from a trial of VRT guided by a virtual coach found that the presence of a member of staff helped to reinforce learning, which may suggest that certain elements of therapy require a certain level of formulation [40]. Our data suggest that the public view personalization as an important component of therapy and that VRTs can be improved by offering a certain level of formulation.

Another novelty of this study is the exploration of presence in relation to VRTs. Previous research has suggested that increasing presence can increase the effectiveness of VRT [41]. Newer VR-enabled headsets provide a greater level of presence as the quality of graphics and functionality, such as interactivity and sensitivity of sensors (eg, eye tracking), have improved significantly. Therefore, we sought to explore the importance of this sense of presence in the general population. Our findings indicate that those with a greater expectation of presence are more positive and more likely to show a preference for VRTs. Notably, our findings indicate that even those who are familiar with VR share concerns regarding the lack of presence and immersion in VRTs. This suggests that developers and researchers must continue to develop and update their intervention designs to ensure that VRTs do not become stagnant and continue to elicit a sense of presence.

Most of those who viewed VRTs positively emphasized the need for a choice to help increase access to mental health treatment. More automated VRTs have been designed considering the pressure to deliver psychological therapies in mental health services and the lack of resources to meet this need [10,11]. Our findings indicate that the public is aware of this and views VRTs guided by a virtual coach as an acceptable solution. Our respondents also indicated that VRT guided by a virtual coach would be suitable for delivery at home, further alleviating resource pressures. However, this was not the case for all participants, with a notable proportion wanting to access VRTs in a location that was safe, familiar, free of distractions, and large enough to use the VRT. The flexibility of location in delivering VRT is an important consideration for increasing access and meeting the needs of service users, especially when considering the strong links between poor mental health and housing quality [42]. The delivery location for VRT should be considered on a case-by-case basis.

A further consideration for the implementation of VRT found in our study is the importance of information. Our qualitative...
findings indicate that people were keen to better understand what was involved in VRT and the virtual coach. Preintervention expectations are key in managing service users’ expectations while also fostering hopefulness, which in turn improves engagement [43]. This information can also help to allay any concerns service users might have about VRT and help developers to understand their needs to improve the design and implementation of VRTs. For example, a small number of participants opposed the use of VRTs, which when expanded in our qualitative data collection, indicated certain ethical and moral concerns about its use in mental health care. All these views were valid, but a few may be rooted in misconceptions about VR or expectations about how it will be implemented. Therefore, potentially increasing the public’s awareness and understanding of VR and VRT may help to appease them and improve how it is deployed.

We found mixed findings regarding the impact of cybersickness on willingness to engage in VRT. For some participants, cybersickness would dissuade people from using VRT. However, this finding was not ubiquitous, with some saying that they were still interested in trying VRT even if they had experienced cybersickness. This contradicts previous research, suggesting cybersickness is a considerable barrier [17]. As technology progresses, cybersickness might become less important. We also included a question on hygiene, as our questionnaire was shared during the COVID-19 pandemic. However, this was excluded, suggesting that it was not a significant concern for the public.

**Limitations**

The questionnaire has only been validated using the EFA. Further validation is required before we can confidently recommend its widespread use. Specifically, we must confirm the factor structure in a new sample using confirmatory factor analysis and assess its concurrent and discriminant validity. If we are able to replicate the strong psychometric properties found in this study, this questionnaire can be used to understand attitudes toward VRTs delivered by virtual coaches. The scale will also need to be adapted to contexts outside the United Kingdom, for example, by amending items and further validation.

Most of our respondents were female and had no previous experience with mental health conditions or therapy. However, men and those with more experience with mental health conditions or therapy may have different perceptions of VRTs. A recent review found that gender differences might affect the use and acceptability of VR, specifically that women are more susceptible to cybersickness and therefore may be less willing to use VR [44]. On the basis of this, it may be assumed that if we were to conduct a survey with male participants, the attitudes toward VRT guided by a virtual coach could be even more positive. We do not have any available evidence to indicate whether those who are living with or have lived with mental health conditions are likely to be more or less accepting of VRTs. Purposive sampling should be used in future studies to ensure that the views of these groups are included in future validation studies.

Furthermore, we sought text responses for strong agreement or disagreement with certain items. Notably, those with more neutral views may have offered additional insights, but we weighed this against the additional burden on respondents. This may also have led some participants to neutralize their views to avoid triggering a free-text question. However, there were no instructions in the questionnaire that responding differently removed the free-text responses. In addition, the range of scores indicated that this did not deter the participants from giving extreme answers. The qualitative data we captured were sufficient for our analysis.

Finally, the analysis of the relationship between familiarity and attitudes toward VR and VRT was correlational. We could not make any claim regarding the direction or causal nature of these associations. For example, those with more positive attitudes and a better understanding are more likely to become familiar with VR through continued use. However, our qualitative findings indicate that negative experiences with VR do not factor in a willingness to use VRT.

**Recommendations**

Future research should further validate this questionnaire. Once this has been accomplished, the questionnaire could be used to investigate the factors that improve or worsen attitudes toward VRT and VRT guided by a virtual coach. For example, asking questions such as whether trying VR improves perceptions or whether increasing sales of domestic VR kits is associated with improved attitudes. It is also important to explore how these attitudes translate into behavior, that is, whether positive attitudes predict patient preferences and engagement with VRTs. The impact of the level of automation versus the formulation of the virtual coach on attitudes should also be explored, as this was a salient concept within the qualitative data. The AVRT Scale could be adapted and applied to other areas where VR is used to deliver interventions, such as behavior change interventions, or as a training tool. The questionnaire can also be used alongside treatment development, evaluation, and implementation to explore the barriers and facilitators specific to VRT and VRT guided by a virtual coach or the perceptions of certain populations to aid the translation of research into practice [45]. In the long term, any research that considers barriers to the uptake, engagement, and adoption of VRT has the potential to alleviate the demand for trained therapists in clinical settings, thus improving access to psychological therapies.

**Acknowledgments**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. ADGB was funded by the National Institute for Health and Care Research MindTech MedTech Co-operative. The views expressed are those of the authors and not necessarily those of the National Health Service, the National Institute for Health and Care Research,
or the Department of Health. These organizations had no role in the study design, including collection, management, analysis, and interpretation of data, writing of the report, or the decision to submit the report for publication.

ADGB would like to acknowledge the support of the National Institute for Health and Care Research Nottingham Biomedical Research Centre.

Data Availability

All relevant data have been included in this publication. Researchers who would like to access the scale for further validation or adaptation may contact the corresponding author with a methodologically sound proposal.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Description of virtual reality therapy and the virtual coach.

[PDF File (Adobe PDF File), 18 KB - mental_v11i1e48537_app1.pdf ]

Multimedia Appendix 2

Factor structure and loadings.

[PDF File (Adobe PDF File), 172 KB - mental_v11i1e48537_app2.pdf ]

References


21. JISC: online surveys. JISC. 2022. URL: https://www.jisc.ac.uk [accessed 2023-12-14]


Abbreviations

- AVRT: Attitudes Towards Virtual Reality Therapy
- EFA: exploratory factor analysis
- VR: virtual reality
- VRT: virtual reality therapy

Please cite as:
Gomez Bergin AD, Allison AM, Hazell CM Understanding Public Perceptions of Virtual Reality Psychological Therapy Using the Attitudes Towards Virtual Reality Therapy (AVRT) Scale: Mixed Methods Development Study JMIR Ment Health 2024;11:e48537 URL: https://mental.jmir.org/2024/1/e48537 doi:10.2196/48537 PMID:
The Frequency of Design Studies Targeting People With Psychotic Symptoms and Features in Mental Health Care Innovation: Secondary Analysis of a Systematic Review

Lars Veldmeijer, MA; Gijs Terlouw, PhD; Jim Van Os, PhD; Job Van 't Veer, PhD; Nynke Boonstra, PhD

1Department of Psychiatry, Utrecht University Medical Center, Utrecht, Netherlands
2Department of Healthcare and Welfare, NHL Stenden University of Applied Sciences, Leeuwarden, Netherlands
3KieN VIP Mental Health Care Services, Leeuwarden, Netherlands

Corresponding Author:
Lars Veldmeijer, MA
Department of Psychiatry
Utrecht University Medical Center
Heidelberglaan 100
Utrecht, 3584 CX
Netherlands
Phone: 31 68 287 4768
Email: lars.veldmeijer@nhlstenden.com

Abstract
This study examined and reflected on the frequency of people with psychotic symptoms and features as the target population in design studies for mental health care innovation.

(Keywords: design approaches; design; innovation; innovative; innovate; innovations; psychiatry; mental health care; mental health; mental illness; mental disease; involvement; service users; people with lived experience; people with lived experiences; lived experience; lived experiences; co-creation; cocreation; psychosis; psychotic; schizophrenia; schizoid; schizotypal; paranoia; neurosis; hallucinosis; hallucination; hallucinations)

Introduction
There is growing evidence highlighting the importance of involving people with lived experience in design processes in mental health care [1,2]. Particular attention should be directed toward the engagement of people with psychotic symptoms and features [3], as they often feel misunderstood due to their altered perceptions and subjective experiences [4,5]. A bottom-up review of the lived experience of psychosis emphasizes the complexity of psychotic symptoms and features and recommends including lived experience in designing mental health services to address these experiences and needs [6]. Design approaches can promote the involvement of people with firsthand experiences in the development of treatment, therapy, and recovery interventions for mental health care innovation [2]. Currently, it is unknown how frequent design studies specifically target people with psychotic symptoms and features. There is a scoping review of coproducing research on psychosis [7], but coproduction and design approaches are distinct methodologies. Design approaches facilitate designing initiatives that prioritize participants’ needs, expertise, and knowledge whereas coproduction facilitates collaborative delivery and knowledge production. In this research letter, we present findings on the frequency of design studies targeting people with psychotic symptoms by analyzing a prior systematic review data set that focused on involving people with firsthand experiences in designing mental health care innovations. The primary objective of this secondary data analysis was to elucidate how often design studies in mental health care target people with psychotic symptoms and features.

Methods
Primary Data Set and Secondary Data Analysis
We conducted a secondary data analysis using a data set from a prior systematic review that assessed the involvement of service users and people with lived experience in the design processes of mental health care innovation. In the screening process and study selection of the prior systematic review, 33
papers met the inclusion criteria [2]. All included papers were original reports or papers that (1) involved service users, people with lived experience, or both; (2) mentioned design approaches; (3) involved an empirical study; and (4) conducted the study in settings including mental health care services or psychiatry programs. In this secondary analysis, we examined the primary data set to provide an overview of the frequency of design studies in mental health care focusing on people with psychotic symptoms and features. This data set is suitable for this analysis since the search strategy of the systematic review did not target specific mental health conditions.

Data Extraction and Categorization

Studies were categorized based on their primary target population as reported in the studies (Multimedia Appendix 1). We categorized broad terms like psychosis, which encompasses various symptoms and features like altered perceptions, as well as mental health conditions in which psychotic symptoms and features are prevalent, such as schizophrenia. In the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR), these symptoms, features, and conditions fall under the category “schizophrenia spectrum and other psychotic disorders,” covering a spectrum of related mental health conditions [8], also referred to as the psychosis spectrum [9]. Studies addressing psychotic symptoms and features or related conditions alongside unrelated mental health conditions were labeled “various mental health conditions” due to their comorbid nature. We did not count these studies as primarily focusing on people with psychotic symptoms and features.

Table 1. Target populations in mental health design studies.

<table>
<thead>
<tr>
<th>Target population</th>
<th>Count, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychosis</td>
<td>6 (18)</td>
</tr>
<tr>
<td>Depression</td>
<td>4 (12)</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>3 (9)</td>
</tr>
<tr>
<td>Self-harm</td>
<td>2 (6)</td>
</tr>
<tr>
<td>Eating disorders</td>
<td>2 (6)</td>
</tr>
<tr>
<td>Substance use disorders</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Borderline</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Attention-deficit/hyperactivity disorder</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Autism spectrum disorder</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Various mental health conditions</td>
<td>11 (34)</td>
</tr>
</tbody>
</table>

Results

The Frequency of Design Studies Targeting People With Psychotic Symptoms and Features

Of the studies in the data set that focused on specific target populations, 18% (6/33) centered on psychosis and 9% (3/33) concentrated on schizophrenia. Since psychosis and schizophrenia are considered part of a broader defined spectrum of psychotic-related mental health conditions in which psychotic symptoms and features are prevalent, the total proportion of studies that primarily focused on people with psychotic symptoms and features was 27% (9/33). The largest group of studies, accounting for 34% (11/33) of the data set, did not focus on specific target populations and were classified as “various mental health conditions.” In this category, 12% (4/33) included psychotic symptoms and features as a target in addition to other mental health conditions (Table 1).

Discussion

This secondary data analysis revealed a notable emphasis on studies primarily targeting people with psychotic symptoms and features in mental health care design studies. This is noteworthy given the extensive range of mental health conditions in psychiatry, encompassing 21 categories according to the DSM-5-TR [8]. Although “schizophrenia spectrum and other psychotic disorders” constitutes only 4.67% (1/21) of these categories, 27% (9/33) of studies in our data set focused primarily on psychotic symptoms and features. This percentage is high, considering the lifetime prevalence of psychotic disorders is approximately 1% [10]. Another 12% (4/33) of the studies mention psychotic symptoms and features alongside or as a result of other mental health conditions. Although these studies did not focus primarily on people with psychotic symptoms and features, they have shown that much attention has been given to psychotic experiences in design studies. The substantial research focus on people with psychotic symptoms and features in design studies may be attributed to the limited progress in prognosis for severe cases despite extensive research and treatment efforts [11]. This may prompt designers and
Researchers to look for less conventional strategies to enforce novel promising solutions. Additionally, there is a growing call for attention to the subjective experience of psychotic symptoms and features in clinical care, as these vary from individual to individual (eg, [4-6]). Both factors underscore the urgency of involving people with firsthand experiences to capture the vividness of psychotic experiences in the design of innovative services and interventions, ultimately aiming to improve outcomes for service users.

Comparing the 9 studies that primarily focused on people with psychotic symptoms and features in this secondary data analysis to the results of the prior systematic review, we observed that 44% (4/9) demonstrated a high level of participant involvement in their design processes [2]. This is crucial for the development of new innovations because research shows psychotic symptoms and features can seem very different from a lived experience perspective compared to conventional psychiatric conceptualizations [6]. At the same time, the results stress the ongoing need to engage people with lived experience of psychotic symptoms and features in design studies, as more than half of the studies did not show the substantial involvement that would be expected of design processes that aim to tailor innovations to the needs of the target group. Consequently, we recommend future design studies targeting people with psychotic symptoms and features to adopt the co-design methodology, as co-design shows the highest participant involvement levels in mental health care design studies [2]. Furthermore, researchers are encouraged to use the participation matrix [12] alongside co-design to make intentional methodological decisions regarding the phases and roles in which people with lived experience are involved. To prevent tokenism and cooptation in design processes, researchers and designers are recommended to systematically coreflect with people with lived experience, exploring the roles played and distilling benefits and challenges from both perspectives.

Acknowledgments

We appreciate the financial support of the FAITH Research Consortium, GGZ-VS University of Applied Science, as well as from the NHL Stenden University of Applied Sciences PhD program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Categorization of studies based on their target population.

References


Abbreviations

DSM-5-TR: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision

Edited by J Torous; submitted 01.11.23; peer-reviewed by W Yassine, C Lim, A Hudon; comments to author 24.11.23; revised version received 11.12.23; accepted 13.12.23; published 09.01.24.

Please cite as:

Veldmeijer L, Terlouw G, Van Os J, Van ’t Veer J, Boonstra N
The Frequency of Design Studies Targeting People With Psychotic Symptoms and Features in Mental Health Care Innovation: Secondary Analysis of a Systematic Review
JMIR Ment Health 2024;11:e54202
URL: https://mental.jmir.org/2024/1/e54202
doi: 10.2196/54202
PMID:38194249

©Lars Veldmeijer, Gijs Terlouw, Jim Van Os, Job Van ’t Veer, Nynke Boonstra. Originally published in JMIR Mental Health (https://mental.jmir.org), 09.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on https://mental.jmir.org/, as well as this copyright and license information must be included.