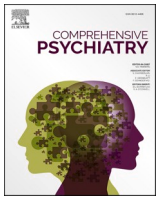
Comprehensive Psychiatry 132 (2024) 152467

Contents lists available at ScienceDirect

Comprehensive Psychiatry

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



When a test is more than just a test: Findings from patient interviews and survey in the trial of a technology to measure antidepressant medication response (the PReDicT Trial)

Susan Brown [a,*], Cornelia Ploeger [b], Boliang Guo [c], Juliana J. Petersen [b], Amy C. Beckenstrom [d], Michael Browning [d,e,f,g], Gerard R. Dawson [d,e], Jürgen Deckert [h], Rebecca Dias [d], Colin T. Dourish [d,e], Philip Gorwood [i,j], Jonathan Kingslake [d,e], Andreas Menke [k,l], Victor Perez Sola [m,n], Andreas Reif [o,p], Henricus Ruhe [q,r], Judit Simon [f,s], Michael Stäblein [o], Anneke van Schaik [t], Dick J. Veltman [u], Richard Morriss [a,c], the PReDicT Group

[a] NIHR MindTech Med Tech Co-operative, University of Nottingham, Nottingham, UK
[b] Institute of General Practice, Goethe-University Frankfurt, Frankfurt am Main, Germany
[c] NIHR ARC East Midlands, University of Nottingham, Nottingham, UK
[d] P1vital Products Limited, Howbery Park, Wallingford, UK
[e] P1vital Limited, Howbery Park, Wallingford, UK
[f] Department of Psychiatry, University of Oxford, Oxford, UK
[g] Oxford Health NHS Trust, Oxford, UK
[h] Department of Psychiatry, Psychosomatics and Psychotherapy, Center of Mental Health, University Hospital of Würzburg, Würzburg, Germany
[i] Université Paris Cité, Institute of Psychiatry and Neuroscience of Paris (IPNP), INSERM U1266, Paris, France
[j] GHU Paris Psychiatrie et Neurosciences, Hôpital Sainte Anne, Paris, France
[k] Medical Park Chiemseeblick, Department of Psychosomatic Medicine and Psychotherapy, Rasthausstr. 25, 83233 Bernau am Chiemsee, Germany
[l] Department of Psychiatry and Psychotherapy, University Hospital, Ludwig Maximilian University of Munich, Nussbaumstr. 7, 80336 Munich, Germany
[m] Hospital del Mar Medical Research Institute, IMIM, Barcelona, Spain
[n] Centro de Investigación Biomédica en Red (CIBERSAM), Madrid, Spain
[o] Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt – Goethe University, Frankfurt am Main, Germany
[p] Fraunhofer Institute for Translational Medicine and Pharmacology ITMP, Theodor-Stern-Kai 7, 60596 Frankfurt am Main, Germany
[q] Department of Psychiatry, Radboudumc, Nijmegen, the Netherlands
[r] Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, the Netherlands
[s] Department of Health Economics, Center for Public Health, Medical University of Vienna, Vienna, Austria
[t] Department of Psychiatry, Amsterdam Public Health Research Institute, Amsterdam UMC, Amsterdam, the Netherlands
[u] Department of Psychiatry, Amsterdam UMC, Amsterdam Neuroscience, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Depression
PRIMARY CARE
Digital Technology
Qualitative
Measurement reactivity

ABSTRACT

Background: A RCT of a novel intervention to detect antidepressant medication response (the PReDicT Test) took place in five European countries, accompanied by a nested study of its acceptability and implementation presented here. The RCT results indicated no effect of the intervention on depression at 8 weeks (primary outcome), although effects on anxiety at 8 weeks and functioning at 24 weeks were found.
Methods: The nested study used mixed methods. The aim was to explore patient experiences of the Test including acceptability and implementation, to inform its use within care. A bespoke survey was completed by trial participants in five countries ($n = 778$) at week 8. Semi-structured interviews were carried out in two countries soon after week 8 (UK $n = 22$, Germany $n = 20$). Quantitative data was analysed descriptively; for qualitative data, thematic analysis was carried out using a framework approach. Results of the two datasets were interrogated together.
Outcomes: Survey results showed the intervention was well received, with a majority of participants indicating they would use it again, and it gave them helpful extra information; a small minority indicated the Test made them feel worse. Qualitative data showed the Test had unexpected properties, including: instigating a process of

* Corresponding author.
E-mail address: sue.brown@nottingham.ac.uk (S. Brown).

https://doi.org/10.1016/j.comppsych.2024.152467

Available online 7 March 2024

reflection, giving participants feedback on progress and new understanding about their illness, and making participants feel supported and more engaged in treatment.
Interpretation: The qualitative and quantitative results are generally consistent. The Test's unexpected properties may explain why the RCT showed little effect, as properties were experienced across both trial arms. Beyond the RCT, the qualitative data sheds light on measurement reactivity, i.e., how measurements of depression can impact patients.

1. Introduction

The introduction of technology into healthcare settings is a complex intervention, benefitting from the collection of quantitative and qualitative data to understand barriers and facilitators to their implementation and impact, including unanticipated and unintended consequences [1–5]. By doing this alongside randomised controlled trials (RCTs) exploring the effectiveness of such technology, interventions can be optimised and refined before they are implemented into clinical practice.

The PReDicT (Predicting Response to Depression Treatment) Test, hereafter referred to as 'the Test', is comprised of the Facial Emotion Recognition Test (FERT) [6] and the Quick Inventory of Depression Symptoms Short Form [SF] (QIDS) 16 [7,8]. It is a new digital technology developed to detect antidepressant medication response at a timepoint that is earlier than usual in a person's care. The Test detects a change in a person's negative emotional bias [9] within 7—10 days of beginning a course of antidepressants. Usual clinical methods take 4—6 weeks, suggesting the Test might significantly reduce the time taken to ascertain response [6]. A RCT was designed to determine whether use of the Test in prescribing decisions led to quicker recovery from depression [10], and examine cost effectiveness. The RCT incorporated a nested mixed methods sub-study (presented here) to capture the experience of using the Test and to investigate the wider factors around implementation, context and acceptability that might affect outcomes and future implementation in clinical practice. Trial results showed that compared to usual care, the Test was not associated with a better response in depression symptoms at week 8 (primary outcome) but anxiety at week 8 and functional outcomes at week 24 (secondary outcomes) were improved [11]. There were also no differences in clinical outcomes by country. Health economic outcomes will be described in a forthcoming paper. The clinical results generated questions around the efficacy of the Test and whether unanticipated factors might have impacted on trial outcomes.

This paper presents results from the nested mixed method sub-study. The sub-study explored both patient and clinician perspectives, but only results from the patients' perspective are discussed here. The aim of this sub-study was to explore the acceptability and implementation of the Test, and explore patient experiences of the Test to inform and refine its use within care.

2. Method

2.1. Overarching design

2.1.1. Design

A nested mixed method study was designed to explore the implementation and acceptability of the Test, how it impacted on care including the experience of delivering and receiving it, and unintended consequences such as changing the doctor—patient relationship. The study was informed by theories from implementation science, in particular normalisation process theory (NPT) [12] and the work of Sekhon et al. [13].

The mixed methods comprised surveys and semi-structured interviews, which were used to explore both the clinician and patient side of the doctor—patient dyad. Mixed methods were selected to allow the potential to identify trends across all participants, but also the potential to uncover nuance within interviews, in-depth experiential insight, and to uncover material that had not been anticipated. As the RCT took place in five countries (France, Germany, Netherlands, Spain and UK), survey methods were used across all countries, however this was not feasible for the qualitative interviews, which took place in the UK and Germany only, chosen for their contrasting healthcare systems and potential for comparison (UK: free at the point of delivery through the National Health Service (NHS) mainly in primary care; Germany: insurance-based system, in primary and secondary care, with greater likelihood of a psychiatrist being involved). In this paper the survey results are utilised to explore the generalisability of the qualitative observation across the whole sample including all the countries in the study.

2.1.2. The trial and trial regimen

The trial protocol [10] provides an overview of the main RCT design. The RCT was carried out within primary care in the UK, and in primary and secondary care in all other countries. Participants were aged 18–70 years with a primary depression starting treatment with a SSRI antidepressant (not fluoxetine). In both arms they completed the Test according to the set regime (see Fig. 1). The Test comprised of two elements, the Facial Emotion Recognition Task (FERT, a computerised test of facial expression recognition) [9] alongside Quick Inventory of Depression Symptoms, Short Form [SF] 16 [7,8], (QIDS, a self-rated measure of depression). Participants received a weekly alert (via text or email) to login to the PReDicT site and complete the Test – either QIDS alone or alongside FERT – indicated for that day. Participants could complete the Test on their technology of choice (smartphone, tablet, laptop, desktop computer) at a time of their choosing. Prescribers (GPs or psychiatrists) only received Test results for participants in the treatment arm. For participants in the treatment arm, results were sent to their prescriber at week 1, for treatment to be adjusted (or not) accordingly. Based on the outcome of the Test, the prescriber decided whether to adjust treatment usually after discussion with the participant. If a change in treatment was initiated at week 1, participants completed an additional FERT at week 2 (alongside the routine QIDS), with Test results sent to the GP, for use in adjusting treatment if required. The main trial outcome was measured at week 8, and then

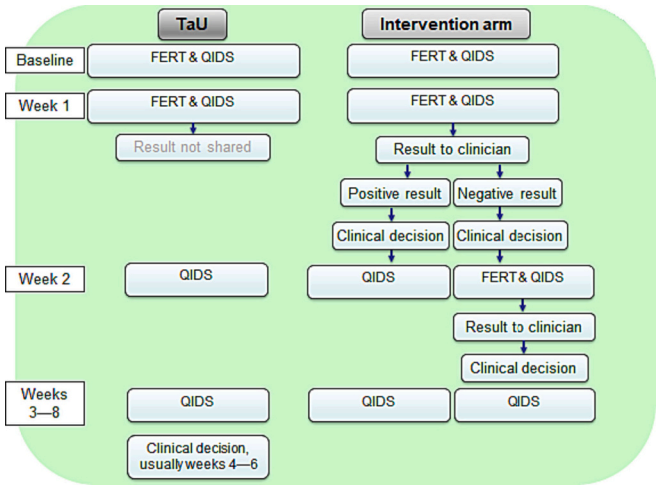


Fig. 1. Trial Test regime for participants.

every 4 weeks until week 48.

2.1.3. Ethics

All participants provided written informed consent. Ethics approval was obtained from the National Research Ethics Service committee, North East York (16/NE/0095), Ile de France Ethics Committee (MDPTRIAL/MM/2016-AO1054–47), Medisch Ethische Toetsingcommissie VU Medisch Centrum (2016.294 NL58027.029.16), CEIC Par de Salut Mar (2016/6795/I), Ethik Komission der Universitat Würzburg (117/16-sc) and Ethik Kommission des Fachbereichs Medizin, Universitätsklinikum der Goethe Universität Frankfurt am Main (34/17B).

2.1.4. Translation of data collection instruments and transcripts

All data collection instruments (surveys, interview schedules) were translated and back-translated, into the language of the country in which they were deployed. Back-translation checked the accuracy and consistency of meaning across languages.

In relation to translation of German transcripts to English, for storage and checking purposes, four interview schedules were translated and back-translated to check for accuracy of the translation process itself. Once that was completed and checked, translation of the remaining transcripts (up to half of the total conducted in Germany) was carried out by the same translator.

2.2. Patient Acceptability Questionnaire (PAQ) survey methods

2.2.1. PAQ design and data collection

This 16-item questionnaire (Appendix 1) was created for the study to explore the acceptability of the Test to participants, including whether they considered it a positive improvement to care. It drew on NPT [12], (which addresses key concepts such as coherence of the intervention, sense making e.g. whether participants understood the purpose of the Test; and appraisal work e.g. potential benefits of use and future intention) and concepts drawn from Sekhon et al. [13] (which addresses concepts around acceptability such as ease of use, willingness to use, and convenience). The PAQ was completed by participants online at week 8 at the same time as trial outcome measures [9]. Two near-identical versions of the questionnaire were created for the two different trial arms, with different wording for question 6 only. Question 6 relates to the role the Test played in treatment, and wording was different because participants were not necessarily blinded to their treatment arm allocation, with some participants having awareness of whether GPs changed their treatment due to Test results.

2.2.2. PAQ: Analysis of survey data

The response with percentages on each PAQ item was categorized into negative (strongly disagree, disagree), neutral (neither agree nor disagree) and positive (strongly agree or agree). The following were explored because of their potential influence on use of technology or depression outcomes [14,15]: country, treatment arm, age, gender, education, income, employment status, level of IT proficiency, living alone or with others, number of previous episodes of depression, and family history of depression. Country influence on the response for each item was explored first by means of multilevel modelling and showed no significant country level variance. Subsequently, a single level multinomial logistic regression with neutral category as reference was performed to explore the influence of each of the above variables on each PAQ item. To explore each influential factor's effect on response to each PAQ item, univariate models were first performed, with all minimization factors (gender, age, and baseline depression severity) adjusted as regression covariates. Then a model including all influential factors with severity of depression additionally adjusted as a covariate were performed to investigate the effects of each influential factor. Finally, the model including all covariates was rerun with a backward approach to select covariate(s) that have the most significant influential effects. Relative risk ratio (RRR) and its 95% confidence interval (95%CI) were presented as the effect estimates for each risk factor's influence. STATA 15 was used to conduct the data analysis. This approach was taken as several factors could influence elements of acceptability of the Test, and such exploration would usefully highlight whether certain groups have a more or less positive experience.

2.3. Qualitative semi-structured interview methods

2.3.1. Interview schedule design and data collection

Qualitative methods were selected to enable exploration of predefined areas, whilst allowing room for the unexpected. A topic guide (Appendix 2) was created for the study, also informed by the NPT constructs (coherence, sense-making, operational work and appraisal work) [12], topics relating to acceptability [13] (e.g. convenience, ease of use, fit within daily life), and open questions around how the Test shaped participants' experience of care, including compared to previous experiences. Additional questions included: 'Did you get any information from doing the Tests' (RM observed in previous studies that participants sometimes inferred unexpected information); challenges relating to recruitment, after initial difficulties in study recruitment. Topic guides evolved iteratively as the study progressed, incorporating unexpected areas that emerged in earlier interviews deemed in need of further exploration; changes were incorporated across both countries. In total, two prompts were added (see Appendix 2, added prompts in italics), after participants reported experiences warranting exploration.

Interviews were carried out either face to face or via telephone between November 2016 and February 2019 by two experienced qualitative interviewers (one with a sociological background and one with a psychology background) and native speakers in UK and Germany respectively, conducted in English and German respectively. Interviews were audio recorded and transcribed verbatim, then transcripts were checked and anonymised, and stored and coded using MaxQDA software.

2.3.2. Interviewee sample and recruitment process

Maximum variation sampling was used, to recruit participants with a range of different characteristics including age, gender, ethnicity, and treatment arm. Attempts were also made to recruit those who reported low satisfaction with the Test (via their scores on the PAQ, expressing any one of the following sentiments: didn't like doing the Test; weren't happy to have Test guide treatment; would not use it again; Test made them feel worse). Table 1 sets out sampling characteristics of the participants interviewed, which reflect the age and gender characteristics of

Table 1
Interviewee sample.

	UK	Germany
Total	22	20
Gender		
Male	7	8
Female	15	12
Age		
18–30	9	12
31–45	6	4
46–60	6	3
60+	1	1
Ethnicity		
White	22	16
Black or African American		
Other	0	1
	0	3
Trial Arm		
PReDicT	15	16
TAU	7	4
PAQ responses		
At least one negative	6	5
None negative		
	16	15

the overall RCT.

Recruitment involved the study team identifying participants who completed their week 8 outcome measures. Researchers in UK and Germany identified potential participants incorporating a range of characteristics (see above) and contacted site-specific staff to check whether consent to be contacted for interview had been granted on the original consent form. If so, study staff would send participants a Participant Information Sheet and Interview Consent Form, and on receipt of a positive response would pass on contact details (and signed consent forms) to the researcher to arrange an interview.

2.3.3. Interviews: Analysis of interview data

Thematic analysis was carried out using a framework approach [16], a structured approach that tracked the development of themes whilst maintaining a close connection to the data that fell within them. A framework approach offers flexibility to focus on both a priori and emergent topics of interest [17,18]. It enabled structured team-working and data management, with analysis on each country's data using a common approach. Elements of constant comparative method [19,20] were used to refine understanding (e.g. negative cases [19,20] were used to challenge and further refine themes, ensuring findings accurately reflected the data).

Analysis was carried out initially by the interviewers in their native language; there was an iterative process of initial coding, with a shared coding framework across the two countries developed through frequent discussions between the UK and German interviewers in consultation with senior researchers who were also clinicians in psychiatry and primary care. The developed framework was then used to code all transcripts. Once coded, content was further interrogated and a range of themes were generated, both inductively and deductively (i.e. shaped by the areas of questioning) [21]. Our framework approach followed a similar analytic path to Braun and Clarke's [22] utilising steps associated with data management (including familiarisation, constructing initial thematic framework, indexing and sorting) and then steps associated with abstraction and interpretation (including categorising and classifying, constructing typologies, mapping linkages) [16]. The highly structured process and outputs facilitated a group-approach to the analysis that allowed cross-checking between codes, themes, respondents and other forms of grouping.

The themes identified related to significant clusters of meaning, that were distinct and prevalent to varying degrees. Not all participants' experiences fell within them, but they were prevalent enough, and had sufficient internal coherence (i.e. the accounts underlying them are similar to each other) to be identified as of significance.

Coding and analysis discussions were held by email and telephone in English. A final analysis discussion was held by telephone with a bilingual interpreter to allow for full and detailed relaying and discussion of findings in each country's native language.

3. Results

3.1. Survey

3.1.1. Sample

The trial had a total cohort of 913 participants, and within that cohort 778 participants (85·2%) completed the PAQ. Table 2 shows there were no important differences between the two treatment arms by any of the variables used in the analysis.

3.1.2. Responses to survey

Fig. 2 shows the responses to the PAQ. There were no statistical differences across any of the questions between countries (data not shown) or treatment arm. Overall, the data showed a positive experience of completing the Test. Acceptability of the Test was high: over 80% agreed or strongly agreed that the Test was easy to use (90·6%, $n = 705$) and understood the purpose of doing it (88·8%, $n = 691$), and less than

Table 2
Baseline demographic and clinical details of patients.

		PReDicT Arm ($n = 460$)	TaU Arm ($n = 453$)
Country n(%)			
	France	39 (8%)	37 (8%)
	Germany	63 (15%)	67 (14%)
	Spain	82 (18%)	82 (18%)
	The Netherlands	28 (6%)	26 (6%)
	UK	248 (53%)	241 (54%)
Age mean(sd)		38.7 (13.53)	39.21 (14)
Sex n(%)			
	Female	285 (62%)	282 (62%)
	Male	175 (38%)	171 (38%)
Ethnicity[a] n(%)			
	White	374 (90%)	374 (89%)
	Ethnic Minority	47 (10%)	42 (11%)
Years of Education mean (SD)		14.14 (3.66)	14.10 (3.47)
Recruited from Primary or Secondary Care			
	Primary	359 (78%)	345 (76%)
	Secondary	101 (22%)	108 (24%)
Employment Status			
	Full or part time work or student	165 (36%)	152 (34%)
	Unemployed, sickness leave or retired	295 (64%)	301 (66%)
Living Situation			
	Living with spouse or partner	265 (58%)	244 (54%)
	Not living with spouse or partner	195 (42%)	209 (46%)
Present or past relationships			
	Ever been married, lived with partner or had children	319 (69%)	318 (70%)
	Never married, lived with partner or had children	141 (31%)	135 (30%)
Family History of Depression n(%)			
	No	225 (49%)	215 (47%)
	Yes	235 (51%)	238 (53%)
QIDS-SR-16 mean (SD)		15.67 (4.53)	15.65 (4.19)
MADRS mean (SD)		28.04 (7.45)	28.07 (7.08)
GAD-7 mean (SD)		13.59 (4.92)	13.77 (4.85)
SAS-SR mean (SD)		63.06 (11.16)	62.88 (11.28)

[a] A local ethical requirement prevented collection of data on ethnicity from patients in France. QIDS-SR-16; Quick Inventory of Depressive Symptoms, 16 item self-report version (score range 0–27). MADRS; Montgomery–Åsberg Depression Rating Scale (score range 0–60). GAD-7; Generalised Anxiety Disorder Assessment, seven-item version (score range 0–21). SAS-SR; Social Adjustment Scale, self-report screener form, T-score (note a higher score indicates greater impairment, score range 38–90). Scales are reported as mean (SD).

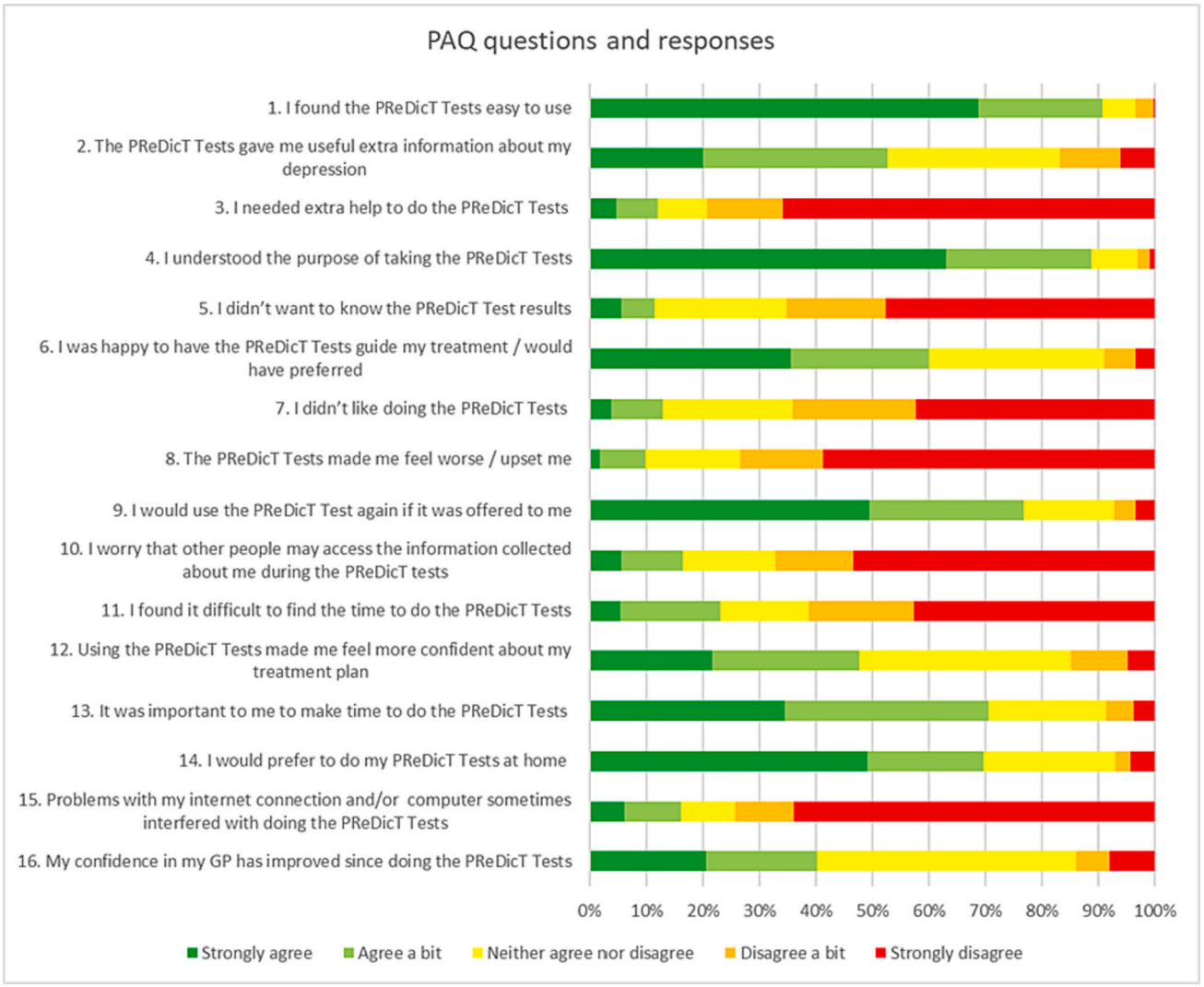


Fig. 2. Responses to survey.

15% agreed or strongly agreed that they did not want to know the Test result (11·3%, $n = 88$) or did not like doing the Test (12·9%, $n = 100$). The PAQ found that 9·8% ($n = 76$) of the sample agreed or strongly agreed that the Test made them feel worse or upset them.

The majority of participants agreed that there was added value in taking the Test. Hence 52·7% ($n = 410$) participants agreed or strongly agreed that the Test gave useful extra information about their depression. Among those allocated to the PReDicT arm, 78·4% ($n = 305$) agreed or strongly agreed that they 'were happy' for the Test to guide their treatment; among those in the TaU arm, 41·7% ($n = 162$) agreed or strongly agreed that they 'would have preferred' that the Test guided their treatment. Although 61·3% ($n = 179$) agreed or strongly agreed it was difficult to find time to do the Test, 70·6% ($n = 549$) agreed or strongly agreed it was important to make time to do them, 76·7% ($n = 597$) would use the Test again, 47·6% ($n = 370$) were more confident about their treatment, and 40·2% ($n = 312$) had more confidence in their doctor after doing the Test.

Further analysis showed that there were only minor differences in the PAQ items by gender (females were less concerned that other people might access information about them), living alone (more likely to want to know test results, less likely to feel confident about the treatment plan using the Test), previous episodes of depression (more likely to understand the purpose of taking the Test), or higher level of IT proficiency and more years of education (both less likely to feel confident about the treatment plan using the Test (see Appendix 3 for full results).

3.2. *Patient interview results*

The interviews indicate that from a patient perspective, the Test had an impact on them that had not been anticipated by the investigators and was mostly the same across both countries. The themes suggest that whilst the Test was intended to act as a tool in detecting medication response, it had additional inherent properties that may have acted as an intervention. These effects were noted across both trial arms. Many of the effects relate to QIDS, though some to FERT as well (as specified below).

3.2.1. *Theme 1 – Catalyst for self-reflection*

For some, the Test served as a catalyst for self-reflection, a process of examining oneself that would not otherwise have taken place. People had to take time out of their day to do the Test, thus carving out time to consider their wellbeing, albeit through the lens of QIDS and/or FERT. The questions (QIDS, but also FERT) directed people's thoughts to specific areas which formed the basis of their reflection.

Theme 1
Catalyst for self-reflection

UK participant 14, TaU arm

It gives you that time to sit back and think about the week that you have had and why you are feeling certain ways even though the tests aren't hugely elaborative it kind of causes you to think more about it. Mentally it is not kind of open questions it is closed like did you go to sleep, how long did it take you to get to sleep, all the things like that I think even if you don't realise it sub-consciously you start thinking about other things some others might struggle with that but as you think that is linked to that and that is linked to that and that again makes it more of a medical illness rather than emotional.

German participant 05, Intervention arm

Participant: It was like a self-control so to speak, that is to think about myself and reflect on things: How am I doing? Has it gotten better? And simply to exchange information with myself. That was good.
Interviewer: Yes. To what extent were the tests helpful for you?
Participant: Just to review the whole thing and start reflecting on things: OK, how have I really been in the last few weeks? Have I noticed any improvement, any deterioration? What were the experiences that could possibly influence it? How can I improve on it? What can I avoid so that I do not get in a situation where things are worse, and simply to have an overview because sometimes experiences blur. I no longer had a feeling of time and the tests helped me to gain it back. I knew: OK, I am now in the last few weeks and I just have to reflect on it: Was it last week and not three weeks ago or so? That was just good for self-monitoring.

3.2.2. Theme 2 – Getting 'information' from the Test

Participants reported being able to infer different forms of information by taking the Test on a regular basis, organised into two sub-themes.

3.2.2.1. Sub-theme – Feedback. Many people could see their responses changing over time; the predominant pattern was gradual improvement, mostly through QIDS (e.g. noticing their response to items differed from previous weeks) but also through FERT (e.g. noticing they recognised more positive expressions). By seeing a change in their pattern of responses, participants received feedback relating to their progress; participants were able to infer positive progress, providing encouragement through tangible markers of improvement.

Theme 2
sub theme 1, Feedback

UK participant 15, TaU arm

Interviewer: and in terms of the QIDS the same thing, did they trigger any emotional responses while doing them?
Participant: I guess at the beginning it made you realise you know when you start antidepressants there can be a dip before you get better?
Interviewer: Yeah?
Participant: The QIDS made you realise the difference but in the same extent it also helped you realise the difference when you were getting better as well. Like have you been feeling restless oh yeah I have been feeling restless, but you also note that compared to last time I am a lot better. It is a nice way of tracking

German participant 14, Intervention arm

Interviewer: And what is your impression, how did the use of the test affect your treatment?
Participant: I would say it was positive. In the sense that I had small successes, if you like. In the sense of: You're better at recognizing the faces, recognizing the emotions. You can answer the questionnaires a bit more positively. And that itself was a small motivation to continue with the therapy. And-, exactly. So that had a rather positive effect.

3.2.2.2. Sub-theme – Psychoeducation. In both countries, regular use of the Test helped some participants to better understand their illness and its impact on them. Seeing the different domains of QIDS – and the range of responses within the domains – also provided insight into potential routes towards making improvements e.g. increasing levels of activities with friends, starting to slowly engage in their usual activities, trying to address changes in appetite. It also gave participants some understanding of how the illness affected them, for example on their ability to do everyday tasks, sleep changes, and numbing of emotions (this was the case in relation to FERT), and a sense of what 'normal' might look like and therefore 'goals' for their recovery. In this sense, participants experienced a form of psychoeducation i.e. they learned more about depression, its severity and how their therapeutic progress might be gauged.

Theme 2
sub theme 2, Psychoeducation

UK participant 9, Intervention arm

Interviewer: What information did you take from doing the tests because it sounds like you got information from it?

Participant: Absolutely and as I said I think it was more from the worded questionnaire tests.. because they are key areas where.. I think a clinical practitioner would be able to see and assess how somebody is coping or not, how they're improving.. but just by the nature of the questions, it makes you think about.. oh these are the areas that.. I'm not able to carry out, why am I not able to carry them out? Is there any way that I can break that activity up so, there were times when the thought of doing the housework at the weekend was just like 'oh god no I just can't do it' you know.. me and housework, that's top to bottom, every single room all done in a day… well, try and break it up into little pieces. But I think the.. worded questions make you think about you know other ways that you could help yourself

German participant 10, TaU arm

I think that the fact that the test deals so much with the consequences of the disease helped me to get to know the disease better. And that's why I can categorize the symptoms in a certain sense. So when I notice, when I come from work at noon and notice: Okay, I don't feel anything anymore. So I can't even say that I'm happy that the working day is over or anything else: "Okay, that's because of my depression". And not somehow because I overworked me at work or something. Yes I think it's reasonably understandable too (laughs). Yes so in that sense I think it helped quite well to diagnose this and separate it from other things. Fatigue or boredom or stress or anything else that could play into the whole thing. Yes.

3.2.3. Theme 3 – Enhanced treatment experience: Commitment, support and engagement

Many participants reported an enhanced treatment experience, arranged into three sub-themes.

3.2.3.1. Sub-theme – Feeling supported (predominantly UK). This subtheme was prevalent in the UK and expressed by one German participant. They felt more supported and closely monitored. Their weekly tests (mainly QIDS) were experienced as additional contacts with their primary care practice. The extra activity, that served to support their care, was valued particularly because starting antidepressants can be a frightening experience. There was a potential safety implication to this sense of extra support; some participants thought their practice would be alerted if they started to decline, and thus felt there was a safety net in place. By implication, no contact from their practice meant that they were doing OK.

Theme 3
sub theme 1, Feeling supported

UK participant 11, Intervention arm

I liked that there was that additional support and I kind of felt that if I did suddenly decline really badly, somebody would be watching and they would like 'ok we need to do something else now', so I actually found that very helpful and it was good that.. it kind of gave me a bit of a support system, 'cause I was like… I am filling this in and I know that I am not doing very well because I know that myself, but… as no one else has come to me and said 'you're not doing very well' I must be doing better than I think I am, or like… there was always that fall back of.. if it is not going well, I

(*continued on next page*)

Theme 3 (*continued*)

have this person that I can contact and… they're kind of making sure that what I am answering is.. you know, going up the scale rather than down, so I actually find it really helpful and really useful.

German participant 18, Intervention arm

Interviewer: Yes. Maybe you can describe what you mean by a sense of control?
Participant: I had the feeling that I was being dealt with. And I could express myself on the PC how I feel. […] I had the feeling that I was somehow making such progress. Or had a control with me.

3.2.3.2. Sub-theme – Being active in own healing. In both the UK and Germany, a substantial proportion reported that making time to do the Test meant they were playing an active part in the process of helping themselves to get better and were taking some responsibility in that process (improved self-motivation). A small number of participants reported feeling as though it created more of a process for their healing, i.e. a programme of actions in which they played a role (improved sense of control).

Theme 3
sub theme 2, Being active in own healing

UK participant 14, TaU arm

I think it is just because that you have to do it [Test] quite regularly so it is kind of functionally taking part in a study, which obviously I was but doing things like that makes you feel that yeah ok, I am actually on a process to start to get better, I have not just been sent away with tablets. I am doing something to get better and doing the tests and going through it.

German participant 13, Intervention arm

Interviewer: What was it like for you that the test was part of your therapy?
Participant: I actually thought it was pretty good. Because it gives you the feeling that you can do something more for yourself.
Interviewer: You can do something, maybe you can describe it a bit more precisely. What exactly do you mean by that?
Participant: I don't know how, you don't have direct influence, but you can contribute to how the therapy develops.

3.2.3.3. Sub-theme – Patients' engagement with antidepressant medication. In Germany, the Test encouraged a commitment by participants to taking their medication. They understood better why they were taking it, and that they may initially experience side-effects. Some continued medication despite not feeling any benefit from it, with the Test serving as a reminder of voluntary obligation to the prescriber. Within the UK, commitment to medication was mentioned less frequently, but did occur. The overall extra support of the Test and monitoring helped participants to feel more comfortable with their treatment.

Theme 3
sub theme 3, Patients' engagement with antidepressant medication

UK participant 9, Intervention arm

It is quite something taking medication particularly when you know it's going to be affecting the chemical balance inside of you.. You know, it's quite scary, I don't like taking medication anyway but I knew that I had to and.. to know that there was somebody, albeit through surveys watching me.. and assessing if I was going to be ok that was a comfort to me, particularly in those early weeks, and I found, for me, that.. that was a good thing and it certainly helped when I was experiencing those side-effects […]it was good to know that behind the scenes they were watching my progress. Whereas I think if I'd have maybe just been taking the medication and it's like well… "here's a months' worth here's two months' worth we'll review it again you know.. when you're coming to the end of your tablets that you have available..

(*continued on next column*)

Theme 3 (*continued*)

come back in and we'll review it then"… I would have felt quite alone and actually maybe a bit scary suffering the side-effects as well..

German participant 08, Intervention arm

Good. I mean, it may not have affected me in anyway. It may really have made no difference to me, but I felt like I was doing something. And I also felt like I had some extra help, almost, you know? That there was a little bit more pressure for me to hundred percent stay on the line of like taking medication, because these tests were relying on me taking my medication. Cause I'm very forgetful. I would forget to take a tablet or whatever. So, I think it was good, that I had that extra bit of almost support and pressure, to do these things. For me to stay on these courses of tablets. I think it definitely-. My psyche helped like, helped me kind of get myself sorted. But the tests themselves I don't think had any effect on me.

3.2.4. Theme 4 – Adverse experience of the Test

A small number of participants expressed negative experiences of the Test, organised into two sub-themes.

3.2.4.1. Sub-theme – Feeling worse. Across both countries, a small number of respondents indicated in the process of answering QIDS, their attention was brought to the severity of their situation. Within the UK, this included the topic of suicidality, although in Germany the stated severity was less serious. Where this occurred in the UK, this included drawing participants' attention to their inner state and causing them to dwell, or indeed to acknowledge feelings they hadn't previously acknowledged, unexpectedly realising how unwell they were.

Theme 4
sub theme 1, Feeling worse

UK participant 7, Intervention arm

Interviewer: in terms of actually doing the tests how did it make you feel while you were doing them?
Participant: Sometimes worse
Interviewer: Right
Participant: Yeah they just got me down I thought 'god do I really feel like that?'… and then another time I'd think 'oh thank god I've done it'
Interviewer: Yeah? …. so can you tell me a bit more about that bit? When you say sometimes it made you feel a bit down?
Participant: Just reading what it said.. um how fidgety I am, you don't notice it till you think 'oh I am a bit today aren't I' and moody and stuff like that. But I must admit, the suicide one when that came up it was never… I used to imagine myself.. not being here.. but *not enough to do anything*. So reading through them, it was easy but.. I'm blameful, you know and that's going down [through] the questions, and that's how I felt.

German participant 02, TaU arm

Interviewer: Did you also have negative feelings?
Participant: Sure. Occasionally. I guess that's normal. Now, not negative feelings about the test, negative feelings in general. When I've thought about what happened.

3.2.4.2. Sub-theme – Experiencing the Test as reductionist. A small number of participants in both countries indicated they felt the Test was reductionist reducing a complex phenomenon to a series of short 'blunt' questions. The questions did not relate to their sense of wellness, and replaced the prescriber asking them how they felt which was perceived to be a better indicator of treatment response. In two UK participants, perceived experience of medication was at odds with the Test results.

Theme 4
sub theme 2, Experiencing the Test as reductionist

UK participant 3, Intervention arm

(*continued on next page*)

Theme 4 (*continued*)

Interviewer: what did you think of the idea of the tests?

Participant: I thought it was a nice idea, I think truthfully it was flawed. I think that when I fired off my first set of answers to an emailed set of questions, I had been on top of the world. I hadn't been happy for years possibly decades. But within 3 or 4 days of the new medication I was on top of world, I really was, I was feeling happy, I didn't want to punch anyone, you know I wasn't get wound up with people's bad behaviour, I was on top of the world really, and the first set of results said "well no, this [antidepressant] isn't working" and you think "it is, it is working", so I have to say that my faith in the test was.. I wouldn't say irrevocably broken, but I did groan a bit next time the tests came through.

4. Discussion

In this large sample of participants seeking treatment for depression, the survey shows that the Test was highly acceptable, and the majority gained benefit in terms of information or confidence. Over 40% stated that they gained confidence in their treatment and/or their doctor from doing the Test and were willing for the Test to guide their treatment. Most found the time to do the Test and would do it again. There were no variables that had a profound impact on the acceptability of the Test; those that did have an impact were small in terms of acceptability and utility of the Test. Generalisations around the effects of age, IT proficiency, gender, education, and income on acceptability of the digital test were not supported. There were no differences by country.

The qualitative data highlighted reasons for both the benefits and adverse experiences of the Test with no perceived differences between the treatment arms. Not all participants experienced all themes, pointing to a heterogeneity of responses to the Test and its role in care. Unexpectedly the qualitative data identified unanticipated effects including that many participants engaged in an active healing process they attributed to the Test. The majority of effects relate to QIDS as opposed to FERT. The Test acted as a catalyst for self-reflection, providing feedback to the participant, including a better understanding of their depression and their response to treatment, and/or more commitment to the treatment such as tolerating side-effects and feeling more supported. It created more structure to the process of healing, due to more frequent quasi-contacts with their GP practice, and a sense of agency in their own healing. Knowing that the Test would feed into decisions relating to their prescribing (and whether their treatment was effective), participants became active agents in generating data that could be used and relied on by clinicians. These properties reflect a shift in care dynamic, where patients move from a role of receiving treatment advice and implementing it, to one where their ongoing activities work in tandem with the clinician. When discussing using the Test, participants frequently cited it as better than being prescribed antidepressants and being told to 'come back in a month', which felt more frightening and isolating. Not all themes were found in both countries. In the UK, participants were shown the technology by a nurse within the primary care practice, so the Test and practice were seen as connected. In Germany, this was done by a study leader, usually not within the practice, so the Test may have been seen as less closely connected with the practice or prescriber.

This study provides important data regarding the conceptualisation of measurement reactivity defined as 'where measurement results in changes in the people being measured' [23, p.454]. The phenomenon is well evidenced through systematic reviews and meta-analyses [24–28], although it is not definitively understood, with the very recent research highlighting the need for greater understanding of the mechanisms at work [27,28]. In the current study, the Test may have increased emotional self-awareness, which has previously been identified as potentially bringing about change in people experiencing depression [29]. Whilst few studies ask in-depth about the experience of completing questionnaires, our findings are consistent with other research [30–33], though may not generalise to measurement situations with a less clear clinical purpose. This study explores the experience of completing the Test in the context of participants expecting their results to be used by clinicians in deciding their treatment; participants made sure to dedicate time to do the Test and give it proper attention. Therefore, the results inform understanding of Measurement Based Care (the use of outcome monitoring to inform therapeutic decisions), which is associated with increased clinical remission rates, reduced depression severity, and improved medication adherence [34]. The findings have implications for the wider field, including potential relevance to other measures. Such measurement reactivity might apply to both self-rated measures of depression (or other mental health conditions) such as QIDS and also interviews such as the Hamilton Depression Rating Scale [35](HDRS-17) where there is a chance to discuss the nature of some of their experiences with an interviewer, even if such discussion is not in a directed way towards self-management or other treatment.

Feedback and self-monitoring are two tools that are understood to be active ingredients of behaviour change interventions [36] drawing on Carver and Scheier's [37] control theory. Feedback and self-monitoring have previously been associated with increased emotional awareness [29,38] and empowerment [39], consistent with the themes found in this study. The design of the Test may have inadvertently introduced these active components, contributing to measurement reactivity in both treatment arms, promoting wider understanding and behaviour change that contributed to higher than anticipated rates of recovery in both treatment arms [11]. These unexpected active components may have interacted or confounded the intended use of the Test in informing prescribing decisions. Both inflation of the treatment effect and confounding of the purpose of the Test may have reduced the chances of demonstrating differences in effectiveness between the treatment arms [27,28].

Of some concern was that some participants in the UK had unrealistic expectations that GPs would be informed of a worsening of mood or suicidal ideation by looking at their weekly Test results when in fact GPs were not given this information (except when QIDS revealed severe suicidal ideation). In both countries, some participants dwelled on how bad their depression was, became distressed by realising they had suicidal ideas or found the Test reductionist turning a complex experience into a too simplistic one. On the PAQ, the Test was disliked by around 13% of participants and about 10% were upset by it, further suggesting the need for clinical oversight of the use of the Test in guiding treatment of depression.

The study provides insight into potential effects that could influence trial outcomes. Findings show the value of this type of exploration among control groups alongside intervention groups to identify unexpected effects. Whilst we don't advocate interviewing large numbers of control participants without good reason, not conducting any interviews with this group closes off the possibility for this type of insight. More trials, especially of complex interventions such as this, would benefit from qualitative exploration of patients' experiences to aid the interpretation of findings. This is especially needed to interpret complex intervention trials [2] where there are several facets to the intervention and changes to usual care, as was the case in the PReDiCT study. In clinical care, the completion of measures of depression at home followed by the opportunity for decision-making with a prescriber or therapist might provide the opportunity for more active involvement of the patient in both decision-making with their health professional and self-management with the potential for greater improvement in their depression.

4.1. Strengths and limitations

It is helpful to use Lincoln and Guba's [40] four criteria to reflect on trustworthiness of the findings (credibility, transferability, dependability, confirmability). Themes were predominantly found in both countries, demonstrating generalisability and that potentially inherent properties were being identified (i.e. triangulation that enhances credibility, transferability and confirmability) [41]. The process of analysis was rigorous and structured, following an iterative process where

constructs were identified and discussed as interpretation proceeded, retaining an auditable link between data and findings that enhances dependability [42]. Reflection upon researcher expectations throughout helped to highlight that the main findings were unexpected. Negative case analysis was used to challenge and refine interpretation, contributing to credibility. Analysis was conducted by separate researchers in their own native language with use of translation services during key discussions to ensure meaning identified within the data was understood by both teams (enhancing credibility and dependability, as findings were not unique to a single researcher or body of data). Findings potentially fit with quantitative data from both the PAQ (i.e. methodological triangulation) and trial outcomes showing some consistency in the findings. The thick description [43] that aids understanding of context and transferability of findings is limited due to the need for brevity, however sufficient nuance and explanation is provided to assist with understanding potential transferability to other situations.

Researchers were unable to carry out respondent validation (i.e. discussing findings with participants to check researchers' interpretation, enhancing credibility) due to a lack of time. Due to cost, only half of German transcripts could be fully translated but transcripts were coded in the native language with the same process of analysis and translated coding scheme. Some of the findings on feedback and self-monitoring may have been magnified by being part of a RCT, which provided structure and additional interaction with research staff. However, the staff were not involved in decision-making about treatment, and data collection was self-reported and remote after baseline assessment and consent. The survey and qualitative interviews were conducted with participants who completed follow up data; those who did not complete follow up may have been more dissatisfied or distressed with the Test. The survey also did not differentiate effects of the QIDS and FERT although the qualitative data was able to do so.

5. Conclusion

The in-depth experience of participants provides potential explanation why the results of the RCT were unexpected. The data sheds light on possible additional mechanisms of action of integrating the Test within care beyond the actual test information. It is a useful demonstration of the value of qualitative research nested within a RCT.

PReDicT group membership

The PReDicT group consists of the named authors and also: Catherine J. Harmer, Guy M. Goodwin, Lisa Pearce Collins, Samantha Campbell, Garima Sharma, Mark Dziedzic, Hannah Alker, Eulalia Esquerra Tuñí, Luis Antonio Barboza Alvitez, Elena Martínez Prats, José Manuel Cruz Domenech, Roberto Mourelle, Jesús Pujol, Nataša Perić.

Funding and disclosure

This research was funded by the European Union (European Research Council, Horizon 2020, grant number 696802). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

ACB, RD and JK are employees of P1vital Products Ltd. (PPL), which owns the PReDicT algorithm and was the study sponsor. GD is an employee of P1vital Ltd. JK, CTD and GD own shares in PPL. GD, JK and CTD own shares in P1vital Ltd. During the conduct of the study MB was employed by, and owned shares in, PPL. MB declares grants from the MRC and Wellcome Trust. He has worked as a consultant for J&J, Boeringher and CHDR and has accepted travel funds from Lundbeck. JD reports grants from the DFG, BMBF and Vogel Foundation during the conduct of the study. He is Co-PI with BioVariance in a study financed by the Bavarian Secretary of Commerce. PG has received, over the last 5 years, fees for presentations at congresses or participation in scientific boards from Alcediag-Alcen, Angelini, GSK, Janssen, Lundbeck, Otsuka, SAGE and Servier. AM has received honoraria for lectures from Medice and Neuraxpharm. RM has received research funding from NIHR, MRC, Wellcome Trust for research and funding for research from Magstim and Electromedical Products Inc. He has received fees for serving on a data monitoring committee for Novartis plc. AR has received honoraria for lectures and/or advisory boards from Medice, Shire/Takeda, Janssen, SAGE/Biogen, COMPASS, Boehringer Ingelheim and LivaNova. HGR has received speaking fees from Lundbeck and Janssen and is coordinator of the Esketamine Nasal-spray Consortium Netherlands (ENC-NL) which is supported by an unrestricted educational grant from Janssen. SB, CP, JP, BG, AvS, JS, VPS, DJV, and MS declare no conflict of interest.

CRediT authorship contribution statement

Susan Brown: Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft. **Cornelia Ploeger:** Data curation, Formal analysis, Investigation, Project administration, Writing – original draft. **Boliang Guo:** Data curation, Formal analysis, Methodology, Validation, Writing – original draft. **Juliana J. Petersen:** Supervision, Validation, Writing – review & editing. **Amy C. Beckenstrom:** Project administration, Resources, Writing – review & editing. **Michael Browning:** Conceptualization, Funding acquisition, Writing – review & editing. **Gerard R. Dawson:** Conceptualization, Funding acquisition, Writing – review & editing. **Jürgen Deckert:** Funding acquisition, Resources, Writing – review & editing. **Rebecca Dias:** Funding acquisition, Project administration, Resources, Writing – review & editing. **Colin T. Dourish:** Conceptualization, Funding acquisition, Writing – review & editing. **Philip Gorwood:** Funding acquisition, Resources, Writing – review & editing. **Jonathan Kingslake:** Conceptualization, Funding acquisition, Resources, Writing – review & editing. **Andreas Menke:** Funding acquisition, Resources, Writing – review & editing. **Victor Perez Sola:** Funding acquisition, Resources, Writing – review & editing. **Andreas Reif:** Funding acquisition, Resources, Writing – review & editing. **Henricus Ruhe:** Funding acquisition, Resources, Writing – review & editing. **Judit Simon:** Conceptualization, Funding acquisition, Writing – review & editing. **Michael Stäblein:** Funding acquisition, Resources, Writing – review & editing. **Anneke van Schaik:** Funding acquisition, Resources, Writing – review & editing. **Dick J. Veltman:** Funding acquisition, Resources, Writing – review & editing. **Richard Morriss:** Conceptualization, Funding acquisition, Methodology, Supervision, Validation, Writing – original draft.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.comppsych.2024.152467.

References

[1] Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. Br. Med. J. 2000;321(7262):694–6.
[2] Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. Bmj-British Med. J. 2015:350.
[3] Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. Bmj-British Med. J. 2021:374.
[4] Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J. Med. Internet Res. 2017;19(11).
[5] Blandford A, Gibbs J, Newhouse N, Perski O, Singh A, Murray E. Seven lessons for interdisciplinary research on interactive digital health interventions. Digital. Health. 2018:4.

[6] Browning M, Kingslake J, Dourish CT, Goodwin GM, Harmer CJ, Dawson GR. Predicting treatment response to antidepressant medication using early changes in emotional processing. Eur. Neuropsychopharmacol. 2019;29(1):66–75.
[7] Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol. Psychiatry 2003;54(5):573–83.
[8] Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al. The inventory of depressive symptomatology, clinician rating (IDS-C) and self-report (IDS-SR), and the quick inventory of depressive symptomatology, clinician rating (QIDS-C) and self-report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol. Med. 2004;34(1):73–82.
[9] Browning M, Kingslake J, Dourish CT, Goodwin GM, Harmer CJ, Dawson GR. Predicting treatment response to antidepressant medication using early changes in emotional processing. Eur. Neuropsychopharmacol. 2019;29(1).
[10] Kingslake J, Dias R, Dawson GR, Simon J, Goodwin GM, Harmer CJ, et al. The effects of using the PReDicT test to guide the antidepressant treatment of depressed patients: study protocol for a randomised controlled trial. Trials. 2017;18.
[11] Browning M, Bilderbeck AC, Dias R, Dourish CT, Kingslake J, Deckert J, et al. The clinical effectiveness of using a predictive algorithm to guide antidepressant treatment in primary care (PReDicT): an open-label, randomised controlled trial. Neuropsychopharmacology. 2021;46:1307–14.
[12] Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, Dowrick C, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. BMC Med. 2010:8.
[13] Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. Bmc. Health Serv. Res. 2017:17.
[14] Trivedi M, Morris D, Pan J, Grannemann B, Rush A. What moderator characteristics are associated with better prognosis for depression? Neuropsychiatr. Dis. Treat. 2005;1(1):51–7.
[15] Hollon SD, Shelton RC, Wisniewski S, Warden D, Biggs MM, Friedman ES, et al. Presenting characteristics of depressed outpatients as a function of recurrence: preliminary findings from the STAR*D clinical trial. J. Psychiatr. Res. 2006;40(1):59–69.
[16] Spencer L, Ritchie J, O'Conor W, Morrell G, Ormston R. Analysis in practice. In: Ritchie J, Lewis J, McNaughton Nicholls C, Ormston R, editors. Qualitative research practice : A guide for social science students and researchers. Los Angeles: SAGE; 2014.
[17] Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess R, editors. Analyzing qualitative data. London: Routledge; 1994. p. 173–94.
[18] Parkinson S, Eatough V, Holmes J, Stapley E, Midgley N. Framework analysis: a worked example of a study exploring young people's experiences of depression. Qual. Res. Psychol. 2016;13(2):109–29.
[19] Glaser BG. The constant comparative method of qualitative analysis. Soc. Probl. 1965;12(4):436–45.
[20] Glaser BG, Strauss AL. The discovery of grounded theory : Strategies for qualitative research. First edition. ed.vol. x. New York: Aldine de Gruyter; 1967. p. 271. 22 cm p.
[21] Spencer L, Ritchie J, Ormston R, O'Connor W, Barnard M. Analysis: Principles and processes. In: Ritchie J, Lewis J, McNaughton Nicholls C, Ormston R, editors. Qualitative research practice : A guide for social science students and researchers. 2nd ed. Los Angeles: SAGE Publications; 2014.
[22] Braun V, Clarke V. Using thematic analysis in psychology. Qual. Res. Psychol. 2006;3(2):77–101.
[23] French DP, Sutton S. Reactivity of measurement in health psychology: how much of a problem is it? What can be done about it? Br. J. Health Psychol. 2010;15:453–68.
[24] Miles LM, Elbourne D, Farmer A, Gulliford M, Locock L, McCambridge J, et al. Bias due to MEasurement Reactions In Trials to improve health (MERIT): protocol for research to develop MRC guidance. Trials. 2018;19.
[25] Miles LM, Rodrigues AM, Sniehotta FF, French DP. Asking questions changes health -related behavior: an updated systematic review and meta -analysis. J. Clin. Epidemiol. 2020;123:59–68.
[26] Wilding S, Conner M, Sandberg T, Prestwich A, Lawton R, Wood C, et al. The question-behaviour effect: a theoretical and methodological review and meta-analysis. Eur. Rev. Soc. Psychol. 2016;27(1):196–230.
[27] French DP, Miles LM, Elbourne D, Farmer A, Gulliford M, Locock L, et al. Reducing bias in trials due to reactions to measurement: experts produced recommendations informed by evidence. J. Clin. Epidemiol. 2021;139:130–9.
[28] French DP, Miles LM, Elbourne D, Farmer A, Gulliford M, Locock L, et al. Reducing bias in trials from reactions to measurement: the MERIT study including developmental work and expert workshop. Health Technol. Assess. 2021;25(55):1-+.
[29] Kauer SD, Reid SC, Crooke AHD, Khor A, Hearps SJC, Jorm AF, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. J. Med. Internet Res. 2012;14(3):15–31.
[30] Midgley N, Isaacs D, Weitkamp K, Target M. The experience of adolescents participating in a randomised clinical trial in the field of mental health: a qualitative study. Trials. 2016;17.
[31] Carey MP, Morrison-Beedy D, Carey KB, Maisto SA, Gordon CM, Pedlow CT. Psychiatric outpatients report their experiences as participants in a randomized clinical trial. J. Nerv. Ment. Dis. 2001;189(5):299–306.
[32] Dowrick C, Leydon GM, McBride A, Howe A, Burgess H, Clarke P, et al. Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. Bmj-British Med. J. 2009:338.
[33] Jacomb PA, Jorm AF, Rodgers B, Korten AE, Henderson AS, Christensen H. Emotional response of participants to a mental health survey. Soc. Psychiatry Psychiatr. Epidemiol. 1999;34(2):80–4.
[34] Zhu M, Hong RH, Yang T, Yang XR, Wang X, Liu J, et al. The Efficacy of Measurement-Based Care for Depressive Disorders: Systematic Review and Meta-Analysis of Randomized Controlled Trials. J Clin Psychiatry 2021;82(5).
[35] Hamilton M. A rating scale for depression. J. Neurol. Neurosurg. Psychiatry 1960;23(1):56–62.
[36] Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a Meta-regression. Health Psychol. 2009;28(6):690–701.
[37] Carver C, Scheier M. Control theory: a useful conceptual framework for personality-social, clinical and health psychology. Psychol. Bull. 1982;92(1):111–35.
[38] Bastiaansen JA, Meurs M, Stelwagen R, Wunderink L, Schoevers RA, Wichers M, et al. Self-monitoring and personalized feedback based on the experiencing sampling method as a tool to boost depression treatment: a protocol of a pragmatic randomized controlled trial (ZELF-i). Bmc. Psychiatry. 2018:18.
[39] Simons CJP, Hartmann JA, Kramer I, Menne-Lothmann C, Hohn P, Van Bemmel AL, et al. Effects of momentary self-monitoring on empowerment in a randomized controlled trial in patients with depression. Eur Psychiatry 2015;30(8):900–6.
[40] Lincoln YS, Guba EG. Naturalistic inquiry. A.: Sage Publications; 1985. 416 ill ; 23cm. p.
[41] Stahl NA, King JR. Expanding approaches for research: understanding and using trustworthiness in qualitative research. J. Dev. Educ. 2020;44(1):26–8.
[42] Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. International journal of qualitative. Methods. 2017;16(1).
[43] Geertz C. The interpretation of cultures. Third edition. ed. New York, NY: Basic Books; 1973. p. 470.